Asymptotic Expansions of Processes
with Extreme Value Random Variable Innovations

by

Christopher David O'Neal

(Under the direction of William P. McCormick and Lynne Seymour)

Abstract

Recently there has been an interest in asymptotic expansions of the tail probabilities of a variety of processes that are ubiquitous in statistics. However, little to no work has been done when the AR(1) process is built upon extreme value random variables. This process appears when the distribution of the current maximum is dependent on the previous. The goal of this dissertation is to explore asymptotic expansions of tail probabilities on this topic, in particular using the Gumbel distribution. In each of the theoretical projects we build second-order expansions, many of which are improvements over the already known first-order ones. We also examine exactly when each of the expansions should and should not be used through simulation studies. Finally, we perform a data analysis in the extreme value theory setting on riverflow data, and as much as possible connect this same data set to the theoretical results.

Index words: Extreme value theory, Gumbel random variable, Taylor series, Streamflow, Autoregressive process, Asymptotic expansion, Regular variation, Markov chain

Asymptotic Expansions of Processes with Extreme Value Random Variable Innovations

by

Christopher David O'Neal

B.S., University of Georgia, 2007

M.S., University of Georgia, 2009

A Dissertation Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2012

Asymptotic Expansions of Processes with Extreme Value Random Variable Innovations

by

Christopher David O'Neal

Approved:

Major Professors: William P. McCormick
Lynne Seymour

Committee: Nicole Lazar
T.N. Sriram
Jaxk Reeves
Xiangrong Yin

Electronic Version Approved:

Dean's Name Here
Dean of the Graduate School
The University of Georgia
August 2012

Acknowledgements

First of all, I would like to say a huge thank you to Dr. William P. McCormick for guiding me on this fascinating project. He provided me with some great ideas to explore, along with plenty of helpful guidance and hints on the proofs. Furthermore, the two classes I took from him gave me a solid background on understanding how to write the proofs.

I would also like to thank Dr. Lynne Seymour for all her help on the data analysis portion of the dissertation. She provided me with a very interesting data set, some creative ideas, and all the help I needed to build the algorithms and graphs. The two programming classes I took from her taught me how to build computer code efficiently, and without this knowledge the data analysis and simulation studies in the dissertation would not have been possible.

In addition, I would like to thank my committee - Dr. Nicole Lazar, Dr. Jaxk Reeves, Dr. T. N. Sriram, and Dr. Xiangrong Yin - for reading my dissertation and for all their helpful suggestions. I also would like to say thank you to Dr. Yehua Li for serving on my committee when I gave my pre-dissertation defense.

Finally, my sincerest and deepest gratitude goes to my Mum, Liza, for all her lifelong love and support for me (and for teaching me fractions!), and my girlfriend, Sandra, for all her love and support for me during the final challenging year of graduate school.

Asymptotic Expansions of Processes with Extreme Value Random Variable Innovations

Christopher David O'Neal

Contents

List of Figures

List of Tables

Chapter 1 An Overview of the Dissertation

1.1 Introduction

Extreme value theory is the study of the behavior of observations that fall extremely far away from the mean of a distribution. Typically one studies asymptotic properties of the upper tail, or in other words the high percentiles. One could also perform analysis on the lower tail, which would be equivalent to studying the limiting behavior of the minimum values. For a basic, nontheoretical introduction to the ideas of extreme value theory, we recommend Coles (2001).

The literature is rich with results. Fisher and Tippett (1928) proved that the maximum of a sample followed one of three limiting distributions. Similar work was confirmed and studied in Gumbel (1958), Fréchet (1927), and Weibull (1951), after whom the three extreme value distributions were named. Balkema and de Haan (1974) and Pickands (1975) studied the distribution of the distance above a specified level, known as Peaks over Threshold (POT). Leadbetter *et al.* (1983) explored extremes of sequences of random variables which are not necessarily stationary.

As for applications of extreme value theory, many disciplines require the asymptotic results of high percentiles. Tail area approximations are important since they shed light on possible behaviors of extreme events and when to expect them to occur. We list the following references as examples relating to environmental statistics. Butler *et al.* (2007) studied trends in wave heights, in particular the North Sea. For earthquake magnitudes, see Caers *et al.* (1999) and Pisarenk and Sornette (2003). Examples of applications to ozone levels may be found in Smith and Shivey (1995). Finally, Smith (1999) studied extremes in rainfall at four different stations across the United States.

There are also applications to nonenvironmental disciplines. One example is in insurance risk, where companies need to study the probabilities of rare events or of large claims. See Asmussen (2001) for further discussion on this topic. For applications to queueing theory, see Borokov (1976), Norros (2003), and Whitt (2002).

This dissertation consists largely of asymptotic expansions of upper tail probabilities of various extreme value formulas. The distribution we focus on for the majority of the dissertation is the Gumbel, for two reasons. Firstly, the Gumbel is a special case of the generalized extreme value model and easier to work with than the Fréchet and Weibull distributions. Secondly, and more importantly, often a likelihood ratio test will allow one to reduce the generalized extreme value model to the Gumbel case, making interpretation a little easier. It is therefore of interest to have results on tail probabilities in such an event.

We are primarily interested in building an AR(1) process where the innovations are Gumbel random variables. This process would be a useful contribution to extreme value theory and also time series. The motivation for considering this process is that it may arise when the distribution of the current annual maximum is dependent on the previous year's maximum, and the random variables have been shown to satisfy the Gumbel distribution, possibly with covariates in the parameters. We consider a variety of cases that mainly revolve around the choice of weights. In certain situations, particular choices of weights could completely change the optimal approximation, and these instances will be discussed later.

There are several motivations for considering higher order expansions of tail probabilities. First, processes like those that will appear in the dissertation are ubiquitous in many disciplines, especially the ones mentioned above, and accurate estimations of their probabilities are needed. Second, further terms would likely provide more precise probabilities than a first-order alone by introducing correction terms. And third, while the literature contains numerous first-order results, much fewer exist on second-order, and even fewer to none on further terms. Therefore the results that follow in this dissertation will be valuable contributions to extreme value theory.

As examples of known results, we recommend Resnick (1986), Bingham *et al.* (1989), and Broniatowski and Fuchs (1995) for the first-order analysis. These texts cover topics ranging from convolutions to the subexponential setting. For second-order results, see Omey (1988), Geluk *et al.* (1997), and Barbe and McCormick (2005). The only paper we know of that contains higher order expansions is Barbe and McCormick (2009). The latter two papers cover approximations in the heavy tail distribution setting.

The dissertation takes the following path. In the next section, we review the distributions, formulas, definitions, and inequalities needed to understand the material to be discussed in later chapters. These statements will come into play in the proofs. Afterwards, a total of five projects relating to extreme value theory will be investigated.

The first project, discussed in Chapter 2, is a streamflow data analysis. For reasons to be discussed later, we shall condense the data into the seasonal maxima (and therefore four observations) per year, and then we will fit a generalized extreme value model through all four seasons, using time as a covariate. The results will come into play in future chapters as we use them to illustrate the later theoretical results.

The second project is covered in Chapters 3 and 4. We shall study the upper tail behavior of the AR(1) process with Gumbel innovations. Chapter 3 establishes the groundwork for the convolution of just two variables, and in Chapter 4 the results are extended to the possibly infinite series. Chapter 3 also provides an introduction to how the situation would be handled if we were working with variables from the Types II or III families. We close the project with some examples, a simulation study, and an application to the Peachtree Creek data.

Chapter 5 contains the third project, which can be viewed as a special case of the second in that it is a convolution, but where all weights are equal. One particular case would be finding the distribution of the sum of n Gumbel random variables. This situation, while common in practice, must be treated separately since the rules for deriving the asymptotics are very different from what we do in the previous project. The steps are rather involved, and so this topic deserves its own project. Again, we list several examples of the main result, as well as conduct a simulation study. Finally, we use the approximation on the winter observations from the streamflow data set.

In Chapter 6, the fourth project borrows ideas from the second and third, but introduces the possibility of ties occurring in the weights used in the convolution. There are actually several ways in which this may happen, but we derive the expansion for only one of those cases to give an idea of how the general problem would be solved. Some examples are provided, and we use as one of the examples an application to the Peachtree Creek again.

Lastly, the fifth project in Chapter 7 explores a different topic, the upper tail behavior of the convolution of weighted regularly varying random variables that are Markov chain dependent. Variables that are regularly varying fall into the Type II extreme value family, and our results are extensions of published ones, not all of which were necessarily chain dependent.

At the end of the dissertation, after the concluding remarks and the bibliography, a series of appendices contain supplemental material. This chapter includes second derivatives and Hessian matrices from Chapter 2, extra proofs from Chapter 4, and tables of numerical values from Chapters 4 and 5.

1.2 Useful Formulas

Before beginning the dissertation, we provide this short section of the formulas used most frequently during the data analysis and the proofs that follow.

Definition 1.1. *The Generalized Extreme Value distribution is defined as*

$$F(x) = \begin{cases} \exp\left\{-\left[1+\frac{\xi}{\sigma}(x-\mu)\right]^{-\frac{1}{\xi}}\right\}, 1+\frac{\xi(x-\mu)}{\sigma} > 0, & \xi \neq 0 \\ \exp\left\{-\exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\}, -\infty < x < \infty, & \xi = 0. \end{cases} \tag{1.1}$$

Here ξ is said to be the shape parameter, μ the location parameter, and $\sigma > 0$ the scale parameter. We abbreviate the distribution as GEV(ξ, μ, σ).

Depending on the choice of ξ, the generalized extreme value distribution falls into one of three families - the Gumbel, the Fréchet, or the Weibull. These three families are defined next.

Definition 1.2. *The Gumbel distribution is also the Type I Extreme Value Distribution. We denote it as Λ, and it is the limiting case as $\xi \to 0$. The distribution is*

$$\Lambda(x) = \exp\left\{-\exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\}, -\infty < x < \infty. \tag{1.2}$$

Note that $\Lambda = GEV(0, \mu, \sigma)$.

Definition 1.3. *The Fréchet distribution is also the Type II Extreme Value Distribution. We denote it as Φ, and it occurs when $\xi > 0$. The distribution is*

$$\Phi(x) = \exp\left\{-\left[1+\frac{\xi}{\sigma}(x-\mu)\right]^{-\frac{1}{\xi}}\right\}, x > \mu - \frac{\sigma}{\xi}. \tag{1.3}$$

Alternatively, for $\alpha > 0$ *we may instead define the Fréchet distribution using the sometimes more convenient formula*

$$\Phi_\alpha(x) = e^{-x^{-\alpha}}, x > 0. \tag{1.4}$$

In this case, $\Phi_\alpha = GEV(\frac{1}{\alpha}, 1, \frac{1}{\alpha})$.

Definition 1.4. *The Weibull distribution is also the Type III Extreme Value Distribution. We denote it as* Ψ*, and it occurs when* $\xi < 0$*. The distribution is*

$$\Psi(x) = \exp\left\{-\left[1 + \frac{\xi}{\sigma}(x - \mu)\right]^{-\frac{1}{\xi}}\right\}, x < \mu - \frac{\sigma}{\xi}. \tag{1.5}$$

Alternatively, for $\alpha > 0$ *we may instead define the Weibull distribution using the sometimes more convenient formula*

$$\Psi_\alpha(x) = \begin{cases} e^{-(-x)^\alpha}, & x < 0 \\ 1, & x \geq 0. \end{cases} \tag{1.6}$$

In this case, $\Psi_\alpha = GEV(-\frac{1}{\alpha}, -1, \frac{1}{\alpha})$.

Next, Euler's constant γ is defined by

$$\gamma = \lim_{n\to\infty}\left[\sum_{k=1}^{n}\frac{1}{k} - \log(n)\right] \cong 0.5772156649.... \tag{1.7}$$

Also known as the Euler-Mascheroni constant, γ appears regularly in the analysis of Gumbel random variables. In particular, if X is a standard Gumbel variable, then $E(X) = \gamma$. There are numerous published integrals and sums that equal γ; see Gradshteyn and Ryzhik (1980), Seo *et al.* (1997), and Choi and Seo (1998).

Now we define the gamma, digamma, and beta functions. The gamma integral will appear extensively in our two-term expansion analyses, while the beta function will arise when we turn our attention to the Weibull cases in Section 3.5. The digamma function appears in Chapter 6 when we discuss ties in weights of convolutions.

Definition 1.5. *For $\alpha > 0$, the gamma function $\Gamma(\alpha)$ is represented by the integral*

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt. \tag{1.8}$$

In particular, if n is a positive integer, then $\Gamma(n) = (n-1)!$. We also have the recursion $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, which holds for all real values of α except for 0 and the negative integers.

Definition 1.6. *For $\alpha > 0$, the digamma function $\Psi(\alpha)$ is represented by*

$$\Psi(\alpha) = \frac{d}{dx} \log \Gamma(x)\Big|_{x=\alpha} = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}. \tag{1.9}$$

Definition 1.7. *For $\alpha, \beta > 0$, the beta function $B(\alpha, \beta)$ is represented by*

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}. \tag{1.10}$$

Next, we state the equations for the mean, median, and variance of the generalized extreme value distribution. We need to use these formulas when conducting the data analysis in Chapter 2.

Definition 1.8. *If $X \sim GEV(\xi, \mu, \sigma)$, then the expected value is*

$$E(X) = \begin{cases} \mu + \sigma\left(\frac{\Gamma(1-\xi)-1}{\xi}\right), & \xi \in (-\infty, 1) \setminus \{0\} \\ \mu + \sigma\gamma, & \xi = 0 \\ \infty, & \xi \geq 1. \end{cases} \tag{1.11}$$

Definition 1.9. *If $X \sim GEV(\xi, \mu, \sigma)$, then the median is*

$$\mathrm{Med}(X) = \begin{cases} \mu + \sigma \left(\frac{(\log 2)^{-\xi} - 1}{\xi} \right), & \xi \neq 0 \\ \mu - \sigma \log(\log 2), & \xi = 0. \end{cases} \tag{1.12}$$

Definition 1.10. *If $X \sim GEV(\xi, \mu, \sigma)$, then the variance is*

$$\mathrm{Var}(X) = \begin{cases} \frac{\sigma^2}{\xi^2} \left[\Gamma(1 - 2\xi) - (\Gamma(1 - \xi))^2 \right], & \xi \in (-\infty, 0.5) \setminus \{0\} \\ \frac{\pi^2 \sigma^2}{6}, & \xi = 0 \\ \infty, & \xi \geq 0.5. \end{cases} \tag{1.13}$$

Having established the basic formulas, we also need to state definitions associated with regular variation, and how they are connected to the generalized extreme value distribution. These statements will be used in the expansions of the Fréchet mixture in Section 3.4, as well as the regular varying variables with Markov chains in Chapter 7.

Definition 1.11. *The tail distribution $\overline{F} = 1 - F$ is said to be regularly varying at ∞ with index $-\alpha$, denoted as $RV_{-\alpha}$, if for $x > 0$*

$$\lim_{t \to \infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-\alpha}. \tag{1.14}$$

Definition 1.12. *A function $L(\cdot)$ is slowly varying at ∞ if for any $c > 0$,*

$$L\left(\frac{x}{c}\right) \Big/ L(x) \to 0 \quad \text{as} \quad x \to \infty.$$

Finally, we state some theorems and inequalities that will be used in the proofs and derivations to come in the dissertation.

Definition 1.13. (Fubini's Theorem) *Let A and B be measure spaces and f be $A \times B$ measurable. Further suppose that*

$$\int_{A\times B} |f(x,y)| d(x,y) < \infty.$$

Then

$$\int_A \int_B f(x,y) dx dy = \int_B \int_A f(x,y) dy dx.$$

Definition 1.14. (Chapman-Kolmogorov Equation - see Resnick (1992)) *If $P_{i,j}(m)$ denotes the Markov chain transition probability of moving from state i to state j in m steps, or $P_{i,j}(m) = P(X_m = j | X_o = i)$, then $P_{i,j}(m+n) = \sum_k P_{i,k}(m) P_{k,j}(n)$.*

Definition 1.15. (Markov's Inequality) *Suppose X is a random variable and $a > 0$. Then*

$$P(|X| > a) \leq \frac{E(|X|)}{a}. \tag{1.15}$$

Definition 1.16. (Chernoff's Inequality) *Suppose X is a random variable and $y > 0$. Then for any $a > 0$,*

$$P(|X| > y) \leq E\left(e^{a|X|}\right) e^{-ay}. \tag{1.16}$$

Definition 1.17. (Jensen's Inequality) *Suppose X is a random variable and that $f : R \to R$ is a convex function. Further suppose that $E(|X|) < \infty$ and $E(|f(X)|)$. Then*

$$f(E(X)) \leq E(f(X)). \tag{1.17}$$

Definition 1.18. (Hölder's Inequality - see Resnick (1999)) *Suppose p, q satisfy*

$$p > 1, q > 1, \qquad \text{and} \qquad \frac{1}{p} + \frac{1}{q} = 1,$$

and that X, Y are random variables satisfying

$$E(|X|^p) < \infty \quad \textit{and} \quad E(|Y|^q) < \infty.$$

Then

$$|E(XY)| \le E(|XY|) \le (E|X|^p)^{1/p}(E|Y|^q)^{1/q}. \tag{1.18}$$

Definition 1.19. (Triangle Inequality) *For any real numbers a_i,*

$$\left|\sum_i a_i\right| \le \sum_i |a_i|.$$

Definition 1.20. (Boole's Inequality) *For a countable set of events $A_i, i = 1, 2, \cdots,$*

$$P\left(\bigcup_i A_i\right) \le \sum_i P(A_i).$$

Chapter 2 An Analysis of the Peachtree Creek

2.1 Introduction

Over recent years, there has been considerable interest in fitting statistical models to riverflow data. To give a few examples, Hollis (1975) established some trends between urbanization and the frequency of floods. Katz *et al.* (2002) discussed a variety of practises for detecting trends in hydrologic extremes, including peaks-over-threshold and block maxima. Villarini *et al.* (2009) fit a nonstationary time series to data from Little Sugar Creek in North Carolina.

The United States Geological Survey (USGS) has continuously monitored the Peachtree Creek in Atlanta, GA since 1958. Daily average measurements of variables such as discharge (hereafter referred to as streamflow), gage height, water temperature, and pH have been recorded. Since 1989, more frequent measurements have been taken, and the data are available for free download at *http://waterdata.usgs.gov/ga/ nwis.* The water statistics need to be monitored in order to detect any changes in water quality, to establish baseline information about the creek, and to detect problems with bacteria and sediment during storms.

Streamflow is the volume of water that flows past a predetermined point in a fixed amount of time. Naturally, higher levels of streamflow are positively correlated with higher levels of the creek and increased chances of a flood occurring. The Peachtree Creek floods when the gage height peaks above 17 feet. Over the years this creek appears to be flooding more frequently and with greater magnitude of water and destruction, such as the September 2009 flood, as explained on the USGS Georgia Water Science Center web site. The goal of this paper is to investigate the statistical significance of this observation and study trends by season.

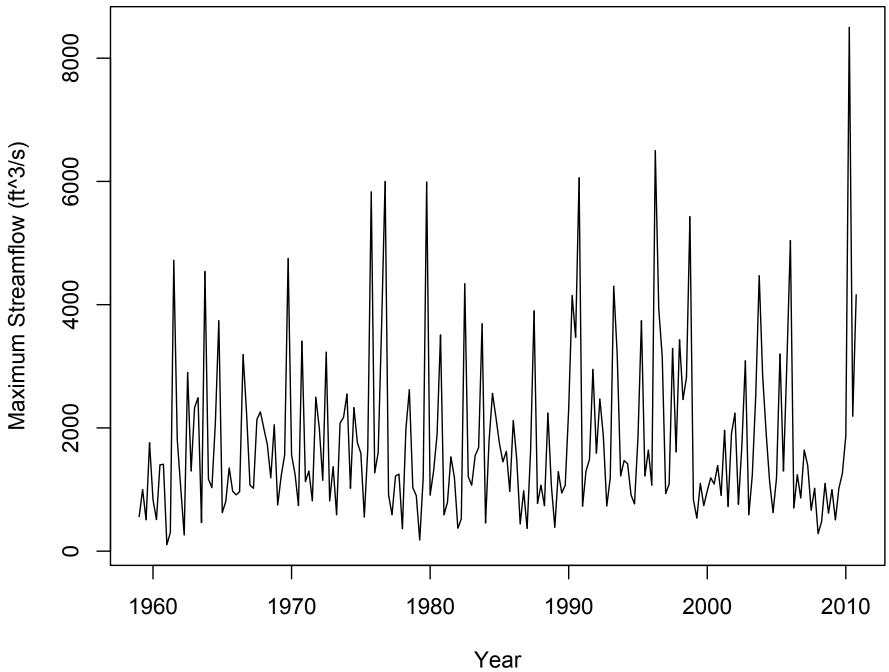
2.2 The Data

There are two data sets we use in the analysis. The first contains the daily mean streamflow measurements recorded every day beginning June 20, 1958, and ending on June 22, 2010. This makes 18996 total days, and only one day (June 16, 2010) has a missing value. This data set is the one we shall use to fit the generalized extreme value model. The station that records the measurements is located in Fulton Country at latitude 33°49'10"N and longitude 84°24'28"W (USGS). This location is in the heart of downtown Atlanta, a couple of miles north from where the interstates I-75 and I-85 merge.

The second data set consists of streamflow and gage heights measured every 15 minutes and therefore up to 96 recordings per day. Observations were recorded at these intervals between October 1, 1989 and September 30, 2009. Using this data set, we condense the data into the daily maximum heights, provided all 96 observations are available. In the event of a day having only some observations recorded, we take the maximum of what was available, provided a reasonable number (two-thirds or more) was recorded.

Next, we compute the seasonal maximum average streamflow for the 52 years of data. For climatology purposes, summer is defined to be the months June through August; fall is September through November; winter is December through February, and spring is March through May. Because the first observation in the data set is from June 20, 1958, it is convenient to define a year as running from June 1 to the following May 31. The 52 years therefore provide 208 seasonal maxima. Figure 2.1 illustrates year versus the 208 seasonal maximum streamflow observations, one for every three months. We indeed observe more severe streamflow measurements in recent years.

Figure 2.1: Year versus Seasonal Maximum Streamflow

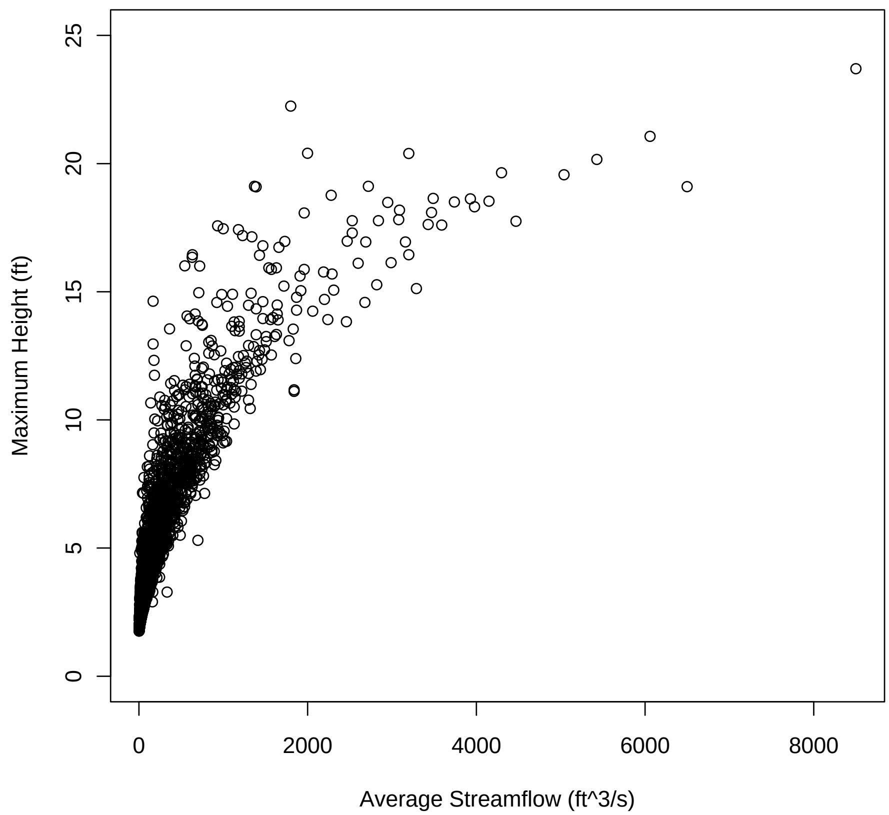


For ease of word usage, from this point onward we shall refer to the maximum mean streamflow as simply the maximum streamflow (and thereby drop the word "mean"). By these terms we mean the maximum of a series of average recordings.

Obviously fitting a model through the maximum heights is the desirable approach, since a flood is much easier to interpret given the height rather than the streamflow, but we choose to work with the recorded streamflow because there are 52 years of data - there are only 20 years for the second data set. However, after fitting the generalized extreme value model we shall use nonparametric techniques to construct estimates for trends in the maximum gage height. Had we estimated the missing maximum heights from 1958 to 1988 first and then fitted a model, there would have been a much greater amount of bias.

It is known that there is a positive trend between streamflow and height, and graphically the relationship looks exponential. However, our data set shows the average streamflow per day rather than individual measurements, so using the graph to estimate heights will not work. In addition, the river changes shape slightly over the years due to erosion, so in the end our nonparametric method is a safer way of estimating heights. Figure 2.2 illustrates an example of the exponential pattern from the second data set.

Figure 2.2: Average Streamflow versus Maximum Height, 1989 - 2009



The first summer (1958) has nineteen missing days because the data set did not start on June 1, so for that one season we just computed the maximum streamflow of the remaining days. Also the data set ends on June 22, 2010, but we stopped at May 31 in order to have an equal number of maxima per season, and also because this final partial summer season had only just begun.

We suspect different behaviors in each of the four seasons. Had we considered only the yearly maxima, any significant seasonal patterns would have been lost. Because of this motivation, we shall fit a generalized extreme value model with effects from season and time. Figure 2.3 shows box-plots of the seasonal maximum streamflow per season, while Table 2.1 displays the summary statistics. Spring and fall clearly have the greatest variation, while summer and winter have the lowest.

Figure 2.3: Box-Plots of Seasonal Maximum Streamflow

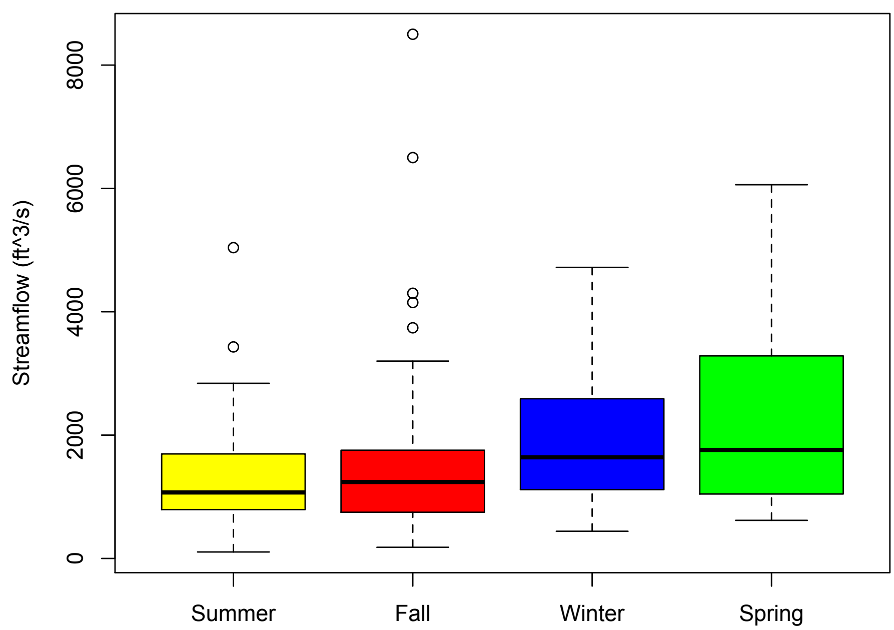
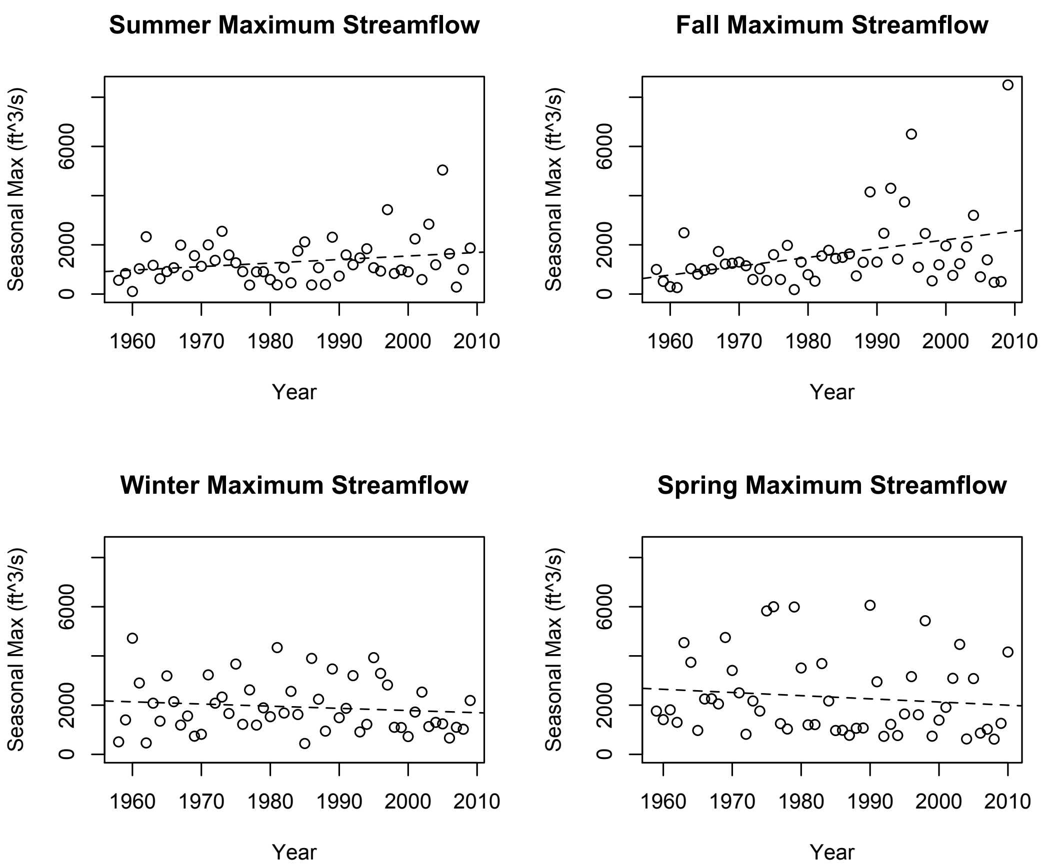


Table 2.1: Seasonal Maxima Summary Statistics

Season	**Mean**	**SD**	**Min.**	**Q1**	**Median**	**Q3**	**Max.**
Summer	1309.5192	878.8647	105	813	1070	1668	5040
Fall	1613.3654	1512.2333	181	755	1240	1742	8500
Winter	1927.1538	1075.7654	441	1122	1640	2575	4720
Spring	2327.4615	1617.2457	618	1052	1760	3222	6060

Figure 2.4 shows the year versus the seasonal maximum streamflow for each season. In each plot, the dotted lines represent the Ordinary Least Squares regression lines fitting streamflow against year.

Figure 2.4: Year versus Seasonal Maximum Streamflow per Season



Note that while OLS is a naïve procedure since the data are maxima, it nevertheless gives us a reasonable idea about linear trends per season. For example, we can see that fall observations exhibit a distinct upward trend in later years. In addition, some of the plots (fall, for instance) appear to have significantly higher variances in the later years. Based on these observations, we can specialize the model accordingly.

2.3 Statistical Methods

Coles (2001) suggests testing for linear trend in time in the location and scale parameters. In a study on annual maximum sea levels at Fremantle, Western Australia, he also tested for effects from the Southern Oscillation Index in the location parameter. Graphs illustrating the effects of this index may be seen in Coles (2008). We initially tested this index as a covariate in our model, but it did not contribute significant effects. While it is theoretically possible to test for time trends in the shape parameter ξ, for practical purposes it is unrealistic, so we keep ξ stationary.

The full model is defined as follows. Let $i = 1, 2, 3, 4$ denote the seasons summer, fall, winter, and spring, respectively. Denote by X_{ti} the seasonal maximum average streamflow in the tth year and ith season, and x_{ti} the recorded observation, $t = 1, ..., 52$. Here $t = 1$ denotes the year from June 1, 1958 through May 31, 1959. Define $F(x_{ti}) = P(X_{ti} \leq x_{ti})$, then

$$F(x_{ti}) = \exp\left\{ - \left[1 + \frac{\xi_i(x_{ti} - \mu_i(t)}{\sigma_i(t)} \right]^{-\frac{1}{\xi_i}} \right\}, \xi_i \neq 0, 1 + \frac{\xi_i(x_{ti} - \mu_i(t))}{\sigma_i(t)} > 0, \tag{2.1}$$

where $\mu_i(t) = \mu_i + \eta_i t$ and $\sigma_i(t) = \sigma_i + \phi_i t$. If any of the $\xi_i = 0$, then use the appropriate limiting Gumbel distribution instead. Let lnL be the log likelihood to be maximized, using a total of twenty parameters, five per season. Then

$$\begin{aligned} \text{lnL} = -\sum_{i=1}^{4} \sum_{t=1}^{52} & \left\{ \log(\sigma_i(t)) + \left(\frac{1}{\xi_i} + 1 \right) \log \left(1 + \frac{\xi_i(x_{ti} - \mu_i(t))}{\sigma_i(t)} \right) \right. \\ & \left. + \left(1 + \frac{\xi_i(x_{ti} - \mu_i(t))}{\sigma_i(t)} \right)^{-\frac{1}{\xi_i}} \right\}, \end{aligned} \tag{2.2}$$

provided that each of the $1 + [\xi_i(x_{ti} - \mu_i(t))]/\sigma_i(t) > 0$, otherwise $\text{lnL} = -\infty$. Again, if any of the ξ_i estimates are very close to 0, then we use the appropriate limiting distribution instead. We used an evolutionary algorithm to fit the model because it was easier to control the stipulation that $1 + [\xi_i(x_{ti} - \mu_i(t))]/\sigma_i(t) > 0$.

A common practice is to set $\sigma_i(t) = \exp\{\sigma_i + \phi_i t\}$ to ensure that the scale parameter remains positive, but we chose not to do this for three reasons. First, our procedure yields estimates that never put the scale parameters close to zero. Second, such a definition would make interpretation much harder than a linear one. And third, for comparison purposes we tried refitting the model with this alternative scale and found that the estimated values were extremely similar, so we decided to use the more interpretable linear definition $\sigma_i(t) = \sigma_i + \phi_i t$.

The selection process worked as follows. Let M_0 and M_1 be nested models such that $M_0 \subset M_1$, and $\ln L_0$ and $\ln L_1$ be the corresponding log-likelihoods. Then $-2(\ln L_0 - \ln L_1))$ is asymptotically a chi-square distribution with degrees of freedom equal to the number of parameters dropped by going to M_0. If the corresponding p-value is below the significance level, then M_1 fits the model significantly better; otherwise M_0 provides just as good a fit. This process is described extensively in the literature, with Coles (2001) and Coles (2008) providing several examples from data sets.

Our full model contained a total of five parameters per season. We considered dropping combinations of the η_i and ϕ_i, and for better interpretability we examined these parameters in groups of four rather than individually. For example, the final model will contain either all four ϕ_i or none, the latter case if none of those parameters are significant. Thus, there were a total of four possible models to consider, and each chi-square test was on degrees of freedom 4 or 8. We used $\alpha = .10$ as a selection criteria because the data were observational rather than experimental.

Our selection procedure chose the full model itself. Having arrived at a candidate, the next step was to investigate whether letting the shape parameters ξ_i vary from zero was appropriate per season. A reasonable way of checking this was to test the hypotheses $H_0 : \xi_i = 0$ against $H_1 : \xi_i \neq 0, i = 1, ..., 4$ and computing the p-values. Table 2.2 contains the results for each shape parameter.

Table 2.2: Shape Parameter Summary

Season	**Param.**	**Estimate**	**SE**	**90% Confidence**	**T-Stat**	**P-Value**
Summer	ξ_1	0.0898	0.1203	(-0.1120, 0.2917)	0.7468	0.4589
Fall	ξ_2	0.2335	0.1317	(0.0125, 0.4544)	1.7728	0.0827
Winter	ξ_3	0.1293	0.1475	(-0.1182, 0.3768)	0.8763	0.3853
Spring	ξ_4	0.7455	0.2249	(0.2931, 1.1979)	3.3151	0.0018

At significance level $\alpha = .10$, the shape parameters for summer and winter were not significantly different from 0, and so for those seasons we refitted the data setting $\xi_i = 0, i = 1, 3$. Fall and spring, on the other hand, had shape parameters that were significantly positive.

Table 2.3 compares the log-likelihoods for summer and winter with $\xi_i \neq 0$ and with $\xi_i = 0$, and then does the nested likelihood test. In the table $-2\Delta\text{lnL}_i$ represents the doubled difference in log-likelihoods in season i. Also the log-likelihoods displayed are for summer and winter alone rather than the sum of all four seasonal likelihoods.

Table 2.3: Summer and Winter Log-Likelihoods

Season	$\xi_i \neq 0$	$\xi_i = 0$	$-2\Delta\textbf{lnL}_i$	**P-Value**
Summer	-413.8720	-414.1838	0.6236	0.4297
Winter	-429.8705	-430.2704	0.7998	0.3712

Both the confidence intervals presented earlier and the above nested model analysis suggest that summer and winter may indeed have the shape parameter set equal to 0. To be clear, the final model treats summer and winter as having a Type I extreme value distribution (where $\xi_1 = \xi_3 = 0$), whereas fall and spring have a Type II distribution (where ξ_2 and ξ_4 are significantly positive).

The final step, as described in Section 2.4, is to make inference about the behavior of the maximum heights, given the established behavior of the maximum streamflow.

2.4 The Selected Models

Table 2.4 summarizes the estimated parameters per season.

Table 2.4: Estimated Parameters per Season

Season	Param.	Estimate	SE	90% Confidence	T-Stat	P-Value
Summer	ξ_1	0	—	—	—	—
	μ_1	727.6025	148.7033	(478.1937, 977.0112)	4.8930	1.2e-5
	σ_1	429.7833	100.9200	(260.5178, 599.0488)	4.2587	9.5e-5
	η_1	8.4951	5.6177	(-0.9271, 17.9173)	1.5122	0.1370
	ϕ_1	5.9982	3.9117	(-0.5626, 12.5591)	1.5334	0.1317
Fall	ξ_2	0.2335	0.1317	(0.0125, 0.4544)	1.7728	0.0827
	μ_2	561.6459	139.5658	(327.4646, 795.8271)	4.0242	0.0002
	σ_2	298.9993	116.2551	(103.9318, 494.0669)	2.5719	0.0133
	η_2	17.1289	6.6717	(5.9343, 28.3234)	2.5674	0.0135
	ϕ_2	13.7862	5.8231	(4.0155, 23.5569)	2.3675	0.0221
Winter	ξ_3	0	—	—	—	—
	μ_3	1687.4669	282.1902	(1214.1707, 2160.7631)	5.9799	2.7e-7
	σ_3	1106.9248	235.9170	(711.2390, 1502.6106)	4.6920	2.3e-5
	η_3	-8.8806	7.9343	(-22.1883, 4.4270)	-1.1193	0.2686
	ϕ_3	-11.3350	6.9608	(-23.0098, 0.3399)	-1.6284	0.1100
Spring	ξ_4	0.7455	0.2249	(0.3682, 1.1229)	3.3151	0.0018
	μ_4	1693.3611	237.2369	(1295.2950, 2091.4272)	7.1378	5.1e-9
	σ_4	855.2209	271.3575	(399.9029, 1310.5388)	3.1516	0.0028
	η_4	-14.7433	6.8138	(-26.1763, -3.3102)	-2.1637	0.0356
	ϕ_4	-6.8812	8.4173	(-21.0049, 7.2424)	-0.8175	0.4178

Interpretation will be easier if we analyze the expected value of the streamflow per season, but this involves reparametrizing the variables and recomputing standard errors and p-values. We begin by considering the expected mean of the seasonal maximum streamflow. For our data, it is easily checked using (1.11) that

$$E(X_{ti}) = \begin{cases} \mu_i + \sigma_i\gamma + (\eta_i + \phi_i\gamma)t, & \text{i = 1,3} \\ \mu_i + \sigma_i\Xi_i^{(1)} + (\eta_i + \phi_i\Xi_i^{(1)})t, & \text{i = 2,4,} \end{cases} \tag{2.3}$$

where $\Xi_i^{(1)} = [\Gamma(1-\xi_i) - 1]/\xi_i$. Interpretation is easy here because we have written the expected value as a linear function of time. In order to make inference about the slopes, we use the delta method to approximate the standard error. Before doing that, we also wish to investigate the effects of the median streamflow per year. The reason we want to consider the median as well is because a generalized extreme value model is skewed, so the median may provide a more robust result. Using (1.12),

$$\text{Med}(X_{ti}) = \begin{cases} \mu_i - \sigma_i \log(\log 2) + (\eta_i - \phi_i \log(\log 2))t, & \text{i = 1,3} \\ \mu_i + \sigma_i\Xi_i^{(2)} + (\eta_i + \phi_i\Xi_i^{(2)})t, & \text{i = 2,4,} \end{cases} \tag{2.4}$$

where $\Xi_i^{(2)} = \left[(\log 2)^{-\xi_i} - 1\right]/\xi_i$. Finally, we also make inference about the seasonal variance through time. For spring, the estimated shape parameter is $\hat{\xi}_4 = 0.7455$, indicating that spring's variance does not exist. (We shall address this issue shortly.) Interpretation will be easier if we instead think in terms of the seasonal standard deviations. Using (1.13), it can be shown that

$$\text{SD}(X_{ti}) = \begin{cases} \frac{\pi}{\sqrt{6}}\sigma_i + \left(\frac{\pi}{\sqrt{6}}\phi_i\right)t, & \text{i = 1,3} \\ \left(\Xi_i^{(3)}\right)^{\frac{1}{2}}\sigma_i + \left(\Xi_i^{(3)}\phi_i\right)t, & \text{i = 2} \\ \infty, & \text{i = 4,} \end{cases} \tag{2.5}$$

where $\Xi_i^{(3)} = [\Gamma(1-2\xi_i) - (\Gamma(1-\xi_i))^2]/\xi_i^2$. Again, we can easily make inference because the standard deviation has been written in terms of a linear function of time. The delta method now works as follows. For seasons summer and winter, define $A_i = \eta_i + \gamma\phi_i$, $B_i = \eta_i - \phi_i \log(\log 2)$, and $C_i = \frac{\pi}{\sqrt{6}}\phi_i, i = 1, 3$. Then the gradient matrix is

$$G_i = \begin{pmatrix} 0 & 0 & 1 & \gamma \\ 0 & 0 & 0 & -\log(\log 2) \\ 0 & 0 & 0 & \frac{\pi}{\sqrt{6}} \end{pmatrix}.$$

Meanwhile, for fall we have $A_2 = \eta_2 + \Xi_2^{(1)}\phi_2$, $B_2 = \eta_2 + \Xi_2^{(2)}\phi_2$, and $C_2 = \Xi_2^{(3)}\phi_2$. The gradient matrix is

$$G_2 = \begin{pmatrix} D_{12} & 0 & 0 & 1 & \Xi_2^{(1)} \\ D_{22} & 0 & 0 & 1 & \Xi_2^{(2)} \\ D_{32} & 0 & 0 & 0 & \Xi_2^{(3)} \end{pmatrix},$$

where D_{ri} is the derivative of $\Xi_i^{(r)}$ with respect to ξ_i; $i = 2$; $r = 1, 2, 3$. Finally, for spring we define A_4 and B_4 similarly to the reparameterizations in fall, and with similar gradient matrix but without the third row:

$$G_4 = \begin{pmatrix} D_{14} & 0 & 0 & 1 & \Xi_4^{(1)} \\ D_{24} & 0 & 0 & 1 & \Xi_4^{(2)} \end{pmatrix}.$$

The complete variance-covariance matrix is therefore $\Delta = G\Sigma G^T$, where

$$G = \begin{pmatrix} G_1 & 0 & \cdots & 0 \\ 0 & G_2 & \ddots & \vdots \\ \vdots & \ddots & G_3 & 0 \\ 0 & \cdots & 0 & G_4 \end{pmatrix}.$$

The eleven square-rooted diagonal entries of Δ give the approximate standard errors, which in turn are used to obtain the two-sided p-values. Note that the p-values are all computed on 49 degrees of freedom, except for spring's, which are on 50. The G_i matrices are

$$G_1 = \begin{pmatrix} 46.2533 & 42.6408 & 21.9891 \\ 42.6408 & 39.7076 & 17.8541 \\ 21.9891 & 17.8541 & 25.1700 \end{pmatrix},$$

$$G_2 = \begin{pmatrix} 114.1901 & 86.3071 & 112.0590 \\ 86.3071 & 70.9387 & 65.6263 \\ 112.0590 & 65.6263 & 174.6047 \end{pmatrix},$$

$$G_3 = \begin{pmatrix} 112.9623 & 100.8884 & 73.4939 \\ 100.8884 & 90.9656 & 60.4001 \\ 73.4939 & 60.4001 & 79.7024 \end{pmatrix},$$

and

$$G_4 = \begin{pmatrix} 2183.2748 & 372.5666 \\ 372.5666 & 100.1301 \end{pmatrix}.$$

To study the significance of variance in spring, we temporarily refitted the spring data with the condition that $\{\xi_4 \in (-\infty, 0.47)\}$. The chosen upper bound of 0.47, while arbitrary, allowed us to study the significance of the variance when forced to be finite. The resulting adjusted model gave the following results for spring, shown in Table 2.5.

Table 2.5: Adjusted Mean, Median, and Standard Deviation for Spring

Season	Param.	Estimate	SE	90% Confidence	T-Stat	P-Value
Spring	A_4	-19.8839	18.3424	(-50.6360, 10.8681)	-1.0840	0.2837
	B_4	-15.8563	9.8943	(-32.4447, 0.7320)	-1.6026	0.1155
	C_4	-19.8839	133.7022	(-244.0424, 204.2746)	-0.1487	0.8824

The effect from the variance on these adjusted estimators is clearly insignificant, and so we decide to drop the restriction on ξ_4 and stick with our original outcome, with infinite spring variance. Table 2.6 shows the final results for the slopes of the means, medians, and standard deviations per season.

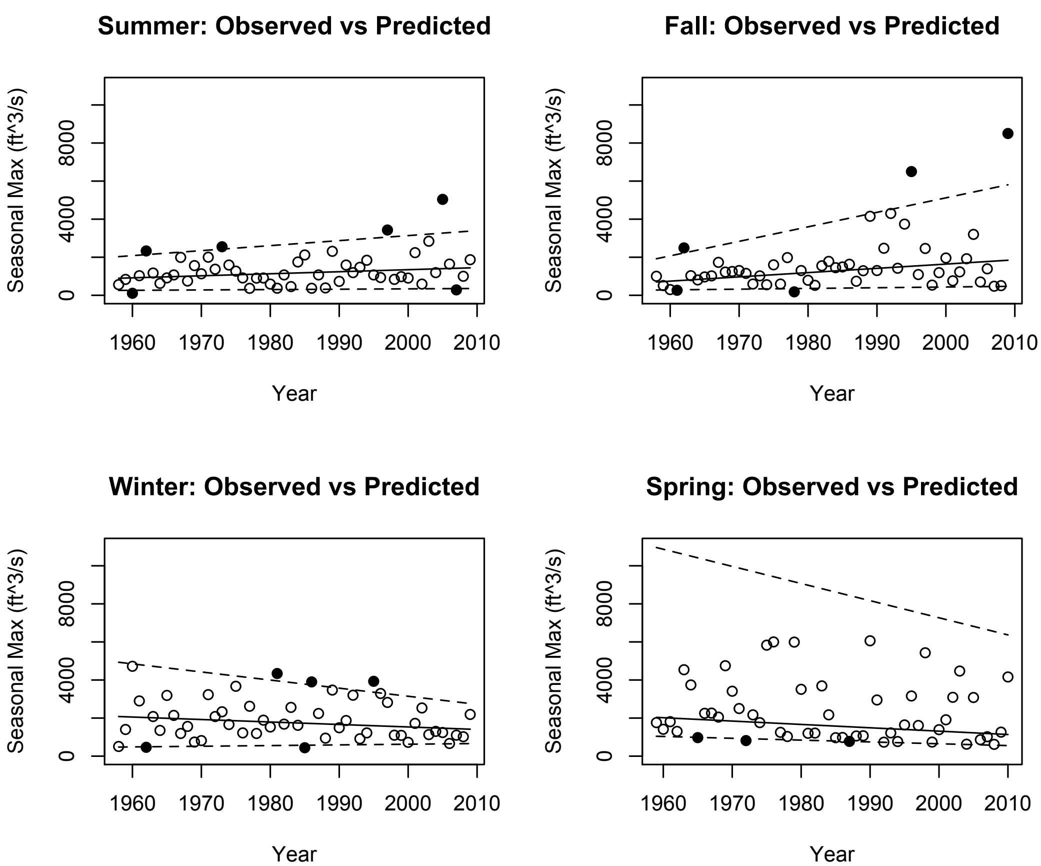
Table 2.6: Means, Medians, and Variances per Season

Season	Param.	Estimate	SE	90% Confidence	T-Stat	P-Value
Summer	A_1	11.9574	6.8010	(0.5552, 23.3596)	1.7582	0.0850
	B_1	10.6936	6.3014	(0.1289, 21.2582)	1.6970	0.0960
	C_1	7.6930	5.0170	(-0.7182, 16.1042)	1.5334	0.1316
Fall	A_2	29.1767	10.6860	(11.2612, 47.0923)	2.7304	0.0088
	B_2	22.4042	8.4225	(8.2834, 36.5249)	2.6600	0.0105
	C_2	29.1767	13.2138	(7.0231, 51.3304)	2.2080	0.0320
Winter	A_3	-15.4233	10.6284	(-33.2423, 2.3957)	-1.4511	0.1531
	B_3	-13.0350	9.5376	(-29.0253, 2.9552)	-1.3667	0.1780
	C_3	-14.5376	8.9276	(-29.5053, 0.4300)	-1.6284	0.1099
Spring	A_4	-38.3563	46.7255	(-116.6638, 39.9513)	-0.8209	0.4156
	B_4	-17.6435	10.0065	(-34.4135, -0.8736)	-1.7632	0.0840

The results show that for each season, the median is a more reliable measure of center than the mean, since it is robust to unusually high observations and the standard errors are smaller. In particular, the standard error for spring's median is 21.4% that of the mean, causing the median to be decreasing significantly, unlike the mean. Therefore from this point forward, we take the median as the better measure of center.

Figure 2.5 shows the observed data with the expected median value per season. The dotted lines represent the 90% prediction intervals, which indeed capture the majority of all observations. The black dots are the few observations that fall outside their prediction intervals.

Figure 2.5: Predicted Median Streamflow per Season



Finally, we give a practical interpretation. The following numbers are reported in cubic feet per second, and for simplicity we just say "units."

1. In summer, the maximum daily average streamflow is increasing on average by 10.6936 units each year, but its standard deviation is not changing significantly.

2. In fall, the maximum daily average streamflow is increasing on average by 22.4042 units each year, and its standard deviation is also increasing by 29.1767 units per year.

3. In winter, the maximum daily average streamflow appears stationary in both the median and the standard deviation.

4. In spring, the maximum average streamflow is decreasing on average by 17.6435 units each year, but its standard deviation is not changing significantly.

2.4.1 Estimated Patterns in Seasonal Maximum Heights

We have successfully determined that the average streamflow is increasing significantly in summer and fall. However, insurance companies and hydrologists would rather know what patterns exist with the river height itself as this quantity is much easier to understand. A flood is classified by its gage height, and so we now pursue establishing what the maximum height each season most likely would have been. That is, given the average streamflow (in cubic feet per second), we wish to predict the corresponding maximum height (in feet).

The second data set mentioned in Section 2.2 contains the average streamflow per day, as well as the maximum height measured per day, spanning the period October 1, 1989 through September 20, 2009. The number of days considered is therefore 7305. Examining Figure 2.2, there is clearly a positive nonlinear trend between the variables. We could not find any reasonable models that accurately described this data set, so instead we proceed nonparamatrically to estimate the heights. To utilize a more robust approach, we compute medians rather than means as follows. Define $\tilde{D}$ to be the vector of ordered unique daily mean streamflow observations rounded to the nearest integer, and note that the values of $\tilde{D}$ range from 3 to 8500. Given a rounded streamflow, find $\tilde{H}$, the median of all maximum heights whose corresponding streamflow is the current considered value. These median heights are reasonable predictions given recorded average streamflow.

Recall that we have used our seasonal extreme value model to predict the median streamflow, given year t and season i. Call these $\text{medD}_{ti}, t = 1, ..., 52$ and $i = 1, ..., 4$. We now compute the predicted maximum height using the following nonparametric procedure:

1. Find the coordinates $(\tilde{D}_a, \tilde{H}_a)$ and $(\tilde{D}_{a+1}, \tilde{H}_{a+1})$ such that $\tilde{D}_a \leq \text{medD}_{ti} < \tilde{D}_{a+1}$.

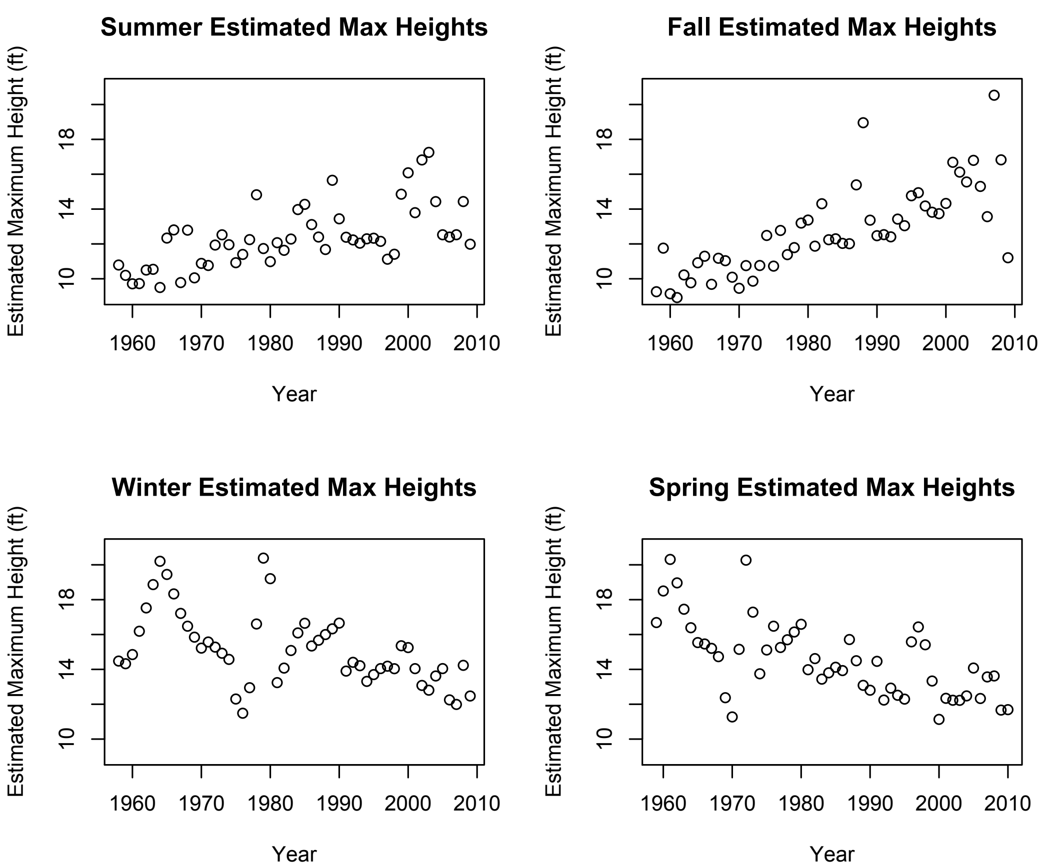
2. Compute $\alpha = \left[\tilde{H}_{a+1} - \tilde{H}_a\right] / \left[\tilde{D}_{a+1} - \tilde{D}_a\right]$.

3. Compute $\text{medH}_{ti} = \alpha\,(\text{medD}_{ti}) + \tilde{H}_a - \alpha\tilde{D}_a$.

4. Repeat for all years t and seasons i.

Figure 2.6, divided by season, shows year versus predicted maximum height.

Figure 2.6: Year versus Predicted Maximum Height



There is a distinct upward trend in summer and fall, and a downward trend in winter and spring. We can smooth the predictions by (a) fitting an ordinary least squares line through the points or (b) fitting a least median-of-squares line. The latter approach is more desirable for our purposes for two reasons. First, we have been using median analyses throughout the chapter for robustness. Second, this approach relaxes the usual assumptions imposed on the residuals. We shall also bootstrap the confidence intervals for the median slopes. Tables 2.7 and 2.8 summarize the results for OLS and median regression, respectively. Define $\theta_{0,i}$ and $\theta_{1,i}$ to be the intercept and slope parameters respectively for the OLS lines, and $\tilde{\theta}_{0,i}$ and $\tilde{\theta}_{1,i}$ to be those for the median lines.

Table 2.7: OLS Estimates for Predicted Heights

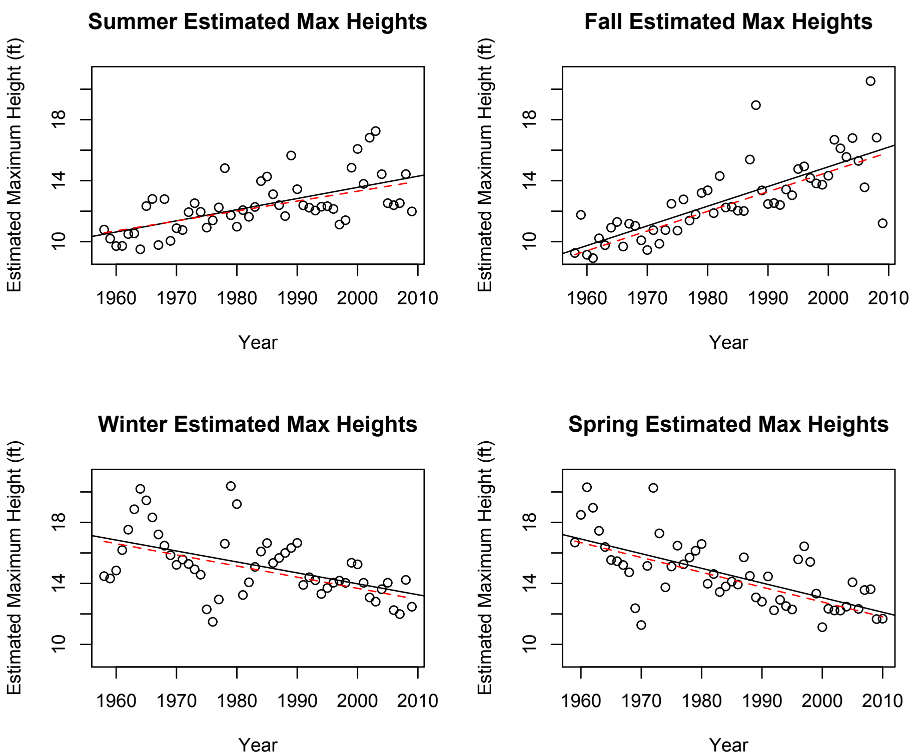
Season	Parameter	Estimate	SE	T-Stat	P-Value
Summer	$\theta_{0,1}$	-132.7745	25.8338	-5.1396	4.9e-6
	$\theta_{1,1}$	0.0732	0.0130	5.6179	8.5e-7
Fall	$\theta_{0,2}$	-243.3921	29.2504	-8.3210	5.3e-11
	$\theta_{1,2}$	0.1292	0.0147	8.7581	1.1e-11
Winter	$\theta_{0,3}$	157.4425	32.8438	4.7937	1.5e-5
	$\theta_{1,3}$	-0.0717	0.0166	-4.3322	7.1e-5
Spring	$\theta_{0,4}$	205.1521	30.2971	6.7714	1.4e-8
	$\theta_{1,4}$	-0.0960	0.0153	-6.2910	7.7e-8

Table 2.8: Median Estimates for Predicted Heights

Season	Parameter	Estimate	SE	95% Confidence	P-Value
Summer	$\tilde{\theta}_{0,1}$	-116.7596	—	—	—
	$\tilde{\theta}_{1,1}$	0.0650	0.0041	(0.0578, 0.0728)	6.4e-81
Fall	$\tilde{\theta}_{0,2}$	-243.3388	—	—	—
	$\tilde{\theta}_{1,2}$	0.1290	0.0030	(0.1229, 0.1359)	$<$ 1.0e-99
Winter	$\tilde{\theta}_{0,3}$	160.1106	—	—	—
	$\tilde{\theta}_{1,3}$	-0.0732	0.0057	(-0.0826, -0.0621)	4.0e-54
Spring	$\tilde{\theta}_{0,4}$	207.3821	—	—	—
	$\tilde{\theta}_{1,4}$	-0.0973	0.0039	(-0.1032, -0.0883)	$<$ 1.0e-99

To bootstrap, we first compute all 1326 possible slopes per season and then resample with replacement from these slopes, recording the median in each case. The 95% confidence interval follows from the 2.5th and 97.5th percentiles of the vector of recorded medians. Note that Table 2.8 only displays the analysis for the slopes since the intercepts are not of interest. Figure 2.7 summarizes the four seasons' results. The solid lines are the OLS estimates, while the dotted lines represent the nonparametric estimates.

Figure 2.7: Height Trends per Season



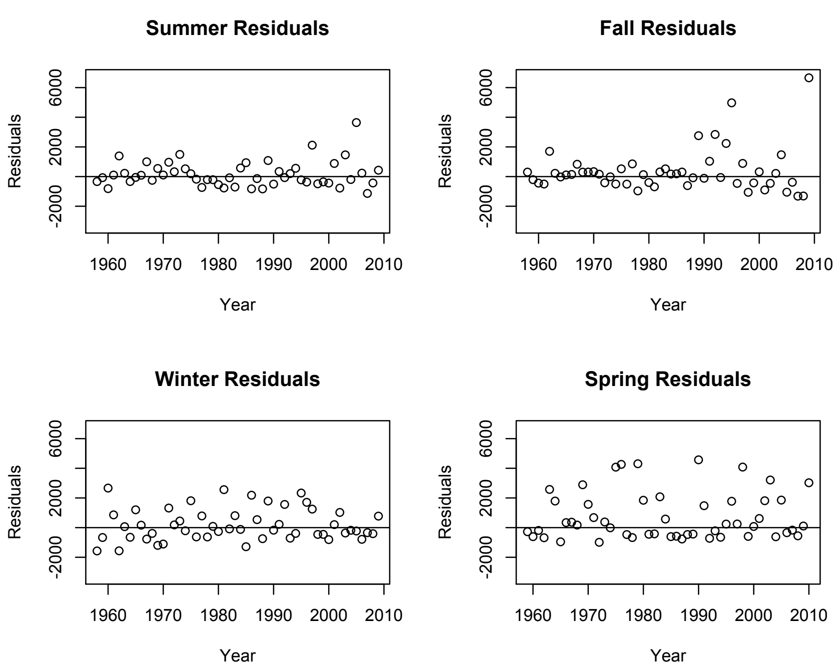
Two observations are immediately apparent from Tables 2.7 and 2.8. First, all the confidence intervals exclude 0, indicating that summer and fall heights are increasing significantly, while winter and spring heights are decreasing significantly. And second, all the standard errors for the median slopes are much smaller than those for the OLS slopes. We therefore state the following conclusions, where the median heights have been converted into inches.

1. In summer, the maximum height is increasing on average by 0.78 inches per year.
2. In fall, the maximum height is increasing on average by 1.55 inches per year.
3. In winter, the maximum height is decreasing on average by 0.88 inches per year.
4. In spring, the maximum height is decreasing on average by 1.17 inches per year.

2.4.2 Diagnostic Checks

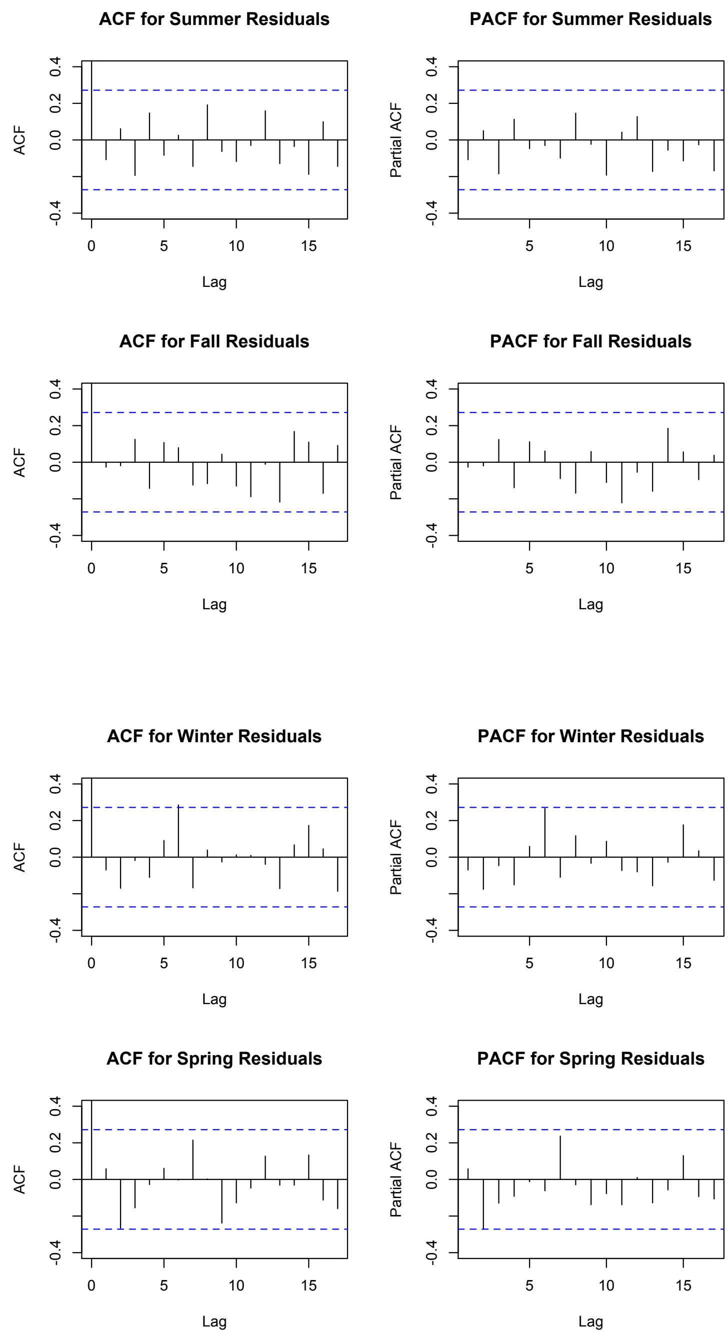
We now check whether the final model fits the streamflow data adequately, and that no further adjustments are necessary. We begin by computing the seasonal residuals, defined by the observed streamflow minus the median predicted streamflow per season. In Figure 2.8, they all appear random but not necessarily centered around 0. This characteristic is to be expected, however, since we are estimating extreme values.

Figure 2.8: Residuals from Final Streamflow Model



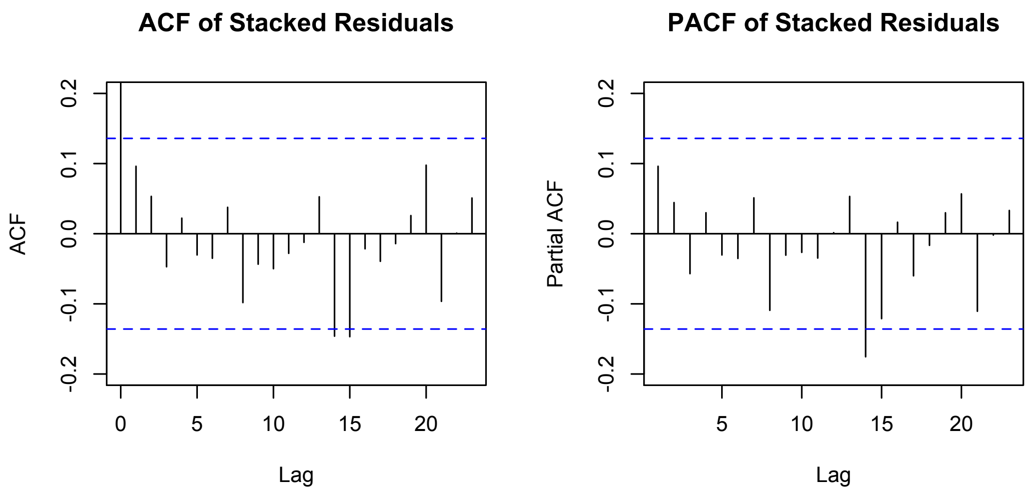
We then examine Figure 2.9, the ACF and PACF plots of the residuals, again broken down by season. There is no evidence of further correlation in time.

Figure 2.9: ACF and PACF Plots by Season



Next, we stack the residuals in order of occurrence. That is, each group of four residuals comes from the summer, fall, winter, and spring, respectively, of the same year. Figure 2.10 shows ACF and PACF plots of the resulting vector, again suggesting no further time trends. The only spikes are at the fourteenth and fifteenth lags, but these are not interpretable.

Figure 2.10: ACF and PACF Plots of Entire Data



Finally, we conduct a goodness of fit test to verify that our chosen model provides a suitable fit to the data. The presence of nonstationarity means that some modification to the data is necessary. Coles (2001) suggests the following transformation that maps the data to standard Gumbel random variables. Define Z_{ti} such that

$$Z_{ti} = \begin{cases} \frac{1}{\hat{\xi}_i} \log \left[1 + \frac{\hat{\xi}_i (X_{ti} - \hat{\mu}_i - \hat{\eta}_i t)}{\hat{\sigma}_i + \hat{\phi}_i t} \right], & \text{i} = 1,3 \\ \frac{X_{ti} - \hat{\mu}_i - \hat{\eta}_i t}{\hat{\sigma}_i + \hat{\phi}_i t}, & \text{i} = 2,4. \end{cases} \tag{2.6}$$

That is, each of the Z_{ti} should have probability distribution function

$$P(Z_{ti} \leq z) = \exp\left\{ -\exp(-z) \right\}, -\infty < z < \infty. \tag{2.7}$$

Before conducting a goodness of fit test, we examine the probability and quantile plots using the following procedure.

1. Compute the Z_{ti} and denote their order statistics by $Z_{(1)}, ..., Z_{(N)}$ where $Z_{(j)} \leq Z_{(j+1)}$, $j = 1, ..., N-1$ and $N = 208$.

2. Draw the probability plot with the pairs

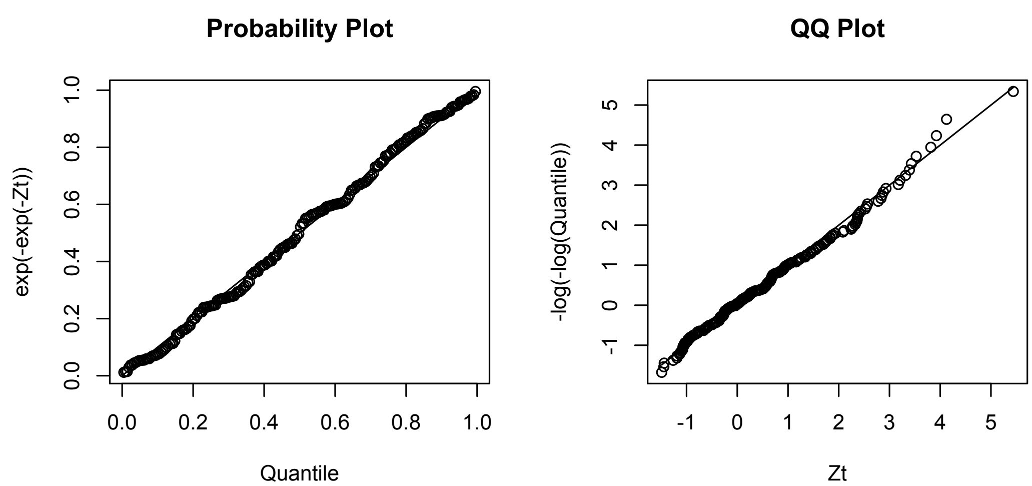
$$\left\{\frac{j}{N+1}, \exp\left\{-\exp(-Z_{(j)})\right\}; j = 1, ..., N\right\}.$$

3. Draw the quantile plot with the pairs

$$\left\{Z_{(j)}, -\log\left[-\log\left(\frac{j}{N+1}\right)\right]; j = 1, ..., N\right\}.$$

Both plots in Figure 2.11 suggest a reasonable fit.

Figure 2.11: Probability and QQ Plots



Most of the goodness-of-fit tests in the literature, however, are for testing for fit to the normal distribution. There are considerably fewer results available in the extreme value context.

As examples of references that do address this matter, Chandra *et al.* (1981) discuss a Kolmogorov-type statistic, while Hosking (1984) suggests a Wald-type test. Kinnison (1985) introduced a table for conducting a goodness-of-fit test to the Gumbel distribution. However, Kinnison (1989) published an improved table based on 100,000 simulations rather than 5,000 as was the case in his earlier work. His improved test proceeds as follows.

1. Compute the probability pairs

$$\left\{\frac{j}{N+1}, \exp\left\{-\exp(-Z_{(j)})\right\}; j = 1, ..., N\right\}.$$

2. Calculate the Pearson correlation coefficient through these pairs.

3. Examine the table whose entries represent the critical points for each combination of sample size (rows) and probability (columns).

4. Find the correct row, then scroll across to find the lower and upper critical points for the test statistic. Then look at the column headings to put a bound on the probability.

The correlation coefficient between the quantiles and the $\exp\left\{-\exp(-Z_{(j)})\right\}$ is 0.9988. We consult Kinnison's table in the row with the nearest sample size of 200, then note that the coefficient is beyond the last column. Thus, the probability of a good fit is above 0.95, suggesting the model fits the data extremely well. The row is duplicated in Table 2.9.

Table 2.9: Gumbel GOF Test for Sample Size n = 200

Prob.	**0.01**	**0.025**	**0.05**	**0.10**	**0.25**	**0.50**	**0.75**	**0.95**
Corr.	0.9702	0.9785	0.9838	0.9883	0.9930	0.9957	0.9972	0.9983

2.4.3 Forecasting The 2010 - 2011 Season

Having verified that our model has significant results and that the model fits the data well, we now conduct a forecast. To give the most reliable results, we use the original model with the daily mean streamflow to predict the seasonal maximum mean streamflow for the next year. That is, we forecast the maximum mean streamflow for Summer 2010 through Spring 2011. The preceding data analysis was conducted during Spring 2011, which means that now we have an extra year of data. Rather than incorporating this extra year into the model, we instead use the model to forecast the next four streamflow measurements and then compare with the actual values to see how accurate our model is. Table 2.10 summarizes the results.

Table 2.10: Forecasted Medians for the 2010 - 2011 Year

Season	**Predicted**	**90% Prediction Interval**	**Actual**	**In Interval?**
Summer 2010	1451.8823	(357.4878, 3398.6284)	545	Yes
Fall 2010	1863.4795	(472.8293, 5882.4474)	1350	Yes
Winter 2010	1402.3123	(661.4273, 2720.2239)	730	Yes
Spring 2011	1118.7097	(544.3879, 6277.8608)	866	Yes

In all four seasons, the prediction intervals successfully captured the observed values. Fall 2010 and Spring 2011 had predicted streamflow closest to the actual measurements. Also in each season, the actual value was located closer to the lower endpoint of the prediction interval.

2.5 Conclusions

The goal of the project was to establish that the Peachtree Creek's flood rates are increasing over time. Seasonality was the critical step in the analysis. The river's streamflow is significantly increasing in both summer and fall, and the higher the streamflow, the more likely a flood will occur. Thus, the probability of a flood is increasing in summer and fall, but decreasing in winter and spring.

The Peachtree Creek is flooding with greater magnitudes today than fifty-two years ago. The seasonal maximum height is increasing significantly in summer and fall, the latter having the greater rate. As a result, a flood is more likely to occur in fall and summer. The reasons for increased probability may vary from climate change to urbanization. Future studies on the maximum gage height could incorporate additional covariates such as population density, which most likely would be positively correlated with the gage height. The techniques presented in this paper may also be used on other bodies of water; for instance, the Mississippi River which flooded most recently in May 2011.

Chapter 3 Extreme Value Mixtures of Two Variables

3.1 Introduction

In environmental statistics, it is often the case that the current annual maximum measurement of temperature, sea level, river gage height, etc. is influenced in some way by the previous year's maximum. For example, if the temperature in a certain location is unusually high in spring 2010, it could be used as a covariate to predict the seasonal maximum temperature for spring 2011. Such cycles often happen in environmental statistics.

Building a reasonable model for this situation would work as follows. We first fit a generalized extreme value model through the data set, possibly choosing to introduce seasonal and temporal effects and other covariates if so desired. We would then add the most recent observation as an additional covariate on the location parameter and then find the maximum likelihood estimates of all parameters. Finally, we would compute confidence intervals and p-values.

In this dissertation we focus primarily on the Type I distribution, the Gumbel, but similar studies may be carried out on the Types II and III. We chose to study the Gumbel distribution in depth because of its ease in comparison to the Fréchet and the Weibull. In addition, the Gumbel arises very frequently in data analysis. It is often the case that the shape parameter in the generalized extreme value distribution is not significantly different from 0, in which case one would reduce the model to a Gumbel. We did just that for summer and winter in Chapter 2.

Let us now define the generalized extreme value model. For simplicity we assume that the only covariate is the most recent observation. The cumulative distribution function of the Gumbel distribution is defined in (1.2), and recall that it is denoted as $\text{GEV}(0, \mu, \sigma)$.

Now introduce the previous observation as a covariate on the location with parameter β. Throughout the dissertation, unless stated otherwise we take $-1 < \beta < 1$. Thus $X_1|X_0 \sim \text{GEV}(0, \mu_1 + \beta X_0, \sigma_1)$ and $X_0 \sim \text{GEV}(0, \mu_0, \sigma_0)$ with $\sigma_0, \sigma_1 > 0$. Evidence for the necessity of inclusion of the previous observation as a covariate can best be detected by examining ACF and PACF plots and checking for spikes at the first lag.

The question of interest in this chapter is that $X_1|X_0$ is a Gumbel random variable, but we want to know the unconditional distribution of X_1. Once answered, we will have taken the first step to building probability approximations for the AR(1) process with Gumbel innovations. This is a distribution that appears regularly in environmental statistics, so our result will be a valuable contribution to extreme value theory.

We should point out that for the Peachtree Creek data, there was no visible evidence from Figures 2.9 and 2.10 that the previous year's observation would contribute significantly to the model. One can guess this by noting that in none of the plots does the first lag stretch outside the 95% confidence bounds. However, in other data sets we may very well observe such a lag, in which case it is worth investigating this interesting new research question. Nevertheless, at the end of Chapter 4 we will refit the Peachtree Creek data with the previous year's observations to illustrate an example of how such a model would work.

Chapters 3 and 4 together take up the second project in the dissertation. In this chapter we first consider the standard Gumbel case where $\mu_i = 0$ and $\sigma_i = 1, i = 0, 1$. We derive two-term expansions for the tail probability along with precise error terms for various values of β. We later check the effectiveness of our results in a simulation study. Finally, in Chapter 4 we extend the results to a full AR(1) process. Also, at the end of Chapter 3 we give a brief introduction to the studied two variable mixture but for the Fréchet and Weibull cases.

3.2 Mixture of Gumbel Random Variables

In this section we derive some theoretical results about upper tail probabilities for the unconditional distribution, as well as some results on the error term involved. We shall see that, depending on the choice of β, there are four different approximations for the probability. Recall that for the standard Gumbel,

$$\Lambda(y) = \exp\left(-e^{-y}\right), \qquad -\infty < y < \infty. \tag{3.1}$$

We have $Z_1|Z_0 \sim GEV(0, \beta Z_0, 1)$ and $Z_0 \sim \Lambda$. Notice that

$$P(Z_1 - \beta Z_0 \leq y | Z_0 = z) = P(Z_1 \leq y + \beta z | Z_0 = z) = \Lambda(y).$$

As a consequence, $Z_1 - \beta Z_0$ and Z_0 are independent and $Z_1 - \beta Z_0 \sim \Lambda$. In other words, the probability we are estimating is $P(\beta Z_1 + \beta Z_0 > y)$. This observation is important because now the model may take any weights. In particular, let $c_1 > c_0$ be any positive constants. Then

$$P(c_1 Z_1 + c_0 Z_0 > y) = P\left(Z_1 + \frac{c_0}{c_1} Z_0 > \frac{y}{c_1}\right), \tag{3.2}$$

and by choosing $\beta = \frac{c_0}{c_1}$ we get the same situation. As a side note, if $c_1 = c_0$ then the probability must be approximated using a different technique. We delay this discussion until Chapter 5.

To be clear, we explain all the possible cases in the following list and at what point we answer them.

1. When $c_0 < c_1 < 0$, the probability is negligible as $y \to \infty$. This is because most of the Gumbel's mass is on the positive half line, and therefore $\{c_0 Z_0 + c_1 Z_1 > y\} =$

$\left\{Z_0 + \frac{c_1}{c_0} Z_1 < \frac{y}{c_0}\right\}$ is extremely unlikely to occur. Lemma 3.2 establishes that the answer is $o\left(e^{y/c_0}\right)$.

2. When $c_0 < 0 < c_1$, $-1 < \beta < 0$ and the approximation is (3.5). Lemma 3.3 provides the remainder term.

3. When $\beta = 0$, the problem reduces to the probability of just one Gumbel. The approximation is the same as the previous item.

4. When $0 < 2c_0 < c_1$, $0 < \beta < \frac{1}{2}$ and the approximation is (3.5). Lemma 3.4 provides the remainder term.

5. When $0 < 2c_0 = c_1$, $\beta = \frac{1}{2}$ and the approximation is (3.12), given at the end of Lemma 3.5.

6. When $0 < c_1 < 2c_0$, $\frac{1}{2} < \beta < 1$ and the approximation is derived in Lemma 3.7.

7. When $0 < c_0 = c_1$, $\beta = 1$ and the asymptotics need to be treated very differently. This discussion is rather involved and, for that reason, takes up its own project. We postpone that analysis until Chapter 5.

The exact integral is

$$P(Z_1 > y) = \int_{-\infty}^{\infty} \left\{1 - \exp\left(-e^{-(y-\beta x)}\right)\right\} e^{-x} \exp\left(-e^{-x}\right) dx. \tag{3.3}$$

A change of variables rewrites (3.3) as

$$P(Z_1 > y) = \int_{0}^{\infty} \left\{1 - \exp\left(-t x^{-\beta}\right)\right\} e^{-x} dx, \tag{3.4}$$

where $t = e^{-y}$. For large y, $1 - \exp\left(-tx^{-\beta}\right)$ is small and can be approximated using a two-term Taylor series by

$$1 - \exp\left(-tx^{-\beta}\right) \approx tx^{-\beta} - \frac{1}{2}t^2x^{-2\beta}.$$

Therefore for a suitable choice of β and t small enough,

$$P(Z_1 > y) = \int_0^\infty \left[tx^{-\beta} - \frac{1}{2}t^2x^{-2\beta}\right] e^{-x}dx + R,$$

where R is a remainder term to be analyzed shortly. Consequently for y large enough and certain choices of β,

$$P(Z_1 > y) = \Gamma(1-\beta)e^{-y} - \frac{1}{2}\Gamma(1-2\beta)e^{-2y} + R. \tag{3.5}$$

Definition 3.1. *Let $\rho(x)$ be a probability function on (x_l, x_u) and $A_i(x), i = 1, 2, 3$ be three approximation formulas to the probability. Also let $R_i(x)$ be the error terms associated with each $A_i(x)$. Then β is said to be a pivot point if*

$$\rho(x) = \begin{cases} A_1(x) + R_1(x), & x_l < x < \beta \\ A_2(x) + R_2(x), & x = \beta \\ A_3(x) + R_3(x), & \beta < x < x_u. \end{cases} \tag{3.6}$$

We now establish some lemmas that explain the possible cases.

Lemma 3.1. *For $w > 0$, $0 \leq 1 - e^{-w} - w + \frac{1}{2}w^2 \leq \min\left(\frac{1}{2}w^2, \frac{1}{6}w^3\right)$.*

Proof. The inequality is easily established using multiple integrals:

$$0 \le 1 - e^{-w} - w + \frac{1}{2}w^2 = \int_0^w [e^{-u} - 1 + u] du = \int_0^w \int_0^u [1 - e^{-v}] dv du$$
$$= \int_0^w \int_0^u \int_0^v e^{-s} ds dv du \le \int_0^w \int_0^u \int_0^v 1 ds dv du = \frac{1}{6}w^3.$$

Also since $1 - w \le e^{-w}$, $0 \le 1 - e^{-w} - w + \frac{1}{2}w^2 \le \frac{1}{2}w^2$. □

We begin with the negligible case, where $c_0 < c_1 < 0$, item 1 in the beginning of the section.

Lemma 3.2. *For* $-c_0 < c_1 < 0$, $P(c_0 Z_0 + c_1 Z_1 > y) = o\left(e^{y/c_0}\right)$ *as* $y \to \infty$.

Proof. Let $F_{\frac{c_1}{c_0}Z_1}$ denote the distribution of $\frac{c_1}{c_0}Z_1$. Choose $\zeta < 0$ to be large in the negative direction. Then

$$P(c_0 Z_0 + c_1 Z_1 > y) = P\left(Z_0 + \frac{c_1}{c_0}Z_1 < \frac{y}{c_0}\right)$$
$$= \int_\zeta^\infty P\left(Z_0 + \frac{c_1}{c_0}Z_1 < \frac{y}{c_0}\right) dF_{\frac{c_1}{c_0}Z_1}(z) + \int_{-\infty}^\zeta P\left(Z_0 + \frac{c_1}{c_0}Z_1 < \frac{y}{c_0}\right) dF_{\frac{c_1}{c_0}Z_1}(z)$$
$$= (I) + (II).$$

Note that

$$\sup_{\zeta < z < \infty} P\left(Z_0 + \frac{c_1}{c_0}Z_1 < \frac{y}{c_0}\right) = \exp\left[-e^{-\left(\frac{y - c_1\zeta}{c_0}\right)}\right],$$

and therefore as $y \to \infty$

$$(I) \le \exp\left[-e^{-\left(\frac{y - c_1\zeta}{c_0}\right)}\right] \to 0$$

because for ζ fixed, $\frac{y - c_1\zeta}{c_0} \to -\infty$. Now observe that as $y \to \infty$

$$\frac{\exp\left[-e^{-\left(\frac{y - c_1\zeta}{c_0}\right)}\right]}{e^{y/c_0}} = \exp\left[-\frac{y}{c_0} - \exp\left[-e^{-\left(\frac{y - c_1\zeta}{c_0}\right)}\right]\right] \to 0,$$

and so $(I) = o\left(e^{y/c_0}\right)$. Next,

$$(II) \leq F_{\frac{c_1}{c_0}Z_1}(\zeta) = \exp\left[-e^{-c_0\zeta/c_1}\right] \to 0 \qquad \text{as} \qquad \zeta \downarrow -\infty,$$

and therefore $P(c_0 Z_0 + c_1 Z_1 > y) = o\left(e^{y/c_0}\right)$ as $y \to \infty$. □

We turn to the case where $-1 < \beta \leq 0$. The expansion is the same as (3.5), and Lemma 3.3 establishes the error term.

Lemma 3.3. *If $-1 < \beta \leq 0$, then $R = o(e^{-2y})$ as $y \to \infty$.*

Proof. Set $w = tx^{-\beta}$. By Lemma 3.1,

$$R = \int_0^\infty \left[1 - e^{-w} - w + \frac{1}{2}w^2\right] e^{-x} dx \leq \frac{1}{6}t^3 \int_0^\infty x^{-3\beta} e^{-x} dx = \frac{1}{6}\Gamma(1-3\beta)t^3.$$

Therefore $R = O(t^3) = o(t^2) = o(e^{-2y})$. □

From this point forward, we assume that $c_0, c_1 > 0$, which makes $0 < \beta < 1$.

Lemma 3.4. *If $0 < \beta < \frac{1}{2}$, then $R = O(e^{-y/\beta})$ as $y \to \infty$.*

Proof. Write $w = tx^{-\beta}$ with $t = e^{-y}$, and observe that

$$\begin{aligned} R &= \int_0^\infty \left[1 - e^{-w} - w + \frac{1}{2}w^2\right] e^{-x} dx \\ &= \int_0^\infty \left[1 - e^{-w} - w + \frac{1}{2}w^2\right] e^{-t^{1/\beta}w^{-1/\beta}} t^{1/\beta} \times \frac{1}{\beta} w^{-\frac{1}{\beta}-1} dw, \end{aligned}$$

because $x = \left(\frac{t}{w}\right)^{1/\beta}$ and therefore $\left|\frac{dx}{dw}\right| = \frac{1}{\beta}t^{1/\beta}w^{-\frac{1}{\beta}-1}$. We have that $3 - \frac{1}{\beta} - 1 = 2 - \frac{1}{\beta} < 0$

since $0 < \beta < \frac{1}{2}$. The goal is to show that

$$\int_0^\infty \left[1 - e^{-w} - w + \frac{1}{2}w^2\right] e^{-t^{1/\beta}w^{-1/\beta}} w^{-\frac{1}{\beta}-1} dw < \infty.$$

By Lemma 3.1, $0 \le \left|1 - e^{-w} - w + \frac{1}{2}w^2\right| \le \frac{1}{2}w^2$, and also $2 - \frac{1}{\beta} - 1 = 1 - \frac{1}{\beta} < -1$. Hence

$$\int_1^\infty \left[1 - e^{-w} - w + \frac{1}{2}w^2\right] e^{-t^{1/\beta}w^{-1/\beta}} w^{-\frac{1}{\beta}-1} dw < \int_1^\infty w^{-1-\frac{1}{\beta}} dw < \infty.$$

Next,

$$\begin{aligned}
\int_0^1 \left[1 - e^{-w} - w + \frac{1}{2}w^2\right] e^{-t^{1/\beta}w^{-1/\beta}} w^{-\frac{1}{\beta}-1} dw &\le \int_0^1 w^{2-\frac{1}{\beta}} e^{-t^{1/\beta}w^{-1/\beta}} dw \\
&= \int_1^\infty w^{\frac{1}{\beta}-2} e^{-t^{1/\beta}w^{1/\beta}} w^{-2} dw \qquad \text{(by putting } w^{-1} \text{ in for } w) \\
&= \int_1^\infty w^{-\beta\left(\frac{1}{\beta}-4\right)} e^{-t^{1/\beta}w} \beta w^{-\beta-1} dw \qquad \text{(by putting } w^{\beta} \text{ in for } w) \\
&= \int_1^\infty \beta w^{-3\beta} e^{-t^{1/\beta}w} dw < \infty.
\end{aligned}$$

Therefore for some $C > 0$, $R < Ct^{1/\beta}$ and thus $R = O\left(t^{1/\beta}\right) = O\left(e^{-y/\beta}\right)$. □

When $\beta = \frac{1}{2}$, the second term in (3.5) needs to be treated differently. Split the probability into two integrals with a two-term and a one-term expansion, respectively:

$$\begin{aligned}
P(Z_1 > y) &= \int_{t^2}^\infty \left[tx^{-\beta} - \frac{1}{2}t^2x^{-2\beta}\right] e^{-x} dx + \int_0^{t^2} tx^{-\beta} e^{-x} dx + R \\
&= t\int_0^\infty x^{-\frac{1}{2}} e^{-x} dx - \frac{t^2}{2}\int_{t^2}^\infty x^{-1} e^{-x} dx + R \\
&= \sqrt{\pi}t - \frac{t^2}{2}\int_{t^2}^\infty x^{-1} e^{-x} dx + R.
\end{aligned}$$

Observe that for some constant C,

$$\int_{t^2}^{\infty} x^{-1}e^{-x}dx = \int_{t^2}^{1} x^{-1}dx + \int_{t^2}^{1} x^{-1}\left[e^{-x} - 1\right] dx + \int_{1}^{\infty} x^{-1}e^{-x}dx$$
$$= -2\log(t) + C.$$

Therefore the approximation is

$$P(Z_1 > y) = \sqrt{\pi}t + t^2 \log(t) + Ct^2 + R. \tag{3.7}$$

The extra term Ct^2, as we shall see in Lemma 3.5, can be absorbed into the remainder.

Lemma 3.5. *When* $\beta = \frac{1}{2}$, *as* $y \to \infty$ $R = O\left(e^{-2y}\right)$ *in (3.7).*

Proof. Let $w = tx^{-\beta}$, $t = e^{-y}$, and $C_i, i = 1, 2, 3$ be positive constants. Split R into three integrals via

$$\begin{aligned} R = &\int_{1}^{\infty} \left[1 - e^{-w} - w + \frac{1}{2}w^2\right] e^{-x}dx + \int_{t^2}^{1} \left[1 - e^{-w} - w + \frac{1}{2}w^2\right] e^{-x}dx \\ &+ \int_{0}^{t^2} \left[1 - e^{-w} - w\right] e^{-x}dx \\ = &R_1 + R_2 + R_3. \end{aligned} \tag{3.8}$$

Examining the first integral in (3.8),

$$R_1 \le \int_{1}^{\infty} \frac{1}{6}w^3e^{-x}dx = \frac{1}{6}t^3 \int_{1}^{\infty} x^{-3\beta}e^{-x}dx = C_1t^3. \tag{3.9}$$

Next,

$$
\begin{aligned}
|R_2| &\le \left| \int_{t^2}^{1} \left[1 - e^{-w} - w + \frac{1}{2} w^2 \right] dx \right| = 2t^2 \left| \int_{t}^{1} \left[1 - e^{-w} - w + \frac{1}{2} w^2 \right] w^{-3} dw \right| \\
&\le \frac{1}{3} t^2 \int_t^1 1 dw = \frac{1}{3} t^2 (1-t) \le C_2 t^2,
\end{aligned} \tag{3.10}
$$

because the term involving t^3 is negligible. Finally, noting that $|1 - e^{-w} - w| \le w$ for $w > 0$,

$$
\begin{aligned}
|R_3| &\le \left| \int_0^{t^2} \left[1 - e^{-w} - w \right] dx \right| = 2t^2 \int_1^{\infty} |1 - e^{-w} - w| w^{-3} dw \le 2t^2 \int_1^{\infty} w^{-2} dw \\
&= C_3 t^2.
\end{aligned} \tag{3.11}
$$

Putting (3.9) through (3.11) in (3.8), $R = O(t^2) = O(e^{-2y})$. Consequently, as $y \to \infty$ the expansion (3.8) is

$$
P(Z_1 > y) = \sqrt{\pi} e^{-y} - y e^{-2y} + O(e^{-2y}). \tag{3.12}
$$

□

When $\frac{1}{2} < \beta < 1$, we cannot approximate the integral using the ordinary two-term expansion because the choice of β puts $1 - 2\beta < 0$. While $\Gamma(\alpha)$ is defined for negative nonintegers via $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$, the corresponding integral that would represent the gamma function in this case is improper. Instead, we employ an alternative usage of the one-term expansion to get a second term. Before presenting the answer, we begin with a lemma that will help simplify an integral that appears in the expansion.

Lemma 3.6. *Suppose* $\frac{1}{2} < \beta < 1$. *Then*

$$
\int_0^{\infty} \left(e^{-x} - 1 + x \right) x^{-\frac{1}{\beta} - 1} dx = \frac{\beta^2}{1 - \beta} \Gamma\left(2 - \beta^{-1} \right).
$$

Proof. We first compute the integral over the interval (δ, ∞) where $\delta > 0$. Integration by parts gives

$$\beta\left(e^{-\delta}-1+\delta\right)\delta^{-\frac{1}{\beta}}+\beta\int_{\delta}^{\infty}\left(1-e^{-x}\right)x^{-\frac{1}{\beta}}dx.$$

Integrating by parts a second time results in

$$\begin{aligned}
&\beta\left(e^{-\delta}-1+\delta\right)\delta^{-\frac{1}{\beta}}+\frac{\beta^2}{1-\beta}(1-e^{-\delta})\delta^{1-\frac{1}{\beta}}+\frac{\beta^2}{1-\beta}\int_{\delta}^{\infty}x^{1-\frac{1}{\beta}}e^{-x}dx\\
&=\frac{\beta}{1-\beta}\left(\frac{\delta+(1-\beta)e^{-\delta}-\beta\delta e^{-\delta}-(1-\beta)}{\delta^{\frac{1}{\beta}}}\right)+\frac{\beta^2}{1-\beta}\int_{\delta}^{\infty}x^{1-\frac{1}{\beta}}e^{-x}dx \qquad (3.13)\\
&=(I)+(II).
\end{aligned}$$

Now we send $\delta \downarrow 0$. (II) gives the gamma function:

$$\lim_{\delta\downarrow 0}(II)=\frac{\beta^2}{1-\beta}\int_{0}^{\infty}x^{\left(2-\frac{1}{\beta}\right)-1}e^{-x}dx=\frac{\beta^2}{1-\beta}\Gamma\left(2-\beta^{-1}\right). \qquad (3.14)$$

Using L'Hopital's Rule twice on (I),

$$\begin{aligned}
\lim_{\delta\downarrow 0}(I)&=\lim_{\delta\downarrow 0}\frac{\beta^2}{1-\beta}\left(\frac{1-(1-\beta)e^{-\delta}-\beta e^{-\delta}+\beta\delta e^{-\delta}}{\delta^{\frac{1}{\beta}-1}}\right)\\
&=\lim_{\delta\downarrow 0}\frac{\beta^2}{1-\beta}\left(\frac{1-e^{-\delta}+\beta\delta e^{-\delta}}{\delta^{\frac{1}{\beta}-1}}\right)=\lim_{\delta\downarrow 0}\frac{\beta^3}{(1-\beta)^2}\left(\frac{e^{-\delta}+\beta e^{-\delta}-\beta\delta e^{-\delta}}{\delta^{\frac{1}{\beta}-2}}\right) \qquad (3.15)\\
&=\lim_{\delta\downarrow 0}\frac{\beta^3}{(1-\beta)^2}\left(\frac{(1+\beta-\beta\delta)\delta^{2-\frac{1}{\beta}}}{e^{\delta}}\right)=0.
\end{aligned}$$

The fact that $0 < 2-\frac{1}{\beta} < 1$ was used in the final calculation. Putting (3.14) and (3.15) into (3.13),

$$\int_{\delta}^{\infty}\left(e^{-x}-1+x\right)x^{-\frac{1}{\beta}-1}dx\to\frac{\beta^2}{1-\beta}\Gamma\left(2-\beta^{-1}\right)\qquad\text{as}\qquad\delta\downarrow 0.$$

Next, we examine $\int_0^\delta (e^{-x} - 1 + x) x^{-\frac{1}{\beta}-1} dx$. Observe that

$$\lim_{x \downarrow 0} \frac{e^{-x} - 1 + x}{x^2} = \lim_{x \downarrow 0} \frac{-e^{-x} + 1}{2x} = \lim_{x \downarrow 0} \frac{e^{-x}}{2} = \frac{1}{2},$$

and so $e^{-x} - 1 + x \sim \frac{1}{2} x^2$ as $x \downarrow 0$. Then

$$\int_0^\delta \left(e^{-x} - 1 + x\right) x^{-\frac{1}{\beta}-1} dx \sim \frac{1}{2} \int_0^\delta x^{1-\frac{1}{\beta}} dx \to 0,$$

because $1 - \frac{1}{\beta} > -1$. The result follows. □

We are now ready to derive the expansion.

Lemma 3.7. *Suppose $\frac{1}{2} < \beta < 1$. Then as $y \to \infty$*

$$P(Z_1 > y) = \Gamma(1-\beta) e^{-y} - \frac{\beta}{1-\beta} \Gamma\left(2 - \beta^{-1}\right) e^{-y/\beta} + o\left(e^{-y/\beta}\right).$$

Proof. Write $t = e^{-y}$ and observe that

$$\frac{P(Z_1 > y) - \Gamma(1-\beta) t}{t^{1/\beta}} = \int_0^\infty \frac{1 - \exp(-t x^{-\beta}) - t x^{-\beta}}{t^{1/\beta}} e^{-x} dx. \tag{3.16}$$

Use the change of variables $w = t^{-1/\beta} x$ to get the integral

$$\int_0^\infty \left[1 - \exp(-w^{-\beta}) - w^{-\beta}\right] \exp(-t^{1/\beta} w) dw.$$

Now use another change of variables $x = w^{-\beta}$ to turn the integral into

$$-\frac{1}{\beta} \int_0^\infty \left(e^{-x} - 1 + x\right) \exp(-t^{1/\beta} x^{-1/\beta}) x^{-\frac{1}{\beta}-1} dx.$$

Next, observe that by Lemma 3.6

$$\int_0^\infty \left(e^{-x} - 1 + x\right) \exp(-t^{1/\beta} x^{-1/\beta}) x^{-\frac{1}{\beta}-1} dx \leq \int_0^\infty \left(e^{-x} - 1 + x\right) x^{-\frac{1}{\beta}-1} dx$$
$$= \frac{\beta^2}{1-\beta} \Gamma\left(2 - \beta^{-1}\right) < \infty,$$

so by dominated convergence

$$\lim_{t \downarrow 0} -\frac{1}{\beta} \int_0^\infty \left(e^{-x} - 1 + x\right) \exp(-t^{1/\beta} x^{-1/\beta}) x^{-\frac{1}{\beta}-1} dx = -\frac{\beta}{1-\beta} \Gamma\left(2 - \beta^{-1}\right).$$

Hence as $t \downarrow 0$

$$\frac{P(Z_1 > y) - \Gamma(1-\beta)t}{t^{1/\beta}} = -\frac{\beta}{1-\beta} \Gamma\left(2 - \beta^{-1}\right) + o(1). \tag{3.17}$$

Rearranging terms, as $t \downarrow 0$

$$P(Z_1 > y) = \Gamma(1-\beta)t - \frac{\beta}{1-\beta} \Gamma\left(2 - \beta^{-1}\right) t^{1/\beta} + o(t^{1/\beta}),$$

and finally, as $y \to \infty$

$$P(Z_1 > y) = \Gamma(1-\beta)e^{-y} - \frac{\beta}{1-\beta} \Gamma\left(2 - \beta^{-1}\right) e^{-y/\beta} + o\left(e^{-y/\beta}\right),$$

as required. □

We have completed the proof for each value of $-1 < \beta < 1$. Theorem 3.1 summarizes the results from this section.

Theorem 3.1. *Let $Z_1|Z_0 \sim GEV(0, \beta Z_0, 1)$ and $Z_0 \sim \Lambda$. Then as $y \to \infty$*

$$P(Z_1 > y) = \begin{cases} \Gamma(1-\beta)e^{-y} - \frac{1}{2}\Gamma(1-2\beta)e^{-2y} + o(e^{-2y}), & -1 < \beta \leq 0 \\ \Gamma(1-\beta)e^{-y} - \frac{1}{2}\Gamma(1-2\beta)e^{-2y} + O(e^{-y/\beta}), & 0 < \beta < \frac{1}{2} \\ \sqrt{\pi}e^{-y} - ye^{-2y} + O(e^{-2y}), & \beta = \frac{1}{2} \\ \Gamma(1-\beta)e^{-y} - \frac{\beta}{1-\beta}\Gamma\left(2-\beta^{-1}\right)e^{-y/\beta} + o\left(e^{-y/\beta}\right), & \frac{1}{2} < \beta < 1. \end{cases}$$

Having an alternative statement when two weights are involved would be useful as well, as in the probability in (3.2).

Corollary 3.1. *Let $Z_1, Z_0 \sim \Lambda$ and $c_1 > c_0 > 0$. Set $V = c_1 Z_1 + c_0 Z_0$. Then as $y \to \infty$*

$$P(V > y) = \begin{cases} o\left(e^{y/c_0}\right) & , c_0 < c_1 < y \\ \Gamma(1-\frac{c_0}{c_1})e^{-y/c_1} - \frac{1}{2}\Gamma(1-\frac{2c_0}{c_1})e^{-2y/c_1} + o(e^{-2y}), & c_0 < 0 < c_1 \\ \Gamma(1-\frac{c_0}{c_1})e^{-y/c_1} - \frac{1}{2}\Gamma(1-\frac{2c_0}{c_1})e^{-2y/c_1} + O(e^{-y/c_0}), & 0 < 2c_0 < c_1 \\ \sqrt{\pi}e^{-y/c_1} - \frac{y}{c_1}e^{-2y/c_1} + O(e^{-2y}), & 0 < 2c_0 = c_1 \\ \Gamma(1-\frac{c_0}{c_1})e^{-y/c_1} - \frac{c_0}{c_1-c_0}\Gamma\left(2-\frac{c_1}{c_0}\right)e^{-y/c_0} + o\left(e^{-y/c_0}\right), & 0 < c_1 < 2c_0. \end{cases}$$

We have now laid the groundwork for when the random variables are standard Gumbels. Now we begin generalizing our results. First, we assume that the scale and original location parameters are fixed. Let $X_i \sim \text{GEV}(0, \mu, \sigma)$, or put another way $X_i = \mu + \sigma Z_i$ where $Z_i \sim \Lambda, i = 0, 1$ and $\sigma > 0$. The easiest way to address the probability is to rewrite

$$P(X_1 + \beta X_0 > y) = P(\mu + \sigma Z_1 + \beta(\mu + \sigma Z_0) > y) = P\left(Z_1 + \beta Z_0 > \frac{y - (1+\beta)\mu}{\sigma}\right).$$

We have successfully rewritten the probability in a familiar format. The updated theorem and corollary are stated below, and their proofs are similar to those of Lemmas 3.3 through 3.7.

Theorem 3.2. *Let* $X_1|X_0 \sim GEV(0, \mu + \beta X_0, \sigma)$ *and* $X_0 \sim GEV(0, \mu, \sigma)$*. Define* $y^* = \frac{y-(1+\beta)\mu}{\sigma}$*. Then as* $y \to \infty$

$$P(X_1 > y) = \begin{cases} \Gamma(1-\beta)e^{-y^*} - \frac{1}{2}\Gamma(1-2\beta)e^{-2y^*} + o(e^{-2y/\sigma}), & -1 < \beta < 0 \\ \Gamma(1-\beta)e^{-y^*} - \frac{1}{2}\Gamma(1-2\beta)e^{-2y^*} + O(e^{-y/\sigma\beta}), & 0 < \beta < \frac{1}{2} \\ \sqrt{\pi}e^{-y^*} - \frac{y}{\sigma}e^{-2y^*} + O(e^{-2y/\sigma}), & \beta = \frac{1}{2} \\ \Gamma(1-\beta)e^{-y^*} - \frac{\beta}{1-\beta}\Gamma\left(2-\beta^{-1}\right)e^{-y^*/\beta} + o\left(e^{-y/\beta\sigma}\right), & \frac{1}{2} < \beta < 1. \end{cases}$$

Corollary 3.2. *Let* $X_i = \mu + \sigma Z_i, Z_i \sim \Lambda, i = 0, 1,$ *and* $c_1 > c_0 > 0$*. Define* $\beta = \frac{c_1}{c_0}$*,* $y^* = \frac{y-(1+\beta)\mu}{\sigma}$*, and* $V = c_1X_1 + c_0X_0$*. Then as* $y \to \infty$

$$P(V > y) = \begin{cases} o\left(e^{y/c_0\sigma}\right), & c_0 < c_1 < 0 \\ \Gamma(1-\frac{c_0}{c_1})e^{-y^*/c_1} - \frac{1}{2}\Gamma(1-\frac{2c_0}{c_1})e^{-2y^*/c_1} + o(e^{-2y/\sigma}), & c_0 < 0 < c_1 \\ \Gamma(1-\frac{c_0}{c_1})e^{-y^*/c_1} - \frac{1}{2}\Gamma(1-\frac{2c_0}{c_1})e^{-2y^*/c_1} + O(e^{-y/c_0\sigma}), & 2c_0 < c_1 \\ \sqrt{\pi}e^{-y^*/c_1} - \frac{y}{c_1\sigma}e^{-2y^*/c_1} + O(e^{-2y/\sigma}), & 2c_0 = c_1 \\ \Gamma(1-\frac{c_0}{c_1})e^{-y^*/c_1} - \frac{c_0}{c_1-c_0}\Gamma\left(2-\frac{c_1}{c_0}\right)e^{-y^*/c_0} + o\left(e^{-y/c_0\sigma}\right), & 2c_0 > c_1. \end{cases}$$

Now suppose $X_1|X_0 \sim \text{GEV}(0, \mu_1 + \beta X_0, \sigma_1)$ and $X_0 \sim \text{GEV}(0, \mu_0, \sigma_0)$ with $\sigma_0, \sigma_1 > 0$. To work with the probability,

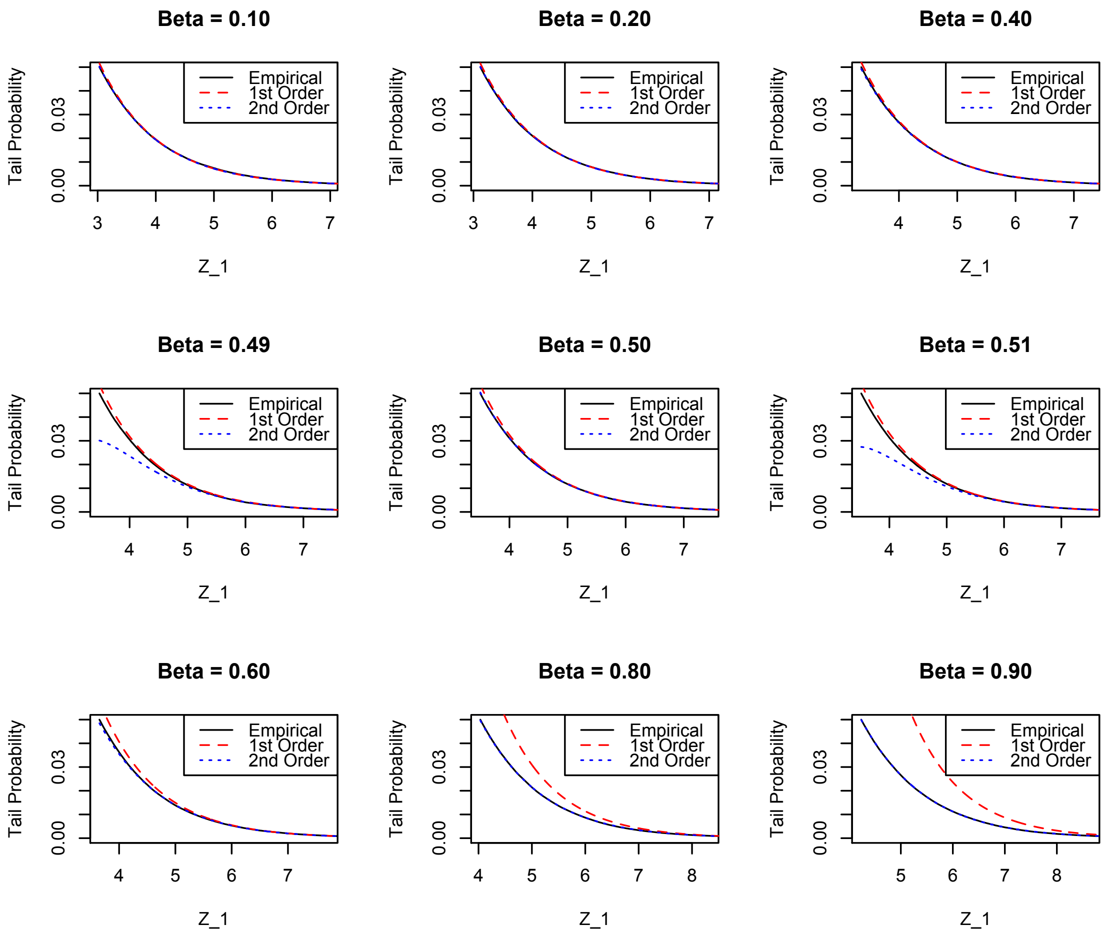
$$\begin{aligned}P(X_1 + \beta X_0 > y) &= P(\mu_1 + \sigma_1 Z_1 + \beta(\mu_0 + \sigma_0 Z_0) > y) \\ &= P\left(\sigma_1 Z_1 + \beta\sigma_0 Z_0 > y - \mu_1 - \beta\mu_0\right).\end{aligned}$$

One could then proceed to derive an extension of Theorem 3.1 and Corollary 3.1 by dividing the Z_i by the larger of σ_1 and $\beta\sigma_0$.

3.3 Simulation Results for Gumbel Mixture

Before building the complete AR(1) process in Chapter 4, we check how our three approximations behave for various values of β. We simulate $N = 10$ million values of Z_1 and graph the empirical tail probability, as well as our second-order approximation in Theorem 3.1. For comparison purposes we also graph the first-order approximation. We investigate what happens at $\beta = \{0.10, 0.20, 0.40, 0.49, 0.50, 0.51, 0.60, 0.80, 0.90\}$. The reason for considering $\beta = 0.49$ and 0.51 is to see what is happening near the $\beta = \frac{1}{2}$ pivot point. Figure 3.1 displays the empirical probability (solid black line), the first-order approximation (dashed red line), and the second-order approximation (dotted blue line) for the 95th percentile and higher.

Figure 3.1: Approximations for Various Values of β



Here are some observations. First, for $0 < \beta < 0.40$ the two approximations are virtually indistinguishable and estimate extremely accurately. Second, when $\beta = 0.49$ the first-order approximation estimates the empirical probability very well, while the second-order approximation underestimates. Only around the 99th percentile does the latter finally catch up. The same observation applies to when $\beta = 0.51$, suggesting that the second-order does not behave very well when β is very close to 0.50 on either side. This is not a surprise, since 0.50 is the pivot point at which the approximation changes terms. Next, when $\beta = 0.50$ both formulas estimate the probability reasonably well. Finally, as β moves upward to 1 the second-order approximation estimates extremely well, but the first-order overestimates. The discrepancy becomes worse as β gets closer to 1.

In Table 3.1 we look at the errors in both approximations. We define an error to be the empirical probability minus the estimated probability. Therefore a positive error indicates an underestimate, and a negative error an overestimate.

Table 3.1: Errors in Approximations of Theorem 3.1

β	**Approx**	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
0.10	1st	-0.0013	-0.0003	-0.0002	-8.0e-5	-3.9e-5	-1.1e-5	-1.9e-6
	2nd	1.1e-5	3.2e-5	9.1e-7	-2.8e-5	-2.6e-5	-1.1e-5	-1.9e-6
0.20	1st	-0.0013	-0.0003	-0.0002	-4.0e-5	-1.5e-5	-3.4e-5	-9.4e-6
	2nd	0.0002	5.2e-5	2.7e-5	1.6e-5	-1.2e-6	-1.3e-5	-9.4e-6
0.40	1st	-0.0024	-0.0006	-0.0004	-0.0001	-4.7e-5	-8.5e-6	6.1e-6
	2nd	0.0004	3.2e-5	1.6e-5	-1.6e-5	-2.1e-5	-7.5e-6	6.1e-6
0.49	1st	-0.0037	-0.0011	-0.0007	-0.0002	-5.5e-5	1.5e-6	6.3e-7
	2nd	0.0199	0.0045	0.0028	0.0006	0.0002	9.6e-6	7.1e-7
0.50	1st	-0.0039	-0.0011	-0.0007	-0.0002	-4.4e-5	-4.5e-6	-2.4e-6
	2nd	-0.0006	-0.0002	-7.9e-5	-2.3e-5	3.4e-6	-2.1e-6	-2.4e-6
0.51	1st	-0.0039	-0.0011	-0.0008	-0.0002	-7.5e-5	6.7e-7	-2.1e-6
	2nd	0.0226	0.0053	0.0033	0.0008	0.0002	1.1e-5	-2.0e-6
0.60	1st	-0.0070	-0.0021	-0.0015	-0.0004	-0.0001	-6.4e-6	4.5e-6
	2nd	0.0020	0.0005	0.0003	8.3e-5	3.1e-5	4.3e-6	4.7e-6
0.80	1st	-0.0309	-0.0114	-0.0084	-0.0032	-0.0013	-0.0001	-4.3e-6
	2nd	0.0006	0.0002	0.0001	3.1e-5	1.8e-5	2.7e-5	3.4e-6
0.90	1st	-0.0865	-0.0344	-0.0258	-0.0107	-0.0045	-0.0006	-3.3e-5
	2nd	0.0003	5.5e-5	2.6e-5	-1.9e-5	-7.9e-6	6.3e-6	6.0e-6

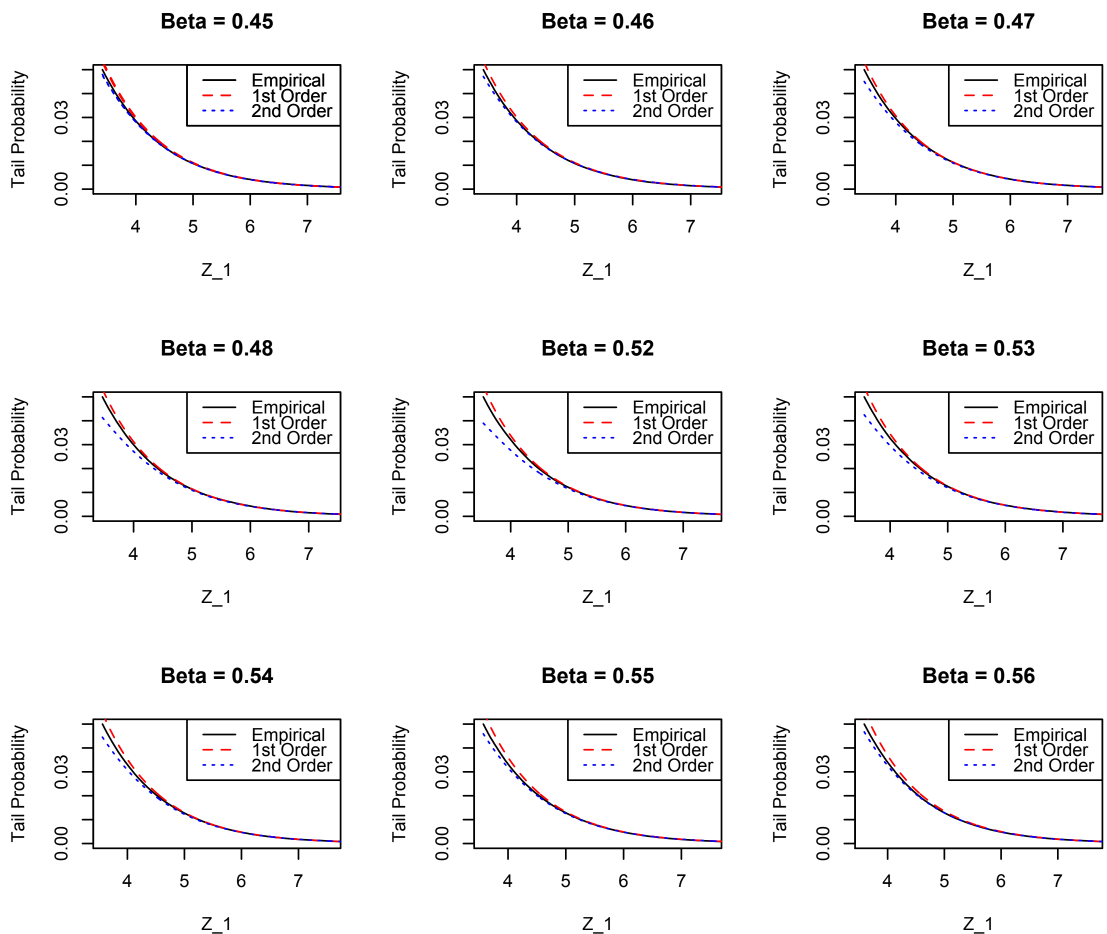
The results support the observations we have already made from the graphs. We now take a closer look by examining the relative errors in Table 3.2. Define a relative error to be the error from Table 3.1 divided by the corresponding approximated probability.

Table 3.2: Relative Errors in Approximations of Theorem 3.1

β	**Approx**	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
0.10	1st	-0.0259	-0.0116	-0.0103	-0.0079	-0.0078	-0.0112	-0.0184
	2nd	0.0002	0.0013	0.0000	-0.0028	-0.0052	-0.0107	-0.0183
0.20	1st	-0.0251	-0.0118	-0.0098	-0.0040	-0.0030	-0.0138	-0.0863
	2nd	0.0031	0.0021	0.0013	0.0016	-0.0002	-0.0132	-0.0863
0.40	1st	-0.0462	-0.0253	-0.0204	-0.0121	-0.0094	-0.0085	0.0653
	2nd	0.0086	0.0013	0.0008	-0.0016	-0.0042	-0.0074	0.0654
0.49	1st	-0.0682	-0.0409	-0.0355	-0.0223	-0.0110	0.0015	0.0063
	2nd	0.6608	0.2190	0.1615	0.0670	0.0317	0.0097	0.0072
0.50	1st	-0.0718	-0.0418	-0.0331	-0.0190	-0.0087	-0.0045	-0.0239
	2nd	-0.0126	-0.0070	-0.0039	-0.0023	0.0007	-0.0021	-0.0235
0.51	1st	-0.0732	-0.0427	-0.0363	-0.0217	-0.0147	0.0007	-0.0205
	2nd	0.8245	0.2680	0.1994	0.0864	0.0380	0.0114	-0.0193
0.60	1st	-0.1225	-0.0785	-0.0686	-0.0429	-0.0256	-0.0064	0.0475
	2nd	0.0418	0.0195	0.0149	0.0084	0.0063	0.0043	0.0499
0.80	1st	-0.3821	-0.3137	-0.2953	-0.2451	-0.2023	-0.1093	-0.0409
	2nd	0.0114	0.0073	0.0058	0.0031	0.0037	0.0279	0.0354
0.90	1st	-0.6337	-0.5791	-0.5629	-0.5169	-0.4740	-0.3850	-0.2495
	2nd	0.0069	0.0022	0.0013	-0.0019	-0.0016	0.0064	0.0639

Comparing the relative error sizes for both formulas, Table 3.2 shows that the second-order approximation gives more accurate estimation than the first-order for all $0 < \beta < 1$, except for a small neighborhood around but not including $\beta = 0.50$. That is, for some $\delta_1, \delta_2 > 0$ the first-order approximation is more accurate for $\beta \in (0.50 - \delta_1, 0.50 + \delta_2) \setminus \{0.50\}$. The following graphs in Figure 3.2 suggest that such a neighborhood may be $[0.47, 0.54]$. As a side note, we also examined the results for $-1 < \beta < 0$, and again the second-order approximates more accurately. Therefore, we exclude those results from the dissertation.

Figure 3.2: Approximations in a Neighborhood of $\beta = 0.50$



To conclude, when working with just two Gumbel random variables, we recommend the following procedure:

1. Use the first-order approximation for $\beta \in [0.47, 0.50) \cup (0.50, 0.54]$.

2. Use the second order approximation for all other values, namely $\beta \in (-1, 0.47) \cup \{0.50\} \cup (0.54, 1)$.

3.4 Mixture of Fréchet Random Variables

The idea of two-term expansions for Gumbel mixtures can be extended to the Fréchet and Weibull families. Such an extension is useful because all three extreme value families would be complete. However, Chapters 4 through 6 return to the Gumbel case.

Let $\Phi_\alpha(x) = e^{-x^{-\alpha}}, x > 0, \alpha > 0$ denote the Fréchet model, as in (1.4). In terms of the generalized extreme value family, this definition is given by GEV $\left(\frac{1}{\alpha}, 1, \frac{1}{\alpha}\right)$.

Suppose $X_1 \sim \Phi_\alpha$ and

$$(X_2|X_1 = x) \sim \Phi_\alpha(\cdot - \beta x), \qquad \cdot > \beta x.$$

That is, $X_2|X_1 \sim$ GEV $\left(\frac{1}{\alpha}, \beta x + 1, \frac{1}{\alpha}\right)$ on $\{x : x - \beta X_1 > 0\}$. Next, notice that for $z > 0$

$$P\left(X_2 - \beta X_1 \le z|X_1 = x\right) = P\left(X_2 \le z + \beta x|X_1 = x\right) = \Phi_\alpha(z).$$

Thus $X_2 - \beta X_1$ and X_1 are independent and $X_2 - \beta X_1 \sim \Phi_\alpha$. In other words, the model is $X_2 = \beta X_1 + Z$ where Z and X_1 are independent and $Z \sim \Phi_\alpha$. Also by definition, $1 - \Phi_\alpha$ is $RV_{-\alpha}$.

We shall use a result from Barbe and McCormick (2005) to obtain a two-term expansion for $P(X_2 > y)$. First we look at a couple of definitions.

Definition 3.2. *A distribution function F is said to be asymptotically smooth with index $-\alpha$ if*

$$\lim_{\delta \to 0} \limsup_{y \to \infty} \sup_{0 \le |x| \le \delta} \left| \frac{F(y(1-x)) - F(y)}{xF(y)} - \alpha \right| = 0.$$

Definition 3.3. *A distribution function F is right-tail dominant if for any $\delta > 0$,*

$$\lim_{y \to \infty} \frac{F(-y\delta)}{\overline{F}(y)} = 0.$$

The theorem is stated below.

Theorem 3.3. *Let F and G be two distribution functions such that $\overline{F}, \overline{G} \in RV_{-\alpha}$ with $\alpha \geq 1$. It is supposed that F and G are asymptotically smooth and right-tail dominant with $\int_{-\infty}^{0} ydF(y)$ and $\int_{-\infty}^{0} ydG(y)$ both finite. Denote by $F \star G(y)$ the convolution of F and G. Then as $y \to \infty$*

$$1 - F \star G(y) = \overline{F}(y) + \overline{G}(y) + \frac{\alpha}{y}\left[\overline{F}(y)\mu_G(y) + \overline{G}(y)\mu_F(y)\right](1 + o(1)), \tag{3.18}$$

where $\mu_F(y) = \int_{-y}^{y} xdF(x), y \geq 0$ denotes the truncated mean of F, and similarly for $\mu_G(y)$. In particular, if the two means are finite and equal μ_F and μ_G respectively, then the conclusion is

$$1 - F \star G(y) = \overline{F}(y) + \overline{G}(y) + \frac{\alpha}{y}\left[\overline{F}(y)\mu_G + \overline{G}(y)\mu_F\right](1 + o(1)). \tag{3.19}$$

Provided the assumptions are met, this result will take care of the case where $\alpha \geq 1$. Luckily we need not check asymptotic smoothness on F with the definition above, thanks to the following additional result from Barbe and McCormick (2005).

Lemma 3.8. *If F has an ultimately monotone density, then F is asymptotically smooth.*

The density of F is ultimately decreasing since it can be shown that $F''(y) \sim -\alpha(\alpha+1)y^{-\alpha-2}$ as $y \to \infty$. Thus, F is asymptotically smooth, and further is right-tail dominant because $F(-y\delta) = 0$ for any $y > 0$. As for the last assumption in the theorem, $\int_{-\infty}^{0} ydF(y) = 0$ because $F(y) = 0$ on the negative reals. One more lemma is needed before we derive the expansion.

Lemma 3.9. *If $Z \sim \Phi_\alpha$, then $P(Z > y) = y^{-\alpha} - \frac{1}{2}y^{-2\alpha} + o(y^{-2\alpha})$ as $y \to \infty$.*

Proof. Write

$$1 - \Phi_\alpha(y) - y^{-\alpha} = \int_y^\infty \left[e^{-x^{-\alpha}} - 1\right] \alpha x^{-\alpha-1} dx = -\int_y^\infty \left(\frac{1 - e^{-x^{-\alpha}}}{x^{-\alpha}}\right) \alpha x^{-2\alpha-1} dx.$$

Note that $\phi(z) = \frac{1-e^{-z}}{z} \to 1$ as $z \to 0$. Consider

$$\frac{\displaystyle\int_y^\infty \phi(x^{-\alpha}) \alpha x^{-2\alpha-1} dx}{\int_y^\infty \alpha x^{-2\alpha-1} dx} = \frac{\int_y^\infty \phi(x^{-\alpha}) \alpha x^{-2\alpha-1} dx}{\frac{1}{2} y^{-2\alpha}} = 2\alpha y^{2\alpha} \int_y^\infty \phi(x^{-\alpha}) x^{-2\alpha-1} dx.$$

A change of variables brings

$$2\alpha y^{2\alpha} \int_y^\infty \phi(x^{-\alpha}) x^{-2\alpha-1} dx = 2\alpha \int_1^\infty \phi((yx)^{-\alpha}) x^{-2\alpha-1} dx.$$

Thus by dominated convergence

$$\frac{1 - \Phi_\alpha(y) - y^{-\alpha}}{\frac{1}{2} y^{-2\alpha}} = -1 + o(1) \quad \text{as} \quad y \to \infty,$$

and the result follows. □

We may therefore use Theorem 3.3 and Lemma 3.9 on the Fréchet mixture as follows. If

$\alpha > 1$, then $E(Z) = \mu_F$ is finite and is equal to $\Gamma(1-\alpha^{-1})$. The probability as $y \to \infty$ is

$$\begin{aligned}
P(X_2 > y) &= P(\beta X_1 + Z > y) \\
&= \frac{\alpha}{y}\left[P(\beta X_1 > y)E(Z) + P(Z > y)E(\beta X_1)\right](1+o(1)) + P(\beta X_1 > y) + P(Z > y) \\
&= \beta^\alpha y^{-\alpha} - \frac{1}{2}\beta^{2\alpha} y^{-2\alpha} + y^{-\alpha} - \frac{1}{2} y^{-2\alpha} + o(y^{-2\alpha}) + \alpha y^{-1} E(Z)\Bigg[\beta^\alpha y^{-\alpha} - \frac{1}{2}\beta^{2\alpha} y^{-2\alpha} \\
&\qquad + \beta y^{-\alpha} - \frac{1}{2}\beta y^{-2\alpha} + o(y^{-2\alpha})\Bigg](1+o(1)) \\
&= (1+\beta^\alpha) y^{-\alpha} + \alpha\beta(1+\beta^{\alpha-1})\Gamma\left(1-\alpha^{-1}\right) y^{-\alpha-1} + o(y^{-\alpha-1}).
\end{aligned}$$

Next, if $\alpha = 1$ then $E(Z)$ is infinite and we must use Theorem 3.3 with the truncated mean. The two-term expansion is then

$$P(X_2 > y) = (1+\beta)y^{-1} + \beta\left(\int_0^y x^{-1}e^{-x^{-1}}dx\right) y^{-2} + o(y^{-2}).$$

It is easily checked that $\int_0^y x^{-1}e^{-x^{-1}}dx < \infty$ and that $\left(\int_0^y x^{-1}e^{-x^{-1}}dx\right) y^{-2} \to 0$ as $y \to \infty$. We now turn to the case where $0 < \alpha < 1$ using another result from Barbe and McCormick (2005).

Theorem 3.4. *Let F and G be asymptotically smooth distribution functions with support on the positive reals such that $\overline{F}, \overline{G} \in RV_{-\alpha}, 0 < \alpha < 1$. Define the quantity*

$$I(\alpha) = \int_0^{1/2} \left((1-t)^{-\alpha} - 1\right) \alpha t^{-\alpha-1} dt. \tag{3.20}$$

Then

$$\lim_{y\to\infty} \frac{1 - F \star G(y) - \overline{F}(y) - \overline{G}(y)}{\overline{F}(y)\overline{G}(y)} = 2I(\alpha) + 2^{2\alpha} - 2^{\alpha+1}. \tag{3.21}$$

Setting $\Theta = 2I(\alpha) + 2^{2\alpha} - 2^{\alpha+1}$*, as* $y \to \infty$ *the statement can be restated as*

$$1 - F \star G(y) = \overline{F}(y) + \overline{G}(y) + \Theta\overline{F}(y)\overline{G}(y) + o(\overline{F}(y)\overline{G}(y)). \tag{3.22}$$

We now use Theorem 3.4 and Lemma 3.9 to compute a two-term expansion for $P(X_2 > y)$:

$$\begin{aligned} P(X_2 > Y) &= P(\beta X_1 > y) + P(Z > y) + \Theta P(\beta X_1 > y)P(Z > y) \\ &\quad + o\left(P(\beta X_1 > y)P(Z > y)\right) \\ &= (1+\beta^\alpha)y^{-\alpha} - \frac{1}{2}(1+\beta^{2\alpha})y^{-2\alpha} + o(y^{-2\alpha}) + (II) + (III). \end{aligned}$$

Examining the second piece,

$$\begin{aligned} (II) &= \Theta\left[\beta^\alpha y^{-\alpha} - \frac{1}{2}\beta^{2\alpha}y^{-2\alpha} + o(y^{-2\alpha})\right]\left[y^{-\alpha} - \frac{1}{2}y^{-2\alpha} + o(y^{-2\alpha})\right] \\ &= \Theta\beta^\alpha y^{-2\alpha} + o(y^{-2\alpha}). \end{aligned}$$

Then $(III) = o(y^{-2\alpha})$ and the expansion becomes

$$P(X_2 > y) = (1+\beta^\alpha)y^{-\alpha} - \frac{1}{2}(1+\beta^{2\alpha} - 2\Theta\beta^\alpha)y^{-2\alpha} + o(y^{-2\alpha}).$$

We have finally provided a two-term expansion for the Fréchet mixture for all $\alpha > 0$, summarizing the results in the next theorem.

Theorem 3.5. *Let* $(X_2|X_1 = x) \sim \Phi_\alpha(\cdot - \beta x)$ *and* $X_1 \sim \Phi_\alpha$ *with* $\alpha, \beta > 0$. *Then as* $y \to \infty$

$$P(X_2 > y) = \begin{cases} (1+\beta^\alpha)y^{-\alpha} + \alpha\beta(1+\beta^{\alpha-1})\Gamma\left(1-\alpha^{-1}\right)y^{-\alpha-1} & \\ \qquad + o(y^{-\alpha-1}), & \alpha > 1 \\ (1+\beta)y^{-1} + \beta\left(\int_0^y x^{-1}e^{-x^{-1}}dx\right)y^{-2} + o(y^{-2}), & \alpha = 1 \\ (1+\beta^\alpha)y^{-\alpha} - \frac{1}{2}(1+\beta^{2\alpha} - 2\Theta\beta^\alpha)y^{-2\alpha} + o(y^{-2\alpha}), & 0 < \alpha < 1. \end{cases} \tag{3.23}$$

3.5 Mixture of Weibull Random Variables

Finally, we extend the mixture analysis to the Type III, or Weibull, case. Such an answer would be a useful contribution to extreme value theory since the Weibull distribution is very common in practice. First introduced in Fréchet (1927), it arises in applications relating to decay or failure times. Weibull (1951) provides many examples of data sets in which the Weibull was modeled, including a study on Indian cotton fiber strength and another study concerning the stature of adult males in the British Isles. Pinder *et al.* (1978) fits Weibull models to survivorship curves of various birds.

As stated in (1.5), for $\alpha > 0$, denote the Weibull family by

$$\Psi_\alpha(x) = \begin{cases} e^{-(-x)^\alpha}, & x < 0 \\ 1, & x \geq 0. \end{cases} \tag{3.24}$$

In terms of the generalized extreme value family, $\Psi_\alpha = \text{GEV}\left(-\frac{1}{\alpha}, -1, \frac{1}{\alpha}\right)$. Suppose $X_1 \sim \Psi_\alpha$ and $\{X_2|X_1 = x\} \sim \Psi_\alpha(\cdot - \beta x), \cdot > \beta x$ where $\beta > 0$. That is,

$$X_2|X_1 \sim \text{GEV}\left(-\frac{1}{\alpha}, \beta X_1 - 1, \frac{1}{\alpha}\right) \quad \text{on} \quad x - \beta X_1 < 0.$$

Now observe that for $y < 0$

$$\begin{aligned} P(X_2 > y) &= \int_{-\infty}^{0} P(X_2 > y | X_1 = x) \frac{d}{dx} (\Psi_\alpha(x)) \\ &= \int_{-\infty}^{0} P(X_2 - \beta X_1 > y - \beta x | X_1 = x) \frac{d}{dx} (\Psi_\alpha(x)) \\ &= \int_{-\infty}^{y/\beta} (1-1) \frac{d}{dx} (\Psi_\alpha(x)) + \int_{y/\beta}^{0} [1 - \Psi_\alpha(y - \beta x)] \frac{d}{dx} (\Psi_\alpha(x)) \\ &= \int_{y/\beta}^{0} [1 - \Psi_\alpha(y - \beta x)] \frac{d}{dx} (\Psi_\alpha(x)) . \end{aligned}$$

The last step follows because $\Psi_\alpha(x) = 1$ for $x \in (-\infty, \frac{y}{\beta}]$, and since $0 \leq \frac{d}{dx} (\Psi_\alpha(x)) < \alpha$ the first integral is 0 by dominated convergence. We proof the following lemma before deriving the expansion.

Lemma 3.10. *For $y < 0$,*

$$1 - \Psi_\alpha(y) = (-y)^\alpha - \frac{1}{2}(-y)^{2\alpha} + o((-y)^{2\alpha}) \qquad as \qquad y \uparrow 0. \tag{3.25}$$

Proof. Write

$$1 - \Psi_\alpha(y) - (-y)^\alpha = \int_{y}^{0} \left\{ e^{-(-x)^\alpha} - 1 \right\} \alpha(-x)^{\alpha-1} dx = - \int_{y}^{0} \left[\frac{1 - e^{-(-x)^\alpha}}{(-x)^\alpha} \right] \alpha(-x)^{2\alpha-1} dx.$$

Note that $\phi(z) = \frac{1-e^{-z}}{z} \to 1$ as $z \uparrow 0$. Consider the following quantity:

$$\begin{aligned} \frac{\displaystyle\int_{y}^{0} \phi((-x)^\alpha) \alpha(-x)^{2\alpha-1} dx}{\int_{y}^{0} \alpha(-x)^{2\alpha-1} dx} &= \frac{\int_{y}^{0} \phi((-x)^\alpha) \alpha(-x)^{2\alpha-1} dx}{\frac{1}{2}(-y)^{2\alpha}} \\ &= 2\alpha(-y)^{-2\alpha} \int_{y}^{0} \phi((-x)^\alpha)(-x)^{2\alpha-1} dx. \end{aligned}$$

A change of variables brings

$$2\alpha\int_0^1 \phi((-yx)^\alpha)x^{2\alpha-1}dx = 2\alpha\int_0^1 x^{2\alpha-1}dx + o(1) = 1 + o(1).$$

Thus by dominated convergence

$$\frac{1-\Psi_\alpha(y)-(-y)^\alpha}{\frac{1}{2}(-y)^{2\alpha}} = -1 + o(1) \qquad \text{as} \qquad y \uparrow 0,$$

and the result follows. □

The next theorem establishes the two-term expansion for the Weibull mixture. We shall make use of the beta function defined in (1.7).

Theorem 3.6. *Let* $(X_2|X_1 = x) \sim \Psi_\alpha(\cdot - \beta x)$ *and* $X_1 \sim \Psi_\alpha$ *with* $\alpha, \beta > 0$. *Then as* $y \uparrow 0$

$$P(X_2 > y) = C_2(-y)^{2\alpha} + C_3(-y)^{3\alpha} + o((-y)^{3\alpha}), \tag{3.26}$$

where $C_2 = \frac{\alpha B(\alpha,\alpha+1)}{\beta^\alpha}$ *and* $C_3 = -\frac{\alpha}{\beta^\alpha}\left[\frac{B(\alpha,2\alpha+1)}{2} + \frac{B(2\alpha,\alpha+1)}{\beta^\alpha}\right]$.

Proof. We perform a second-order expansion on the first exponential term in the integral, and a first-order expansion on $e^{-(-x)^\alpha}$. No further advanced expansions are needed since the resulting terms would be negligible. As $y \uparrow 0$, the integral becomes

$$\begin{aligned} P(X_2 > y) &= \int_{y/\beta}^0 \left\{1 - e^{-[-(y-\beta x)]^\alpha}\right\} \alpha(-x)^{\alpha-1}e^{-(-x)^\alpha}dx \\ &= \int_{y/\beta}^0 \left[[-(y-\beta x)]^\alpha - \frac{1}{2}[-(y-\beta x)]^{2\alpha} + o((-y)^{2\alpha})\right] \alpha(-x)^{\alpha-1}\left[1-(-x)^\alpha\right]dx. \end{aligned}$$

Use the change of variables $w = \frac{\beta x}{y}$ and therefore $\left|\frac{dx}{du}\right| = -\frac{y}{\beta}$. The Jacobian has the negative sign because $y < 0$. The integral becomes

$$\begin{aligned}
P(X_2 > y) = &\int_0^1 \left\{(-y)^\alpha (1-x)^\alpha - \frac{1}{2}(-y)^{2\alpha}(1-x)^{2\alpha} + o((-y)^{2\alpha})\right\} \\
&\times \alpha \left(-\frac{y}{\beta}\right)^\alpha x^{\alpha-1} \left[1 - \left(-\frac{yx}{\beta}\right)^\alpha + \frac{1}{2}\left(-\frac{yx}{\beta}\right)^{2\alpha}\right] dx \\
= &\frac{\alpha}{\beta^\alpha}(-y)^\alpha \int_0^1 \left\{(-y)^\alpha (1-x)^\alpha - \frac{1}{2}(-y)^{2\alpha}(1-x)^{2\alpha} + o((-y)^{2\alpha})\right\} \\
&\times \left[x^{\alpha-1} - \left(-\frac{y}{\beta}\right)^\alpha x^{2\alpha-1} + \frac{1}{2}\left(-\frac{y}{\beta}\right)^{2\alpha} x^{3\alpha-1}\right] dx \\
= &\frac{\alpha}{\beta^\alpha}(-y)^\alpha \int_0^1 \left\{(-y)^\alpha x^{\alpha-1}(1-x)^\alpha - \frac{1}{2}(-y)^{2\alpha} x^{\alpha-1}(1-x)^{2\alpha} + o((-y)^{2\alpha})\right. \\
&- \frac{(-y)^{2\alpha}}{\beta^\alpha} x^{2\alpha-1}(1-x)^\alpha + \frac{(-y)^{3\alpha}}{2\beta^\alpha} x^{2\alpha-1}(1-x)^{2\alpha} \\
&\left. + \frac{(-y)^{3\alpha}}{2\beta^{2\alpha}} x^{3\alpha-1}(1-x)^\alpha - \frac{(-y)^{4\alpha}}{4\beta^{2\alpha}} x^{3\alpha-1}(1-x)^{2\alpha}\right\} dx.
\end{aligned}$$

The last three terms are negligible, and the integral becomes

$$\begin{aligned}
P(X_2 > y) = &\frac{\alpha}{\beta^\alpha}(-y)^{2\alpha} \int_0^1 x^{\alpha-1}(1-x)^\alpha dx - \frac{\alpha}{2\beta^\alpha}(-y)^{3\alpha} \int_0^1 x^{\alpha-1}(1-x)^{2\alpha} dx \\
&- \frac{\alpha}{\beta^{2\alpha}}(-y)^{3\alpha} \int_0^1 x^{2\alpha-1}(1-x)^\alpha dx + o((-y)^{3\alpha}) \\
= &\frac{\alpha B(\alpha, \alpha+1)}{\beta^\alpha}(-y)^{2\alpha} - \frac{\alpha}{\beta^\alpha}\left[\frac{B(\alpha, 2\alpha+1)}{2} + \frac{B(2\alpha, \alpha+1)}{\beta^\alpha}\right](-y)^{3\alpha} + o((-y)^{3\alpha}) \\
= &C_2(-y)^{2\alpha} + C_3(-y)^{3\alpha} + o((-y)^{3\alpha}).
\end{aligned}$$

□

The particular case where $\alpha = 1$ can be worked out exactly and is easily derived through direct integration:

Corollary 3.3. *For $\alpha = 1$ and any $\beta > 0$,*

$$P(X_2 > y) = \begin{cases} 1 - \left(\frac{1}{1-\beta}\right) e^{y} + \left(\frac{\beta}{1-\beta}\right) e^{y/\beta}, & \beta \neq 1 \\ 1 - (1-y)e^{y}, & \beta = 1. \end{cases} \tag{3.27}$$

It is interesting to note that in the above corollary, as $\beta \uparrow 1$ the first case limits to $1 - 2e^{y}$ but this does not match the second case. One can easily see that the limiting $1 - 2e^{y}$ is incorrect by noting that for $y \in (-\log(2), 0)$ the claimed probability is negative. Curiously, the term $(-y)^{\alpha}$ vanishes from the expansion, which necessitates getting the $(-y)^{3\alpha}$ term.

Finally, by using the same process one can derive an expansion for the Weibull mixture using as many terms as desired. We shall illustrate this process and then discuss how to extract the nth order term.

Theorem 3.7. *Let X_1, X_2, α, β be defined as above. Then an infinite order expansion for the tail probability of X_2 as $y \uparrow 0$ is*

$$P(X_2 > y) = \frac{\alpha}{\beta^{\alpha}} \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^{i+j+1} B\left((j+1)\alpha, i\alpha + 1\right)}{i! j! \beta^{j\alpha}} (-y)^{(i+j+1)\alpha}. \tag{3.28}$$

Proof. We proceed as before, but this time using the complete Taylor expansions:

$$\begin{aligned} P(X_2 > y) &= \int_{y/\beta}^{0} \left\{1 - e^{-[-(y-\beta x)]^{\alpha}}\right\} \alpha(-x)^{\alpha-1} e^{-(-x)^{\alpha}} dx \\ &= \alpha \int_{y/\beta}^{0} \sum_{i=1}^{\infty} \frac{(-1)^{i+1}[-(y-\beta x)]^{i\alpha}}{i!} (-x)^{\alpha-1} \sum_{j=0}^{\infty} \frac{(-1)^{j}(-x)^{j\alpha}}{j!} dx. \end{aligned}$$

Put an upper bound on the inner integrand:

$$\left|\frac{(-1)^{i+j+1}}{i!j!}[-(y-\beta x)]^{i\alpha}(-x)^{(j+1)\alpha-1}\right| \leq \frac{[-(y-\beta x)]^{i\alpha}}{i!}\frac{(-x)^{(j+1)\alpha-1}}{j!}$$
$$\leq \frac{1}{i!j!}(-y)^{i\alpha}\left(-\frac{y}{\beta}\right)^{(j+1)\alpha-1} \leq \frac{(-y)^{2\alpha-1}}{\beta^{\alpha-1}i!j!}.$$

Thus, the integrand is integrable on both counting measures and the Lebesgue measure, so we may bring both summations to the front:

$$P(X_2 > y) = \alpha\sum_{i=1}^{\infty}\frac{(-1)^{i+1}}{i!}\sum_{j=0}^{\infty}\frac{(-1)^j}{j!}\int_{y/\beta}^{0}(-x)^{(j+1)\alpha-1}[-(y-\beta x)]^{i\alpha}dx.$$

A change of variables produces

$$\begin{aligned}
P(X_2 > y) &= \alpha\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{i+j+1}}{i!j!}\int_0^1\left(-\frac{y}{\beta}\right)^{(j+1)\alpha}x^{(j+1)\alpha-1}(1-x)^{i\alpha}(-y)^{i\alpha}dx \\
&= \alpha\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{i+j+1}}{i!j!}\left(-\frac{y}{\beta}\right)^{(j+1)\alpha}(-y)^{i\alpha}B((j+1)\alpha, i\alpha+1) \\
&= \frac{\alpha}{\beta^{\alpha}}\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{i+j+1}B\left((j+1)\alpha, i\alpha+1\right)}{i!j!\beta^{j\alpha}}(-y)^{(i+j+1)\alpha},
\end{aligned}$$

as required.

It remains to show that the infinite sum converges. First observe that

$$\begin{aligned}
\Omega &= \left|\frac{\alpha}{\beta^{\alpha}}\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{i+j+1}B\left((j+1)\alpha, i\alpha+1\right)}{i!j!\beta^{j\alpha}}(-y)^{(i+j+1)\alpha}\right| \\
&\leq \frac{\alpha(-y)^{2\alpha}}{\beta^{\alpha}}\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\frac{B\left((j+1)\alpha, i\alpha+1\right)}{i!j!\beta^{j\alpha}}.
\end{aligned}$$

We shall use the fact that for $c_1, c_2 > 1, 0 \leq B(c_1, c_2) \leq 1$. This is true for $i \geq 1$ and $j \geq 1$, but when $j=0$ the expression $B(\alpha, i\alpha+1)$ is only guaranteed to satisfy this inequality when $\alpha > 1$. We adjust the beta function accordingly:

$$
\begin{aligned}
\frac{B\left((j+1)\alpha, i\alpha+1\right)}{i!j!\beta^{j\alpha}} &= \frac{\Gamma((j+1)\alpha)\Gamma(i\alpha+1)}{\Gamma((i+j+1)\alpha+1)i!j!\beta^{j\alpha}} \times \frac{(j+1)\alpha}{(j+1)\alpha} \\
&= \frac{\Gamma((j+1)\alpha+1)\Gamma(i\alpha+1)}{\alpha i!(j+1)!\beta^{j\alpha}\Gamma((i+j+1)\alpha+1)} \times \frac{\Gamma((i+j+1)\alpha+2)}{\Gamma((i+j+1)\alpha+2)} \\
&= \frac{B((j+1)\alpha+1, i\alpha+1)}{\alpha i!(j+1)!\beta^{j\alpha}} \times \frac{\Gamma((i+j+1)\alpha+2)}{\Gamma((i+j+1)\alpha+1)} \\
&\leq \frac{(i+j+1)\alpha+1}{\alpha i!(j+1)!\beta^{j\alpha}} = \frac{\alpha+1}{\alpha i!(j+1)!\beta^{j\alpha}} + \frac{1}{(i-1)!(j+1)!\beta^{j\alpha}} + \frac{j}{i!(j+1)!\beta^{j\alpha}}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
\Omega &\leq \frac{\alpha(-y)^{2\alpha}}{\beta^\alpha} \sum_{i=1}^{\infty}\sum_{j=0}^{\infty} \left[\frac{\alpha+1}{\alpha i!(j+1)!\beta^{j\alpha}} + \frac{1}{(i-1)!(j+1)!\beta^{j\alpha}} + \frac{j}{i!(j+1)!\beta^{j\alpha}}\right] \\
&< \frac{3\alpha(-y)^{2\alpha}}{\beta^\alpha} \left[\frac{\alpha+1}{\alpha}\sum_{j=0}^{\infty} \frac{1}{(j+1)!\beta^{j\alpha}} + \sum_{j=0}^{\infty} \frac{1}{(j+1)!\beta^{j\alpha}} + \sum_{j=0}^{\infty} \frac{j}{(j+1)!\beta^{j\alpha}}\right] \\
&= \frac{3\alpha(-y)^{2\alpha}}{\beta^\alpha} \left[\frac{2\alpha+1}{\alpha}\sum_{j=0}^{\infty} \frac{1}{(j+1)!\beta^{j\alpha}} + \sum_{j=0}^{\infty} \frac{j}{(j+1)!\beta^{j\alpha}}\right] \\
&= \frac{3\alpha(-y)^{2\alpha}}{\beta^\alpha} \left[\left(\frac{2\alpha+1}{\alpha}\right)(I) + (II)\right].
\end{aligned}
$$

Compute the two sums:

$$
(I) = \sum_{j=0}^{\infty} \frac{\beta^\alpha}{(j+1)!\beta^{(j+1)\alpha}} = \beta^\alpha \sum_{j=1}^{\infty} \frac{(\beta^{-\alpha})^j}{j!} = \beta^\alpha \left(1 - \exp\left(\beta^{-\alpha}\right)\right)
$$

and

$$
\begin{aligned}
(II) &= \sum_{j=0}^{\infty} \frac{(j+1-1)\beta^{\alpha}}{(j+1)!\beta^{(j+1)\alpha}} = \beta^{\alpha} \sum_{j=1}^{\infty} \frac{j-1}{j!\beta^{j\alpha}} = \beta^{\alpha} \left[\sum_{j=1}^{\infty} \frac{1}{(j-1)!\beta^{j\alpha}} - \sum_{j=1}^{\infty} \frac{1}{j!\beta^{j\alpha}} \right] \\
&= \beta^{\alpha} \left[\frac{1}{\beta^{\alpha}} \sum_{j=1}^{\infty} \frac{1}{(j-1)!\beta^{(j-1)\alpha}} - \sum_{j=1}^{\infty} \frac{1}{j!\beta^{j\alpha}} \right] = \sum_{j=0}^{\infty} \frac{(\beta^{-\alpha})^j}{j!} - \beta^{\alpha} \sum_{j=1}^{\infty} \frac{(\beta^{-\alpha})^j}{j!} \\
&= (1+\beta^{\alpha}) \exp\left(\beta^{-\alpha}\right) - \beta^{\alpha}.
\end{aligned}
$$

Finally

$$
\Omega \le 3\alpha(-y)^{2\alpha} \left[\frac{2\alpha+1}{\alpha} \left(1 - \exp(\beta^{-\alpha})\right) + (1+\beta^{-\alpha}) \exp(\beta^{-\alpha}) - 1 \right],
$$

and so $\Omega \to 0$ as $y \uparrow 0$. Therefore the infinite series converges. □

Now suppose we want C_m, the mth component in the n-term expansion, $m = 2, ..., n+1$. The result of Theorem 3.7 can be used to accomplish this. To "peel off" the mth term, find all i, j such that $i \ge 1, j \ge 0, i+j+1 = m$. Then

$$
\begin{aligned}
C_m &= \frac{\alpha}{\beta^{\alpha}} \sum_{j=0}^{m-2} \frac{(-1)^m B\left((j+1)\alpha, (m-j-1)\alpha + 1\right)}{j!(m-j-1)!\beta^{j\alpha}} \\
&= \frac{\alpha}{\beta^{\alpha}} \sum_{j=0}^{m-2} \frac{(-1)^m B\left((j+1)\alpha, (m-(j+1))\alpha + 1\right)}{((j+1)-1)!(m-(j+1))!\beta^{((j+1)-1)\alpha}} \\
&= \frac{\alpha}{\beta^{\alpha}} \sum_{j=1}^{m-1} \frac{(-1)^m B\left(j\alpha, (m-j)\alpha + 1\right)}{(j-1)!(m-j)!\beta^{(j-1)\alpha}}.
\end{aligned}
$$

As for the remaining pieces where $i+j+1 > m$,

$$
\left| \frac{\alpha}{\beta^{\alpha}} \sum_{i+j+1>m} \frac{(-1)^{m+k} B\left((j+1)\alpha, (m+k-j-1)\alpha + 1\right)}{j!(m+k-j-1)!\beta^{j\alpha}} (-y)^{(m+L)\alpha} \right| = o((-y)^{m\alpha}),
$$

because the series was shown to converge. We therefore have the following result.

Theorem 3.8. *Let* $(X_2|X_1 = x) \sim \Psi_\alpha(\cdot - \beta x)$ *and* $X_1 \sim \Psi_\alpha$ *with* $\alpha, \beta > 0$. *Then an* n*-term expansion for the tail probability of* X_2 *as* $y \uparrow 0$ *is*

$$P(X_2 > y) = \sum_{m=2}^{n+1} C_m(-y)^{m\alpha} + o((-y)^{m\alpha}),$$

where $C_m = \frac{\alpha}{\beta^\alpha} \sum_{j=1}^{m-1} \frac{(-1)^m B\left(j\alpha, (m-j)\alpha + 1\right)}{(j-1)!(m-j)!\beta^{(j-1)\alpha}}$.

Chapter 4 The AR(1) Process with Gumbel Innovations

In the previous chapter we derived a two-term expansion for the mixture of two Gumbel random variables. That is, if $\Lambda(x) = \exp(-e^{-x})$ denotes the standard Gumbel distribution, we considered the distribution of $\beta Z_1 + Z_0$ where $Z_1, Z_0 \sim \Lambda$. In this chapter we generalize the results for the $AR(1)$ process given by $X_0 = Z_0$ and

$$X_n = \beta X_{n-1} + Z_n, \qquad n \geq 1, \tag{4.1}$$

with the $\{Z_n\}$ i.i.d. with distribution Λ. Consequently

$$X_n = \sum_{k=0}^{\infty} \beta^k Z_{n-k}. \tag{4.2}$$

The first step is to ensure that X_n converges for $|\beta| < 1$. Note that

$$\begin{aligned} E|Z_1| &= \int_{-\infty}^{0} (-x)e^{-x} \exp(-e^{-x})dx + \int_{0}^{\infty} xe^{-x} \exp(-e^{-x})dx \\ &= \int_{1}^{\infty} \log(x)e^{-x}dx + \int_{0}^{1} (-\log(x))e^{-x}dx \leq 2, \end{aligned}$$

because

$$\int_0^1 (-\log(x))e^{-x}dx < \int_0^1 (-\log(x))dx = \int_0^{\infty} xe^{-x}dx = 1.$$

Now observe that

$$E\left|\sum_{k=m}^{\infty} \beta^k Z_{1-k}\right| \leq \frac{|\beta|^m}{1-|\beta|} E|Z_1| \to 0 \qquad \text{as} \qquad m \to \infty.$$

Thus we have

$$\left|\sum_{k=m}^{\infty} \beta^k Z_{1-k}\right| \leq \sum_{k=m}^{\infty} |\beta|^k |Z_{1-k}| \xrightarrow{L^1} 0,$$

and therefore $\sum_{k=m}^{\infty} |\beta|^k |Z_{1-k}| \xrightarrow{p} 0$. Set $U_m = \sum_{k=m}^{\infty} |\beta|^k |Z_{1-k}|$ and note that U_m is a monotonically decreasing sequence of nonnegative random variables. Hence $U_m \to U$ as $m \to \infty$ where the limit U exists almost surely. Since the limit must agree with the limit in probability, $U \overset{a.s.}{=} 0$. Thus

$$\sum_{k=m}^{\infty} \beta^k Z_{1-k} \xrightarrow{a.s.} 0 \quad \text{as} \quad m \to \infty.$$

Hence $X_1 = \sum_{k=0}^{\infty} \beta^k Z_{1-k}$ exists as an a.s. limit, and $X_n = \sum_{k=0}^{\infty} \beta^k Z_{n-k}$ is the stationary solution of the $AR(1)$ process.

4.1 Lemmas Needed for the Expansions

Before moving to the two-term expansion for X_n, we state and prove several lemmas that come up in the subsequent proofs. Also in the upcoming sections, we shall use the notation $X = V + W$ where $V = Z_0 + \beta Z_1$ and $W = \sum_{k=2}^{\infty} \beta^k Z_k$, and further $V \sim F_V$ and $W \sim F_W$.

Lemma 4.1. *For any $\lambda\beta < 1$,*

$$E\left[\exp\left(\frac{\lambda}{\beta} W\right)\right] < \infty.$$

Proof. It suffices to show the alternative statement holds:

$$E\left[\exp\left(\lambda\beta \sum_{k=0}^{\infty} \beta^k Z_k\right)\right] < \infty.$$

Set $\xi = \lambda\beta < 1$. Observe that

$$E\left[\exp\left(\xi\sum_{k=0}^{\infty}\beta^k Z_k\right)\right] = \prod_{k=0}^{\infty} E\left[\exp\left(\xi\beta^k Z_1\right)\right].$$

Choose any $0 < a < 1$, then

$$E\left(e^{aZ_1}\right) = \int_{-\infty}^{\infty} e^{az}e^{-z}\exp(-e^{-z})dz = \int_0^{\infty} z^{-a}e^{-z}dz = \Gamma(1-a).$$

Now notice that

$$\frac{1}{a}\left(\int_0^{\infty} z^{-a}e^{-z}dz - 1\right) = \int_0^{\infty}\left(\frac{e^{-a\log(z)}-1}{a}\right)e^{-z}dz.$$

Also

$$\left|\frac{1}{a}\int_0^{a\log(z)} e^{-t}dt\right| \le \log(z) \qquad \text{for} \quad z \ge 1, \quad \text{and}$$
$$\left|\frac{1}{a}\int_0^{a\log(1/z)} e^{-t}dt\right| \le \log\left(\frac{1}{z}\right) \qquad \text{for} \quad 0 < z \le 1.$$

It follows that

$$\left|\frac{1}{a}(z^{-a}-1)e^{-z}\right| \le |\log(z)|e^{-z},$$

which is integrable over the positive reals, and so by dominated convergence

$$\lim_{a\downarrow 0}\frac{1}{a}\left(\int_0^{\infty} z^{-a}e^{-z}dz - 1\right) = \int_0^{\infty}(-\log(z))e^{-z}dz = \int_{-\infty}^{\infty} ze^{-z}\exp(-e^{-z})dz$$
$$= E(Z_1).$$

Thus, as $a \downarrow 0$

$$E\left(e^{aZ_1}\right) = \Gamma(1-a) = 1 + aE(Z_1)(1+o(1)).$$

Of course,

$$E(Z_1) = \lim_{a\downarrow 0} \frac{\Gamma(1-a)-1}{a} = -\Gamma'(1) = \gamma$$

where γ is the Euler constant defined in (1.7). Hence for $k \geq k_0$ large enough,

$$E\left[\exp\left(\xi\beta^k Z_1\right)\right] < 1 + 2\xi\beta^k E|Z_1|.$$

Therefore

$$\prod_{k=k_0}^{\infty} E\left[\exp\left(\xi\beta^k Z_1\right)\right] < \prod_{k=k_0}^{\infty}\left(1+2\xi\beta^k E|Z_1|\right) < \exp\left(2\xi\sum_{k=k_0}^{\infty}\beta^k E|Z_1|\right)$$
$$= \exp\left(\frac{2\xi\beta^{k_0}}{1-\beta}E|Z_1|\right) < \infty.$$

Finally, since $0 < \xi < 1$ we have the usual moment generating function

$$E\left(e^{\xi Z_1}\right) = \Gamma(1-\xi) < \infty,$$

and for $k = 0, 1, ..., k_0 - 1$

$$E\left[\exp\left(\xi\beta^k Z_1\right)\right] = \Gamma(1-\xi\beta^k) < \infty.$$

Putting it all together,

$$E\left[\exp\left(\xi\sum_{k=0}^{\infty}\beta^k Z_k\right)\right] < \infty.$$

□

One important consequence of Lemma 4.1 is that for any $0 < \beta, \xi < 1$, the infinite product of gamma functions converges. That is,

$$\prod_{k=0}^{\infty} \Gamma(1 - \xi\beta^k) < \infty, \tag{4.3}$$

and this product will show up in the expansions in the next three sections.

Lemma 4.2. *Let $0 < \beta < 1$, and choose a such that $\beta < a < 1$. Then as $y \to \infty$*

$$P(X > y) = \int_{-\infty}^{ay} P(V > y - w) dF_W(w) + o\left(e^{-y/\beta}\right).$$

Proof. Observe that

$$P(X > y) = P(V+W > y) = \int_{ay}^{\infty} P(V > y-w) dF_W(w) + \int_{-\infty}^{ay} P(V > y-w) dF_W(w). \tag{4.4}$$

Now for $1 < \lambda < \frac{1}{\beta}$, we have by Chernoff's Inequality

$$\int_{ay}^{\infty} P(V > y - w) dF_W(w) \leq P(W > ay) \leq E\left(e^{\lambda W/\beta}\right) e^{-\lambda ay/\beta}.$$

Additionally choose λ so that $\lambda > \frac{1}{a}$, which is possible because $\frac{1}{a} < \frac{1}{\beta}$. Then by Lemma 4.1,

$$P(W > ay) = O\left(e^{-\lambda ay/\beta}\right) = o\left(e^{-y/\beta}\right). \tag{4.5}$$

□

Lemma 4.3. *Let W be as defined in Lemma 4.1. Choose a and λ such that $1 < \lambda < \frac{1}{\beta}$, $\beta < a < 1$, and $a\lambda > 1$. Then as $y \to \infty$*

$$\int_{ay}^{\infty} e^{-(y-w)} dF_W(w) = o\left(e^{-y/\beta}\right).$$

Proof. By integration by parts,

$$\int_{ay}^{\infty} e^{w} dF_W(w) = -\int_{ay}^{\infty} e^{w} d\overline{F}_W(w) = e^{ay}\overline{F}_W(ay) + \int_{ay}^{\infty} \overline{F}_W(w) e^{w} dw.$$

By Lemma 4.1 and Chernoff's Inequality,

$$\overline{F}_W(ay) \leq E\left(e^{\lambda W/\beta}\right) e^{-\lambda ay/\beta}.$$

Similarly, we find

$$\int_{ay}^{\infty} \overline{F}_W(w) e^{w} dw \leq E\left(e^{\lambda W/\beta}\right) \int_{ay}^{\infty} e^{-\frac{\lambda}{\beta}w + w} dw = E\left(e^{\lambda W/\beta}\right) \left(\frac{\lambda}{\beta} - 1\right)^{-1} e^{-\left(\frac{\lambda}{\beta}-1\right)ay}.$$

Therefore for some $c > 0$ we obtain

$$\int_{ay}^{\infty} e^{-(y-w)} dF_W(w) \leq c \exp\left(-\left[\frac{\lambda a}{\beta} + (1-a)\right] y\right).$$

Finally, since $\beta < a < 1$ and $a\lambda > 1$ we have that $1 - a + \frac{\lambda a}{\beta} > \frac{1}{\beta}$ and therefore

$$\int_{ay}^{\infty} e^{-(y-w)} dF_W(w) = o\left(e^{-y/\beta}\right).$$

□

Lemma 4.4. *For W defined in Lemma 4.1 and $0 < \beta < 1$, $E\left(\left|We^{\frac{1}{\beta}W}\right|\right) < \infty$.*

Proof. Choose any $1 < \lambda < \frac{1}{\beta}$. Then by Lemma 4.1, $E\left(e^{\frac{\lambda}{\beta}W}\right) < \infty$. Let ν be such that $\frac{1}{\lambda} + \frac{1}{\nu} = 1$. Then by Hölder's Inequality in (1.18),

$$E\left(\left|We^{\frac{1}{\beta}W}\right|\right) \leq (E|W|^{\nu})^{1/\nu}\left[E\left(e^{\frac{\lambda}{\beta}W}\right)\right]^{1/\lambda} < \infty.$$

□

Lemmas 4.5 through 4.7 assume that $-1 < \beta < 0$. In this situation, write the sum as

$$\sum_{k=0}^{\infty} \beta^k Z_k \stackrel{d}{=} \sum_{k=0}^{\infty} \beta^{2k} Z_k + \sum_{k=0}^{\infty} \beta^{2k+1} Z_k = S + T.$$

Lemma 4.5. *For $-1 < \beta < 0$, $a > 0$, and T defined above,*

$$P(T > ay) = o\left(e^{-y/\beta^2}\right) \qquad \text{as} \qquad y \to \infty.$$

Proof. We first show that the lemma holds for $\beta Z_0 + \beta^3 Z_1$. Let $F_{\beta^2 Z_1}$ denote the distribution of $\beta^2 Z_1$. Choose $\zeta < 0$ to be large in the negative direction. Then

$$\begin{aligned}
P(\beta Z_0 + \beta^3 Z_1 > ay) &= P\left(Z_0 + \beta^2 Z_1 < \frac{ay}{\beta}\right) \\
&= \int_{\zeta}^{\infty} P\left(Z_0 < \frac{ay}{\beta} - \beta^2 z\right) dF_{\beta^2 Z_1}(z) + \int_{-\infty}^{\zeta} P\left(Z_0 < \frac{ay}{\beta} - \beta^2 z\right) dF_{\beta^2 Z_1}(z) \\
&= (I) + (II).
\end{aligned}$$

Note that

$$\sup_{\zeta < z < \infty} P\left(Z_0 < \frac{ay}{\beta} - \beta^2 z\right) = \exp\left[-e^{-\left(\frac{ay}{\beta} - \beta^2 \zeta\right)}\right],$$

and therefore as $y \to \infty$ $(I) \leq \exp\left[-e^{-\left(\frac{ay}{\beta} - \beta^2 \zeta\right)}\right] \to 0$ because $\frac{ay}{\beta} - \beta^2 \zeta \to -\infty$ for ζ fixed. Now observe that as $y \to \infty$

$$\frac{\exp\left[-e^{-\left(\frac{ay}{\beta} - \beta^2 \zeta\right)}\right]}{e^{-y/\beta^2}} = \exp\left[\frac{y}{\beta^2} - e^{-\left(\frac{ay}{\beta} - \beta^2 \zeta\right)}\right] \to 0,$$

and so $(I) = o\left(e^{-y/\beta^2}\right)$. Next,

$$(II) \leq F_{\beta^2 Z_1}(\zeta) = \exp\left[-e^{-\zeta/\beta^2}\right] \to 0 \qquad \text{as} \qquad \zeta \downarrow -\infty,$$

and therefore $P(\beta Z_0 + \beta^3 Z_1 > ay) = o\left(e^{-y/\beta^2}\right)$ as $y \to \infty$. The next step is to assume that for $n \geq 2$,

$$P\left(\beta \sum_{k=0}^{n-1} \beta^{2k} Z_k > ay\right) = P\left(\beta U_{n-1} > ay\right) = o\left(e^{-y/\beta^2}\right) \qquad \text{as} \qquad y \to \infty.$$

Consider the probability

$$\begin{aligned}
&P(\beta U_{n-1} + \beta^{2n+1} Z_n > ay) = P\left(U_{n-1} + \beta^{2n} Z_n < \frac{ay}{\beta}\right) \\
&= \int_{\zeta}^{\infty} P\left(U_{n-1} < \frac{ay}{\beta} - \beta^{2n} z\right) dF_{\beta^{2n} Z_1}(z) + \int_{-\infty}^{\zeta} P\left(Z_0 < \frac{ay}{\beta} - \beta^{2n} z\right) dF_{\beta^{2n} Z_1}(z) \\
&= (I) + (II).
\end{aligned}$$

We have by assumption that $P\left(U_{n-1} < \frac{ay}{\beta}\right) = o\left(e^{-y/\beta^2}\right)$, so for y large enough

$$\sup_{\zeta < z < \infty} P\left(U_{n-1} < \frac{ay}{\beta} - \beta^{2n} z\right) = P\left(U_{n-1} < \frac{ay}{\beta} - \beta^{2n} \zeta\right) = o\left(e^{-y/\beta^2}\right).$$

Therefore for any $\epsilon > 0$,

$$(I) < \epsilon e^{-y/\beta^2} \int_{\zeta}^{\infty} dF_{\beta^{2n} Z_n}(z) \leq \epsilon e^{-y/\beta^2},$$

and so $(I) = o\left(e^{-y/\beta^2}\right)$ as $y \to \infty$. As for (II),

$$(II) \leq F_{\beta^{2n} Z_n}(\zeta) = \exp\left[-e^{-\zeta/\beta^{2n}}\right] \to 0 \qquad \text{as} \qquad \zeta \downarrow -\infty.$$

Therefore as $y \to \infty$,

$$P\left(\beta \sum_{k=0}^{n} \beta^{2k} Z_k > ay\right) = P\left(\beta U_n > ay\right) = o\left(e^{-y/\beta^2}\right).$$

Finally, we leap to the infinite sum $T = \beta \sum_{k=0}^{\infty}$. For any $\delta > 0$, there exists n_0 large enough such that $\left|\sum_{k=n_0+1}^{\infty} \beta^{2k} Z_k\right| < \delta$, because we know that $\left|\sum_{k=n_0+1}^{\infty} \beta^{2k} Z_k\right| \stackrel{a.s.}{\to} 0$ as $n_0 \to \infty$. Consider

$$P\left(\sum_{k=0}^{n_0} \beta^{2k} Z_k + \sum_{k=n_0+1}^{\infty} \beta^{2k} Z_k < \frac{ay}{\beta}\right) = P\left(U_{n_0} + \tilde{U}_{n_0+1} < \frac{ay}{\beta}\right).$$

We have that $-\delta < \tilde{U}_{n_0+1} < \delta$ almost surely, and so for any $\epsilon > 0$

$$\begin{aligned} P\left(U_{n_0} < \frac{ay}{\beta} - \delta\right) \leq P\left(U_{n_0} + \tilde{U}_{n_0+1} < \frac{ay}{\beta}\right) \leq P\left(U_{n_0} < \frac{ay}{\beta} + \delta\right) \\ \Rightarrow 0 \leq P\left(U_{n_0} + \tilde{U}_{n_0+1} < \frac{ay}{\beta}\right) < \epsilon e^{-y/\beta^2}. \end{aligned}$$

Thus $P\left(U_{n_0} + \tilde{U}_{n_0+1} < \frac{ay}{\beta}\right) = P(T > ay) = o\left(e^{-y/\beta^2}\right)$ as $y \to \infty$. □

Lemma 4.6. *For $-1 < \beta < 0$ and T defined above, $E\left(e^{T/\beta^2}\right) < \infty$.*

Proof. First observe that

$$E\left(e^{T/\beta^2}\right) = E\exp\left(\frac{1}{\beta}\sum_{k=0}^{\infty}\beta^{2k}Z_k\right) = E\left(e^{Z_0/\beta}\right)\prod_{k=0}^{\infty}E\left(e^{\beta^{2k+1}Z_1}\right).$$

In Lemma 4.1, set $\xi = \beta^{k+1}$. This choice works since either $-1 < \beta^{k+1} < 0$ or $0 < \beta^{k+1} < 1$, but in either case $\xi < 1$. Then for k_0 large enough,

$$E\left(e^{T/\beta^2}\right) < \Gamma\left(1-\beta^{-1}\right)\prod_{k=0}^{k_0-1}E\left(e^{\xi\beta^k Z_1}\right)\exp\left(\frac{2\xi\beta^{k_0}}{1-\beta}E|Z_1|\right) < \infty.$$

□

Lemma 4.7. *For* $-1 < \beta < 0$, $0 < a < 1$, *and* T *defined above,*

$$\int_{ay}^{\infty} e^{-(y-t)}dF_T(t) < \infty.$$

Proof. If $F_T(t)$ denotes the distribution of T, then first note that by integration by parts,

$$e^{-y}\int_{ay}^{\infty} e^t dF_T(t) = e^{-y}e^{ay}\overline{F}_T(ay) + e^{-y}\int_{ay}^{\infty} e^t\overline{F}_T(t)dt = (I) + (II).$$

By Lemma 4.5, for y large enough

$$(I) = e^{-(1-a)y}P(T > ay) = e^{-(1-a)y}o\left(e^{-y/\beta^2}\right),$$

and therefore $(I) = o\left(e^{-y/\beta^2}\right)$ as $y \to \infty$. Next, we have that for any $\epsilon > 0$ and y large,

$e^{y/\beta^2}\overline{F}_T(ay) < \epsilon$ and thus $e^{y/a\beta}\overline{F}_T(t) < \epsilon$. Therefore

$$\begin{aligned}(II) &< \epsilon e^{-y} \int_{ay}^{\infty} e^t e^{-t/a\beta^2} dt = \epsilon e^{-y} \int_{ay}^{\infty} \exp\left[-\left(\frac{1}{a\beta^2} - 1\right) t\right] dt \\ &= \frac{\epsilon a\beta^2}{1 - a\beta^2} e^{-y} \exp\left[-\left(\frac{1 - a\beta^2}{\beta^2}\right) y\right] = o\left(\exp\left[-\left(\frac{1}{\beta^2} + (1 - a)\right)\right]\right).\end{aligned}$$

Thus $(II) = o\left(e^{-y/\beta^2}\right)$ as $y \to \infty$, which completes the proof. □

4.2 The AR(1) Process when $0 < \beta < 1$

In the next three lemmas, assuming that $0 < \beta < 1$, we perform calculations that all build upon various pieces of Theorem 3.1. While the steps are similar to one another, due to the subtle differences among Theorem 3.1 we need to examine each lemma separately. There are three cases, treating $\beta = \frac{1}{2}$ as the pivot point.

Lemma 4.8. *If* $0 < \beta < \frac{1}{2}$ *and* Z_k *are i.i.d. standard Gumbel random variables, then as* $y \to \infty$

$$P\left(\sum_{k=0}^{\infty} \beta^k Z_k > y\right) = \prod_{k=1}^{\infty} \Gamma\left(1 - \beta^k\right) e^{-y} - \frac{1}{2} \prod_{k=1}^{\infty} \Gamma\left(1 - 2\beta^k\right) e^{-2y} + O\left(e^{-y/\beta}\right).$$

Proof. Let a and λ be positive reals such that $\beta < a < 1$, $1 < \lambda < \frac{1}{\beta}$, and $\lambda > \frac{1}{a}$. Then by Lemma 4.2,

$$P(X > y) = \int_{-\infty}^{ay} P(V > y - w) dF_W(w) + o\left(e^{-y/\beta}\right). \tag{4.6}$$

By Theorem 3.1, we have for y large enough and any $w < ay$ that

$$P(V > y - w) = \Gamma(1 - \beta) e^{-(y-w)} - \frac{1}{2} \Gamma(1 - 2\beta) e^{-2(y-w)} + R(y - w),$$

where for some $c > 0$

$$\sup_{w<ay} \left| R(y-w)e^{(y-w)/\beta} \right| < c.$$

Set $K_1 = \Gamma(1-\beta)$ and $K_2 = -\frac{1}{2}\Gamma(1-2\beta)$. Then

$$\begin{aligned} &\int_{-\infty}^{ay} \overline{F}_V(y-w)dF_W(w) \\ &\quad = \int_{-\infty}^{ay} \left[\frac{\overline{F}_V(y-w) - K_1 e^{-(y-w)}}{K_2 e^{-2(y-w)}} \right] K_2 e^{-2(y-w)} dF_W(w) + K_1 \int_{-\infty}^{ay} e^{-(y-w)} dF_W(w). \end{aligned} \tag{4.7}$$

Now

$$\frac{\overline{F}_V(y-w) - K_1 e^{-(y-w)}}{K_2 e^{-2(y-w)}} = 1 + \frac{1}{K_2} R(y-w)e^{2(y-w)} = 1 + \epsilon(y-w).$$

We examine the $\epsilon(y-w)$:

$$\begin{aligned} \sup_{w<ay} |\epsilon(y-w)| &= \frac{1}{K_2} \sup_{w<ay} \left| R(y-w)e^{(y-w)/\beta} e^{-(y-w)/\beta} e^{2(y-w)} \right| \\ &< \frac{c}{K_2} \sup_{w<ay} \left| \exp\left(-\left(\frac{1}{\beta} - 2\right)(y-w) \right) \right| \\ &\leq \frac{c}{K_2} \exp\left(-\frac{(1-a)(1-2\beta)y}{\beta} \right). \end{aligned}$$

Because $\frac{(1-a)(1-2\beta)}{\beta} > 0$, we have that $\epsilon(y-w) \to 0$ as $y \to \infty$ uniformly in $w < ay$. Consider the piece

$$\begin{aligned} &\int_{-\infty}^{ay} [1 + \epsilon(y-w)] K_2 e^{-2(y-w)} dF_W(w) \\ &\quad = K_2 e^{-2y} \int_{-\infty}^{ay} e^{2w} dF_W(w) + K_2 \int_{-\infty}^{ay} \epsilon(y-w) e^{-2(y-w)} dF_W(w). \end{aligned} \tag{4.8}$$

Examining the first integral in (4.8),

$$\begin{aligned}
\lim_{y\to\infty}\int_{-\infty}^{ay} K_2 e^{2w} dF_W(w) &= -\frac{1}{2}\Gamma(1-2\beta)\int_{-\infty}^{\infty} e^{2w} dF_W(w) = -\frac{1}{2}\Gamma(1-2\beta)E\left(e^{2W}\right) \\
&= -\frac{1}{2}\Gamma(1-2\beta)E\left[\exp\left(\sum_{k=2}^{\infty} 2\beta^k Z_k\right)\right] = -\frac{1}{2}\Gamma(1-2\beta)\prod_{k=2}^{\infty} E\left(e^{2\beta^k Z_k}\right) \\
&= -\frac{1}{2}\prod_{k=1}^{\infty}\Gamma(1-2\beta^k).
\end{aligned} \tag{4.9}$$

Turning to the second integral in (4.8),

$$\begin{aligned}
&\left|K_2\int_{-\infty}^{ay}\epsilon(y-w)e^{-2(y-w)}dF_W(w)\right| = \left|\int_{-\infty}^{ay} R(y-w)dF_W(w)\right| \\
&\quad = \left|\int_{-\infty}^{ay} R(y-w)e^{(y-w)/\beta}e^{-(y-w)/\beta}dF_W(w)\right| < c\int_{-\infty}^{ay} e^{-(y-w)/\beta}dF_W(w) \\
&\quad \le ce^{-y/\beta}\int_{-\infty}^{\infty} e^{w/\beta}dF_W(w) = cE\left(e^{W/\beta}\right)e^{-y/\beta}.
\end{aligned}$$

By Lemma 4.1, $E\left(e^{W/\beta}\right) < \infty$ and therefore

$$K_2\int_{-\infty}^{ay}\epsilon(y-w)e^{-2(y-w)}dF_W(w) = O\left(e^{-y/\beta}\right). \tag{4.10}$$

Next, note that

$$\int_{-\infty}^{\infty}\Gamma(1-\beta)e^{w}dF_W(w) = \Gamma(1-\beta)E\left(e^{W}\right) = \prod_{k=1}^{\infty}\Gamma(1-\beta^k). \tag{4.11}$$

We now consider the integral

$$K_1 e^{-y}\int_{-\infty}^{ay} e^{w}dF_W(w) = \prod_{k=1}^{\infty}\Gamma(1-\beta^k)e^{-y} - \Gamma(1-\beta)e^{-y}\int_{ay}^{\infty} e^{w}dF_W(w). \tag{4.12}$$

By Lemma 4.3, the last integral in (4.12) is $o(e^{-y/\beta})$. □

Lemma 4.9. *If $\beta = \frac{1}{2}$, then as $y \to \infty$*

$$P\left(\sum_{k=0}^{\infty} (0.5)^k Z_k > y\right) = \prod_{k=1}^{\infty} \Gamma\left(1 - (0.5)^k\right)\left(e^{-y} - ye^{-2y}\right) + O\left(e^{-2y}\right).$$

In particular, to five decimal places $\prod_{k=1}^{\infty} \Gamma\left(1 - (0.5)^k\right) = 2.55501$.

Proof. Let a and λ be positive reals such that $\beta < a < 1$, $1 < \lambda < \frac{1}{\beta}$, and $\lambda > \frac{1}{a}$. Then by Lemma 4.2,

$$P(X > y) = \int_{-\infty}^{ay} P(V > y - w)dF_W(w) + o\left(e^{-2y}\right). \tag{4.13}$$

By Theorem 3.1, for y large enough and any $w < ay$

$$P(V > y - w) = \sqrt{\pi}e^{-(y-w)} - (y - w)e^{-2(y-w)} + R(y - w),$$

where for some $c > 0$

$$\sup_{w<ay} \left|R(y - w)e^{2(y-w)}\right| < c.$$

Observe that

$$\begin{aligned}\int_{-\infty}^{ay} \overline{F}_V(y - w)dF_W(w) &= \sqrt{\pi}\int_{-\infty}^{ay} e^{-(y-w)}dF_W(w)\\ &+ \int_{-\infty}^{ay} \left[\frac{\overline{F}_V(y - w) - \sqrt{\pi}e^{-(y-w)}}{-(y - w)e^{-2(y-w)}}\right](-(y - w))e^{-2(y-w)}dF_W(w).\end{aligned} \tag{4.14}$$

Now

$$\frac{\overline{F}_V(y - w) - \sqrt{\pi}e^{-(y-w)}}{-(y - w)e^{-2(y-w)}} = 1 - \frac{R(y - w)e^{2(y-w)}}{y - w} = 1 + \epsilon(y - w).$$

For y large enough,

$$\sup_{w<ay} |\epsilon(y-w)| = \sup_{w<ay} \left| \frac{R(y-w)e^{2(y-w)}}{y-w} \right| < \frac{c}{(1-a)y},$$

and therefore $\epsilon(y-w) \to 0$ as $y \to \infty$ uniformly in $w < ay$. Now consider

$$\begin{aligned}
\int_{-\infty}^{ay} & [1+\epsilon(y-w)] \left[-(y-w)e^{-2(y-w)}\right] dF_W(w) \\
&= e^{-2y} \int_{-\infty}^{ay} we^{2w} dF_W(w) - ye^{-2y} \int_{-\infty}^{ay} e^{2w} dF_W(w) \\
&\quad - \int_{-\infty}^{ay} \epsilon(y-w)(y-w)e^{-2(y-w)} dF_W(w) \\
&= (I) + (II) + (III).
\end{aligned} \tag{4.15}$$

We examine the three integrals in (4.15). By Lemma 4.4,

$$(I) = E\left(We^{2W}\right) - \int_{ay}^{\infty} we^{2w} dF_W(w) = E\left(We^{2W}\right) + o(1). \tag{4.16}$$

As for (II),

$$\lim_{y\to\infty} \int_{-\infty}^{ay} e^{2w} dF_W(w) = \int_{-\infty}^{\infty} e^{2w} dF_W(w) = E\left(e^{2W}\right) = \prod_{k=1}^{\infty} \Gamma\left(1-(0.5)^k\right). \tag{4.17}$$

Examining (III),

$$\begin{aligned}
&\left| - \int_{-\infty}^{ay} \epsilon(y-w)(y-w)e^{-2(y-w)} dF_W(w) \right| = \left| \int_{-\infty}^{ay} R(y-w) dF_W(w) \right| \\
&\quad = \left| \int_{-\infty}^{ay} R(y-w)e^{2(y-w)}e^{-2(y-w)} dF_W(w) \right| < ce^{-2y} \int_{-\infty}^{ay} e^{2w} dF_W(w) \\
&\quad \le ce^{-2y} \int_{-\infty}^{\infty} e^{2w} dF_W(w) = cE\left(e^{2W}\right) e^{-2y}.
\end{aligned}$$

By Lemma 4.1, $E\left(e^{2W}\right) < \infty$ and so

$$(I) + (III) = e^{-2y}\left[E\left(We^{2W}\right) + o(1)\right] + O\left(e^{-2y}\right) = O\left(e^{-2y}\right). \tag{4.18}$$

Next, note that

$$\int_{-\infty}^{\infty} \sqrt{\pi} e^{w} dF_W(w) = \Gamma(1-0.5)\, E\left(e^{W}\right) = \prod_{k=1}^{\infty} \Gamma\left(1-(0.5)^k\right). \tag{4.19}$$

Collecting (4.15) through (4.19),

$$\int_{-\infty}^{ay} \overline{F}_V(y-w) dF_W(w) = \prod_{k=1}^{\infty} \Gamma\left(1-(0.5)^k\right)\left(e^{-y} - ye^{-2y}\right) + O\left(e^{-2y}\right) - \sqrt{\pi} e^{-y} \int_{ay}^{\infty} e^{w} dF_W(w).$$

By Lemma 4.3, the last integral above is $o\left(e^{-2y}\right)$, and the result follows. □

Finally, we allow $\frac{1}{2} < \beta < 1$ and see that the formula is similar to the earlier two.

Lemma 4.10. *If $\frac{1}{2} < \beta < 1$, then as $y \to \infty$*

$$P\left(\sum_{k=0}^{\infty} \beta^k Z_k > y\right) = \prod_{k=1}^{\infty} \Gamma\left(1-\beta^k\right)\left[e^{-y} - \frac{\beta}{1-\beta}\Gamma\left(2-\beta^{-1}\right) e^{-y/\beta}\right] + o\left(e^{-y/\beta}\right).$$

Proof. Let a and λ be positive reals such that $\beta < a < 1$, $1 < \lambda < \frac{1}{\beta}$, and $\lambda > \frac{1}{a}$. Once again consider the partition

$$P(X > y) = \int_{-\infty}^{ay} P(V > y-w) dF_W(w) + o\left(e^{-y/\beta}\right). \tag{4.20}$$

By Theorem 3.1, for y large enough and any $w < ay$

$$P(V > y - w) = \Gamma(1-\beta)e^{-(y-w)} - \frac{\beta}{1-\beta}\Gamma\left(2-\beta^{-1}\right)e^{-(y-w)/\beta} + R(y-w),$$

where for any $\epsilon > 0$ and y large enough

$$\sup_{w<ay}\left|R(y-w)e^{(y-w)/\beta}\right| < \epsilon.$$

Set $K_1 = \Gamma(1-\beta)$ and $K_2 = -\frac{\beta}{1-\beta}\Gamma\left(2-\beta^{-1}\right)$. Then observe that

$$\begin{aligned}\int_{-\infty}^{ay}\overline{F}_V(y-w)dF_W(w) = \int_{-\infty}^{ay}&\left[\frac{\overline{F}_V(y-w) - K_1e^{-(y-w)}}{K_2e^{-(y-w)/\beta}}\right]K_2e^{-(y-w)/\beta}dF_W(w)\\ &+\int_{-\infty}^{ay}K_1e^{-(y-w)}dF_W(w).\end{aligned} \tag{4.21}$$

Now

$$\frac{\overline{F}_V(y-w) - K_1e^{-(y-w)}}{K_2e^{-(y-w)/\beta}} = 1 + \frac{1}{K_2}R(y-w)e^{(y-w)/\beta} = 1 + \epsilon(y-w).$$

Note that $\sup_{w<ay}|\epsilon(y-w)| < \dfrac{\epsilon}{K_2}$. Consider the piece

$$\begin{aligned}&\int_{-\infty}^{ay}[1+\epsilon(y-w)]K_2e^{-(y-w)/\beta}dF_W(w)\\ &\quad= K_2e^{-y/\beta}\int_{-\infty}^{ay}e^{w/\beta}dF_W(w) + K_2\int_{-\infty}^{ay}\epsilon(y-w)e^{-(y-w)/\beta}dF_W(w).\end{aligned} \tag{4.22}$$

Examining the first integral in (4.22),

$$\lim_{y\to\infty}\int_{-\infty}^{ay} K_2 e^{w/\beta} dF_W(w) = K_2 E\left(e^{W/\beta}\right) = K_2 E\left[\exp\left(\sum_{k=1}^{\infty}\beta^k Z_{k+1}\right)\right]$$
$$= K_2\prod_{k=1}^{\infty} E\left(e^{\beta^k Z_1}\right) = -\frac{\beta}{1-\beta}\Gamma\left(2-\beta^{-1}\right)\prod_{k=1}^{\infty}\Gamma(1-\beta^k). \tag{4.23}$$

Turning to the second integral,

$$\left|K_2\int_{-\infty}^{ay}\epsilon(y-w)e^{-(y-w)/\beta}dF_W(w)\right| = \left|\int_{-\infty}^{ay} R(y-w)dF_W(w)\right|$$
$$= \left|\int_{-\infty}^{ay} R(y-w)e^{(y-w)/\beta}e^{-(y-w)/\beta}dF_W(w)\right| < \epsilon\int_{-\infty}^{ay} e^{-(y-w)/\beta}dF_W(w)$$
$$\leq \epsilon e^{-y/\beta}\int_{-\infty}^{\infty} e^{w/\beta}dF_W(w) = \epsilon E\left(e^{W/\beta}\right)e^{-y/\beta}.$$

By Lemma 4.1, $E\left(e^{W/\beta}\right) < \infty$ and therefore

$$K_2\int_{-\infty}^{ay}\epsilon(y-w)e^{-(y-w)/\beta}dF_W(w) = o\left(e^{-y/\beta}\right). \tag{4.24}$$

Next, note that

$$\int_{-\infty}^{\infty}\Gamma(1-\beta)e^w dF_W(w) = \Gamma(1-\beta)E\left(e^W\right) = \prod_{k=1}^{\infty}\Gamma(1-\beta^k). \tag{4.25}$$

We now consider the integral

$$K_1 e^{-y}\int_{-\infty}^{ay} e^w dF_W(w) = \prod_{k=1}^{\infty}\Gamma(1-\beta^k)e^{-y} - \Gamma(1-\beta)e^{-y}\int_{ay}^{\infty} e^w dF_W(w). \tag{4.26}$$

Collecting (4.23) through (4.26),

$$\int_{-\infty}^{ay} \overline{F}_V(y-w)dF_W(w) = \prod_{k=1}^{\infty} \Gamma(1-\beta^k) \left[e^{-y} - \frac{\beta}{1-\beta} \Gamma\left(2-\beta^{-1}\right) e^{-y/\beta} \right] + o\left(e^{-y/\beta}\right)$$
$$- \Gamma(1-\beta)e^{-y} \int_{ay}^{\infty} e^w dF_W(w).$$

By Lemma 4.3, the last integral above is $o\left(e^{-y/\beta}\right)$, and the result follows. □

4.3 The AR(1) Process when $-1 < \beta < 0$

Up to now we have assumed that $0 < \beta < 1$ when performing the two-term expansions on the AR(1) process with Gumbel innovations. We now turn to the other possibility, when $-1 < \beta < 0$. With a few necessary changes and extensions, the derivation is similar to what we saw in the previous sections. To get started, write

$$\sum_{k=0}^{\infty} \beta^k Z_k \stackrel{d}{=} \sum_{k=0}^{\infty} \beta^{2k} Z_k + \sum_{k=0}^{\infty} \beta^{2k+1} Z_k = S + T. \tag{4.27}$$

Notice that all the weights in T are negative since all the β's have odd exponents. Further note that in S, $0 < \beta^2 < 1$ and so we may utilize the same techniques as earlier, but basing our results on β^2 rather than just β. We shall split into three cases, but this time $\beta = -\frac{\sqrt{2}}{2}$ emerges as the necessary pivot point. Lemma 4.11 examines the first case, and the remaining two are mentioned in Theorem 4.1 and proven in Appendices B.1 and B.2.

Lemma 4.11. *If $-\frac{\sqrt{2}}{2} < \beta < 0$ and Z_k are i.i.d. standard Gumbel random variables, then as $y \to \infty$*

$$P\left(\sum_{k=0}^{\infty} \beta^k Z_k > y\right) = \prod_{k=1}^{\infty} \Gamma\left(1-\beta^k\right) e^{-y} - \frac{1}{2} \prod_{k=1}^{\infty} \Gamma\left(1-2\beta^k\right) e^{-2y} + O\left(e^{-y/\beta^2}\right).$$

Proof. First notice that by Lemma 4.8,

$$P(S > y) = \prod_{k=1}^{\infty} \Gamma\left(1 - \beta^{2k}\right) e^{-y} - \frac{1}{2} \prod_{k=1}^{\infty} \Gamma\left(1 - 2\beta^{2k}\right) e^{-2y} + O\left(e^{-y/\beta^2}\right). \tag{4.28}$$

Let $X = S + T$, and for some $0 < a < 1$ write

$$P(X > y) = \int_{ay}^{\infty} P(S > y - t) dF_T(t) + \int_{-\infty}^{ay} P(S > y - t) dF_T(t). \tag{4.29}$$

First,

$$\int_{ay}^{\infty} P(S > y - t) dF_T(t) \leq P(T > ay),$$

which by Lemma 4.5 is $o\left(e^{-y/\beta^2}\right)$ as $y \to \infty$. Now by (4.28) we have that for y large enough and any $t < ay$

$$P(S > y - t) = \prod_{k=1}^{\infty} \Gamma\left(1 - \beta^{2k}\right) e^{-(y-t)} - \frac{1}{2} \prod_{k=1}^{\infty} \Gamma\left(1 - 2\beta^{2k}\right) e^{-2(y-t)} + R(y - t),$$

where for some $c > 0$

$$\sup_{t<ay} \left| R(y - t) e^{(y-t)/\beta^2} \right| < c.$$

Set $K_1 = \prod_{k=1}^{\infty} \Gamma(1 - \beta^{2k})$ and $K_2 = -\frac{1}{2} \prod_{k=1}^{\infty} \Gamma(1 - 2\beta^{2k})$, and let F_S be the distribution of S. Then observe that

$$\begin{aligned} \int_{-\infty}^{ay} \overline{F}_S(y - t) dF_T(t) = \int_{-\infty}^{ay} \left[\frac{\overline{F}_S(y - t) - K_1 e^{-(y-t)}}{K_2 e^{-2(y-t)}} \right] K_2 e^{-2(y-t)} dF_T(t) \\ + K_1 \int_{-\infty}^{ay} e^{-(y-t)} dF_T(t). \end{aligned} \tag{4.30}$$

Now

$$\frac{\overline{F}_S(y-t) - K_1 e^{-(y-t)}}{K_2 e^{-2(y-t)}} = 1 + \frac{1}{K_2} R(y-t) e^{2(y-t)} = 1 + \epsilon(y-t).$$

We examine the $\epsilon(y-t)$:

$$\begin{aligned}
\sup_{t<ay} |\epsilon(y-t)| &= \frac{1}{|K_2|} \sup_{t<ay} \left| R(y-t) e^{(y-t)/\beta^2} e^{-(y-t)/\beta^2} e^{2(y-t)} \right| \\
&< \frac{c}{|K_2|} \sup_{t<ay} \left| \exp\left(- \left(\frac{1}{\beta^2} - 2 \right) (y-t) \right) \right| \\
&\leq \frac{c}{|K_2|} \exp\left(- \frac{(1-a)(1-2\beta^2) y}{\beta^2} \right).
\end{aligned}$$

Because $\frac{(1-a)(1-2\beta^2)}{\beta^2} > 0$, we have that $\epsilon(y-t) \to 0$ as $y \to \infty$ uniformly in $t < ay$. Consider the piece

$$\begin{aligned}
&\int_{-\infty}^{ay} [1 + \epsilon(y-t)] K_2 e^{-2(y-t)} dF_T(t) \\
&\quad = K_2 e^{-2y} \int_{-\infty}^{ay} e^{2t} dF_T(t) + K_2 \int_{-\infty}^{ay} \epsilon(y-t) e^{-2(y-t)} dF_T(t).
\end{aligned} \tag{4.31}$$

Examining the first integral in (4.31),

$$\begin{aligned}
&\lim_{y\to\infty} \int_{-\infty}^{ay} K_2 e^{2t} dF_T(t) = -\frac{1}{2} \prod_{k=1}^{\infty} \Gamma(1 - 2\beta^{2k}) \int_{-\infty}^{\infty} e^{2t} dF_T(t) \\
&= -\frac{1}{2} \prod_{k=1}^{\infty} \Gamma(1 - 2\beta^{2k}) E\left(e^{2T}\right) = -\frac{1}{2} \prod_{k=1}^{\infty} \Gamma(1 - 2\beta^{2k}) E\left[\exp\left(\sum_{k=0}^{\infty} 2\beta^{2k+1} Z_k \right) \right] \\
&= -\frac{1}{2} \prod_{k=1}^{\infty} \Gamma(1 - 2\beta^{2k}) \prod_{k=0}^{\infty} E\left(e^{2\beta^{2k+1} Z_1} \right) \\
&= \frac{1}{2} \prod_{k=1}^{\infty} \Gamma(1 - 2\beta^{2k}) \prod_{k=0}^{\infty} \Gamma(1 - 2\beta^{2k+1}) = -\frac{1}{2} \prod_{k=1}^{\infty} \Gamma(1 - 2\beta^{k}).
\end{aligned} \tag{4.32}$$

Turning to the second integral in (4.31),

$$\begin{aligned}
&\left| K_2 \int_{-\infty}^{ay} \epsilon(y-t)e^{-2(y-t)} dF_T(t) \right| = \left| \int_{-\infty}^{ay} R(y-t) dF_T(t) \right| \\
&\quad = \left| \int_{-\infty}^{ay} R(y-t) e^{(y-t)/\beta^2} e^{-(y-t)/\beta^2} dF_T(t) \right| < c \int_{-\infty}^{ay} e^{-(y-t)/\beta^2} dF_T(t) \\
&\quad \leq c e^{-y/\beta^2} \int_{-\infty}^{\infty} e^{t/\beta^2} dF_T(t) = cE\left(e^{T/\beta^2}\right) e^{-y/\beta^2}.
\end{aligned}$$

By Lemma 4.6, $E\left(e^{T/\beta^2}\right) < \infty$ and therefore

$$K_2 \int_{-\infty}^{ay} \epsilon(y-t)e^{-2(y-t)} dF_T(t) = O\left(e^{-y/\beta^2}\right). \tag{4.33}$$

Next, note that

$$K_1 \int_{-\infty}^{\infty} e^t dF_T(t) = K_1 E\left(e^T\right) = \prod_{k=1}^{\infty} \Gamma(1-\beta^{2k}) \prod_{k=0}^{\infty} \Gamma(1-\beta^{2k+1}) = \prod_{k=1}^{\infty} \Gamma(1-\beta^k). \tag{4.34}$$

We now consider the integral

$$K_1 e^{-y} \int_{-\infty}^{ay} e^t dF_T(t) = \prod_{k=1}^{\infty} \Gamma(1-\beta^k) e^{-y} - K_1 e^{-y} \int_{ay}^{\infty} e^t dF_T(t). \tag{4.35}$$

Collecting (4.32) through (4.35),

$$\begin{aligned}
\int_{-\infty}^{ay} \overline{F}_S(y-t) dF_T(t) = \prod_{k=1}^{\infty} \Gamma(1-\beta^k) e^{-y} - \frac{1}{2} \prod_{k=1}^{\infty} \Gamma(1-2\beta^k) e^{-2y} + O\left(e^{-y/\beta^2}\right) \\
- K_1 e^{-y} \int_{ay}^{\infty} e^t dF_T(t).
\end{aligned}$$

By Lemma 4.7, the last integral above is $o(e^{-y/\beta^2})$, and the result follows. □

4.4 The Complete AR(1) Expansion Result

In this short section, we state the main result of the chapter. Thereom 4.1 combines Lemmas 4.8 through 4.11, plus it includes the cases $\beta = -\frac{\sqrt{2}}{2}$ and $-1 < \beta < -\frac{\sqrt{2}}{2}$, the proofs of which are in Appendices B.1 and B.2. In addition, Appendices B.5 and B.6 contain tables of values for the infinite products $\prod_{k=1}^{\infty} \Gamma\left(1-\beta^k\right)$ and $\prod_{k=1}^{\infty} \Gamma\left(1-2\beta^k\right)$, respectively.

Theorem 4.1. *Let Z_k be i.i.d. standard Gumbel random variables. If $0 < \beta < \frac{1}{2}$, then as $y \to \infty$*

$$P\left(\sum_{k=0}^{\infty} \beta^k Z_k > y\right) = \prod_{k=1}^{\infty} \Gamma\left(1-\beta^k\right) e^{-y} - \frac{1}{2} \prod_{k=1}^{\infty} \Gamma\left(1-2\beta^k\right) e^{-2y} + O\left(e^{-y/\beta}\right).$$

If $\beta = \frac{1}{2}$,

$$P\left(\sum_{k=0}^{\infty} (0.5)^k Z_k > y\right) = \prod_{k=1}^{\infty} \Gamma\left(1-(0.5)^k\right)\left(e^{-y} - ye^{-2y}\right) + O\left(e^{-2y}\right).$$

When $\frac{1}{2} < \beta < 1$,

$$P\left(\sum_{k=0}^{\infty} \beta^k Z_k > y\right) = \prod_{k=1}^{\infty} \Gamma\left(1-\beta^k\right)\left(e^{-y} - \frac{\beta}{1-\beta}\Gamma\left(2-\beta^{-1}\right) e^{-y/\beta}\right) + o\left(e^{-y/\beta}\right).$$

When $-\frac{\sqrt{2}}{2} < \beta < 0$, and Z_k are i.i.d. standard Gumbel random variables, then

$$P\left(\sum_{k=0}^{\infty} \beta^k Z_k > y\right) = \prod_{k=1}^{\infty} \Gamma\left(1-\beta^k\right) e^{-y} - \frac{1}{2} \prod_{k=1}^{\infty} \Gamma\left(1-2\beta^k\right) e^{-2y} + O\left(e^{-y/\beta^2}\right).$$

If $\beta = -\frac{\sqrt{2}}{2}$*, then*

$$P\left(\sum_{k=0}^{\infty}\beta^k Z_k > y\right) = \prod_{k=1}^{\infty}\Gamma\left(1-\beta^k\right)\prod_{k=0}^{\infty}\Gamma\left(1+\sqrt{2}(0.5)^{k+1}\right)e^{-y}$$
$$-\prod_{k=1}^{\infty}\Gamma\left(1-\beta^k\right)\prod_{k=0}^{\infty}\Gamma\left(1+\sqrt{2}(0.5)^{k}\right)ye^{-2y} + O\left(e^{-2y}\right),$$

and to five decimal places the approximation is

$$P\left(\sum_{k=0}^{\infty}\beta^k Z_k > y\right) = 1.19005e^{-y} - 1.49210ye^{-2y} + O\left(e^{-2y}\right).$$

Finally, if $-1 < \beta < -\frac{\sqrt{2}}{2}$*, then*

$$P\left(\sum_{k=0}^{\infty}\beta^k Z_k > y\right) = \prod_{k=1}^{\infty}\Gamma\left(1-\beta^k\right)\left[e^{-y} - \frac{\beta^2\Gamma(1-\beta^{-1})\Gamma(2-\beta^{-2})}{1-\beta^2}e^{-y/\beta^2}\right]$$
$$+ o\left(e^{-y/\beta^2}\right).$$

As a side note, if we have $n < \infty$ *Gumbel random variables, then the products are taken from* $k = 0$ *or* 1 *up to* $n-1$*.*

4.5 The AR(1) Process in the Non-IID Setting

Finally, we shall derive a two-term expansion for the situation in which the Gumbel random variables take different location and scale parameters. That is, $X_k = \sigma Z_k + \mu$ with $Z_k \sim \Lambda, \sigma_k > 0$. Such a situation would be ideal when a time trend exists in the location and/or scale parameters. For instance, in the Peachtree Creek data set we considered a linear time trend in the scale parameter via $\sigma_k = \sigma + \phi k$. Often one may wish to define $\sigma_k = \sigma e^{\phi k}$ to ensure that the scale remains positive. It may also be desirable to test for quadratic trends in the location or scale, or even using covariate terms. Such possibilities are discussed in Coles (2001) and Coles (2008).

Because of the nonstationarity of the Gumbel time series, careful assumptions need to be imposed if one were to build an infinite series. Instead, we focus on the finite weighted sum. Define $X_k = 0$ for $k < 1$ and $Y_k = \beta Y_{k-1} + X_k$ for $1 \leq k < n$. Then $Y_n = \sum_{k=0}^{n-1} \beta^k X_{n-k}$ where $2 \leq n < \infty$. This situation is preferable because in data analysis, a time series model would be fit to a finite series anyway. Observe that

$$Y_n = \sum_{k=0}^{n-1} \beta^k \left(\sigma_{n-k} Z_{n-k} + \mu_{n-k}\right) \stackrel{d}{=} \sum_{k=0}^{n-1} \beta^k \sigma_{n-k} Z_k + \sum_{k=0}^{n-1} \beta^k \mu_{n-k}.$$

Some new notation is needed before discussing this generalization. First, since the series is finite we only require $\beta > 0$. (The case where $\beta < 0$ will be left as an open question.) Put $\overline{\mu}_n = \sum_{k=0}^{n-1} \beta^k \mu_{n-k}$. Define $\beta^{(0)} = \max\left(\beta^k \sigma_{n-k}\right)$ and $\beta^{(1)} = \max\left(\beta^k \sigma_{n-k} : \beta^k \sigma_{n-k} < \beta^{(0)}\right)$. That is, $\beta^{(0)}$ and $\beta^{(1)}$ are the highest and second highest amongst the $\beta^k \sigma_{n-k}$, and it is assumed that there are no multiplicites of these two quantities. Finally, let $\beta^{(m)}, 2 \leq m \leq n-1$ denote the $(m+1)th$ highest of the $\beta^k \sigma_{n-k}$. To be clear, $\beta^{(0)} > \beta^{(1)} > \beta^{(2)} \geq \beta^{(3)} \geq \cdots \geq \beta^{(n-1)}$. In which case, we may rewrite the series as

$$Y_n \stackrel{d}{=} \sum_{k=0}^{n-1} \beta^{(k)} Z_k + \overline{\mu}_n = \beta^{(0)} \left[\sum_{k=0}^{n-1} \frac{\beta^{(k)}}{\beta^{(0)}} Z_k + \overline{\mu}_n\right].$$

In what follows in Theorem 4.2, define $V = Z_0 + \frac{\beta^{(1)}}{\beta^{(0)}} Z_1$, $W = \sum_{k=2}^{n-1} \frac{\beta^{(k)}}{\beta^{(0)}} Z_k$, and $Y_n = V + W$. Then we shall derive the probability $P(Y_n > y)$, or equivalently $P(Y_n > y_n^*)$ where $y_n^* = \frac{y - \overline{\mu}_n}{\beta^{(0)}}$. This time the pivot point is $\frac{\beta^{(1)}}{\beta^{(0)}} = \frac{1}{2}$.

Theorem 4.2. *Let $\beta > 0$ and $Y_n, \beta^{(k)}, \overline{\mu}_n$, and y_n^* be defined as earlier with no multiplicities* *of* $\beta^{(0)}$ *nor* $\beta^{(1)}$. *If* $0 < \frac{\beta^{(1)}}{\beta^{(0)}} < \frac{1}{2}$, *then as* $y \to \infty$ *a two-term expansion is given by*

$$P(Y_n > y_n^*) = \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right) e^{-y_n^*} - \frac{1}{2} \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{2\beta^{(k)}}{\beta^{(0)}}\right) e^{-2y_n^*} + O\left(e^{-y/\beta^{(1)}}\right).$$

In the particular case where $\frac{\beta^{(1)}}{\beta^{(0)}} = \frac{1}{2}$, *the expansion is*

$$P(Y_n > y_n^*) = \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right) e^{-y_n^*} - \prod_{k=2}^{n-1} \Gamma\left(1 - \frac{2\beta^{(k)}}{\beta^{(0)}}\right) y_n^* e^{-2y_n^*} + O\left(e^{-y/\beta^{(1)}}\right).$$

Lastly, when $\frac{1}{2} < \frac{\beta^{(1)}}{\beta^{(0)}} < 1$, *the expansion is*

$$\begin{aligned} P(Y_n > y_n^*) = &-\frac{\beta^{(1)}}{\beta^{(0)} - \beta^{(1)}} \Gamma\left(2 - \frac{\beta^{(0)}}{\beta^{(1)}}\right) \prod_{k=2}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(1)}}\right) e^{-\beta^{(0)} y_n^*/\beta^{(1)}} \\ &+ \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right) e^{-y_n^*} + o\left(e^{-y/\beta^{(1)}}\right). \end{aligned}$$

Proof. We present the proof of the first equation only; the other two follow similar extensions of Theorem 4.1 and may be found in Appendices B.3 and B.4. Let a be a positive real such that $\frac{\beta^{(2)}}{\beta^{(1)}} < a < 1$. Then observe that

$$P(V + W > y_n^*) = \int_{ay_n^*}^{\infty} P(V > y_n^* - w) dF_W(w) + \int_{-\infty}^{ay_n^*} P(V > y_n^* - w) dF_W(w).$$

For $1 < \lambda < \frac{\beta^{(1)}}{\beta^{(2)}}$ we have

$$\begin{aligned} \int_{ay_n^*}^{\infty} P(V > y_n^* - w) dF_W(w) &\le P(W > ay_n^*) \\ &\le E\left[\exp\left(\frac{\beta^{(0)}\lambda W}{\beta^{(1)}}\right)\right] \exp\left(-\frac{\beta^{(0)}\lambda a y_n^*}{\beta^{(1)}}\right). \end{aligned}$$

Further choose λ so that $\lambda > \frac{1}{a}$, which is possible because $\frac{1}{a} < \frac{\beta^{(1)}}{\beta^{(2)}}$. Then

$$\int_{ay_n^*}^{\infty} P(V > y_n^* - w)dF_W(w) = O\left(\exp\left(-\frac{\beta^{(0)}\lambda a y_n^*}{\beta^{(1)}}\right)\right) = o\left(e^{-y/\beta^{(1)}}\right).$$

By Theorem 3.1, we have for y_n^* large enough and any $w < ay_n^*$ that

$$P(V > y_n^* - w) = \Gamma\left(1 - \frac{\beta^{(1)}}{\beta^{(0)}}\right) e^{-(y_n^* - w)} - \frac{1}{2}\Gamma\left(1 - \frac{2\beta^{(1)}}{\beta^{(0)}}\right) e^{-2(y_n^* - w)} + R(y_n^* - w),$$

where for some $c > 0$

$$\sup_{w < ay_n^*} \left| R(y_n^* - w) e^{\beta^{(0)}(y_n^* - w)/\beta^{(1)}} \right| < c.$$

Set $K_1 = \Gamma\left(1 - \frac{\beta^{(1)}}{\beta^{(0)}}\right)$ and $K_2 = -\frac{1}{2}\Gamma\left(1 - \frac{2\beta^{(1)}}{\beta^{(0)}}\right)$. Then observe that

$$\int_{-\infty}^{ay_n^*} \overline{F}_V(y_n^* - w)dF_W(w) = \int_{-\infty}^{ay_n^*} \left[\frac{\overline{F}_V(y_n^* - w) - K_1 e^{-(y_n^* - w)}}{K_2 e^{-2(y_n^* - w)}}\right] K_2 e^{-2(y_n^* - w)} dF_W(w)$$
$$+ K_1 \int_{-\infty}^{ay_n^*} e^{-(y_n^* - w)} dF_W(w).$$

Now

$$\frac{\overline{F}_V(y_n^* - w) - K_1 e^{-(y_n^* - w)}}{K_2 e^{-2(y_n^* - w)}} = 1 + \frac{1}{K_2} R(y_n^* - w) e^{2(y_n^* - w)} = 1 + \epsilon(y_n^* - w).$$

We examine the $\epsilon(y_n^* - w)$:

$$\begin{aligned}
\sup_{w < ay_n^*} |\epsilon(y_n^* - w)| &= \frac{1}{|K_2|} \sup_{w < ay_n^*} \left| R(y_n^* - w) e^{\beta^{(0)}(y_n^* - w)/\beta^{(1)}} e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}} e^{2(y_n^* - w)} \right| \\
&\leq \frac{c}{|K_2|} \sup_{w < ay_n^*} \left| \exp\left(-\left(\frac{\beta^{(0)}}{\beta^{(1)}} - 2\right)(y_n^* - w)\right)\right| \\
&\leq \frac{c}{|K_2|} \exp\left(-\frac{(1-a)(\beta^{(0)} - 2\beta^{(1)})y_n^*}{\beta^{(1)}}\right).
\end{aligned}$$

Because $\beta^{(0)} - 2\beta^{(1)} > 0$, $\epsilon(y_n^* - w) \to 0$ as $y_n^* \to \infty$ uniformly in $w < ay_n^*$. Consider the piece

$$\int_{-\infty}^{ay_n^*} [1 + \epsilon(y_n^* - w)] K_2 e^{-2(y_n^* - w)} dF_W(w)$$
$$= K_2 e^{-2y_n^*} \int_{-\infty}^{ay_n^*} e^{2w} dF_W(w) + K_2 \int_{-\infty}^{ay_n^*} \epsilon(y_n^* - w) e^{-2(y_n^* - w)} dF_W(w).$$

Examining the first integral,

$$\lim_{y_n^* \to \infty} \int_{-\infty}^{ay_n^*} K_2 e^{2w} dF_W(w) = -\frac{1}{2}\Gamma\left(1 - \frac{2\beta^{(1)}}{\beta^{(0)}}\right) \int_{-\infty}^{\infty} e^{2w} dF_W(w)$$
$$= -\frac{1}{2}\Gamma\left(1 - \frac{2\beta^{(1)}}{\beta^{(0)}}\right) E\left(e^{2W}\right) = -\frac{1}{2}\Gamma\left(1 - \frac{2\beta^{(1)}}{\beta^{(0)}}\right) E\left[\exp\left(\sum_{k=2}^{n-1} \frac{2\beta^{(k)}}{\beta^{(0)}} Z_k\right)\right]$$
$$= -\frac{1}{2}\Gamma\left(1 - \frac{2\beta^{(1)}}{\beta^{(0)}}\right) \prod_{k=2}^{n-1} E\left(e^{2\beta^{(k)} Z_1/\beta^{(0)}}\right) = -\frac{1}{2} \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{2\beta^{(k)}}{\beta^{(0)}}\right).$$

The last calculation was possible because $1 - \frac{2\beta^{(k)}}{\beta^{(0)}} > 0$ for all $1 \leq k \leq n-1$. Turning to the second integral,

$$K_2 \int_{-\infty}^{ay_n^*} \epsilon(y_n^* - w) e^{-2(y_n^* - w)} dF_W(w) = \int_{-\infty}^{ay_n^*} R(y_n^* - w) dF_W(w)$$
$$= \int_{-\infty}^{ay_n^*} R(y_n^* - w) e^{\beta^{(0)}(y_n^* - w)/\beta^{(1)}} e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}} dF_W(w)$$
$$< c \int_{-\infty}^{ay_n^*} e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}} dF_W(w) \leq c e^{-\beta^{(0)} y_n^*/\beta^{(1)}} \int_{-\infty}^{\infty} e^{\beta^{(0)} w/\beta^{(1)}} dF_W(w)$$
$$= cE\left(e^{\beta^{(0)} W/\beta^{(1)}}\right) e^{-\beta^{(0)} y_n^*/\beta^{(1)}}.$$

Therefore

$$K_2 \int_{-\infty}^{ay_n^*} \epsilon(y_n^* - w) e^{-2(y_n^* - w)} dF_W(w) = O\left(e^{-\beta^{(0)} y_n^*/\beta^{(1)}}\right) = O\left(e^{-y/\beta^{(1)}}\right).$$

Next, note that

$$\int_{-\infty}^{\infty} \Gamma\left(1-\frac{\beta^{(1)}}{\beta^{(0)}}\right) e^{w} dF_W(w) = \Gamma\left(1-\frac{\beta^{(1)}}{\beta^{(0)}}\right) E\left(e^{W}\right) = \prod_{k=1}^{n-1} \Gamma\left(1-\frac{\beta^{(k)}}{\beta^{(0)}}\right).$$

We now consider the integral

$$K_1 e^{-y_n^*} \int_{-\infty}^{ay_n^*} e^{w} dF_W(w) = \prod_{k=1}^{n-1} \Gamma\left(1-\frac{\beta^{(k)}}{\beta^{(0)}}\right) e^{-y_n^*} - \Gamma\left(1-\frac{\beta^{(1)}}{\beta^{(0)}}\right) e^{-y_n^*} \int_{ay_n^*}^{\infty} e^{w} dF_W(w).$$

Collecting all the terms,

$$\begin{aligned}\int_{-\infty}^{ay_n^*} \overline{F}_V(y_n^* - w) dF_W(w) &= \prod_{k=1}^{n-1} \Gamma\left(1-\frac{\beta^{(k)}}{\beta^{(0)}}\right) e^{-y_n^*} - \frac{1}{2}\prod_{k=1}^{n-1} \Gamma\left(1-\frac{2\beta^{(k)}}{\beta^{(0)}}\right) e^{-2y_n^*} \\ &\quad + O\left(e^{-y/\beta^{(1)}}\right) - \Gamma(1-\beta) e^{-y_n^*} \int_{ay_n^*}^{\infty} e^{w} dF_W(w).\end{aligned}$$

For the last integral above, note that by integration by parts

$$\int_{ay_n^*}^{\infty} e^{w} dF_W(w) = -\int_{ay_n^*}^{\infty} e^{w} d\overline{F}_W(w) = e^{ay_n^*}\overline{F}_W(ay_n^*) + \int_{ay_n^*}^{\infty} \overline{F}_W(w) e^{w} dw.$$

Observe that

$$e^{ay_n^*}\overline{F}_W(ay_n^*) \leq E\left[\exp\left(\frac{\beta^{(0)}\lambda W}{\beta^{(1)}}\right)\right] \exp\left(-\left(\frac{\beta^{(0)}\lambda}{\beta^{(1)}} - 1\right) ay_n^*\right).$$

Similarly, we find

$$\begin{aligned}\int_{ay_n^*}^{\infty} \overline{F}_W(w) e^{w} dw &\leq E\left[\exp\left(\frac{\beta^{(0)}\lambda W}{\beta^{(1)}}\right)\right] \int_{ay_n^*}^{\infty} \exp\left(-\frac{\beta^{(0)}\lambda w}{\beta^{(1)}} + w\right) dw \\ &= E\left[\exp\left(\frac{\beta^{(0)}\lambda W}{\beta^{(1)}}\right)\right] \left(\frac{\beta^{(0)}\lambda}{\beta^{(1)}} - 1\right)^{-1} \exp\left(-\left(\frac{\beta^{(0)}\lambda}{\beta^{(1)}} - 1\right) ay_n^*\right).\end{aligned}$$

Therefore for some $c > 0$ we obtain

$$\int_{ay_n^*}^{\infty} e^{-(y_n^*-w)} dF_W(w) \leq c \exp\left(-\left(\frac{\beta^{(0)}\lambda a}{\beta^{(1)}} + 1 - a\right) y_n^*\right).$$

Finally, since $\frac{\beta^{(2)}}{\beta^{(1)}} < a < 1$ and $\lambda > 1$ we have that $\frac{\beta^{(0)}\lambda a}{\beta^{(1)}} + 1 - a > \frac{\beta^{(0)}}{\beta^{(1)}}$, and therefore

$$\int_{ay_n^*}^{\infty} e^{-(y_n^*-w)} dF_W(w) = o\left(e^{-\beta^{(0)} y_n^*/\beta^{(1)}}\right) = o\left(e^{-y/\beta^{(1)}}\right).$$

The result follows since $o\left(e^{-y/\beta^{(1)}}\right) + O\left(e^{-y/\beta^{(1)}}\right) = O\left(e^{-y/\beta^{(1)}}\right)$. □

As mentioned earlier, Theorem 4.2 assumes that $\beta^{(0)}$ and $\beta^{(1)}$ each have multiplicity of one. For application purposes, thanks to the choice of $\beta^{(k)}$ this will almost always be the case. However, it may be possible to have ties when arbitrary constants are chosen. For instance, if we were convolving a sequence of random variables from two different Gumbel distributions, there may indeed be multiplicities to consider. Chapter 6 discusses how such a situation would be handled.

In order to ease the notation with $\beta^{(k)}$, we restate Theorem 4.2 using new constants, which are assumed to be positive.

Theorem 4.3. *Let $d_k, k = 1, \cdots, n, n \geq 3$ be positive constants, and define $c_k = d_{(k)}$, the order statistics arranged from largest to smallest. That is, $c_1 > c_2 > c_3 \geq c_4 \geq \cdots \geq c_n$, and in particular $c_1 = \max(d_k)$. Assume that c_1 and c_2 have multiplicities of 1. If $0 < 2c_2 < c_1$, then as $y \to \infty$ a two-term expansion is given by*

$$P\left(\sum_{k=1}^{n} d_k Z_k > y\right) = \prod_{k=2}^{n} \Gamma\left(1 - \frac{c_k}{c_1}\right) e^{-y/c_1} - \frac{1}{2}\prod_{k=2}^{n} \Gamma\left(1 - \frac{2c_k}{c_1}\right) e^{-2y/c_1} + O\left(e^{-y/c_2}\right).$$

In the particular case where $0 < 2c_2 = c_1$, the expansion is

$$P\left(\sum_{k=1}^{n} d_k Z_k > y\right) = \prod_{k=2}^{n} \Gamma\left(1 - \frac{c_k}{c_1}\right) e^{-y/c_1} - \prod_{k=3}^{n} \Gamma\left(1 - \frac{2c_k}{c_1}\right) \frac{y}{c_1} e^{-2y/c_1} + O\left(e^{-y/c_2}\right).$$

Lastly, when $0 < c_1 < 2c_2$, the expansion is

$$\begin{aligned} P\left(\sum_{k=1}^{n} d_k Z_k > y\right) &= -\frac{c_2}{c_1 - c_2} \Gamma\left(2 - \frac{c_1}{c_2}\right) \prod_{k=3}^{n} \Gamma\left(1 - \frac{c_k}{c_2}\right) e^{-y/c_2} \\ &\quad + \prod_{k=2}^{n} \Gamma\left(1 - \frac{c_k}{c_1}\right) e^{-y/c_1} + o\left(e^{-y/c_2}\right). \end{aligned}$$

4.6 Examples of AR(1) Processes

We now present several examples of how the process may appear in data analysis problems. In all of the following examples it is assumed that $\beta > 0$ and that there are $n < \infty$ units of time. We denote the Gumbel realizations as $X_1, ..., X_n$. It can be shown that the mean and variance of the process Y_n are given by

$$E(Y_n) = \gamma \sum_{k=0}^{n-1} \beta^k \sigma_{n-k} + \overline{\mu}_n \quad \text{and} \quad \text{Var}(Y_n) = \frac{\pi^2}{6} \sum_{k=0}^{n-1} \beta^{2k} \sigma_{n-k}^2. \tag{4.36}$$

Example 4.1. First, when all the $\mu_k = 0$ and $\sigma_k = 1$, Y_n reduces to the original process described in Theorem 4.1, taken over a finite time period. In addition,

$$E(Y_n) = \frac{\gamma(1 - \beta^n)}{1 - \beta} \quad \text{and} \quad \text{Var}(Y_n) = \frac{\pi^2(1 - \beta^{2n})}{6(1 - \beta^2)}.$$

If $0 < \beta < 1$, then we may extend the process to an $AR(1)$, in which case as $n \to \infty$

$$E(Y_n) \to \frac{\gamma}{1-\beta} \qquad \text{and} \qquad \mathrm{Var}(Y_n) \to \frac{\pi^2}{6(1-\beta^2)}.$$

Example 4.2. Now let $\mu_k = \mu$ and $\sigma_k = \sigma, \sigma > 0$, the general i.i.d. case. Then $\overline{\mu}_n = \mu(1-\beta^n)/(1-\beta)$ and observe that

$$P(Y_n > y) = P\left(\sum_{k=0}^{n-1} \beta^k Z_{n-k} > \frac{y - \overline{\mu}_n}{\sigma}\right) = P\left(\sum_{k=0}^{n-1} \beta^k Z_{n-k} > y_n^*\right),$$

allowing Theorem 4.1 to be used. Observe that in this situation, $O\left(e^{-y_n^*/\beta}\right) = O\left(e^{-y/\sigma\beta}\right)$ and $O\left(e^{-2y_n^*}\right) = O\left(e^{-2y/\sigma}\right)$. Further note that for Lemma 4.9, $y_n^* e^{-2y_n^*} = \frac{y}{\sigma} e^{-2y_n^*} + O\left(e^{-2y/\sigma}\right)$. Finally,

$$E(Y_n) = \frac{(\gamma\sigma + \mu)(1-\beta^n)}{1-\beta} \qquad \text{and} \qquad \mathrm{Var}(Y_n) = \frac{\pi^2\sigma^2(1-\beta^{2n})}{6(1-\beta^2)}.$$

If $0 < \beta < 1$, then as $n \to \infty$

$$E(Y_n) \to \frac{\gamma\sigma + \mu}{1-\beta} \qquad \text{and} \qquad \mathrm{Var}(Y_n) \to \frac{\pi^2\sigma^2}{6(1-\beta^2)}.$$

Example 4.3. Focusing just on the location parameters, let $\mu_k = \mu + \theta k$. Notice that this definition is how we defined the location parameter given season in the Peachtree Creek project, and $\theta = 0$ corresponds to no significant linear time effect. Then

$$\overline{\mu}_n = \frac{(\mu + \theta n)(1-\beta^n)}{1-\beta} - \frac{\theta\left[\beta - n\beta^n + (n-1)\beta^{n+1}\right]}{(1-\beta)^2}.$$

Example 4.4. Now suppose graphical evidence suggests the location parameter may have a significant quadratic effect. Then define $\mu_k = \mu + \theta_1 k + \theta_2 k^2$. By carefully computing the formula for $\sum_{k=0}^{n-1} k^2\beta^k$, it can be shown that

$$\overline{\mu}_n = \frac{(\mu+\theta_1 n+\theta_2 n^2)(1-\beta^n)}{1-\beta} - \frac{(\theta_1+2n\theta_2)(\beta-n\beta^n+(n-1)\beta^{n+1})}{(1-\beta)^2} + \frac{\theta_2\left[\beta(1+\beta)-n^2\beta^n+(2n^2-2n-1)\beta^{n+1}-(n-1)^2\beta^{n+2}\right]}{(1-\beta)^3}.$$

Example 4.5. We now turn to the scale parameters. In data analysis it is often of interest to define $\sigma_k = \sigma e^{\phi k}$, $\sigma > 0$, which ensures that the scales remain positive. In this example, we assume that $0 < \beta e^{-\phi} < 1$. Since $\beta^k \sigma_{n-k} = \sigma e^{n\phi}(\beta e^{-\phi})^k$ is strictly decreasing in k, it follows that $\beta^{(k)} = \beta^k\sigma_{n-k}, k = 0, 1, ..., n-1$. Then $y_n^* = (y - \overline{\mu}_n)/\sigma$ and the pivot point in Theorem 4.2 is $\beta^{(1)}/\beta^{(0)} = \beta e^{-\phi}$. As $y \to \infty$, the first result in Theorem 4.2 reads

$$P(Y_n > y) = \prod_{k=1}^{n-1}\Gamma\left(1-(\beta e^{-\phi})^k\right)e^{-y_n^*} - \frac{1}{2}\prod_{k=1}^{n-1}\Gamma\left(1-2(\beta e^{-\phi})^k\right)e^{-2y_n^*} + O\left(e^{-y/\beta^{(1)}}\right),$$

provided that $0 < \beta e^{-\phi} < \frac{1}{2}$. Lastly, given appropriately computed $\overline{\mu}_n$, the mean and variance of Y_n are

$$E(Y_n) = \gamma\sigma e^{n\phi}\left(\frac{1-(\beta e^{-\phi})^n}{1-\beta e^{-\phi}}\right) + \overline{\mu}_n \quad \text{and} \quad \mathrm{Var}(Y_n) = \frac{\sigma^2 e^{2\phi n}\pi^2}{6}\left(\frac{1-(\beta e^{-\phi})^{2n}}{1-(\beta e^{-\phi})^2}\right).$$

Example 4.6. Now suppose graphical evidence suggests defining the scale parameters in a linear fashion, namely $\sigma_k = \sigma + \phi k$, $\sigma > 0$. The advantage of such a definition is that linearity makes interpretation of the parameters much easier than an exponential definition. Of course, one has to be careful defining the parameter space of ϕ, for if $\phi < 0$ then at some point the scale parameters may become negative. Nevertheless, for application purposes we will have a finite time period, so in some cases it is possible to fit a model and allow ϕ to take on any real values. This definition is what we used in the Peachtree Creek data set.

The first task is to obtain $\beta^{(0)}$ and $\beta^{(1)}$. If the $\beta^k\sigma_{n-k}$ are strictly decreasing in k, then $\beta^{(0)} = \sigma + n\phi$ and $\beta^{(1)} = \beta(\sigma + (n-1)\phi)$. However, this may not happen, depending on the choice of values for the parameters. In which case, observe that the ratio of two consecutive terms of $\beta^k\sigma_{n-k}$ is

$$\frac{\beta^{k+1}(\sigma + \phi n - \phi(k+1))}{\beta^k(\sigma + \phi n - \phi k)} = \beta\left(1 - \frac{\phi}{\sigma + \phi n - \phi k}\right).$$

The maximum value $\beta^{(0)}$ occurs at $k = L$, where $L \geq 1$ satisfies

$$\beta\left(1 - \frac{\phi}{\sigma + \phi n - \phi L}\right) < 1 \quad \text{and} \quad \beta\left(1 - \frac{\phi}{\sigma + \phi n - \phi(L-1)}\right) > 1.$$

We need not investigate the equality possibility since it is assumed that multiplicities in the $\beta^{(0)}$ and $\beta^{(1)}$ do not happen. Combining the two conditions, locate the L that satisfies

$$\frac{\beta}{1-\beta} + \frac{\sigma}{\phi} + n < L < \frac{\beta}{1-\beta} + \frac{\sigma}{\phi} + n + 1.$$

Then $\beta^{(1)} = \max\left[\beta^{L-1}(\sigma + \phi(n+1-L)), \beta^{L+1}(\sigma + \phi(n-1-L))\right]$. Lastly, given $\overline{\mu}_n$, the mean and variance of Y_n are

$$E(Y_n) = \gamma(\sigma + \phi n)\left(\frac{1-\beta^n}{1-\beta}\right) - \gamma\phi\left(\frac{\beta - n\beta^n + (n-1)\beta^{n+1}}{(1-\beta)^2}\right) + \overline{\mu}_n$$

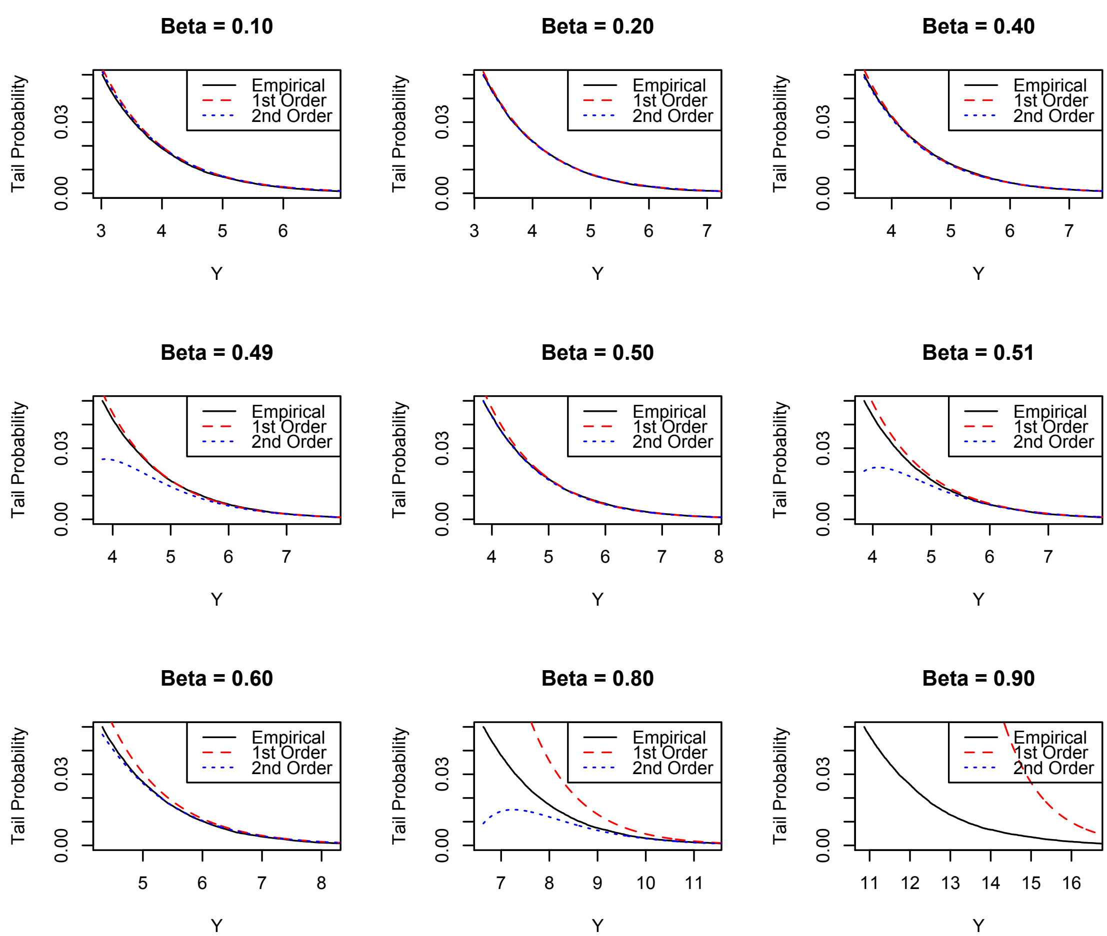
and

$$\begin{aligned}\mathrm{Var}(Y_n) = {} & \frac{\pi^2(\sigma + \phi n)^2}{6}\left(\frac{1-\beta^{2n}}{1-\beta^2}\right) - \frac{\pi^2\phi(\sigma + \phi n)}{3}\left(\frac{\beta^2 - n\beta^{2n} + (n-1)\beta^{2(n+1)}}{(1-\beta^2)^2}\right) \\ & + \frac{\pi^2\phi^2}{6}\left(\frac{\beta^2(1+\beta^2) - n^2\beta^{2n} + (2n^2 - 2n - 1)\beta^{2(n+1)} - (n-1)^2\beta^{2(n+2)}}{(1-\beta^2)^3}\right).\end{aligned}$$

4.7 Simulation Results

Using a similar setup as in Chapter 3, we now check how our six approximations behave for various values of $-1 < \beta < 1$. Let Y represent the AR(1) process. We simulate $N = 10$ million values of Y and graph the empirical tail probability, as well as the appropriate second-order approximation from Theorem 4.1, in Figure 4.1. For comparison purposes we also graph the first-order approximation.

Figure 4.1: First and Second-Order Approximations for $0 < \beta < 1$



We first investigate what happens for various positive values of β, especially near the $\beta = \frac{1}{2}$ pivot point. Figure 4.1 displays the empirical probability (solid black line), the first-order approximation (dashed red line), and the second-order approximation (dotted blue line) for the 95th percentile and higher. Table 4.1 displays the raw errors at specified percentiles, defined as empirical minus estimated probabilities. Table 4.2 contains the relative errors, the raw error divided by the estimated probability. In both tables, for a given β and percentile the better approximation is highlighted.

Table 4.1: Errors in Approximations of Theorem 4.1

β	**Approx**	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
0.10	1st	-0.0013	-0.0003	-0.0002	-6.2e-5	3.7e-7	1.4e-6	-3.9e-7
	2nd	7.9e-6	3.9e-6	-1.3e-5	-1.0e-5	1.3e-5	2.0e-6	-3.8e-7
0.20	1st	-0.0015	-0.0004	-0.0002	-5.2e-5	5.7e-6	1.9e-6	2.6e-6
	2nd	-6.3e-5	-3.8e-5	2.1e-5	3.1e-6	1.9e-5	2.5e-6	2.6e-6
0.40	1st	-0.0024	-0.0007	-0.0004	-7.9e-5	-2.9e-8	-1.5e-6	-1.3e-6
	2nd	0.0006	5.6e-5	4.1e-5	3.3e-5	2.8e-5	-4.6e-7	-1.3e-6
0.49	1st	-0.0042	-0.0012	-0.0008	-0.0003	-0.0001	-2.2e-5	-4.0e-6
	2nd	0.0247	0.0055	0.0034	0.0008	0.0001	-1.2e-5	-3.9e-6
0.50	1st	-0.0044	-0.0012	-0.0008	-0.0002	-8.8e-5	-2.3e-5	-5.0e-7
	2nd	8.1e-5	5.0e-5	2.6e-6	4.0e-6	-2.5e-5	-1.9e-5	-4.6e-7
0.51	1st	-0.0047	-0.0012	-0.0008	-0.0002	-5.0e-5	-2.4e-6	-3.6e-6
	2nd	0.0292	0.0068	0.0043	0.0010	0.0003	1.1e-5	-3.4e-6
0.60	1st	-0.0093	-0.0028	-0.0019	-0.0006	-0.0002	-2.9e-5	1.3e-8
	2nd	0.0039	0.0009	0.0006	0.0001	4.2e-5	-1.4e-5	3.3e-7
0.80	1st	-0.0861	-0.0303	-0.0219	-0.0082	-0.0031	-0.0003	-1.6e-5
	2nd	0.0401	0.0106	0.0070	0.0020	0.0006	5.8e-5	2.7e-6
0.90	1st	-1.5883	-0.5002	-0.3511	-0.1220	-0.0442	-0.0048	-0.0002
	2nd	3.1579	0.8407	0.5604	0.1670	0.0523	0.0042	0.0001

At first, we see very similar results as those from Figure 3.1. Once again, for $0 < \beta < 0.40$ there is virtually no difference in estimation between the approximations, although the second-order is slightly more accurate. Also the second-order approximation is better than the first-order for $\beta = 0.50$. In addition, we once again have a neighborhood around $\beta = 0.50$ in which the first-order is better for lower percentiles. This time, the neighborhood is approximately $[0.47, 0.53]$.

However, after this neighborhood ends we start to see some differences. For $\beta = 0.60$ and 0.80, the second-order is again better than the first, although for the latter both approxi-

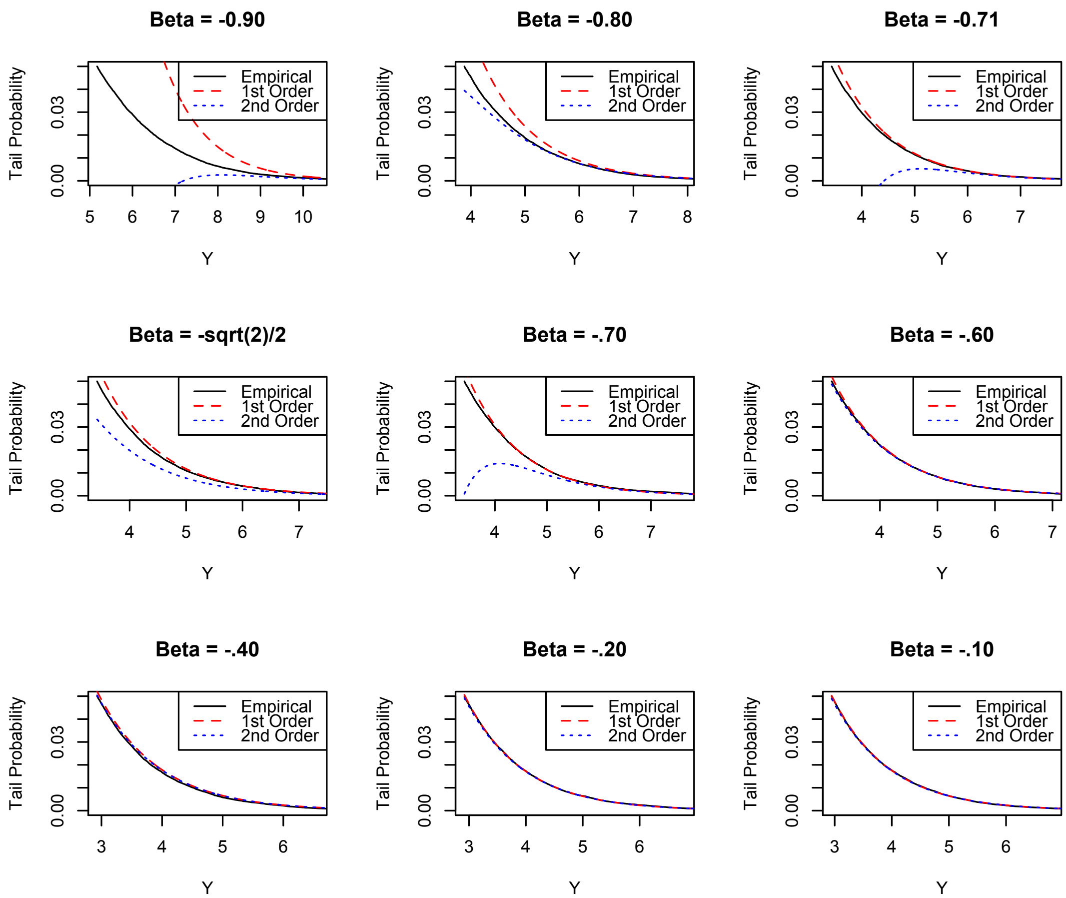
Table 4.2: Relative Errors in Approximations of Theorem 4.1

β	**Approx**	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
0.10	1st	-0.0260	-0.0128	-0.0110	-0.0061	0.0001	0.0014	-0.0039
	2nd	0.0002	0.0002	-0.0007	-0.0010	0.0026	0.0020	-0.0038
0.20	1st	-0.0296	-0.0155	-0.0101	-0.0052	0.0011	0.0019	0.0262
	2nd	-0.0013	-0.0015	0.0010	0.0003	0.0039	0.0025	0.0262
0.40	1st	-0.0466	-0.0261	-0.0205	-0.0079	-0.0000	-0.0016	-0.0132
	2nd	0.0120	0.0023	0.0020	0.0033	0.0055	-0.0005	-0.0131
0.49	1st	-0.0768	-0.0468	-0.0407	-0.0261	-0.0230	-0.0217	-0.0382
	2nd	0.9760	0.2847	0.2068	0.0833	0.0288	-0.0117	-0.0373
0.50	1st	-0.0804	-0.0450	-0.0391	-0.0217	-0.0172	-0.0222	-0.0049
	2nd	0.0016	0.0020	0.0001	0.0004	-0.0049	-0.0191	-0.0045
0.51	1st	-0.0859	-0.0473	-0.0384	-0.0230	-0.0098	-0.0024	-0.0347
	2nd	1.3989	0.3720	0.2727	0.1149	0.0565	0.0111	-0.0333
0.60	1st	-0.1569	-0.1021	-0.0869	-0.0569	-0.0357	-0.0286	0.0001
	2nd	0.0847	0.0375	0.0314	0.0148	0.0086	-0.0139	0.0033
0.80	1st	-0.6327	-0.5483	-0.5231	-0.4497	-0.3855	-0.2479	-0.1346
	2nd	4.0589	0.7410	0.5431	0.2523	0.1349	0.0616	0.0281
0.90	1st	-0.9695	-0.9524	-0.9461	-0.9242	-0.8984	-0.8281	-0.7101
	2nd	-1.0161	-1.0306	-1.0370	-1.0637	-1.1056	-1.3138	-3.2141

mations begin to perform poorly for percentiles less than the 99th. And finally, for $\beta = .90$ both approximations are useless. This is not a surprising result since according to Tables B.5 and B.6 in the appendices, the infinite products in the expansions grow large quickly.

In Chapter 3, we discovered that the two approximations were very accurate and indistinguishable for $-1 < \beta < 0$, and for that reason excluded the results. However, here we actually do have some interesting differences as β gets closer to -1. Figure 4.2 explores various negative values of β. This time we investigate what happens in a neighborhood around the pivot point $\beta = \frac{\sqrt{2}}{2}$. We estimate the neighborhood to be $[-0.73, -0.68]$.

Figure 4.2: First and Second-Order Approximations for $-1 < \beta < 0$



As before, for β not too near the pivot point we see hardly any difference in estimation. In the neighborhood around $\beta = \frac{\sqrt{2}}{2}$, the first-order approximation is better, and this time including the pivot point itself. But for $-1 < \beta < -0.80$, both approximations are poor except for very high percentiles. The conclusions we can draw are as follows:

1. On $\beta \in [-0.73, -0.68] \cup [0.47, 0.53) \setminus \{0.50\}$, use the first-order approximation.

2. On $\beta \in [-0.80, -0.73) \cup (-0.68, 0.47) \cup \{0.50\} \cup (0.53, 0.80]$, use the second-order approximation.

3. Except for very high percentiles, both approximations are useless for $\beta \in (-1, -0.80) \cup (0.80, 1)$.

4.8 Fitting the Model to the Peachtree Creek Data

Having established the theory behind Theorem 4.2, we now fit the model in (4.37) to the Peachtree Creek data set used in Chapter 2. We first review the notation that was used. Recall that x_{ti} was the maximum mean streamflow in year t and season i, $t = 1, ..., 52$ and $i = 1, ..., 4$. The seasons were, respectively, summer, fall, winter, and spring, while $t = 1$ corresponded to the time period June 1, 1958 through May 31, 1959.

To update the cumulative distribution function, we have

$$F(x_{ti}) = \exp\left\{-\left[1 + \frac{\xi_i(x_{ti} - \mu_i(t)}{\sigma_i(t)}\right]^{-\frac{1}{\xi_i}}\right\}, \xi_i \neq 0, 1 + \frac{\xi_i(x_{ti} - \mu_i(t))}{\sigma_i(t)} > 0, \tag{4.37}$$

where $\mu_i(t) = \mu_i + \eta_i t + \beta x_{t-1,i}$ and $\sigma_i(t) = \sigma_i + \phi_i t$, otherwise

$$F(x_{ti}) = \exp\left\{-\exp\left[-\left(\frac{x_{ti} - \mu_i(t)}{\sigma_i(t)}\right)\right]\right\}, \xi_i = 0. \tag{4.38}$$

Recall that summer and winter ($i = 1, 3$) satisfied the Gumbel model, while fall and spring ($i = 2, 4$) had a shape parameter that was significantly different from 0. Therefore we refit the data keeping these same shapes per season, only introducing the β_i. If $\ln L_i$ is the log likelihood per season with a total of either five or six parameters, then for fall and winter

$$\mathrm{lnL}_i = -\sum_{t=2}^{52}\left\{\log(\sigma_i(t)) + \left(\frac{1}{\xi_i}+1\right)\log\left(1+\frac{\xi_i(x_{ti}-\mu_i(t))}{\sigma_i(t)}\right)\right.$$
$$\left. + \left(1+\frac{\xi_i(x_{ti}-\mu_i(t))}{\sigma_i(t)}\right)^{-\frac{1}{\xi_i}}\right\},$$

provided that both of the $1 + [\xi_i(x_{ti} - \mu_i(t))]/\sigma_i(t) > 0$, otherwise $\mathrm{lnL}_i = -\infty$. For summer and winter,

$$\mathrm{lnL}_i = -\sum_{t=2}^{52}\left\{\log(\sigma_i(t)) + \left(\frac{x_{ti}-\mu_i(t)}{\sigma_i(t)}\right) + \exp\left[-\left(\frac{x_{ti}-\mu_i(t)}{\sigma_i(t)}\right)\right]\right\}.$$

The results for all 22 parameters are shown in Table 4.3.

Table 4.3: Estimated Parameters per Season

Season	Param.	Estimate	SE	90% Confidence	T-Stat	P-Value
Summer	ξ_1	0	—	—	—	—
	μ_1	763.7013	190.4298	(444.0344, 1083.3682)	4.0104	0.0002
	σ_1	454.0155	105.7854	(276.4378, 631.5932)	4.2919	9.0e-5
	η_1	7.2813	6.2497	(-3.2098, 17.7723)	1.1651	0.2500
	β_1	0.0023	0.1191	(-0.1977, 0.2023)	0.0196	0.9845
	ϕ_1	5.2739	3.9347	(-1.3312, 11.8790)	1.3403	0.1867
Fall	ξ_2	0.2444	0.1321	(0.0225, 0.4663)	1.8500	0.0709
	μ_2	483.3228	165.3661	(205.6025, 761.0432)	2.9227	0.0054
	σ_2	254.7333	138.1127	(22.7830, 486.6835)	1.8444	0.0717
	η_2	18.0495	7.4977	(5.4577, 30.6413)	2.4073	0.0202
	β_2	0.0342	0.0888	(-0.1149, 0.1832)	0.3851	0.7020
	ϕ_2	15.5458	6.8602	(4.0245, 27.0671)	2.2661	0.0283
Winter	ξ_3	0	—	—	—	—
	μ_3	1718.9809	406.1681	(1037.1625, 2400.7992)	4.2322	0.0001
	σ_3	1097.5508	239.9702	(694.7224, 1500.3793)	4.5737	3.6e-5
	η_3	-11.0343	8.1862	(-24.7760, 2.7075)	-1.3479	0.1843
	β_3	0.0329	0.1176	(-0.1646, 0.2304)	0.2798	0.7809
	ϕ_3	-11.2262	7.0487	(-23.0587, 0.6062)	-1.5927	0.1181
Spring	ξ_4	0.7813	0.2262	(0.4013, 1.1613)	3.4534	0.0012
	μ_4	1716.7448	256.6974	(1285.6402, 2147.8495)	6.6878	3.0e-8
	σ_4	851.9976	280.5456	(380.8416, 1323.1536)	3.0369	0.0040
	η_4	-14.3076	6.8744	(-25.8527, -2.7625)	-2.0813	0.0431
	β_4	-0.0219	0.0361	(-0.0825, 0.0388)	-0.6056	0.5478
	ϕ_4	-7.0763	8.4730	(-21.3061, 7.1536)	-0.8352	0.4080

We should point out that in Figures 2.9 and 2.10, there was no visual evidence of a significant effect from the most recent observation in any of the seasons. That is, in the ACF and PACF plots, there were no spikes at lag 1. Therefore it is not a surprise that none of the β_i estimates are significantly different from 0. We might as well drop them from the model and stick with the one we fit in Section 2.4. The rest of the estimates are fairly similar to those in Table 2.4.

Nevertheless, to illustrate how Theorem 4.2 works in practice, we shall keep working with this model from Table 4.3. First, if $V_i, i = 1, ..., 4$ is the variance-covariance matrix for season i and the 0's are matrices with zero entries and appropriate dimensions, then the complete variance-covariance matrix of all 22 parameters is given by

$$V = \begin{pmatrix} V_1 & 0 & 0 & 0 \\ 0 & V_2 & 0 & 0 \\ 0 & 0 & V_3 & 0 \\ 0 & 0 & 0 & V_4 \end{pmatrix},$$

Furthermore, the V_i are

$$V_1 = \begin{pmatrix} 36263.491 & 7424.394 & -600.696 & -11.959 & -235.107 \\ 7424.394 & 11190.541 & -160.155 & -1.584 & -331.015 \\ -600.696 & -160.155 & 39.058 & -0.248 & 7.046 \\ -11.959 & -1.584 & -0.248 & 0.014 & 0.056 \\ -235.107 & -331.015 & 7.046 & 0.056 & 15.482 \end{pmatrix},$$

$$V_2 = \begin{pmatrix} 0.018 & -0.030 & 2.047 & -0.171 & -0.001 & -0.157 \\ -0.030 & 27345.946 & 17984.451 & -702.497 & -6.601 & -774.701 \\ 2.047 & 17984.451 & 19075.122 & -581.837 & -3.224 & -806.763 \\ -0.171 & -702.497 & -581.837 & 56.215 & -0.181 & 33.720 \\ -0.001 & -6.601 & -3.224 & -0.181 & 0.008 & 0.157 \\ -0.157 & -774.701 & -806.763 & 33.720 & 0.157 & 47.063 \end{pmatrix},$$

$$V_3 = \begin{pmatrix} 164972.554 & 25866.431 & -2719.661 & -34.301 & -690.081 \\ 25866.431 & 57585.694 & -773.508 & 0.928 & -1578.765 \\ -2719.661 & -773.508 & 67.013 & 0.270 & 24.185 \\ -34.301 & 0.928 & 0.270 & 0.014 & -0.044 \\ -690.081 & -1578.765 & 24.185 & -0.044 & 49.685 \end{pmatrix},$$

and

$$V_4 = \begin{pmatrix} 0.051 & -12.851 & 9.006 & -0.0155 & 0.001 & -0.406 \\ -12.851 & 65893.547 & 54686.021 & -1455.030 & -2.991 & -1350.792 \\ 9.006 & 54686.021 & 78705.844 & -1528.451 & 0.799 & -2113.446 \\ -0.016 & -1455.030 & -1528.451 & 47.258 & -0.006 & 49.950 \\ 0.001 & -2.991 & 0.7993 & -0.006 & 0.001 & -0.036 \\ -0.406 & -1350.792 & -2113.446 & 49.950 & -0.036 & 71.793 \end{pmatrix}.$$

We examine summer and winter since these are the only two seasons which contain a Gumbel fit. Starting with summer, $\overline{\mu}_{52} = 1144.9810$, $\beta_1^{(0)} = 728.2586$, and $\beta_1^{(1)} = 1.6860$, so we use the first equation in Theorem 4.2. Then it can be shown that

$$P(Y_{52} > y) \approx 4.8238e^{-0.0014y} - 11.6346e^{-0.0027y}.$$

We can also compute the approximate high percentiles. If $P = P(Y_{52} > y)$ is lower than 0.05, for instance, and y is the corresponding streamflow, then the $100(1-P)$th percentile is

$$y \approx -728.2586 \log\left(0.2073 - \frac{\sqrt{23.2690 - 46.5385P}}{23.2693}\right).$$

Now we turn to winter: $\overline{\mu}_{52} = 1184.5620$, $\beta_3^{(0)} = 513.7875$, and $\beta_3^{(1)} = 17.2798$. Then by the first equation in Theorem 4.2,

$$P(Y_{52} > y) \approx 10.2429e^{-0.0019y} - 52.5613e^{-0.0039y}.$$

Winter's $100(1-P)$th percentile is then

$$y \approx -513.7875 \log\left(0.0974 - \frac{\sqrt{104.9168 - 210.2453P}}{105.1227}\right).$$

Table 4.4 highlights some of the key upper percentiles. Note the missing value in the table; setting $P = .0001$ in winter's percentile equation gave a nonreal result.

Table 4.4: Estimated Percentiles per Season

Season	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
Summer	3309	3824	3988	4497	5005	6187	7975
Winter	2725	3092	3210	3579	3957	4993	–

4.9 Open Questions

There are a couple of questions that are currently left unanswered. The first topic is how to derive the expansions in Section 4.5 for the case where $\beta < 0$. In Section 4.4 we grouped the random variables by positive and negative weights, so we conjecture that the negative expansions to complete Section 4.6 would be grouped in a similar way. The expansion should then be controlled by the two largest positive weights. Therefore the expansions would be extensions of those derived in Section 4.4.

The second topic to investigate is extending the approximations in Section 4.6 to include infinite sums of weighted random variables, such as those summarized in Section 4.5. This generalization would require some additional proofs and assumptions, namely some extensions of the lemmas in Section 4.2. In addition, some sort of summability condition on the $\beta^{(k)}$ would be needed. One common requirement on constants that appears in other disciplines is to assume that $\sum_k \left|\beta^{(k)}\right|^\lambda < \infty$ for some $\lambda > 0$. We conjecture that this stipulation, along with some additional lemmas, would provide the probability approximations for an infinite series with Gumbel convolutions.

Chapter 5 The Convolution of Gumbel Random Variables

5.1 Introduction

In the realm of environmental and nonenvironmental statistics, it is often of interest to study the behavior of the sum of maximum observations. For example, suppose an insurance company wants to find the distribution of the total maximum claim amounts taken over specified blocks of time. If a Gumbel distribution fits the monthly maximum claim amounts, then the company may want to know the distribution of the total claim amounts over a year (or indeed, the average monthly figure). This problem involves deriving the distribution of twelve Gumbel random variables. If the variables all have different location and scale parameters, then we may use Theorem 4.3 to approximate the upper tail probabilities.

However, if the Gumbel random variables are independent and identically distributed, then the weights will all be equal. In which case, the asymptotics need to be studied in a very different manner, and the formula we eventually derive is quite different from the results in Chapter 4. This chapter explores how the approximation changes under such a setting.

As another example relating to the Peachtree Creek data, recall that winter was shown to be stationary in the location and scale parameters. Suppose we refit the data for that season, removing year as a covariate, and then we find the distribution of the total winter maximum streamflow over five years. All weights would be equal, and we would need to use the results from this chapter to answer that question. This topic will be explored in Section 5.8.

Rootzén (1986) carried out a similar study on a more general class of independent random variables. In that paper, he derived the first-order expansion under some assumptions on both the variables and the weights. The Gumbel distribution is a special case of his result.

In this chapter we employ similar techniques that Rootzén used in his proof. We start by finding the two-term expansion for the sum of just two Gumbel random variables using a creative Taylor series approach. Then in Section 5.3 we find the general two-term expansion for the sum of $n \geq 2$ variables which fall into a more general class, then reduce to the Gumbel case. After examining simulation results, we derive further terms in the expansion in Section 5.4. The final result in its entirety is summarized in Section 5.5, with examples, more simulations, and a data analysis in the remainder of the chapter.

5.2 The Convolution of Two Gumbel Random Variables

To begin the procedure, we first derive a two-term expansion for the tail probability of the sum of two Gumbel random variables. We do this by performing an infinite Taylor series expansion and then working out which terms are negligible. While a somewhat lengthy and tedious procedure, our method has the advantage that if one desires more terms in the expansion, one can simply modify the proof to get those extra terms. In the next section, we introduce some theory to get a general two-term expansion.

Let Z_1, Z_2 be independent and identically distributed standard Gumbel random variables with distribution function $F_{Z_i}(x) = \exp\{-e^{-x}\}$ and therefore density function $f_{Z_i}(x) = e^{-x} \exp -e^{-x}$, $-\infty < x < \infty$. The formula for the upper tail probability of such a convolution is

$$P(Z_1 + Z_2 > y) = 1 - \int_{-\infty}^{\infty} f_{Z_1}(x) F_{Z_2}(y - x) dx. \tag{5.1}$$

Adapting (5.1) to the Gumbel case gives

$$\begin{aligned}P(Z_1 + Z_2 > y) &= 1 - \int_{-\infty}^{\infty} e^{-x} \exp\left\{-[e^{-x} + e^{-y}e^{x}]\right\} dx \\ &= 1 - \int_{0}^{\infty} e^{-\left[x + e^{-y}x^{-1}\right]} dx = \int_{0}^{\infty} e^{-x}\left[1 - e^{-e^{-y}x^{-1}}\right] dx.\end{aligned}$$

Define $t = e^{-y}$. Then it is of interest to examine the tail probability as $t \downarrow 0$, which is equivalent to letting $y \to \infty$. Then the tail probability formula is

$$P(Z_1 + Z_2 > y) = \int_{0}^{\infty} e^{-x}\left[1 - e^{-zx^{-1}}\right] dx. \tag{5.2}$$

During the proof of the expansion, we need to compute the integral $\int_{a}^{\infty} x^{-k}e^{-x}dx$ for $k \geq 1$. Integration by parts tells us that when $k = 1$ and $a > 0$

$$\int_{a}^{\infty} x^{-1}e^{-x}dx = -\log(a)e^{-a} + \int_{a}^{\infty} \log(x)e^{-x}dx. \tag{5.3}$$

Lemma 5.1 establishes the recursive answer for $k \geq 2$.

Lemma 5.1. *Define* $\Pi_k(j) = \prod_{L=1}^{j}\left(\frac{1}{k-L}\right)$ *for* $j = 1, ..., k-1$. *In particular,* $\Pi_k(k-2) = \Pi_k(k-1) = \frac{1}{(k-1)!}$. *Then for* $a > 0$ *and* $k \geq 2$,

$$\begin{aligned}\int_{a}^{\infty} x^{-k}e^{-x}dx = &\left[\sum_{j=1}^{k-1}(-1)^{j-1}a^{j-k}\Pi_k(j)\right] e^{-a} + \frac{(-1)^k \log(a)e^{-a}}{(k-1)!} \\ &+ \frac{(-1)^{k-1}}{(k-1)!}\int_{a}^{\infty} \log(x)e^{-x}dx.\end{aligned} \tag{5.4}$$

Proof. We proceed by induction on k. When $k=2$, integration by parts provides

$$\int_a^\infty x^{-2}e^{-x}dx = a^{-1}e^{-a} - \int_a^\infty x^{-1}e^{-x}dx = a^{-1}e^{-a} + \log(a)e^{-a} - \int_a^\infty \log(x)e^{-x}dx.$$

Then (5.4) reduces to

$$\int_a^\infty x^{-2}e^{-x}dx = a^{-1}e^{-a} + \log(a)e^{-a} - \int_a^\infty \log(x)e^{-x}dx,$$

because $\Pi_2(1) = 1$. Now assume that (5.4) is true; we shall do integration by parts on

$$\int_a^\infty x^{-(k+1)}e^{-x}dx.$$

Setting $u = e^{-x}$ and $dv = x^{-(k+1)}dx$, we obtain $du = -e^{-x}dx$ and $v = -\frac{1}{k}x^{-k}$. The integral becomes

$$\begin{aligned}
\int_a^\infty x^{-(k+1)}e^{-x}dx &= \frac{1}{k}a^{-k}e^{-a} - \frac{1}{k}\int_a^\infty x^{-k}e^{-x}dx \\
&= \frac{1}{k}a^{-k}e^{-a} - \frac{1}{k}\left[\sum_{j=1}^{k-1}(-1)^{j-1}a^{j-k}\Pi_k(j)\right]e^{-a} + \frac{(-1)^{k+1}}{k!}\log(a)e^{-a} \\
&\qquad + \frac{(-1)^k}{k!}\int_a^\infty \log(x)e^{-x}dx \\
&= \left[\sum_{j=1}^{(k+1)-1}(-1)^{j-1}a^{j-(k+1)}\Pi_{k+1}(j)\right]e^{-a} + \frac{(-1)^{k+1}}{((k+1)-1)!}\log(a)e^{-a} \\
&\qquad + \frac{(-1)^{(k+1)-1}}{((k+1)-1)!}\int_a^\infty \log(x)e^{-x}dx,
\end{aligned}$$

as required. □

Now recall Euler's constant γ from (1.7), which to five decimals is $\gamma \approx 0.57721$. We introduce two integrals which come from Choi and Seo (1998) and will be used in the next proof:

$$\gamma = -\int_0^\infty \log(x)e^{-x}dx \tag{5.5}$$

and

$$\gamma = \int_0^1 x^{-1}[1-e^{-x}]dx - \int_1^\infty x^{-1}e^{-x}dx,$$

the latter of which can be rewritten via integration by parts as the more useful

$$\gamma = \int_0^1 x^{-1}[1-e^{-x}]dx - \int_1^\infty \log(x)e^{-x}dx. \tag{5.6}$$

We now state and prove the following theorem.

Theorem 5.1. *Let Z_1, Z_2 be independent and identically distributed Gumbel random variables. As $y \to \infty$, a two-term expansion for the tail probability of $Z_1 + Z_2$ is*

$$P(Z_1 + Z_2 > y) = (y + 1 - 2\gamma)\, e^{-y} + o(e^{-y}).$$

Proof. Set $t = e^{-y}$ and split (5.2) into two integrals:

$$P(Z_1 + Z_2 > y) = \int_t^\infty e^{-x}\left[1 - e^{-tx^{-1}}\right]dx + \int_0^t e^{-x}\left[1 - e^{-tx^{-1}}\right]dx = J_1 + J_2. \tag{5.7}$$

Consider J_1 first. An infinite Taylor expansion gives

$$J_1 = \int_t^\infty \sum_{i=1}^\infty \frac{(-1)^{i-1}t^i}{i!}x^{-i}e^{-x}dx.$$

Observe that

$$\left|\frac{(-1)^{i-1}t^i}{i!}x^{-i}e^{-x}\right| \leq \frac{1}{i!}e^{-x},$$

because $0 < x^{-i} \leq t^{-i}$ on $x \in (t, \infty)$. Therefore the integrand is integrable on the counting and Lebesgue measures, and so by Fubini's Theorem,

$$J_1 = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}t^i}{i!} \int_t^{\infty} x^{-i}e^{-x}dx.$$

Split into the $\{i = 1\}$ and $\{i \geq 2\}$ cases via

$$J_1 = t\int_t^{\infty} x^{-1}e^{-x}dx + \sum_{i=2}^{\infty} \frac{(-1)^{i-1}t^i}{i!} \int_t^{\infty} x^{-i}e^{-x}dx,$$

and simplifying using Lemma 5.1,

$$\begin{aligned} J_1 &= -t\log(t)e^{-t} + t\int_t^{\infty} \log(x)e^{-x}dx + \sum_{i=2}^{\infty} \frac{(-1)^{i-1}t^i}{i!} \left\{\sum_{j=1}^{i-1}(-1)^{j-1}t^{j-i}\Pi_i(j)e^{-t}\right. \\ &\qquad \left. + \frac{(-1)^i}{(i-1)!}\log(t)e^{-t} + \frac{(-1)^{i-1}}{(i-1)!}\int_t^{\infty}\log(x)e^{-x}dx\right\} \\ &= -t\log(t)e^{-t} + t\int_t^{\infty}\log(x)e^{-x}dx + \left[\sum_{i=2}^{\infty}\sum_{j=1}^{i-1}\frac{(-1)^{i+j}t^j\Pi_i(j)}{i!}\right]e^{-t} \\ &\qquad - \sum_{i=2}^{\infty}\frac{t^i\log(t)e^{-t}}{i!(i-1)!} + \sum_{i=2}^{\infty}\frac{t^i}{i!(i-1)!}\int_t^{\infty}\log(x)e^{-x}dx \\ &= (I) + (II) + (III) + (IV) + (V). \end{aligned} \tag{5.8}$$

We now examine each of the five pieces in (5.8). To start, $(I) = -t\log(t) + o(t)$ as $t \downarrow 0$, and by (5.5),

$$(II) = -\gamma t - t\int_0^t \log(x)e^{-x}dx.$$

Observe that

$$\left| -t \int_0^t \log(x) e^{-x} dx \right| \leq t \int_0^t \log(x) dx = t^2 \log(t) - t^2 = o(t),$$

and therefore $(II) = -\gamma t + o(t)$ as $t \downarrow 0$. Next, (III) provides terms involving $t^i, i = 1, 2, ...$ We pick off those involving just t and show that what remains is $o(t)$. To do this, it is convenient to change the order of summation. Observe that

$$\left| \frac{(-1)^{i+j} t^j \Pi_i(j)}{i!} \right| \leq \frac{1}{i!} t^j,$$

which is integrable on the two counting measures because $t \downarrow 0$. Thus

$$\sum_{i=2}^{\infty} \sum_{j=1}^{i-1} \frac{(-1)^{i+j} t^j \Pi_i(j)}{i!} = \sum_{j=1}^{\infty} \sum_{i=j+1}^{\infty} \frac{(-1)^{i+j} \Pi_i(j)}{i!} t^j = \Lambda_1 t + \sum_{j=2}^{\infty} \sum_{i=j+1}^{\infty} \frac{(-1)^{i+j} \Pi_i(j)}{i!} t^j,$$

where $\Lambda_1 = \sum_{i=2}^{\infty} \frac{(-1)^{i-1}}{i!(i-1)}$. Λ_1 converges and, to five decimals, is equal to -0.42872. As for the remainder,

$$\left| \sum_{j=2}^{\infty} \sum_{i=j+1}^{\infty} \frac{(-1)^{i+j} \Pi_i(j)}{i!} t^j \right| \leq \left[\sum_{i=2}^{\infty} \frac{1}{i!} \right] t^2 \leq t^2.$$

Thus the remainder converges and is $o(t)$, and therefore

$$(III) = \left[\sum_{i=2}^{\infty} \sum_{j=1}^{i-1} \frac{(-1)^{i+j} t^j \Pi_i(j)}{i!} \right] e^{-t} = \Lambda_1 t e^{-t} + o(te^{-t}) = \Lambda_1 t + o(t). \tag{5.9}$$

Next, (IV) is

$$\left| - \sum_{i=2}^{\infty} \frac{t^i \log(t) e^{-t}}{i!(i-1)!} \right| \leq \left[\sum_{i=2}^{\infty} \frac{1}{i!} \right] t^2 \log(t) e^{-t} \leq t^2 \log(t) e^{-t},$$

and therefore $(IV) = o(t)$. Examining (V),

$$\left| -\sum_{i=2}^{\infty} \frac{t^i}{i!(i-1)!} \int_t^{\infty} \log(x) e^{-x} dx \right| \leq \gamma \left[\sum_{i=2}^{\infty} \frac{1}{i!} \right] t^2 \leq \gamma t^2,$$

so $(V) = o(t)$ as well. Putting all five terms in (5.8) together, as $t \downarrow 0$

$$J_1 = -t \log(t) + (\Lambda_1 - \gamma) t + o(t). \tag{5.10}$$

The next step is to examine J_2 in (5.7). We employ another Taylor series expansion:

$$J_2 = \int_0^t e^{-x} \left[1 - e^{-tx^{-1}} \right] dx = \int_0^t \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} x^i \left[1 - e^{-tx^{-1}} \right] dx.$$

Observe that $1 - e^{-1} \leq 1 - e^{-tx^{-1}} \leq 1$ on $x \in (0, t)$, and therefore

$$\left| \frac{(-1)^i}{i!} x^i \left[1 - e^{-tx^{-1}} \right] \right| \leq \frac{1}{i!} x^i,$$

which is integrable on both measures. Therefore we use Fubini's Theorem again, along with a change of variables:

$$\begin{aligned} J_2 &= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \int_0^t x^i \left[1 - e^{-tx^{-1}} \right] dx = \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \int_1^{\infty} (tx^{-1})^i (tx^{-2}) \left[1 - e^{-x} \right] dx \\ &= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} t^{i+1} \int_1^{\infty} x^{-(i+2)} \left[1 - e^{-x} \right] dx \\ &= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} t^{i+1} \left\{ \int_1^{\infty} x^{-(i+2)} dx - \int_1^{\infty} x^{-(i+2)} e^{-x} dx \right\} \end{aligned}$$

$$= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} t^{i+1} \left\{ \frac{1}{i+1} - \int_1^{\infty} x^{-(i+2)} e^{-x} dx \right\}$$

$$= \left[1 - \int_1^{\infty} x^{-2} e^{-x} dx \right] t + \sum_{i=1}^{\infty} \frac{(-1)^i}{i!} t^{i+1} \left\{ \frac{1}{i+1} - \int_1^{\infty} x^{-(i+2)} e^{-x} dx \right\}.$$

Now we use Lemma 5.1 with $k = i + 2$ and $a = 1$. The last integral above becomes

$$\int_1^{\infty} x^{-(i+2)} e^{-x} dx = e^{-1} \sum_{j=1}^{i+1} (-1)^{j-1} \Pi_{i+2}(j) + \frac{(-1)^{i+1}}{(i+1)!} \int_1^{\infty} \log(x) e^{-x} dx.$$

Put an upper bound on $R = \sum_{i=1}^{\infty} \frac{(-1)^i}{i!} t^{i+1} \left\{ \frac{1}{i+1} - \int_1^{\infty} x^{-(i+2)} e^{-x} dx \right\}$:

$$\begin{aligned}
|R| &\leq t^2 \sum_{i=1}^{\infty} \frac{1}{i!} \left[\frac{1}{i+1} + e^{-1} \sum_{j=1}^{i+1} \Pi_{i+2}(j) + \frac{1}{(i+1)!} \int_1^{\infty} \log(x) e^{-x} dx \right] \\
&\leq t^2 \sum_{i=1}^{\infty} \frac{1}{i!} \left[\frac{1}{i+1} + e^{-1} \sum_{j=1}^{i+1} \frac{1}{i+1} + \frac{1}{(i+1)!} \int_1^{\infty} \log(x) e^{-x} dx \right] \\
&\leq t^2 \sum_{i=1}^{\infty} \frac{1}{i!} \left[\frac{1}{i+1} + e^{-1} + \frac{1}{(i+1)!} \int_1^{\infty} \log(x) e^{-x} dx \right] \\
&= t^2 \left\{ \sum_{i=2}^{\infty} \frac{1}{i!} + e^{-1} \sum_{i=1}^{\infty} \frac{1}{i!} + \sum_{i=1}^{\infty} \frac{1}{i!(i+1)!} \int_1^{\infty} \log(x) e^{-x} dx \right\} \\
&\leq 2t^2 \left\{ 1 + \int_1^{\infty} \log(x) e^{-x} dx \right\} \leq 2t^2.
\end{aligned}$$

Therefore $R = o(t)$, and so

$$J_2 = \left[1 - \int_1^{\infty} x^{-2} e^{-x} \right] t + o(t) = \left[1 - e^{-1} + \int_1^{\infty} \log(x) e^{-x} dx \right] t + o(t). \quad (5.11)$$

Putting (5.10) and (5.11) into (5.7), as $t \downarrow 0$

$$P(Z_1 + Z_2 > y) = -t\log(t) + \left[\Lambda_1 - \gamma + 1 - e^{-1} + \int_1^\infty \log(x)e^{-x}dx\right] t + o(t).$$

It remains to show that $\Lambda_1 - \gamma + 1 - e^{-1} + \int_1^\infty \log(x)e^{-x}dx = 1 - 2\gamma$. The first step is to expand the following integral as a Taylor series:

$$\begin{aligned}\int_0^1 x^{-1}\left[1 - e^{-x}\right]dx &= \int_0^1 x^{-1}\sum_{i=1}^\infty \frac{(-1)^{i-1}x^i}{i!}dx = \sum_{i=1}^\infty \frac{(-1)^{i-1}}{i!}\int_0^1 x^{i-1}dx \\ &= -\sum_{i=1}^\infty \frac{(-1)^i}{i!i}.\end{aligned}$$

Again, the interchanging of integration and summation is possible because

$$\left|\frac{(-1)^{i-1}}{i!}x^{i-1}\right| \leq \frac{x^{i-1}}{i!},$$

which is integrable on the counting measure and on $x \in (0,1)$. Next, we write $\Lambda_1 - e^{-1}$ as a Taylor series by first expanding e^{-1}:

$$\begin{aligned}\Lambda_1 - e^{-1} &= \sum_{i=2}^\infty \frac{(-1)^{i-1}}{i!(i-1)} - \sum_{i=0}^\infty \frac{(-1)^i}{i!} = \sum_{i=2}^\infty \frac{(-1)^{i-1}}{i!(i-1)} + \sum_{i=2}^\infty \frac{(-1)^{i-1}}{i!} \\ &= \sum_{i=2}^\infty \frac{(-1)^{i-1}}{i!}\left[\frac{1}{i-1} + 1\right] = \sum_{i=2}^\infty \frac{(-1)^{i-1}}{(i-1)!(i-1)} = \sum_{i=1}^\infty \frac{(-1)^i}{i!i} \\ &= -\int_0^1 x^{-1}\left[1 - e^{-x}\right]dx.\end{aligned}$$

Now (5.6) provides

$$-\gamma = -\int_0^1 x^{-1}\left[1 - e^{-x}\right]dx + \int_1^\infty \log(x)e^{-x}dx = \Lambda_1 - e^{-1} + \int_1^\infty \log(x)e^{-x}dx,$$

and finally

$$\Lambda_1 - \gamma + 1 - e^{-1} + \int_1^\infty \log(x)e^{-x}dx = 1 - 2\gamma.$$

Converting $t = e^{-y}$, the two-term expansion as $y \to \infty$ is

$$P(Z_1 + Z_2 > y) = (y + 1 - 2\gamma)e^{-y} + o\left(e^{-y}\right).$$

□

Corollary 5.1. *Let Z_1, Z_2 be independent and identically distributed Gumbel random variables, and let $c > 0$. As $y \to \infty$, a two-term expansion for the tail probability of $c(Z_1 + Z_2)$ is*

$$P(c(Z_1 + Z_2) > y) = \left(\frac{y}{c} + 1 - 2\gamma\right) e^{-y/c} + o(e^{-y/c}).$$

5.3 The Two-Term N-Fold Convolution

We were able to derive a two-term expansion for the sum of two Gumbel random variables, albeit in a messy way. It should come as no surprise that for more than two variables, derivations with Taylor series would be too involved and complicated. It would be much more sensible to show the result holds for a more general class of distributions, then use that result to specialize to the Gumbel. The purpose of including the preceding proof was to show an interesting alternative method of checking the distribution for two Gumbels. We now turn to the general setting.

Let F denote a distribution function with support in $[0, \infty)$. It is supposed that $F(0) = 0$ and that for some $\alpha > 1$,

$$1 - F(y) = e^{-y} + o\left(e^{-\alpha y}\right) \qquad \text{as} \qquad y \to \infty. \tag{5.12}$$

Define

$$\theta = -\int_0^\infty e^y d(\overline{F}(y) - e^{-y}) = \int_0^\infty e^y d(F(y) + e^{-y}) = \lim_{y\to\infty}\left[\int_0^y e^x dF(x) - y\right], \qquad (5.13)$$

and so we see that θ provides a measure of how much departure there is from F and the standard exponential distribution. Now define

$$A_n(y) = \left(\frac{1}{(n-1)!}y^{n-1} + \frac{1+n\theta}{(n-2)!}y^{n-2}\right)e^{-y}, \qquad n \geq 2. \qquad (5.14)$$

The goal is to show that $A_n(y)$ serves as a two-term expansion for $\overline{G}_n(y) = 1 - F^{*n}(y)$ as $y \to \infty$. The first task is to demonstrate that A_n approximately satisfies the same convolution equation as $\overline{G}_n$.

Lemma 5.2. *For $n \geq 2$, as $z \to \infty$*

$$\int_0^{z/2} A_n(z-y)dF(y) - \int_0^{z/2}(1 - F(z-y))dA_n(y) + A_n\left(\frac{z}{2}\right)\overline{F}\left(\frac{z}{2}\right)$$
$$= A_{n+1}(z) + O\left(z^{n-2}e^{-z}\right).$$

Proof. For $n \geq 2$, define $k_1^{(n)} = \frac{1}{(n-1)!}$ and $k_2^{(n)} = \frac{1+n\theta}{(n-2)!}$ and write

$$A_n(y) = \left(k_1^{(n)}y^{n-1} + k_2^{(n)}y^{n-2}\right)e^{-y}.$$

We first notice that for any $k \geq 0$, an integration by parts establishes that

$$\int_0^{z/2}(z-y)^k e^y dF(y) = \int_0^{z/2}\overline{F}(y)\left(-k(z-y)^{k-1} + (z-y)^k\right)e^y dy - \left(\frac{z}{2}\right)^k \overline{F}\left(\frac{z}{2}\right)e^{z/2} + z^k, \qquad (5.15)$$

with $-k(z-y)^{k-1}+(z-y)^k=1$ for $k=0$. Now for $k\geq 0$

$$\int_0^{z/2}(z-y)^k\left(e^y\overline{F}(y)-1\right)dy=z^{k+1}\int_0^{1/2}(1-y)^k\left(e^{zy}\overline{F}(zy)-1\right)dy.$$

Next note that for any $\epsilon>0$ and z large enough

$$\int_{1/\sqrt{z}}^{1/2}(1-y)^k\left|e^{zy}\overline{F}(zy)-1\right|dy\leq\epsilon\int_{1/\sqrt{z}}^{1/2}e^{-(\alpha-1)zy}dy\leq\frac{\epsilon}{(\alpha-1)z}e^{-(\alpha-1)\sqrt{z}},$$

and therefore

$$\int_{1/\sqrt{z}}^{1/2}(1-y)^k\left|e^{zy}\overline{F}(zy)-1\right|dy=o\left(\frac{1}{z}e^{-(\alpha-1)\sqrt{z}}\right). \tag{5.16}$$

Note that

$$\begin{aligned}z\int_0^{1/\sqrt{z}}(1-y)^k\left(e^{zy}\overline{F}(zy)-1\right)dy&=\int_0^{\sqrt{z}}\left(1-\frac{y}{z}\right)^k\left(e^y\overline{F}(y)-1\right)dy\\&=\int_0^{\sqrt{z}}\left(e^y\overline{F}(y)-1\right)dy+\int_0^{\sqrt{z}}\left(\left(1-\frac{y}{z}\right)^k-1\right)\left(e^y\overline{F}(y)-1\right)dy.\end{aligned} \tag{5.17}$$

Thus we have as $z\to\infty$

$$\int_0^{\sqrt{z}}\left|\left(1-\frac{y}{z}\right)^k-1\right|\left|e^y\overline{F}(y)-1\right|dy\leq\frac{k}{z}\int_0^{\infty}y\left|e^y\overline{F}(y)-1\right|dy=O\left(\frac{1}{z}\right), \tag{5.18}$$

and further for any $\epsilon>0$ and z large enough

$$\int_{\sqrt{z}}^{\infty}\left|e^y\overline{F}(y)-1\right|dy\leq\epsilon\int_{\sqrt{z}}^{\infty}e^{-(\alpha-1)y}dy=\frac{\epsilon}{\alpha-1}e^{-(\alpha-1)\sqrt{z}}$$

so that

$$\int_{\sqrt{z}}^{\infty}\left|e^y\overline{F}(y)-1\right|dy=o\left(e^{-(\alpha-1)\sqrt{z}}\right). \tag{5.19}$$

Noting that

$$\int_0^\infty \left(e^y\overline{F}(y)-1\right)dy=\int_0^\infty \left(\overline{F}(y)-e^{-y}\right)d(e^y)=-\int_0^\infty e^y d\left(\overline{F}(y)-e^{-y}\right)=\theta,$$

we have from (5.17) to (5.19)

$$z\int_0^{1/\sqrt{z}}(1-y)^k\left(e^{zy}\overline{F}(yz)-1\right)dy=\theta+O\left(\frac{1}{z}\right). \tag{5.20}$$

Hence from (5.16) and (5.20) we have as $z\to\infty$

$$\int_0^{z/2}(z-y)^k\left(e^y\overline{F}(y)-1\right)dy=\theta z^k+O\left(z^{k-1}\right). \tag{5.21}$$

Therefore from (5.15) and (5.21) we obtain

$$\begin{aligned}
\int_0^{z/2}(z-y)^k e^y dF(y) &= -\left(\frac{z}{2}\right)^k\overline{F}\left(\frac{z}{2}\right)e^{z/2}+z^k\\
&\quad+\int_0^{z/2}\left(-k(z-y)^{k-1}+(z-y)^k\right)dy\\
&\quad+\int_0^{z/2}\left(e^y\overline{F}(y)-1\right)\left(-k(z-y)^{k-1}+(z-y)^k\right)dy\\
&=-\left(\frac{z}{2}\right)^k\overline{F}\left(\frac{z}{2}\right)e^{z/2}+z^k+\theta z^k-\left(z^k-\left(\frac{z}{2}\right)^k\right)\\
&\quad+\frac{1}{k+1}\left(z^{k+1}-\left(\frac{z}{2}\right)^{k+1}\right)+O\left(z^{k-1}\right)\\
&=\frac{1}{k+1}\left(z^{k+1}-\left(\frac{z}{2}\right)^{k+1}\right)+\theta z^k+O\left(z^{k-1}\right).
\end{aligned} \tag{5.22}$$

Now we compute

$$-\int_0^{z/2}\overline{F}(z-y)dA_n(y)=-\int_0^{z/2}\left(\overline{F}(z-y)-e^{-(z-y)}\right)dA_n(y)-\int_0^{z/2}e^{-(z-y)}dA_n(y).$$

We obtain

$$-\int_0^{z/2} e^y dA_n(y) = -e^{z/2} A_n\left(\frac{z}{2}\right) + A_n(0) + \int_0^{z/2} \left(k_1^{(n)} y^{n-1} + k_2^{(n)} y^{n-2}\right) dy$$
$$= -k_1^{(n)} \left(\frac{z}{2}\right)^{n-1} + \frac{k_1^{(n)}}{n}\left(\frac{z}{2}\right)^n + \frac{k_2^{(n)}}{n-1}\left(\frac{z}{2}\right)^{n-1} + O\left(z^{n-2}\right). \tag{5.23}$$

Observe that for some $c > 0$

$$\left|\int_0^{z/2} \left(\overline{F}(z-y) - e^{-(z-y)}\right) dA_n(y)\right| = \left|\int_0^{z/2} \left(\overline{F}(z-y) - e^{-(z-y)}\right)\right.$$
$$\left.\times \left((n-1)k_1^{(n)} y^{n-2} + (n-2)k_2^{(n)} y^{n-3} - k_1^{(n)} y^{n-1} - k_2^{(n)} y^{n-2}\right) e^{-y} dy\right| \tag{5.24}$$
$$\leq c e^{-z} \int_0^{z/2} e^{-(\alpha-1)z/2} y^{n-1} dy = O\left(z^n e^{-(\alpha+1)z/2}\right).$$

Finally we have

$$A_n\left(\frac{z}{2}\right)\overline{F}\left(\frac{z}{2}\right) = \left(k_1^{(n)}\left(\frac{z}{2}\right)^{n-1} + O\left(z^{n-2}\right)\right) e^{-z}. \tag{5.25}$$

Use $k = n-1$ and $n-2$ in (5.22) to obtain

$$\int_0^{z/2} A_n(z-y) dF(y) = \left\{\frac{k_1^{(n)}}{n}\left(z^n - \left(\frac{z}{2}\right)^n\right) + k_1^{(n)} \theta z^{n-1}\right.$$
$$\left. + \frac{k_2^{(n)}}{n-1}\left(z^{n-1} - \left(\frac{z}{2}\right)^{n-1}\right) + O\left(z^{n-2}\right)\right\} e^{-z}. \tag{5.26}$$

From (5.23) and (5.24), we obtain

$$-\int_0^{z/2} \overline{F}(z-y) dA_n(y) = \left(-k_1^{(n)}\left(\frac{z}{2}\right)^{n-1} + \frac{k_1^{(n)}}{n}\left(\frac{z}{2}\right)^n + \frac{k_2^{(n)}}{n-1}\left(\frac{z}{2}\right)^{n-1}\right) e^{-z}$$
$$+ O\left(z^{n-2} e^{-z}\right). \tag{5.27}$$

Summing (5.25) to (5.27),

$$\int_0^{z/2} A_n(z-y)dF(y) - \int_0^{z/2} \overline{F}(z-y)dA_n(y) + A_n\left(\frac{z}{2}\right)\overline{F}\left(\frac{z}{2}\right)$$
$$= \left(\frac{k_1^{(n)}}{n}z^n + \left(k_1^{(n)}\theta + \frac{k_2^{(n)}}{n-1}\right)z^{n-1} + O\left(z^{n-2}\right)\right)e^{-z}.$$

The lemma has now been proven since, going back to our definitions $k_1^{(n)} = \frac{1}{(n-1)!}$ and $k_2^{(n)} = \frac{1+n\theta}{(n-2)!}$, we have

$$\frac{k_1^{(n)}}{n} = \frac{1}{n!} = k_1^{(n+1)}$$

and

$$k_1^{(n)}\theta + \frac{k_2^{(n)}}{n-1} = \frac{(n-1)!k_1^{(n)}\theta + (n-2)!k_2^{(n)}}{(n-1)!} = \frac{\theta+1+n\theta}{(n-1)!} = \frac{1+(n+1)\theta}{(n-1)!} = k_2^{(n+1)}.$$

□

For the next theorem we assume that the random variables are nonnegative, but we will remove this restriction later. The next lemma establishes the expansion for the convolution of two random variables.

Lemma 5.3. *Let F be a distribution function with support in $[0,\infty)$ satisfying condition (5.12), and let θ and A_n be defined as in (5.13) and (5.14) respectively. In particular, $A_2(y) = (y+1+2\theta)\,e^{-y}$. Then for $n=2$, as $y\to\infty$*

$$1 - F^{*2}(y) = (y+1+2\theta)\,e^{-y} + o\left(ye^{-(\alpha+1)y/2}\right).$$

Proof. We have

$$1 - F^{*2}(z) = 2\int_0^{z/2} \overline{F}(z-y)dF(y) + \left[\overline{F}\left(\frac{z}{2}\right)\right]^2.$$

Next, write

$$\int_0^{z/2} \overline{F}(z-y)dF(y) = e^{-z}\int_0^{z/2} \left(e^{(z-y)}\overline{F}(z-y) - 1\right) e^y dF(y) + e^{-z}\int_0^{z/2} e^y dF(y). \quad (5.28)$$

We have

$$\int_0^{z/2} e^y dF(y) = \frac{z}{2} - \int_0^{z/2} e^y d\left(\overline{F}(y) - e^{-y}\right) = \frac{z}{2} + \theta + \int_{z/2}^{\infty} e^y d\left(\overline{F}(y) - e^{-y}\right). \quad (5.29)$$

Next,

$$\begin{aligned}\int_{z/2}^{\infty} e^y d\left(\overline{F}(y) - e^{-y}\right) &= 1 - \overline{F}\left(\frac{z}{2}\right)e^{z/2} - \int_{z/2}^{\infty}\left(\frac{\overline{F}(y) - e^{-y}}{e^{-\alpha y}}\right)e^{-(\alpha-1)y}dy \\ &= o\left(e^{-(\alpha-1)z/2}\right).\end{aligned}$$

Thus we have

$$\int_0^{z/2} e^y dF(y) = \frac{z}{2} + \theta + o\left(e^{-(\alpha-1)z/2}\right). \quad (5.30)$$

Moreover, for any $\epsilon > 0$ and z large enough, using (5.30)

$$\begin{aligned}\int_0^{z/2} \left|e^{(z-y)}\overline{F}(z-y) - 1\right| e^y dF(y) &\le \epsilon e^{-(\alpha-1)z/2}\int_0^{z/2} e^y dF(y) \\ &\le \epsilon e^{-(\alpha-1)z/2}\left(\frac{z}{2} + \theta + 1\right).\end{aligned} \quad (5.31)$$

Thus from (5.31) we find

$$e^{-z}\int_0^{z/2}\left|e^{(z-y)}\overline{F}(z-y)-1\right|e^y dF(y)=o\left(ze^{-(\alpha+1)z/2}\right). \tag{5.32}$$

Therefore from (5.28), (5.30), and (5.32) we obtain

$$\int_0^{z/2}\overline{F}(z-y)dF(y)=\left(\frac{z}{2}+\theta\right)e^{-z}+o\left(ze^{-(\alpha+1)z/2}\right).$$

Lastly, $\left[\overline{F}\left(\frac{z}{2}\right)\right]^2 = e^{-z}+o\left(e^{-(\alpha+1)z/2}\right)$, and therefore $1-F^{*2}(z) = (z+1+2\theta)e^{-z}+o\left(ze^{-(\alpha+1)z/2}\right)$. □

We now use induction to derive the expansion for the convolution of n variables.

Theorem 5.2. *Let F be a distribution function with support in $[0,\infty)$ satisfying condition (5.12), and let θ and A_n be defined as in (5.13) and (5.14) respectively. Then for $n\geq 2$, as $y\to\infty$*

$$\overline{G}_n(y)=1-F^{*n}(y)=A_n(y)+O\left(y^{n-3}e^{-y}\right).$$

Proof. Lemma 5.3 takes care of the case $n=2$. Next, suppose that as $y\to\infty$, $\overline{G}_n(y)=A_n(y)+O\left(y^{n-3}e^{-y}\right)$, and we shall show that the case $n+1$ holds. By Lemma 5.2,

$$\begin{aligned}\overline{G}_{n+1}(z)-A_{n+1}(z)&=\int_0^{z/2}\overline{G}_n(z-y)dF(y)+\int_0^{z/2}\overline{F}(z-y)dG_n(y)\\&\quad-\left\{\int_0^{z/2}A_n(z-y)dF(y)-\int_0^{z/2}\overline{F}(z-y)dA_n(y)+A_n\left(\frac{z}{2}\right)\overline{F}\left(\frac{z}{2}\right)\right\}\\&\quad+\overline{G}_n\left(\frac{z}{2}\right)\overline{F}\left(\frac{z}{2}\right)+O\left(z^{n-2}e^{-z}\right)\\&=\int_0^{z/2}\left[\overline{G}_n(z-y)-A_n(z-y)\right]dF(y)+\int_0^{z/2}\overline{F}(z-y)d\left(G_n(y)+A_n(y)\right)\\&\quad+\left(\overline{G}_n\left(\frac{z}{2}\right)-A_n\left(\frac{z}{2}\right)\right)\overline{F}\left(\frac{z}{2}\right)+O\left(z^{n-2}e^{-z}\right).\end{aligned} \tag{5.33}$$

Now we use the induction hypothesis to find that for some $c > 0$

$$\begin{aligned}\left|\int_0^{z/2} \left[\overline{G}_n(z-y) - A_n(z-y)\right] dF(y)\right| &\le c\int_0^{z/2} (z-y)^{n-3}e^{-(z-y)}dF(y) \\ &\le ce^{-z}z^{n-3}\int_0^{z/2} e^y dF(y) \le cz^{n-3}e^{-z}\left(\frac{z}{2}+\theta+1\right) = O\left(z^{n-2}e^{-z}\right),\end{aligned} \tag{5.34}$$

where we have used (5.30) to bound the last integral. Next, through integration by parts

$$\begin{aligned}\int_0^{z/2} e^y d\left(G_n(y) + A_n(y)\right) = -e^{z/2}\left(\overline{G}_n\left(\frac{z}{2}\right) - A_n\left(\frac{z}{2}\right)\right) + 1 - A_n(0) \\ + \int_0^{z/2}\left(\overline{G}_n(y) - A_n(y)\right)e^y dy.\end{aligned}$$

By the induction hypothesis, we find that for $n \ge 2$ the above expression can be written as

$$\int_0^{z/2} e^y d\left(G_n(y) + A_n(y)\right) = O\left(z^{n-2}\right). \tag{5.35}$$

Next, we write

$$\begin{aligned}\int_0^{z/2} \overline{F}(z-y) d\left(G_n(y) + A_n(y)\right) = e^{-z}\int_0^{z/2}\left[e^{z-y}\overline{F}(z-y) - 1\right]e^y d\left(G_n(y) + A_n(y)\right) \\ + e^{-z}\int_0^{z/2} e^y d\left(G_n(y) + A_n(y)\right).\end{aligned}$$

Note that

$$\begin{aligned}e^{-z}\left|\int_0^{z/2}\left(e^{z-y}\overline{F}(z-y) - 1\right)e^y d\left(G_n(y) + A_n(y)\right)\right| \\ \le e^{-(\alpha+1)z/2}\left(\int_0^{z/2} e^y dG_n(y) - \int_0^{z/2} e^y dA_n(y)\right).\end{aligned}$$

Further, note that if $X_i, 1 \le i \le n$, denotes a sample from distribution function F, then for sufficiently large z we have the upper bound

$$\int_0^{z/2} e^y dG_n(y) \le E\left(\exp\left(\sum_{i=1}^n X_i\right) I\left\{\bigvee X_i \le z/2\right\}\right) = \left(\int_0^{z/2} e^y dF(y)\right)^n$$
$$= \left(1 - e^{z/2}\overline{F}(z/2) + \int_0^{z/2} e^z \overline{F}(x)dx\right)^n \le z^n.$$

Similarly, one establishes that $-\int_0^{z/2} e^y dA_n(y) = O(z^n)$. Therefore we obtain

$$\int_0^{z/2} \overline{F}(z-y)d\left(G_n(y) + A_n(y)\right) = O\left(z^{n-2}e^{-z}\right). \tag{5.36}$$

Finally

$$\left|\overline{G}_n\left(\frac{z}{2}\right) - A_n\left(\frac{z}{2}\right)\right| \overline{F}\left(\frac{z}{2}\right) = O\left(z^{n-3}e^{-z}\right). \tag{5.37}$$

From (5.33), (5.34), (5.36), and (5.37) we have

$$\overline{G}_{n+1}(z) - A_{n+1}(z) = O\left(z^{n-2}e^{-z}\right).$$

Hence the induction step, along with Lemmas 5.2 and 5.3, establishes the theorem. □

To generalize this result, we now relax the restriction that F has support in $[0, \infty)$. In order to allow for mass on the negative half line, we define the conditional distributions F_+ and F_- by

$$\overline{F}_+(y) = \frac{\overline{F}(y)}{\overline{F}(0)} \quad \text{for} \quad y > 0$$
$$\text{and} \quad F_-(y) = \frac{F(y)}{F(0)} \quad \text{for} \quad y \le 0.$$

Note that if $X \sim F$, then

$$(X|X > 0) \sim F_+ \quad \text{and} \quad (X|X \le 0) \sim F_-.$$

Let $X_1, ..., X_n$ be iid from F. Denote a random subset of $\{1, ..., n\}$ by $I = \{i : 1 \le i \le n, X_i \le 0\}$. Let $(X_i^-, i \ge 1)$ and $(X_i^+, i \ge 1)$ be two independent sequences of iid random variables with $X_i^- \sim F_-$ and $X_i^+ \sim F_+$. Observe that

$$((X_i, i \in I), (X_i, i \notin I)|I) \stackrel{d}{=} \left((X_i^-, 1 \le i \le |I|), (X_i^+, 1 \le i \le n - |I|)\right).$$

That is, conditional on the set I, the nonpositive random variables $(X_i, i \in I)$ and the positive random variables $(X_i, i \notin I)$ are independent groups of variables with common distribution F_- within the nonpositive group and F_+ within the positive group. Therefore conditionally on I,

$$P\left(\sum_{i=1}^{n} X_i \le y \middle| I\right) = P\left(\sum_{i=1}^{|I|} X_i^- + \sum_{i=1}^{n-|I|} X_i^+ \le y\right).$$

Let $H_m(y) = F_-^{*m}(y)$ and $G_l(y) = F_+^{*l}(y)$. Then

$$P\left(\sum_{i=1}^{n} X_i \le y \middle| I\right) = H_{|I|} * G_{n-|I|}(y).$$

Taking expectations over I, we obtain $F^{*n}(y) = EH_{|I|} * G_{n-|I|}(y)$, and therefore

$$1 - F^{*n}(y) = \sum_{k=0}^{n} \binom{n}{k} (F(0))^k \left(\overline{F}(0)\right)^{n-k} \overline{H_k * G_{n-k}}(y). \tag{5.38}$$

Now we will be able to use the results on distribution functions with support on the positive half line to achieve the desired extension to distributions with support over all reals. In addition to condition (5.12), we make the assumption that for some $\beta > 1$

$$F(y) = o\left(e^{\beta y}\right) \quad \text{as} \quad y \to -\infty. \tag{5.39}$$

Theorem 5.3. *Suppose F is a distribution function such that $\overline{F}(y) = e^{-y} + o\left(e^{-\alpha y}\right)$ as $y \to \infty$ for some $\alpha > 1$ and $F(y) = o\left(e^{\beta y}\right)$ as $y \to -\infty$ for some $\beta > 1$. Then for y large enough,*

$$1 - F^{*n}(y) = \left(\frac{1}{(n-1)!}y^{n-1} + \frac{1+n\theta}{(n-2)!}y^{n-2} + O(y^{n-3})\right) e^{-y}, \tag{5.40}$$

where

$$\theta = -\int_0^\infty e^y d\left(\overline{F}(y) - e^{-y}\right) + \int_{-\infty}^0 e^y dF(y). \tag{5.41}$$

Proof. We first need to make an observation. Let K be a distribution function with support over the positive half line such that for some $c > 0$

$$c\overline{K}(y) = e^{-y} + o\left(e^{-\alpha y}\right) \quad \text{as} \quad y \to \infty,$$

where $\alpha > 1$. Then an inspection of the proof when $c = 1$ shows that

$$c^n \overline{K}^{*n}(y) = \left(\frac{1}{(n-1)!}y^{n-1} + \frac{1+n\theta}{(n-2)!}y^{n-2} + O\left(y^{n-3}\right)\right) e^{-y}, \tag{5.42}$$

where

$$\theta = -\int_0^\infty e^y d\left(c\overline{K}(y) - e^{-y}\right) = \lim_{z\to\infty}\left(c\int_0^z e^y dK(y) - z\right). \tag{5.43}$$

Now consider the contribution from the $k = 0$ term in the sum in (5.38) given by

$$\left(\overline{F}(0)\right)^n \overline{G_n}(y) = \left(\frac{1}{(n-1)!}y^{n-1} + \frac{1+n\theta_+}{(n-2)!}y^{n-2} + O\left(y^{n-3}\right)\right)e^{-y},$$

where

$$\theta_+ = -\int_0^\infty e^y d\left(\overline{F}(y) - e^{-y}\right)$$

and where we have used (5.42) and (5.43) with $c = \overline{F}(0)$ and $K = F_+$. Similarly, we find that the contribution from the $k = 1$ term in (5.38) is given by

$$\begin{aligned} nF(0)\left(\overline{F}(0)\right)^{n-1} &\int_{-\infty}^0 \overline{G_{n-1}}(y-u)dF_-(u) \\ &= n\int_{-\infty}^0 \frac{1}{(n-2)!}(y-u)^{n-2}e^{-(y-u)}dF(u) + O\left(y^{n-3}e^{-y}\right) \\ &= \frac{n}{(n-2)!}y^{n-2}e^{-y}\int_{-\infty}^0 e^u dF(u) + O\left(y^{n-3}e^{-y}\right). \end{aligned} \tag{5.44}$$

Finally, if $2 \leq k \leq n$ then note that for $y > 0$

$$\begin{aligned} (F(0))^k\left(\overline{F}(0)\right)^{n-k}(1 - H_k * G_{n-k}(y)) \leq \left(\overline{F}(0)\right)^{n-k}\overline{G}_{n-k}(y) &= O\left(y^{n-k-1}e^{-y}\right) \\ &= O\left(y^{n-3}e^{-y}\right). \end{aligned} \tag{5.45}$$

Hence from (5.38) and (5.42) through (5.45) we obtain

$$1 - F^{*n}(y) = \left(\frac{1}{(n-1)!}y^{n-1} + \frac{1+n\theta}{(n-2)!}y^{n-2} + O(y^{n-3})\right)e^{-y},$$

where

$$\theta = \theta_+ + \int_{-\infty}^0 e^y dF(y) = -\int_0^\infty e^y d\left(\overline{F}(y) - e^{-y}\right) + \int_{-\infty}^0 e^y dF(y).$$

□

From this theorem we may finally specialize to the Gumbel distribution.

Corollary 5.2. *Let* $F(y) = \exp(-e^{-y}), -\infty < y < \infty$ *denote the Gumbel distribution. Then*

$$1 - F^{*n}(y) = \left(\frac{1}{(n-1)!} y^{n-1} + \frac{1 - n\gamma}{(n-2)!} y^{n-2} + O(y^{n-3}) \right) e^{-y},$$

where γ *is Euler's constant and equals the mean of the Gumbel distribution.*

Proof. First observe that $F(y)$ satisfies the two hypotheses of the theorem. In particular,

$$\overline{F}(y) = 1 - \exp\left(-e^{-y}\right) = e^{-y} + O\left(e^{-2y}\right) \qquad \text{as} \qquad y \to \infty,$$

and so $\overline{F}(y) = e^{-y} + o(e^{-\alpha y})$ for some $1 < \alpha < 2$. Next, for any $\beta > 1$

$$\frac{F(y)}{e^{\beta y}} = \frac{\exp(-e^{-y})}{e^{\beta y}} \to 0 \qquad \text{as} \qquad y \to -\infty,$$

and therefore $F(y) = o\left(e^{\beta y}\right)$. Lastly, it can be shown that $\theta = -\gamma$. □

Observe that Theorem 5.1 is a special case of Corollary 5.2 with $n = 2$.

5.3.1 Simulation Results

We now want to check how well our second order approximation performs in simulation studies, as well as to what extend the approximation is an improvement over the first order result from Rootzén (1986). To run the simulation, we test Corollary 5.2 using $n = 2, 5, 10$. Letting $Z_{ij} \sim \Lambda, i = 1, ..., N, j = 1, ..., n$ where $N = 10$ million, we compute $X_i = \sum_{j=1}^{n} Z_{ij}$, thereby building the empirical cumulative distribution function of X, the convolution. In the graphs we focus on the upper one percent of the distributions. The two approximations we are testing are, for large enough y,

$$\overline{F}_1(y) = \frac{1}{(n-1)!} y^{n-1} e^{-y} \quad \text{and} \quad \overline{F}_2(y) = \left(\frac{1}{(n-1)!} y^{n-1} + \frac{1 - n\gamma}{(n-2)!} y^{n-2} \right) e^{-y}.$$

Figures 5.1, 5.2, and 5.3 show the results for $n = 2, 5$, and 10, respectively.

Figure 5.1: Gumbel Convolution, n = 2

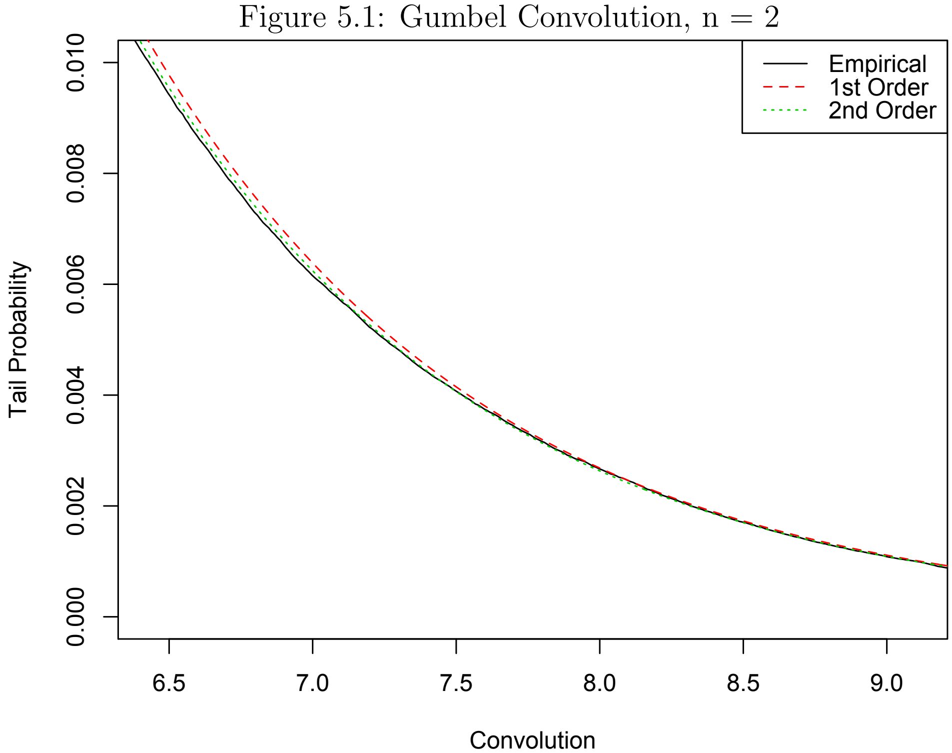


Figure 5.2: Gumbel Convolution, n = 5

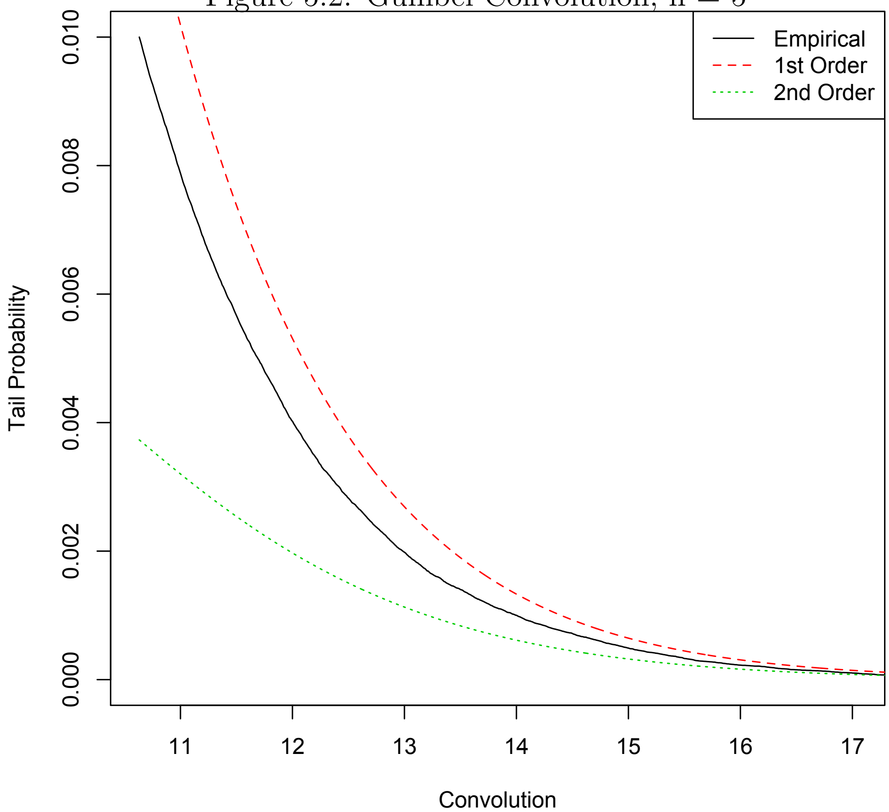
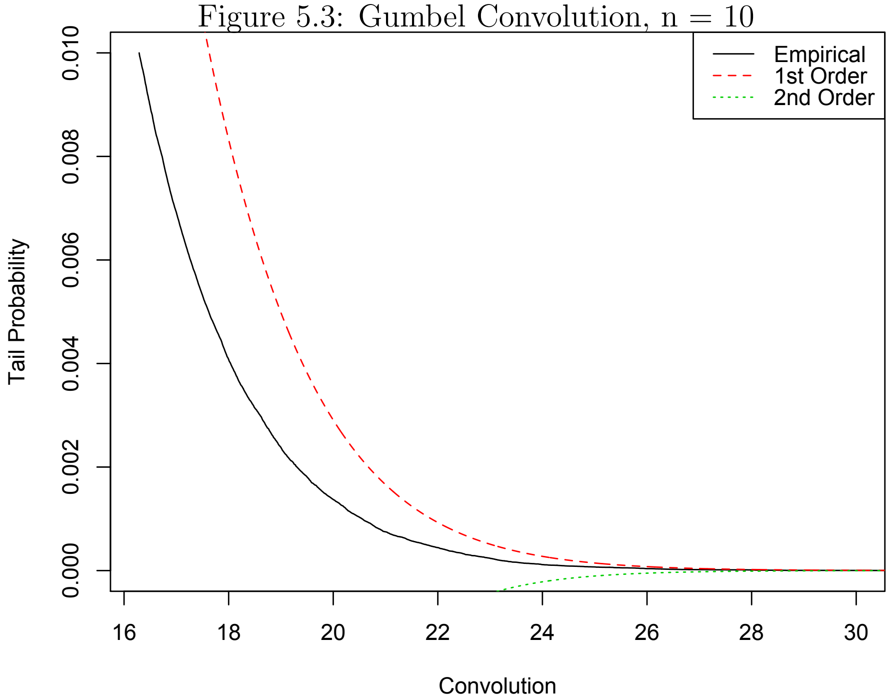


Figure 5.3: Gumbel Convolution, n = 10



Here are some immediate observations. First, for $n > 2$ the first-order approximation overestimates the true probability, while the second-order underestimates. This should not be surprising, since the $(1 - n\gamma)$ in the approximation's second term is always negative, which pulls the estimated probabilities back down. Second, both approximations eventually approach the empirical probability as y grows large. And third, for n at least 5, the second-order approximation is actually worse than the first-order for most of the percentiles. Only after a certain high percentile does the second-order approximation finally overtake the first-order, but this threshold seems to grow with n.

To check how the two functions behave at specific percentiles, we examine Tables 5.1 and 5.2. Table 5.1 displays the raw errors at each percentile, defined as empirical probability minus the approximated probability. Therefore a negative difference means the approximation overestimated the probability, while a positive difference indicates an underestimate. For each sample size and percentile we highlight the cell that gives the more precise approximation; in other words the smaller error in absolute value.

Table 5.1: Errors in Approximations of Probabilities

n	**Approx**	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
2	1st	-0.0015	-0.0007	-0.0005	-0.0003	-0.0001	-2.2e-5	2.3e-6
	2nd	0.0003	7.6e-5	3.3e-5	-6.8e-6	-3.0e-5	-4.6e-6	3.6e-6
5	1st	-0.0082	-0.0057	-0.0049	-0.0028	-0.0015	-0.0003	-3.3e-5
	2nd	0.0469	0.0196	0.0148	0.0063	0.0027	0.0004	2.6e-5
10	1st	-0.0196	-0.0159	-0.0140	-0.0086	-0.0049	-0.0011	-0.0001
	2nd	0.2136	0.1062	0.0843	0.0404	0.0192	0.0033	0.0003

Table 5.2 displays the relative errors of each approximation, defined as the error from Table 5.1 divided by the approximated probability. Again, a negative relative error denotes an overestimate, and a positive relative error an underestimate. For each sample size and percentile we highlight the cell that gives the smaller relative error.

Table 5.2: Relative Errors in Approximations of Probabilities

n	**Approx**	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
2	1st	-0.0285	-0.0260	-0.0259	-0.0246	-0.0271	-0.0215	0.0234
	2nd	0.0064	0.0030	0.0017	-0.0007	-0.0059	-0.0046	0.0371
5	1st	-0.1410	-0.1864	-0.1973	-0.2164	-0.2264	-0.2366	-0.2495
	2nd	15.1958	3.5963	2.8472	1.6915	1.1771	0.6524	0.3429
10	1st	-0.3375	-0.5190	-0.5639	-0.6755	-0.7595	-0.8732	-0.9101
	2nd	69.1836	19.5277	16.2080	10.8818	8.3623	5.5143	3.6136

With $n = 2$, the second-order approximation provides a reasonable improvement even for as low as the 95th percentile. Around the 99.99th percentile the first-order overtakes the second in precision. But for larger sample sizes, even as low as $n = 5$, the first-order provides a more accurate estimate up to a certain percentile, after which the second-order finally provides an improvement. In the case of $n = 5$, the second-order is better only after the 99.99th percentile. When $n = 10$, we do not see any evidence of an improvement even at the 99.99th percentile, although presumably it eventually happens.

The conclusion we draw from this simulation study is that as sample size increases, the second-order approximation is actually worse than the first-order up to higher percentiles, after which the improvement may finally be noticeable. Therefore these results motivate getting more terms in the approximation. In the next section we establish theory for just that. In fact, our upcoming general result is that for the convolution on n Gumbel random variables, one can derive an n-term expansion.

5.4 The General N-Fold Convolution Expansion

After examining the surprising simulation results, we now turn to the task of deriving a more general expansion for the convolution of Gumbel random variables. Whereas in the previous section we made some assumptions about the distribution function, obtained a general result, and then specialized to the particular case of the Gumbel, here for simplicity we assume that the distribution is the Gumbel. We introduce the necessary notation, and then derive the formula with the necessary theory before conducting further simulation results. Finally, we include in Appendices C.1 through C.4 tables of numerical values that are needed in the formula.

Let F denote a distribution function. We ultimately want F to be the standard Gumbel distribution, but we proceed to obtain our expansions by considering distribution functions with support on $(0, \infty)$ and $(-\infty, 0]$ separately.

For the $(0, \infty)$ support case, suppose

$$F(x) = \Lambda(x) - \Lambda(0), \qquad x \geq 0, \tag{5.46}$$

where $\Lambda(x)$ denotes the standard Gumbel distribution. Then the tail distribution

$$\overline{F}(x) = F(\infty) - F(x) = 1 - \Lambda(x), \qquad x \geq 0$$

has the same tail area as the Gumbel, but F is a defective distribution since $F(\infty) = 1 - \Lambda(0)$. Since convolution is defined for functions of bounded variations, including defective distributions, we shall proceed by first working with the defective F and adjusting later to include the negative half line.

For $k \geq 2$, let $A_k(x)$ be an approximation to $\overline{F^{*k}}(x) = F^{*k}(\infty) - F^{*k}(x)$. We shall assume that the error term in the approximation is exponential, namely for some $\alpha > 1$

$$\overline{F^{*k}}(x) = A_k(x) + o(e^{-\alpha x}) \qquad \text{as} \qquad x \to \infty.$$

Note that

$$F^{*k}(\infty) = \int_0^\infty F^{*(k-1)}(\infty) dF(x) = F^{*(k-1)}(\infty)F(\infty) = \cdots = [F(\infty)]^k.$$

Decomposing the approximation, write

$$A_k(x) = \left(\sum_{i=0}^{k-1} a_{k,i} x^i \right) e^{-x}, \qquad x > 0. \tag{5.47}$$

Before proving the main result, we need to introduce three more symbols that will be used in the expansion. Define θ_k and ζ_k as

$$\theta_k = \int_0^\infty x^k e^x d\left(-\overline{F}(x) + e^{-x}\right), \qquad k \geq 0 \tag{5.48}$$

and

$$\zeta_k = \int_{-\infty}^0 x^k e^x d\Lambda(x), \qquad k \geq 0. \tag{5.49}$$

Also for $0 \leq i \leq m$, define $\mu_{i,m}$ as

$$\mu_{i,m} = \sum \left(\frac{i!}{k_1! \cdots k_i!} \right) \left(\frac{m!}{(m - [k_1 + \cdots + k_i])!} \right) \zeta_0^{m-(k_1+\cdots+k_i)} \prod_{L=1}^{i} \left(\frac{\zeta_L}{L!} \right)^{k_L}, \tag{5.50}$$

where the sum is taken over all nonnegative integers $k_1, \cdots, k_i$ such that $k_1 + 2k_2 + \cdots + ik_i = i$.

5.4.1 The Proof on the Positive Half Line

We begin with a useful lemma that lays the groundwork for the expansion.

Lemma 5.4. *If G and K are two improper distribution functions with support on the positive half line, then*

$$\overline{G*K}(x) = \int_0^{x/2} \overline{K}(x-y)dG(y) + \int_0^{x/2} \overline{G}(x-y)dK(y) + \overline{K}\left(\frac{x}{2}\right)\overline{G}\left(\frac{x}{2}\right).$$

Proof. Observe that

$$G*K(\infty) = \int_0^{\infty} K(\infty)dG(x) = K(\infty)G(\infty)$$

and

$$\begin{aligned}\overline{G*K}(x) &= K(\infty)G(\infty) - \int_0^{\infty} K(x-y)dG(y) = \int_0^{\infty}\left[K(\infty) - K(x-y)\right]dG(y) \\ &= \int_0^{x/2}\left[K(\infty) - K(x-y)\right]dG(y) + \int_{x/2}^{\infty}\left[K(\infty) - K(x-y)\right]dG(y).\end{aligned}$$

Next, we have

$$\int_{x/2}^{\infty}\left[K(\infty) - K(x-y)\right]dG(y) = K(\infty)\left[G(\infty) - G\left(\frac{x}{2}\right)\right] - \int_{x/2}^{\infty} K(x-y)dG(y)$$

and

$$\begin{aligned}\int_{x/2}^{\infty} K(x-y)dG(y) &= \int_{x/2}^{x} K(x-y)dG(y) = \int_0^{x/2} K(y)d\left(G(\infty) - G(x-y)\right) \\ &= K(y)\left[G(\infty) - G(x-y)\right]\Big|_0^{x/2} - \int_0^{x/2}\left[G(\infty) - G(x-y)\right]dK(y) \\ &= K\left(\frac{x}{2}\right)\left[G(\infty) - G\left(\frac{x}{2}\right)\right] - \int_0^{x/2}\left[G(\infty) - G(x-y)\right]dK(y).\end{aligned}$$

Substituting, we find

$$\int_{x/2}^{\infty} [K(\infty) - K(x-y)]\, dG(y) = K(\infty)\left[G(\infty) - G\left(\frac{x}{2}\right)\right]$$
$$- K\left(\frac{x}{2}\right)\left[G(\infty) - G\left(\frac{x}{2}\right)\right] + \int_0^{x/2} [G(\infty) - G(x-y)]\, dK(y)$$
$$= K(\infty)G(\infty) - K(\infty)G\left(\frac{x}{2}\right) - G(\infty)K\left(\frac{x}{2}\right) + K\left(\frac{x}{2}\right)G\left(\frac{x}{2}\right)$$
$$+ \int_0^{x/2} [G(\infty) - G(x-y)]\, dK(y)$$
$$= \left(K(\infty) - K\left(\frac{x}{2}\right)\right)\left(G(\infty) - G\left(\frac{x}{2}\right)\right) + \int_0^{x/2} [G(\infty) - G(x-y)]\, dK(y).$$

Thus we find

$$\overline{G * K}(x) = \int_0^{x/2} \overline{K}(x-y)dG(y) + \int_0^{x/2} \overline{G}(x-y)dK(y) + \overline{K}\left(\frac{x}{2}\right)\overline{G}\left(\frac{x}{2}\right), \tag{5.51}$$

as required. □

Now we have, with F given in (5.46) and with support in $(0, \infty)$ using approximation (5.47) in (5.51), that for y large

$$\overline{F^{*(k+1)}}(y) = \int_0^{y/2} \overline{F}^{*k}(y-x)dF(x) + \int_0^{y/2} \overline{F}(y-x)dF^{*k}(x) + \overline{F}^{*k}\left(\frac{y}{2}\right)\overline{F}\left(\frac{y}{2}\right)$$
$$= \int_0^{y/2} [A_k(y-x) + R_k(y-x)]\, dF(x) + \int_0^{y/2} \overline{F}(y-x)d\left(-A_k(x) - R_k(y-x)\right)$$
$$+ \left[A_k\left(\frac{y}{2}\right) + o\left(e^{-\alpha y/2}\right)\right]\overline{F}\left(\frac{y}{2}\right),$$

where for y large enough and any $\epsilon > 0$, $\sup\limits_{0<x<\frac{y}{2}} |R_k(y-x)e^{\alpha y}| < \epsilon$ and $R_k(x) = \overline{F^{*k}}(x) -$

$A_k(x)$. Dividing into six pieces, we have

$$\overline{F^{*(k+1)}}(y) = \int_0^{y/2} A_k(y-x)dF(x) + \int_0^{y/2} \overline{F}(y-x)d\left(-A_k(x)\right) + A_k\left(\frac{y}{2}\right)\overline{F}\left(\frac{y}{2}\right)$$
$$+ \int_0^{y/2} R_k(y-x)dF(x) + \int_0^{y/2} \overline{F}(y-x)d(-R_k(y-x)) + o\left(e^{-\alpha y/2}\overline{F}\left(\frac{y}{2}\right)\right) \quad (5.52)$$
$$= (I) + (II) + (III) + (IV) + (V) + (VI).$$

Examining (I) first,

$$\begin{aligned}\int_0^{y/2} A_k(y-x)dF(x) &= \int_0^{y/2} \left[\sum_{i=0}^{k-1} a_{k,i}(y-x)^i\right] e^{-(y-x)}dF(x) \\ &= \int_0^{y/2} \left[\sum_{i=0}^{k-1} a_{k,i}\sum_{j=0}^{i}\binom{i}{j}y^j(-1)^{i-j}x^{i-j}\right] e^{-y}e^x dF(x) \qquad (5.53) \\ &= \sum_{i=0}^{k-1}\sum_{j=0}^{i} a_{k,i}\binom{i}{j}(-1)^{i-j}y^j e^{-y}\int_0^{y/2} x^{i-j}e^x dF(x).\end{aligned}$$

Now using the definition of θ_{i-j} given in (5.48),

$$\begin{aligned}\int_0^{y/2} x^{i-j}e^x dF(x) &= \int_0^{y/2} x^{i-j}e^x d(-\overline{F}(x) + e^{-x}) - \int_0^{y/2} x^{i-j}e^x d(e^{-x}) \\ &= \theta_{i-j} + \int_0^{y/2} x^{i-j}dx - \int_{y/2}^{\infty} x^{i-j}e^x d(-\overline{F}(x) + e^{-x}) \qquad (5.54) \\ &= \theta_{i-j} + \frac{1}{i-j+1}\left(\frac{y}{2}\right)^{i-j+1} - \int_{y/2}^{\infty} x^{i-j}e^x d(-\overline{F}(x) + e^{-x}).\end{aligned}$$

Noting that $d(-\overline{F}(x) + e^{-x}) = -e^{-x}\left[1 - \exp(-e^{-x})\right]dx$, the last integral in (5.54) can be rewritten as

$$-\int_{y/2}^{\infty} x^{i-j} e^{x} d(-\overline{F}(x) + e^{-x}) = \int_{y/2}^{\infty} x^{i-j}[1 - \exp(-e^{-x})]dx$$
$$= \int_{y/2}^{\infty} x^{i-j} \sum_{L=1}^{\infty} \frac{(-1)^{L-1} e^{-Lx}}{L!} dx = \sum_{L=1}^{\infty} \frac{(-1)^{L-1}}{L!} \int_{y/2}^{\infty} x^{i-j} e^{-Lx} dx.$$

By induction, it can be shown that for $n \geq 1$ and $L = 1, 2, \cdots,$

$$\int_{y/2}^{\infty} x^{n} e^{-Lx} dx = \frac{1}{L}\left(\frac{y}{2}\right)^{n} e^{-Ly/2} + \sum_{t=1}^{n} \frac{1}{L^{t+1}} \left(\frac{y}{2}\right)^{n-t} \prod_{p=0}^{t-1}(n-p) e^{-Ly/2} = O(y^{n} e^{-Ly/2}),$$

and therefore as $y \to \infty$

$$\int_{y/2}^{\infty} x^{i-j} e^{-Lx} dx = O(y^{i-j} e^{-Ly/2}).$$

Even in the case where $i = j$, the asymptotics still hold in that the error term is $O(e^{-Ly/2})$. Thus as $y \to \infty$

$$\int_{y/2}^{\infty} x^{i-j}[1 - \exp(-e^{-x})]dx = \sum_{L=1}^{\infty} \frac{(-1)^{L-1}}{L!} \int_{y/2}^{\infty} x^{i-j} e^{-Lx} dx = O(y^{i-j} e^{-y/2}).$$

Equation (5.54) now reads

$$\int_{0}^{y/2} x^{i-j} e^{x} dF(x) = \theta_{i-j} + \frac{1}{i-j+1}\left(\frac{y}{2}\right)^{i-j+1} + O(y^{i-j} e^{-y/2}). \tag{5.55}$$

Substituting (5.55) into (5.53),

$$\begin{aligned}
&\int_0^{y/2} A_k(y-x)dF(x) \\
&\quad = \sum_{i=0}^{k-1}\sum_{j=0}^{i} a_{k,i}\binom{i}{j}(-1)^{i-j}y^j e^{-y}\left[\theta_{i-j} + \frac{1}{i-j+1}\left(\frac{y}{2}\right)^{i-j+1} + O(y^{i-j}e^{-y/2})\right] \\
&\quad = \sum_{i=0}^{k-1}\sum_{j=0}^{i} a_{k,i}\binom{i}{j}(-1)^{i-j}y^j e^{-y}\left[\theta_{i-j} + \frac{1}{i-j+1}\left(\frac{y}{2}\right)^{i-j+1}\right] + O(y^i e^{-3y/2}).
\end{aligned}$$

Note that provided $\alpha < \frac{3}{2}$,

$$\lim_{y\to\infty} \frac{y^i e^{-3y/2}}{e^{-\alpha y}} = \lim_{y\to\infty} y^i e^{-(\frac{3}{2}-\alpha)y} = 0,$$

so now we require $1 < \alpha < \frac{3}{2}$. To conclude,

$$(I) = \sum_{i=0}^{k-1}\sum_{j=0}^{i} a_{k,i}\binom{i}{j}(-1)^{i-j}y^j e^{-y}\left[\theta_{i-j} + \frac{1}{i-j+1}\left(\frac{y}{2}\right)^{i-j+1}\right] + O(y^i e^{-3y/2}). \tag{5.56}$$

Moving on to (II) in (5.52),

$$\begin{aligned}
-\frac{d}{dx}A_k(x) &= -\frac{d}{dx}\left(\sum_{i=0}^{k-1} a_{k,i}x^i\right)e^{-x} = -\left(\sum_{i=1}^{k-1} i a_{k,i}x^{i-1} - \sum_{i=0}^{k-1} a_{k,i}x^i\right)e^{-x} \\
&= \sum_{i=1}^{k-1}(a_{k,i}x^i - i a_{k,i}x^{i-1})e^{-x} + a_{k,0}e^{-x}.
\end{aligned} \tag{5.57}$$

Then

$$
\begin{aligned}
&\int_0^{y/2} \overline{F}(y-x)d(-A_k(x)) \\
&\quad = \int_0^{y/2} \left[e^{-(y-x)} + R_k(y-x)\right] \left[\sum_{i=1}^{k-1}(a_{k,i}x^i - ia_{k,i}x^{i-1}) + a_{k,0}\right] e^{-x}dx \\
&\quad = e^{-y} \int_0^{y/2} \left[\sum_{i=1}^{k-1}(a_{k,i}x^i - ia_{k,i}x^{i-1}) + a_{k,0}\right] dx \\
&\qquad\quad + \int_0^{y/2} \left[\sum_{i=1}^{k-1}(a_{k,i}x^i - ia_{k,i}x^{i-1}) + a_{k,0}\right] R_k(y-x)e^{-x}dx \\
&\quad = (IIA) + (IIB).
\end{aligned}
\tag{5.58}
$$

Examining (IIA),

$$
\begin{aligned}
(IIA) &= e^{-y}\left\{\sum_{i=1}^{k-1}\int_0^{y/2}(a_{k,i}x^i - ia_{k,i}x^{i-1})dx + \int_0^{y/2} a_{k,0}dx\right\} \\
&= e^{-y}\left\{\sum_{i=1}^{k-1} a_{k,i}\left[\frac{1}{i+1}\left(\frac{y}{2}\right)^{i+1} - \left(\frac{y}{2}\right)^{i}\right] + a_{k,0}\left(\frac{y}{2}\right)\right\}.
\end{aligned}
\tag{5.59}
$$

Turning to (IIB),

$$
\begin{aligned}
|(IIB)| &= \left|\int_0^{y/2}\left[\sum_{i=1}^{k-1}(a_{k,i}x^i - ia_{k,i}x^{i-1}) + a_{k,0}\right] R_k(y-x)e^{\alpha(y-x)}e^{-x}e^{-\alpha(y-x)}dx\right| \\
&< \epsilon e^{-\alpha y}\left|\int_0^{y/2}\left[\sum_{i=1}^{k-1}a_{k,i}(x^i - ix^{i-1}) + a_{k,0}\right] e^{(\alpha-1)x}dx\right| \\
&\quad = \epsilon e^{-\alpha y}\left|\left\{\sum_{i=1}^{k-1}a_{k,i}\int_0^{y/2}(x^i - ix^{i-1})e^{(\alpha-1)x}dx + a_{k,0}\int_0^{y/2} e^{(\alpha-1)x}dx\right\}\right|
\end{aligned}
$$

$$= \epsilon e^{-\alpha y} \left| \left\{ \sum_{i=1}^{k-1} a_{k,i} \left[\int_0^{y/2} x^i e^{(\alpha-1)x} dx - i \int_0^{y/2} x^{i-1} e^{(\alpha-1)x} dx \right] + a_{k,0} \int_0^{y/2} e^{(\alpha-1)x} dx \right\} \right| .$$

By induction and integration by parts, it can be shown that for $\alpha > 1$ and $n \geq 1$

$$\int_0^{y/2} x^n e^{(\alpha-1)x} dx = (\alpha-1)^{-1} \left(\frac{y}{2}\right)^n e^{(\alpha-1)y/2} + (-1)^{n+1} n! (\alpha-1)^{-(n+1)}$$
$$+ e^{(\alpha-1)y/2} \sum_{L=0}^{n-1} (-1)^{L+1} (\alpha-1)^{-(L+2)} \prod_{p=0}^{L} (n-p) \left(\frac{y}{2}\right)^{n-1-L},$$

and therefore as $y \to \infty$

$$e^{-\alpha y} \int_0^{y/2} x^n e^{(\alpha-1)x} dx = O\left(y^n e^{-(\alpha+1)y/2}\right).$$

Using this result,

$$|(IIB)| < \epsilon \left| \left\{ O\left(y^n e^{-(\alpha+1)y/2}\right) + a_{k,0} (\alpha-1)^{-1} \left[e^{-(\alpha+1)y/2} - e^{-\alpha y} \right] \right\} \right| .$$

If we choose $1 < \beta < \frac{\alpha+1}{2}$, then

$$\lim_{y\to\infty} \frac{e^{-(\alpha+1)y/2} - e^{-\alpha y}}{e^{-\beta y}} = \lim_{y\to\infty} \left\{ \exp\left(\frac{(2\beta - \alpha - 1)y}{2} \right) - e^{-(\alpha-\beta)y} \right\} = 0.$$

Note that the condition $1 < \beta < \alpha$ is automatically satisfied since $\alpha > 1$. Similarly,

$$\lim_{y\to\infty} \frac{y^n e^{-(\alpha+1)y/2}}{e^{-\beta y}} = 0,$$

and therefore $(IIB) = o(e^{-\beta y})$ as $y \to \infty$. Combining this result with (5.59) in (5.58), (II) in (5.52) reads

$$(II) = e^{-y}\left\{\sum_{i=1}^{k-1} a_{k,i}\left[\frac{1}{i+1}\left(\frac{y}{2}\right)^{i+1} - \left(\frac{y}{2}\right)^{i}\right] + a_{k,0}\left(\frac{y}{2}\right)\right\} + o(e^{-\beta y}). \tag{5.60}$$

As a side note, since we must select $1 < \alpha < \frac{3}{2}$, it follows that we should choose $1 < \beta < \frac{\alpha+1}{2} < \frac{5}{4}$. Now consider (III) in (5.52). Since $\overline{F}(x) = e^{-x} + O(e^{-2x})$, it is certainly true that $\overline{F}(x) = e^{-x} + o(e^{-\alpha x})$ for x large. Therefore

$$\begin{aligned}(III) = A_k\left(\frac{y}{2}\right)\overline{F}\left(\frac{y}{2}\right) &= \left[\sum_{i=0}^{k-1} a_{k,i}\left(\frac{y}{2}\right)^{i}\right] e^{-y/2}\left[e^{-y/2} + o(e^{-\alpha y/2})\right] \\ &= \left[\sum_{i=0}^{k-1} a_{k,i}\left(\frac{y}{2}\right)^{i}\right]\left[e^{-y} + o(e^{-(\alpha+1)y/2})\right] \\ &= \left[\sum_{i=0}^{k-1} a_{k,i}\left(\frac{y}{2}\right)^{i}\right] e^{-y} + O\left(y^{k-1}e^{-(\alpha+1)y/2}\right),\end{aligned}$$

which means that as $y \to \infty$

$$A_k\left(\frac{y}{2}\right)\overline{F}\left(\frac{y}{2}\right) = \left[\sum_{i=0}^{k-1} a_{k,i}\left(\frac{y}{2}\right)^{i}\right] e^{-y} + o\left(e^{-\beta y}\right). \tag{5.61}$$

It remains to derive the overall error term in (5.52); namely expressions (IV), (V), and (VI). Lemma 5.5 does just that.

Lemma 5.5. *In (5.52), $(IV) + (V) + (VI) = o(e^{-\beta y})$ as $y \to \infty$ for β previously defined.*

Proof. In what follows we assume that y is large. We examine (IV) first:

$$\begin{aligned}
\left|\int_0^{y/2} R_k(y-x)dF(x)\right| &= \left|\int_0^{y/2} R_k(y-x)e^{\alpha(y-x)}e^{-\alpha(y-x)}dF(x)\right| \\
&< \epsilon e^{-\alpha y}\int_0^{y/2} e^{\alpha x}dF(x) = \epsilon e^{-\alpha y}\int_0^{y/2} e^{\alpha x}e^{-x}\exp(-e^{-x})dx \\
&= \epsilon e^{-\alpha y}\int_{e^{-y/2}}^{1} x^{-\alpha}e^{-x}dx \le \epsilon e^{-\alpha y}\int_{e^{-y/2}}^{1} x^{-\alpha}dx = \frac{\epsilon}{\alpha-1}e^{-\alpha y}\left[e^{(\alpha-1)y/2}-1\right] \\
&= \frac{\epsilon}{\alpha-1}\left[e^{-(\alpha+1)y/2}-e^{-\alpha y}\right] = o(e^{-\beta y}).
\end{aligned}$$

Next, observe that in (V)

$$\left|\int_0^{y/2} \overline{F}(y-x)d(-R_k(y-x))\right| \le \overline{F}\left(\frac{y}{2}\right)\left[R(y)-R\left(\frac{y}{2}\right)\right],$$

because for $y>0$, $0<\overline{F}(y)\le\overline{F}\left(\frac{y}{2}\right)$. Hence

$$\begin{aligned}
|(V)| &\le \left|\left[e^{-y/2}+o(e^{-\alpha y/2})\right]\left[R_k(y)e^{\alpha y}e^{-\alpha y}-R_k\left(\frac{y}{2}\right)e^{\alpha y/2}e^{-\alpha y/2}\right]\right| \\
&< \epsilon\left[e^{-y/2}+o(e^{-\alpha y/2})\right]\left[e^{-\alpha y}-e^{-\alpha y/2}\right] \\
&= \epsilon\left[e^{-(2\alpha+1)y/2}-e^{-(\alpha+1)y/2}+o\left(e^{-3\alpha y/2}\right)+o\left(e^{-\alpha y}\right)\right] \\
&= \epsilon\left[o\left(e^{-\alpha y}\right)-e^{-(\alpha+1)y/2}\right] = o(e^{-\beta y}).
\end{aligned}$$

Finally, consider $(VI) = o\left(e^{-\alpha y/2}\overline{F}\left(\frac{y}{2}\right)\right)$. Observe that

$$e^{-\alpha y/2}\overline{F}\left(\frac{y}{2}\right) = e^{-\alpha y/2}\left[e^{-y/2}+o\left(e^{-\alpha y/2}\right)\right] = e^{-(\alpha+1)y/2}+o(e^{-\alpha y}) = o(e^{-\beta y}),$$

and therefore $(VI) = o(e^{-\beta y})$. The result follows. □

Before stating our result for the positive half line, it is worth pointing out that $(II)+(III)$

has a somewhat simplified form:

$$\begin{aligned} & e^{-y}\left\{\sum_{i=1}^{k-1} a_{k,i}\left[\frac{1}{i+1}\left(\frac{y}{2}\right)^{i+1}-\left(\frac{y}{2}\right)^{i}\right]+a_{k,0}\left(\frac{y}{2}\right)+\sum_{i=0}^{k-1} a_{k,i}\left(\frac{y}{2}\right)^{i}\right\} \\ & \quad = e^{-y}\left\{\sum_{i=0}^{k-1} \frac{a_{k,i}}{i+1}\left(\frac{y}{2}\right)^{i+1}+a_{k,0}\right\}. \end{aligned} \tag{5.62}$$

Using the results of (5.56), (5.62), and Lemma 5.5 in (5.52), we finally obtain that as $y \to \infty$

$$\begin{aligned} \overline{F^{*(k+1)}}(y) = e^{-y} & \left\{\sum_{i=0}^{k-1}\sum_{j=0}^{i} a_{k,i}\binom{i}{j}(-1)^{i-j}y^{j}\left[\theta_{i-j}+\frac{1}{i-j+1}\left(\frac{y}{2}\right)^{i-j+1}\right]\right. \\ & \left. +\sum_{i=0}^{k-1}\frac{a_{k,i}}{i+1}\left(\frac{y}{2}\right)^{i+1}+a_{k,0}\right\}+o(e^{-\beta y}). \end{aligned} \tag{5.63}$$

In order to "peel off" the terms involving $y^i, i = 0, 1, \cdots, k$, it would be more convenient to change the order of summation in the first term of (5.63) as in Lemma 5.6.

Lemma 5.6. *The approximation in (5.63) can be restated as*

$$\begin{aligned} \overline{F^{*(k+1)}}(y) = e^{-y} & \left\{\sum_{j=0}^{k-1}\left[\sum_{i=j}^{k-1} a_{k,i}\binom{i}{j}(-1)^{i-j}\theta_{i-j}\right]y^{j}+\sum_{i=0}^{k-1}\left[\frac{a_{k,i}}{(i+1)2^{i+1}}\right]y^{i+1}+a_{k,0}\right. \\ & \left. +\sum_{i=0}^{k-1}\left[\sum_{j=0}^{i}\left(\frac{a_{k,i}}{i-j+1}\right)\binom{i}{j}(-1)^{i-j}\left(\frac{1}{2}\right)^{i-j+1}\right]y^{i+1}\right\}+o\left(e^{-\beta y}\right). \end{aligned}$$

We first need to find the starting constants $a_{1,0}$, $a_{2,1}$, and $a_{2,0}$. Since $A_n(y) \sim F(y) \sim e^{-y}$ for y large enough, $a_{1,0} = 1$. Then (5.14) provides $A_2(y) = (y + 1 + 2\theta_0)e^{-y}$, and therefore $a_{2,1} = 1$ and $a_{2,0} = 1 + 2\theta_0$. These initial constants will lay the groundwork for recovering the necessary remaining constants.

To utilize Lemma 5.6, we simply isolate the exponents of interest. That is, we know that

$e^y\overline{F^{*(k+1)}}(y) \approx \sum_{i=0}^{k} a_{k+1,i}y^i$, so to solve for $a_{k+1,i}$ we need to find the constants associated with y^i. For instance, we first derive $a_{k+1,0}$ by picking off the y^0 terms. Set $j = 0$ in the first summation:

$$a_{k+1,0} = \sum_{i=0}^{k-1} a_{k,i}(-1)^i\theta_i + a_{k,0}. \tag{5.64}$$

Next, for $r = 1, \cdots, k-1$ we solve for $a_{k+1,r}$. This requires setting $j = r-1$ in the first summation, $i = r-1$ in the second, and $i = r-1$ in the last:

$$a_{k+1,r} = \sum_{i=r}^{k-1} \binom{i}{r}(-1)^{i-r}a_{k,i}\theta_{i-r} + \sum_{j=0}^{r-1} \binom{r-1}{j}(-1)^{r-1-j}\frac{a_{k,r-1}}{(r-j)2^{r-j}} + \frac{a_{k,r-1}}{r2^r}.$$

However, this equation can be simplified. Noting that $\binom{r-1}{j} = \frac{r-j}{r}\binom{r}{j}$,

$$\begin{aligned}
a_{k+1,r} &= \sum_{i=r}^{k-1} \binom{i}{r}(-1)^{i-r}a_{k,i}\theta_{i-r} - \frac{a_{k,r-1}}{r}\left[\sum_{j=0}^{r}\binom{r}{j}1^j\left(-\frac{1}{2}\right)^{r-j} - \binom{r}{r}1^r\right] + \frac{a_{k,r-1}}{r2^r} \\
&= \sum_{i=r}^{k-1} \binom{i}{r}(-1)^{i-r}a_{k,i}\theta_{i-r} - \frac{a_{k,r-1}}{r}\left[\left(1-\frac{1}{2}\right)^r - 1\right] + \frac{a_{k,r-1}}{r2^r} \\
&= \sum_{i=r}^{k-1} \binom{i}{r}(-1)^{i-r}a_{k,i}\theta_{i-r} + \frac{a_{k,r-1}}{r}.
\end{aligned}$$

Therefore for $r = 1, \cdots, k-1$,

$$a_{k+1,r} = \frac{a_{k,r-1}}{r} + \sum_{i=r}^{k-1} \binom{i}{r}(-1)^{i-r}a_{k,i}\theta_{i-r}. \tag{5.65}$$

The recursion (5.65) is useful for the purposes of writing computer code to generate the next constant. However, for writing out the explicit formulas for expansions we may also find the alternative statement handy: for $r = 2, \cdots, k$,

$$a_{k+1,k+1-r} = \frac{a_{k,k-r}}{k+1-r} + \sum_{i=k+1-r}^{k-1} \binom{i}{k+1-r} (-1)^{i-(k+1-r)} a_{k,i} \theta_{i-(k+1-r)}. \tag{5.66}$$

Lastly, we derive $a_{k+1,k}$. Choose $i = k-1$ in the second and last sums:

$$\begin{aligned}
a_{k+1,k} &= \frac{a_{k,k-1}}{k2^k} + \sum_{j=0}^{k-1} \left(\frac{a_{k,k-1}}{k-j}\right) \binom{k-1}{j} (-1)^{k-1-j} \left(\frac{1}{2}\right)^{k-j} \\
&= a_{k,k-1} \left\{ \frac{1}{k2^k} + \sum_{j=0}^{k-1} \frac{1}{k} \binom{k}{j} (-1)^{k-1-j} \left(\frac{1}{2}\right)^{k-j} \right\} \\
&= a_{k,k-1} \left\{ \frac{1}{k}\left(\frac{1}{2}\right)^k - \frac{1}{k}\left[\sum_{j=0}^{k-1} \binom{k}{j} (1)^j \left(-\frac{1}{2}\right)^{k-j} + \binom{k}{k}\left(-\frac{1}{2}\right)^{k-k} - \binom{k}{k} \right] \right\} \\
&= a_{k,k-1} \left\{ \frac{1}{k}\left(\frac{1}{2}\right)^k - \frac{1}{k}\left(1 - \frac{1}{2}\right)^k + \frac{1}{k} \right\} = \frac{1}{k} a_{k,k-1}.
\end{aligned}$$

Therefore

$$a_{k+1,k} = \frac{1}{k} a_{k,k-1} = \cdots = \frac{1}{k!} a_{1,0} = \frac{1}{k!}. \tag{5.67}$$

We have now successfully found recursion formulas for the $a_{k,i}$, and these will be used in computing the final approximation formula in the next section.

5.4.2 The Proof Over All Reals

Here we shall build upon (5.63) and extend it to include the negative half line. Define

$$H(x) = \frac{\Lambda(x)}{\Lambda(0)} \quad \text{and} \quad G(x) = \frac{\Lambda(x) - \Lambda(0)}{1 - \Lambda(0)},$$

which denote the conditional distribution functions $P(X \le x | X \le 0)$ and $P(X \le x | X > 0)$, respectively, where $X \sim \Lambda$. Observe that for $x > 0$ and $N = 0, \cdots, n$,

$$\begin{aligned}\frac{P\left(\sum_{i=1}^{n} X_i \le x, N = m\right)}{P(N = m)} &= [F(0)]^m [\overline{F}(0)]^{n-m} \\ &\quad \times \left[\frac{\binom{n}{m} P\left(\sum_{i=1}^{n} X_i \le x, X_1 \le 0, \cdots, X_m \le 0, X_{m+1} > 0, \cdots, X_n > 0\right)}{\binom{n}{m}[F(0)]^m[\overline{F}(0)]^{n-m}}\right] \\ &= P([U_1 + \cdots + U_m] + [V_{m+1} + \cdots + V_n] \le x).\end{aligned}$$

Here $U_i = X_i I_{X_i \le 0}$ and $V_i = X_i I_{X_i > 0}$. Thus

$$\begin{aligned}P\left(\sum_{i=1}^{n} X_i > x\right) &= \sum_{m=0}^{n} P\left(\sum_{i=1}^{n} X_i > x \middle| N = n\right)\binom{n}{m}[F(0)]^m[\overline{F}(0)]^{n-m} \\ &= \sum_{m=0}^{n} \overline{H^{*m} * G^{*(n-m)}}(x)\binom{n}{m}[F(0)]^m[\overline{F}(0)]^{n-m}, \qquad x > 0.\end{aligned}$$

Therefore for our context,

$$P\left(\sum_{i=1}^{n} X_i > x\right) = \sum_{m=0}^{n}\binom{n}{m}\overline{H^{*m} * G^{*(n-m)}}(x)[\Lambda(0)]^m[\overline{\Lambda}(0)]^{n-m}, \qquad x > 0. \tag{5.68}$$

Further

$$G^{*(n-m)}(x)[\overline{\Lambda}(0)]^{n-m} = F^{*(n-m)}(x). \tag{5.69}$$

Thus from (5.68), for $x > 0$ we have

$$\begin{aligned}P\left(\sum_{i=1}^{k+1} X_i > x\right) &= \sum_{m=1}^{k}\binom{k+1}{m}\overline{H^{*m} * G^{*(k+1-m)}}(x)[\Lambda(0)]^m[\overline{\Lambda}(0)]^{k+1-m} \\ &\quad + \overline{G^{*(k+1)}}(x)[\overline{\Lambda}(0)]^{k+1}.\end{aligned} \tag{5.70}$$

The summation in (5.60) does not include the $m = k + 1$ term. The reason is because here we have $(k+1)$ variables all coming from H, which is defined only on the negative half line, and therefore their sum is positive with probability 0.

We first examine the second term in (5.70) for $y > 0$:

$$\begin{aligned}\overline{G^{*(k+1)}}(y)\overline{\Lambda}(0)^{k+1} &= [\overline{\Lambda}(0)]^{k+1}\left(1 - G^{*(k+1)}(y)\right) = [\overline{\Lambda}(0)]^{k+1} - F^{*(k+1)}(y) \\ &= \overline{F^{*(k+1)}}(y).\end{aligned} \tag{5.71}$$

This term was already derived since (5.71) = (5.63). Let $\rho_k(y-x)$ represent an error term to be analyzed later, one that is analogous to $R_k(y-x)$ from earlier, and define $(II) = \int_{-\infty}^{0} \rho_k(y-x)dH^{*m}(x)[\Lambda(0)]^m$. We now turn to the first piece in (5.70):

$$\begin{aligned}&\overline{H^{*m} * G^{*(k+1-m)}}(y)[\Lambda(0)]^m[\overline{\Lambda}(0)]^{k+1-m} \\ &\quad = \int_{-\infty}^{0} \overline{G^{*(k+1-m)}}(y-x)dH^{*m}(x)[\Lambda(0)]^m[\overline{\Lambda}(0)]^{k+1-m} \\ &\quad = \int_{-\infty}^{0} \overline{F^{*(k+1-m)}}(y-x)dH^{*m}(x)[\Lambda(0)]^m \\ &\quad = \int_{-\infty}^{0} \left[A_{k+1-m}(y-x) + \rho_k(y-x)\right] dH^{*m}(x)[\Lambda(0)]^m \\ &\quad = \int_{-\infty}^{0} \left[\sum_{i=0}^{k-m} a_{k+1-m,i}(y-x)^i\right] e^{-(y-x)} dH^{*m}(x)[\Lambda(0)]^m + (II) \\ &\quad = e^{-y}\sum_{i=0}^{k-m} a_{k+1-m,i}\int_{-\infty}^{0}\sum_{j=0}^{i}\binom{i}{j} y^j(-x)^{i-j}e^x dH^{*m}(x)[\Lambda(0)]^m + (II) \\ &\quad = e^{-y}\sum_{i=0}^{k-m} a_{k+1-m,i}\sum_{j=0}^{i}\binom{i}{j}(-1)^{i-j}y^j\int_{-\infty}^{0} x^{i-j}e^x dH^{*m}(x)[\Lambda(0)]^m + (II) \\ &\quad = (I) + (II).\end{aligned} \tag{5.72}$$

We analyze in (I)

$$\int_{-\infty}^{0} x^{i-j} e^{x} dH^{*m}(x)[\Lambda(0)]^{m} = \frac{d^{(i-j)}}{dt^{(i-j)}} \int_{-\infty}^{0} e^{tx} dH^{*m}(x)[\Lambda(0)]^{m}\Big|_{t=1}. \tag{5.73}$$

Now if $U_1, \cdots, U_m \overset{iid}{\sim} H$, then

$$\int_{-\infty}^{0} e^{tx} dH^{*m}(x) = E\left(e^{t(U_1+\cdots+U_m)}\right) = \left(E\left(e^{tU_1}\right)\right)^{m} = \left(\int_{-\infty}^{0} e^{tx} dH(x)\right)^{m},$$

and therefore

$$\int_{-\infty}^{0} e^{tx} dH^{*m}(x)[\Lambda(0)]^{m} = \left(\int_{-\infty}^{0} e^{tx} d\Lambda(x)\right)^{m}. \tag{5.74}$$

Then

$$\frac{d}{dt}\left(\int_{-\infty}^{0} e^{tx} d\Lambda(x)\right)^{m} = m\left(\int_{-\infty}^{0} e^{tx} d\Lambda(x)\right)^{m-1} \int_{-\infty}^{0} x e^{tx} d\Lambda(x),$$

which evaluated at $t = 1$ yields

$$m\left(\int_{-\infty}^{0} e^{x} d\Lambda(x)\right)^{m-1} \int_{-\infty}^{0} x e^{x} d\Lambda(x).$$

Before going further, we need to state a result from Roman (1980).

Theorem 5.4. *Let $g(t)$ and $f(t)$ be two differentiable functions, and let D denote the derivative operator and D^L the Lth fold derivative. Also set $k = k_1 + \cdots + k_n$. Then by the formula of Faà di Bruno,*

$$D^{n}\left(f \circ g(t)\right) = \sum \frac{n!}{k_1! \cdots k_n!} (D^{k} f)(g(t)) \left(\frac{Dg(t)}{1!}\right)^{k_1} \cdots \left(\frac{D^{n} g(t)}{n!}\right)^{k_n},$$

where the sum is over all nonnegative integers $k_1, \cdots, k_n$ such that $k_1 + 2k_2 + \cdots + nk_n = n$.

Define $f(x) = x^m$ and $g(t) = \int_{-\infty}^{0} e^{tx} d\Lambda(x)$. We have that

$$(D^k f)(g(t)) = m(m-1)\cdots(m-k+1)(g(t))^{m-k} \qquad \text{and}$$
$$D^L g(t) = \int_{-\infty}^{0} x^L e^{tx} d\Lambda(x).$$

Therefore summing over $k_1, \cdots, k_n$ as described in Theorem 5.4,

$$\begin{aligned} &D^n \left(\int_{-\infty}^{0} e^{tx} d\Lambda(x) \right)^m \Big|_{t=1} \\ &= \sum \left(\frac{n!}{k_1! \cdots k_n!} \right) \frac{m!}{(m-k)!} \left(\int_{-\infty}^{0} e^x d\Lambda(x) \right)^{m-k} \\ &\quad \times \left(\frac{1}{1!} \int_{-\infty}^{0} x e^x d\Lambda(x) \right)^{k_1} \cdots \left(\frac{1}{n!} \int_{-\infty}^{0} x^n e^x d\Lambda(x) \right)^{k_n} \\ &= \sum \left(\frac{n!}{k_1! \cdots k_n!} \right) \frac{m!}{(m-k)!} \zeta_0^{m-k} \left(\frac{1}{1!} \zeta_1 \right)^{k_1} \left(\frac{1}{2!} \zeta_2 \right)^{k_2} \cdots \left(\frac{1}{n!} \zeta_n \right)^{k_n}, \end{aligned} \tag{5.75}$$

where ζ_k is defined in (5.49). Thus by (5.73), (5.74), and (5.75), we find

$$\int_{-\infty}^{0} x^n e^x dH^{*m}(x) [\Lambda(0)]^m = \text{ expression given in (5.75).} \tag{5.76}$$

Hence by (5.72) and (5.76),

$$\begin{aligned} &\overline{H^{*m} * G^{*(k+1-m)}}(y) [\Lambda(0)]^m [\overline{\Lambda}(0)]^{k+1-m} \\ &\quad = e^{-y} \sum_{i=0}^{k-m} a_{k+1-m,i} \sum_{j=0}^{i} \binom{i}{j} (-1)^{i-j} \mu_{i-j,m} y^j + (II), \end{aligned} \tag{5.77}$$

where $\mu_{i-j,m}$ is defined in (5.50). It remains to investigate the piece (II) in (5.72). Observe that for $y > 0$,

$$\sup_{-\infty < x < 0} e^{-\alpha(y-x)} \leq e^{-\alpha y}.$$

Then $\rho_k(y-x) = o\left(e^{-\alpha(y-x)}\right)$ and for any $\epsilon > 0$,

$$\begin{aligned}\int_{-\infty}^{0} \rho_k(y-x) dH^{*m}(x)[\Lambda(0)]^m &< \epsilon \int_{-\infty}^{0} e^{-\alpha(y-x)} dH^{*m}(x)[\Lambda(0)]^m \\ &\leq \epsilon \int_{-\infty}^{0} e^{-\alpha y} dH^{*m}(x)[\Lambda(0)]^m = \epsilon e^{-\alpha y} \left(\int_{-\infty}^{0} d\Lambda(x) \right)^m \\ &= \epsilon [\Lambda(0)]^m e^{-\alpha y} \leq \epsilon e^{-\alpha y}.\end{aligned}$$

Therefore for β defined earlier,

$$\int_{-\infty}^{0} \rho_k(y-x) dH^{*m}(x)[\Lambda(0)]^m = o(e^{-\alpha y}) = o(e^{-\beta y}). \tag{5.78}$$

Thus from (5.63), (5.70), (5.71), (5.77) , and (5.78) we obtain

$$\begin{aligned}1 - \Lambda^{*(k+1)}(y) = &\left\{ \sum_{j=0}^{k-1} \left[\sum_{i=j}^{k-1} a_{k,i} \binom{i}{j} (-1)^{i-j} \theta_{i-j} \right] y^j + \sum_{i=0}^{k-1} \left[\frac{a_{k,i}}{(i+1)2^{i+1}} \right] y^{i+1} + a_{k,0} \right. \\ &+ \sum_{i=0}^{k-1} \left[\sum_{j=0}^{i} \left(\frac{a_{k,i}}{i-j+1} \right) \binom{i}{j} (-1)^{i-j} \left(\frac{1}{2} \right)^{i-j+1} \right] y^{i+1} \\ &\left. + \sum_{m=1}^{k} \binom{k+1}{m} \sum_{i=0}^{k-m} a_{k+1-m,i} \sum_{j=0}^{i} \binom{i}{j} (-1)^{i-j} \mu_{i-j,m} y^j \right\} e^{-y} + o(e^{-\beta y}).\end{aligned} \tag{5.79}$$

The triple-sum in the last term can be rewritten as the more convenient

$$\sum_{j=0}^{k-1} \left[\sum_{m=1}^{k-j} \binom{k+1}{m} \sum_{i=j}^{k-m} a_{k+1-m,i} \binom{i}{j} (-1)^{i-j} \mu_{i-j,m} \right] y^j,$$

which we shall find useful for picking off specific terms.

5.5 The Final N-Fold Gumbel Expansion

We have finally fully derived the n-term expansion for the convolution of n Gumbel random variables. For the sake of having an easy reference, we restate the final result in this brief section along with all the necessary formulas needed to implement it. For $k \geq 0$, define the constants

$$\theta_k = \int_0^\infty x^k e^x d\left(-\overline{F}(x) + e^{-x}\right) = -\int_0^\infty x^k \left[1 - \exp\left(-e^{-x}\right)\right] dx$$

and

$$\zeta_k = \int_{-\infty}^0 x^k e^x d\Lambda(x) = \int_{-\infty}^0 x^k \exp\left(-e^{-x}\right) dx.$$

Also for $0 \leq i \leq m$, define

$$\mu_{i,m} = \sum_{\tau_i} \left(\frac{i!}{k_1! \cdots k_i!}\right)\left(\frac{m!}{(m - [k_1 + \cdots + k_i])!}\right) \zeta_0^{m-(k_1+\cdots+k_i)} \prod_{L=1}^{i} \left(\frac{\zeta_L}{L!}\right)^{k_L},$$

where the sum is taken over all nonnegative integers $k_1, \cdots, k_i$ such that $k_1+2k_2+\cdots+ik_i = i$. Moving to the $a_{k+1,i}$ constants, we have $a_{1,0} = a_{2,1} = 1$ and $a_{2,0} = 1 + 2\theta_0$. For $k \geq 2$,

$$a_{k+1,r} = \begin{cases} \sum_{i=0}^{k-1} a_{k,i}(-1)^i\theta_i + a_{k,0}, & r = 0 \\ \frac{a_{k,r-1}}{r} + \sum_{i=r}^{k-1} \binom{i}{r} (-1)^{i-r} a_{k,i}\theta_{i-r}, & r = 1, \cdots, k-1 \\ \frac{1}{k!}, & r = k. \end{cases}$$

Alternatively, we can say that for $r = 2, \cdots, k$,

$$a_{k+1,k+1-r} = \frac{a_{k,k-r}}{k+1-r} + \sum_{i=k+1-r}^{k-1} \binom{i}{k+1-r} (-1)^{i-(k+1-r)} a_{k,i}\theta_{i-(k+1-r)}.$$

Theorem 5.5. *For all defined constants, as $y \to \infty$*

$$1 - \Lambda^{*(k+1)}(y) = \Bigg\{ \sum_{j=0}^{k-1} \left[\sum_{i=j}^{k-1} a_{k,i} \binom{i}{j} (-1)^{i-j} \theta_{i-j} \right] y^j + \sum_{i=0}^{k-1} \left[\frac{a_{k,i}}{(i+1)2^{i+1}} \right] y^{i+1} + a_{k,0}$$
$$+ \sum_{i=0}^{k-1} \left[\sum_{j=0}^{i} \left(\frac{a_{k,i}}{i-j+1} \right) \binom{i}{j} (-1)^{i-j} \left(\frac{1}{2} \right)^{i-j+1} \right] y^{i+1}$$
$$+ \sum_{j=0}^{k-1} \left[\sum_{m=1}^{k-j} \binom{k+1}{m} \sum_{i=j}^{k-m} a_{k+1-m,i} \binom{i}{j} (-1)^{i-j} \mu_{i-j,m} \right] y^j \Bigg\} e^{-y} + o(e^{-\beta y}).$$

Table C.1 provides values for θ_k and ζ_k, $k = 0, \cdots, 16$, and note that the values of θ_k are fairly close to $-k!$. Table C.2 shows numerical values for $\mu_{i,m}, 0 \le i \le m \le 7$. Appendix C.2 also shows explicit expressions for these $\mu_{i,m}$ written in terms of the ζ_k. Table C.3 gives values for the positive half-line constants $a_{k+1,i}$.

Finally, we need to distinguish between the values of $a_{k+1,i}$ and those belonging to the expansion over all reals. Call the latter $A_{k+1,i}$. Then before turning to examples in the next section, we need to have recursive formulas for the $A_{k+1,i}$ like we did the $a_{k+1,i}$. Using the same technique of picking off the terms with the y^j of interest, we discover the two equations

$$\begin{aligned} A_{k+1,k+1-r} &= \sum_{m=1}^{r-1} \binom{k+1}{m} \sum_{i=k+1-r}^{k-m} a_{k+1-m,i} \binom{i}{k+1-r} (-1)^{i-(k+1-r)} \mu_{i-(k+1-r),m} \\ &\quad + a_{k+1,k+1-r}, \qquad r = 2, \cdots, k+1, \qquad \text{and} \\ A_{k+1,k} &= \frac{1}{k!}. \end{aligned} \tag{5.80}$$

The following alternative theorem provides the same result, but stated more succinctly.

Theorem 5.6. *For all constants defined earlier and for $A_{k+1,i}$ defined in (5.80), as $y \to \infty$*

$$1 - \Lambda^{*(k+1)}(y) = \left\{ \sum_{i=0}^{k} A_{k+1,i} y^i \right\} e^{-y} + o(e^{-\beta y}).$$

Table C.4 lists numerical values of the $A_{k+1,i}$.

5.6 Examples of the N-Fold Gumbel Expansion

Now that the theory has been established, we provide several examples of how the expansion is implemented. Some of the examples derive more explicit formulas for the $A_{k+1,i}$, which in turn will be used in another simulation study.

Example 5.1. Corollary 5.2 is a special case of Theorem 5.6, taking $k = 1$ and noting that $\theta_0 + \zeta_0 = -\gamma$.

Example 5.2. We derive $A_{k+1,k-1}$, the general secondary term in the expansion, for $k \geq 1$. Of course, we already know the answer from Corollary 5.2, but in this example we obtain it from (5.80). First note that

$$\begin{aligned} a_{k+1,k-1} &= \frac{a_{k,k-2}}{k-1} + a_{k,k-1}\theta_0 = \frac{1}{k-1}\left[\frac{a_{k-1,k-3}}{k-2} + \frac{\theta_0}{(k-2)!}\right] + \frac{\theta_0}{(k-1)!} \\ &= \frac{a_{k-1,k-3}}{(k-1)(k-2)} + \frac{2\theta_0}{(k-1)!} = \cdots = \frac{a_{2,0} + (k-1)\theta_0}{(k-1)!}. \end{aligned}$$

Using the fact that $a_{2,0} = 1 + 2\theta_0$,

$$a_{k+1,k-1} = \frac{1 + (k+1)\theta_0}{(k-1)!}.$$

Take $r = 2$ in (5.80), and therefore $m = 1$ and $i = k - 1$. From Appendix C.2 we have that

$\mu_{0,1} = \zeta_0$. Thus

$$\begin{aligned} A_{k+1,k-1} &= a_{k+1,k-1} + \binom{k+1}{1} a_{k,k-1}\mu_{0,1} = \frac{1+(k+1)\theta_0+(k+1)\zeta_0}{(k-1)!} \\ &= \frac{1+(k+1)(\theta_0+\zeta_0)}{(k-1)!} = \frac{1-(k+1)\gamma}{(k-1)!}. \end{aligned} \tag{5.81}$$

Example 5.3. We next derive the three-term expansion for the case $k = 2$, or the sum of 3 standard Gumbel random variables. We do this using Theorem 5.5. The formula reduces to

$$\begin{aligned} 1-\Lambda^{*3}(y) &\approx \Bigg\{ \sum_{j=0}^{1} \left[\sum_{i=j}^{1} a_{2,i} \binom{i}{j} (-1)^{i-j} \theta_{i-j} \right] y^j + \sum_{i=0}^{1} \left[\frac{a_{2,i}}{(i+1)2^{i+1}} \right] y^{i+1} + a_{2,0} \\ &\quad + \sum_{i=0}^{1} \left[\sum_{j=0}^{i} \left(\frac{a_{2,i}}{i-j+1} \right) \binom{i}{j} (-1)^{i-j} \left(\frac{1}{2}\right)^{i-j+1} \right] y^{i+1} \\ &\quad + \sum_{j=0}^{1} \left[\sum_{m=1}^{2-j} \binom{3}{m} \sum_{i=j}^{2-m} a_{3-m,i} \binom{i}{j} (-1)^{i-j} \mu_{i-j,m} \right] y^j \Bigg\} e^{-y} \\ &= \Bigg\{ \left[\frac{a_{2,1}}{2} \right] y^2 + [a_{2,1}\theta_0 + a_{2,0} + 3a_{2,1}\mu_{0,1}] y \\ &\qquad + a_{2,0}\theta_0 - a_{2,1}\theta_1 + a_{2,0} + 3a_{2,0}\mu_{0,1} - 3a_{2,1}\mu_{1,1} + 3a_{1,0}\mu_{0,2} \Bigg\} e^{-y}. \end{aligned}$$

Using the fact that $a_{2,1} = a_{1,0} = 1$, $a_{2,0} = 1 + 2\theta_0$, $\mu_{0,1} = \zeta_0$, $\mu_{0,2} = \zeta_0^2$, and $\mu_{1,1} = \zeta_1$,

$$\begin{aligned} 1-\Lambda^{*3}(y) &= \Bigg\{ \frac{1}{2} y^2 + [1 + 3(\theta_0 + \zeta_0)] y \\ &\qquad + (1+2\theta_0)\theta_0 - \theta_1 + 1 + 2\theta_0 + 3(1+2\theta_0)\zeta_0 - 3\zeta_1 + 3\zeta_0^2 \Bigg\} e^{-y} \\ &= \left\{ \frac{1}{2} y^2 + [1-3\gamma] y + 1 - 3\gamma + 2\theta_0^2 - \theta_1 + 6\theta_0\zeta_0 - 3\zeta_1 + 3\zeta_0^2 \right\} e^{-y} \\ &\approx \left\{ \frac{1}{2} y^2 - 0.73165 y + 0.81806 \right\} e^{-y}. \end{aligned} \tag{5.82}$$

This is the tertiary expansion for the particular case of 3 Gumbel variables.

Example 5.4. Now we find the general third term in the expansion, assuming that $k \geq 2$.

The benefit of doing the previous example first is that we can use it to check our third term. Unfortunately, unlike the general second term derivation, this one is a lot more involved. To start,

$$\begin{aligned} a_{k+1,k-2} &= \frac{a_{k,k-3}}{k-2} + \sum_{i=k-2}^{k-1} \binom{i}{k-2} (-1)^{i-(k-2)} a_{k,i} \theta_{i-(k-2)} \\ &= \frac{a_{k,k-3}}{k-2} + a_{k,k-2}\theta_0 - (k-1)a_{k,k-1}\theta_1 = \frac{a_{k,k-3}}{k-2} + \frac{\theta_0 + k\theta_0^2 - \theta_1}{(k-2)!}. \end{aligned}$$

We need to go at least a couple of steps further into the recursion before we can spot the pattern. Assume that k is large enough so that the following steps may be performed:

$$\begin{aligned} a_{k+1,k-2} &= \frac{a_{k-1,k-4}}{(k-2)(k-3)} + \frac{2(\theta_0 - \theta_1) + (2k-1)\theta_0^2}{(k-2)!} \\ &= \frac{a_{k-2,k-5}}{(k-2)(k-3)(k-4)} + \frac{3(\theta_0 - \theta_1) + (3k-3)\theta_0^2}{(k-2)!} \\ &= \frac{a_{k-3,k-6}}{(k-2)(k-3)(k-4)(k-5)} + \frac{4(\theta_0 - \theta_1) + (4k-6)\theta_0^2}{(k-2)!}. \end{aligned}$$

After some pattern recognition (in particular, the 1, 3, and 6 are triangular numbers), we have that

$$a_{k+1,k-2} = \frac{a_{3,0} + (k-2)(\theta_0 - \theta_1) + \left[(k-2)k - \frac{1}{2}(k-2)(k-3)\right]\theta_0^2}{(k-2)!}.$$

Using (5.64), it can be shown that $a_{3,0} = 1 + 3\theta_0 + 2\theta_0^2 - \theta_1$, and therefore

$$a_{k+1,k-2} = \frac{1 + (k+1)\theta_0 - (k-1)\theta_1 + \frac{1}{2}(k-1)(k+2)\theta_0^2}{(k-2)!}. \tag{5.83}$$

As a quick check, setting $k = 2$ in this equation returns $a_{3,0} = 1 + 3\theta_0 + 2\theta_0^2 - \theta_1$. Turning to $A_{k+1,k-2}$, set $r = 3$ in (5.80) to obtain

$$\begin{aligned} A_{k+1,k-2} &= a_{k+1,k-2} + \sum_{m=1}^{2} \binom{k+1}{m} \sum_{i=k-2}^{k-m} a_{k+1-m,i} \binom{i}{k-2} (-1)^{i-(k-2)} \mu_{i-(k-2),m} \\ &= a_{k+1,k-2} + (k+1) a_{k,k-2} \mu_{0,1} - (k+1) \binom{k-1}{k-2} a_{k,k-1} \mu_{1,1} + \binom{k+1}{2} a_{k-1,k-2} \mu_{0,2} \\ &= a_{k+1,k-2} + \frac{(k+1)(1+k\theta_0)\zeta_0}{(k-2)!} - \frac{(k+1)(k-1)\zeta_1}{(k-1)!} + \frac{k(k+1)\zeta_0^2}{2(k-2)!}. \end{aligned}$$

Substituting in (5.83) and simplifying, we finally have that $A_{k+1,k-2}$ is equal to

$$\frac{1 + (k+1)(k\theta_0\zeta_0 - \zeta_1 - \gamma) - (k-1)\theta_1 + \frac{1}{2}(k-1)(k+2)\theta_0^2 + \frac{1}{2}k(k+1)\zeta_0^2}{(k-2)!}. \tag{5.84}$$

As a check, setting $k = 2$ provides $A_{3,0} = 1 - 3\gamma + 6\theta_0\zeta_0 - 3\zeta_1 - \theta_1 + 2\theta_0^2 + 3\zeta_0^2$, which agrees with our constant term in (5.82).

Example 5.5. We have derived the general forms for the first, second, and third-order terms. Based on our extensive analysis for the latter, it should be no surprise that explicit forms for higher terms, while possible to derive, are very complicated. If one needed further terms in the expansion for k large enough, we recommend the computational values given in Table 5.3, duplicated in Appendix C.4. In order to make the table more user-friendly, we reindex as $A_{n,i}$ where n is the number of variables in the convolution.

Table 5.3: Values for $A_{n,i}$

n \ i	0	1	2	3	4	5	6
1	1.0000	—	—	—	—	—	—
2	-0.1544	1.0000	—	—	—	—	—
3	0.8181	-0.7316	0.5000	—	—	—	—
4	-1.7642	2.2294	-0.6544	0.1667	—	—	—
5	4.3381	-5.2887	1.9870	-0.3143	0.0417	—	—
6	-12.2325	14.1488	-5.1955	1.0086	-0.1026	0.0083	—
7	34.5721	-40.0139	14.7598	-2.8773	0.3526	-0.0253	0.0014

Now we list the expansions for the convolution of $2, \cdots, 7$ Gumbel random variables, based on the results of Table 5.3:

$$
\begin{aligned}
\overline{\Lambda^{*2}}(y) &\approx (y - 0.1544)\, e^{-y} \\
\overline{\Lambda^{*3}}(y) &\approx \left(\frac{1}{2!}y^2 - 0.7316y + 0.8181\right) e^{-y} \\
\overline{\Lambda^{*4}}(y) &\approx \left(\frac{1}{3!}y^3 - 0.6544y^2 + 2.2294y - 1.7642\right) e^{-y} \\
\overline{\Lambda^{*5}}(y) &\approx \left(\frac{1}{4!}y^4 - 0.3143y^3 + 1.9870y^2 - 5.2887y + 4.3381\right) e^{-y} \\
\overline{\Lambda^{*6}}(y) &\approx \left(\frac{1}{5!}y^5 - 0.1026y^4 + 1.0086y^3 - 5.1955y^2 + 14.1488y - 12.2325\right) e^{-y} \\
\overline{\Lambda^{*7}}(y) &\approx \left(\frac{1}{6!}y^6 - 0.0253y^5 + 0.3526y^4 - 2.8773y^3 + 14.7598y^2\right. \\
&\qquad \left. -40.0139y + 34.5721\right) e^{-y}
\end{aligned}
$$

We next state a corollary that establishes the general three-term expansion.

Corollary 5.3. *For $n \geq 3$, a general tertiary expansion as $y \to \infty$ is given by*

$$
1 - \Lambda^{*n}(y) = \left(\frac{1}{(n-1)!}y^{n-1} + \frac{1 - n\gamma}{(n-2)!}y^{n-2} + A_{n.n-3}y^{n-3}\right) e^{-y} + o(e^{-\beta y}),
$$

where

$$A_{n,n-3} = \frac{1 + n[(n-1)\theta_0\zeta_0 - \zeta_1 - \gamma] - (n-2)\theta_1 + \frac{1}{2}(n-2)(n+1)\theta_0^2 + \frac{1}{2}n(n-1)\zeta_0^2}{(n-3)!}.$$

5.7 Simulation Results

The previous example shows that we can computationally derive the complete expansion for a given convolution. That is, we can compute as many terms in the expansion as there are variables to add. But the question is, should we? It may be possible to get a reasonable approximation without resorting to the full expansion. The goal of this next section is to conduct a simulation to see whether the full expansion is needed, or if, say, three or four terms is sufficient enough.

While including more terms may result in a more accurate approximation, there are a couple of reasons we may wish not to do so. The formulas get complicated, and one would need to keep careful track of not only the constants in the formula, but also constants for all previous expansions. If we convolve n Gumbel variables, we would need to store $\binom{n+1}{2}$ total values. And second, the θ_k, ζ_k, and $\mu_{i,m}$ are tough to compute for large k. A better alternative would be to use Corollary 5.3, which would eliminate the necessity of having to compute a large number of constants in advance.

In the following simulation, we focus on the convolution of $n = 3, \cdots, 7$ Gumbel random variables and check how the approximation behaves for $L = 3, \cdots, n$ terms in each expansion. Note that Figures 5.4 through 5.8 are best viewed in color.

Figure 5.4: Gumbel Convolution, n = 3

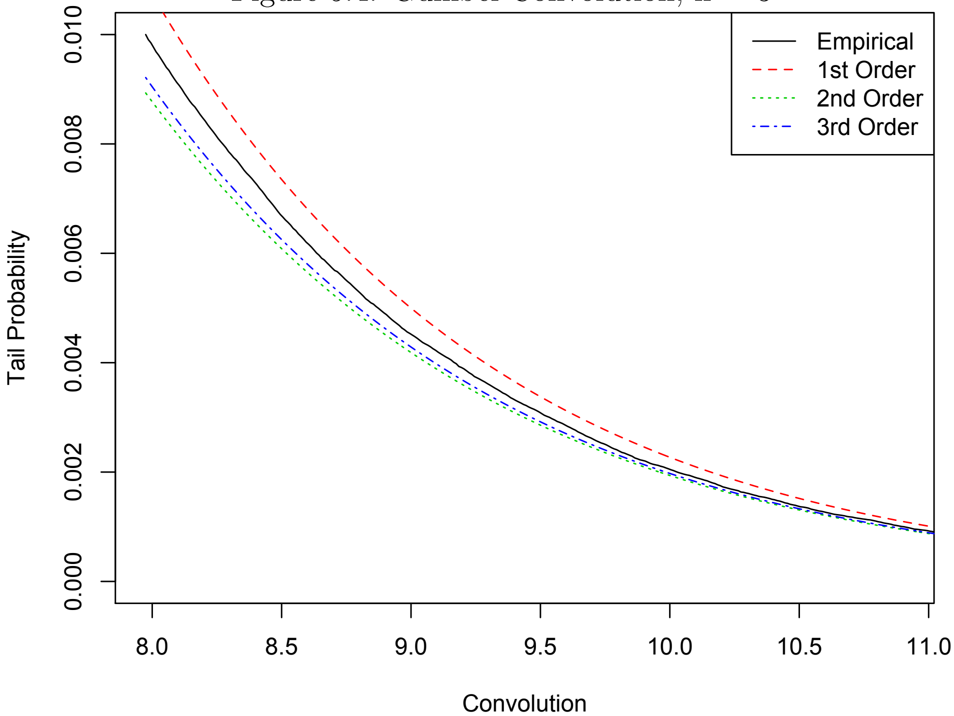


Figure 5.5: Gumbel Convolution, n = 4

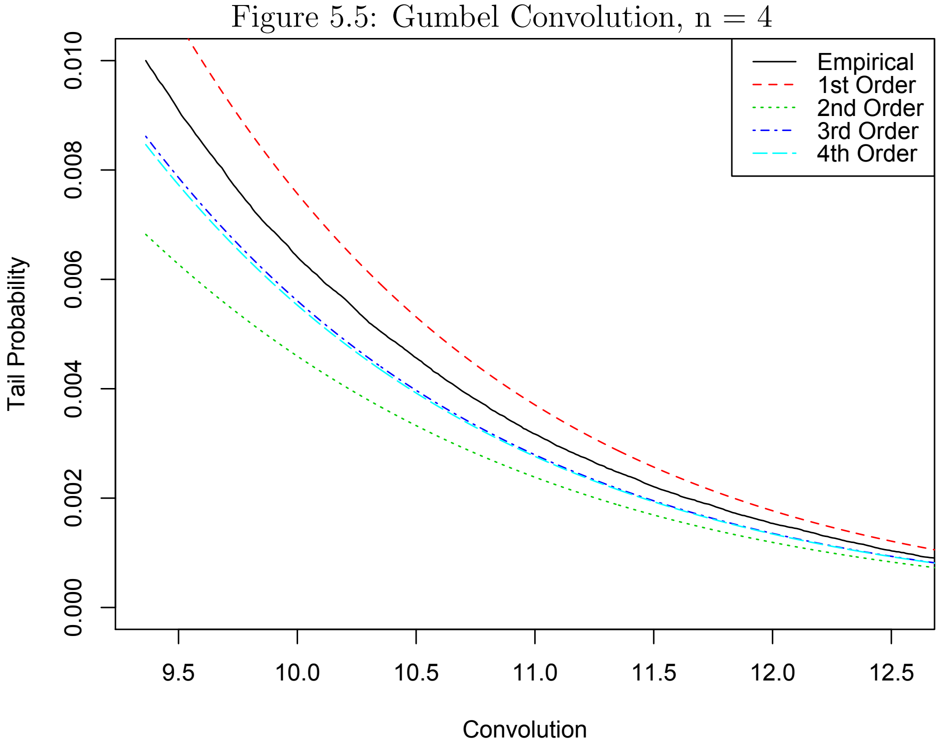


Figure 5.6: Gumbel Convolution, n = 5

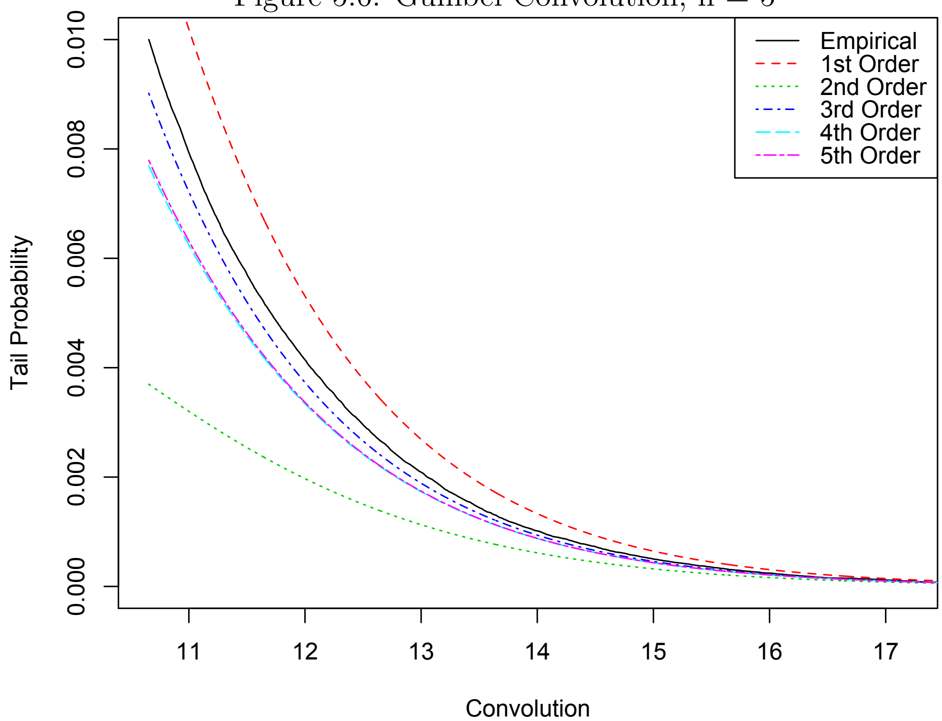


Figure 5.7: Gumbel Convolution, n = 6

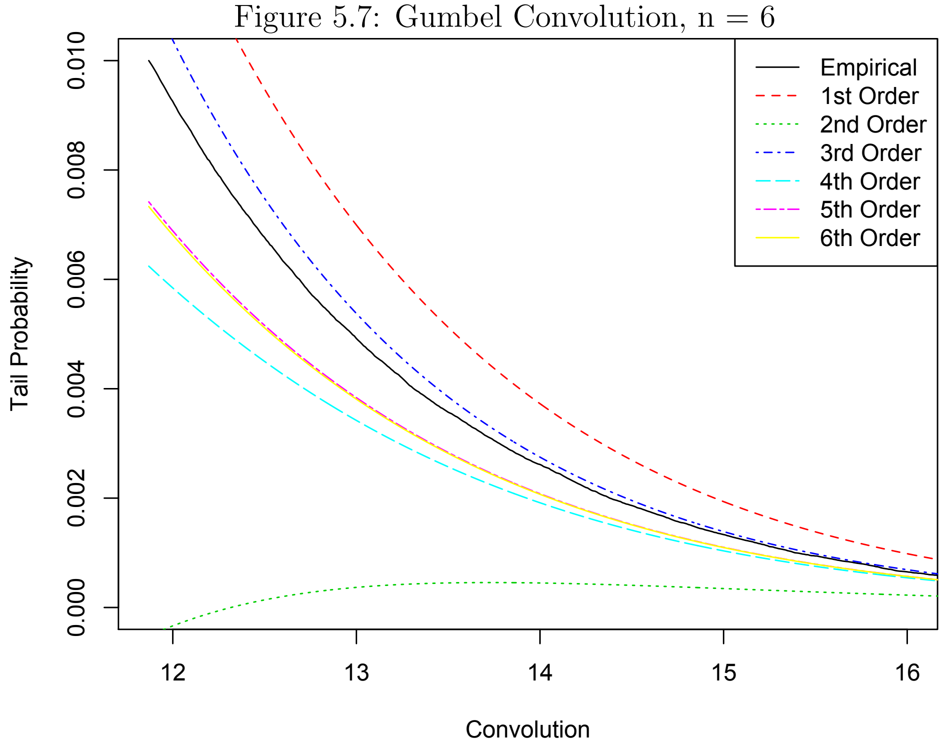
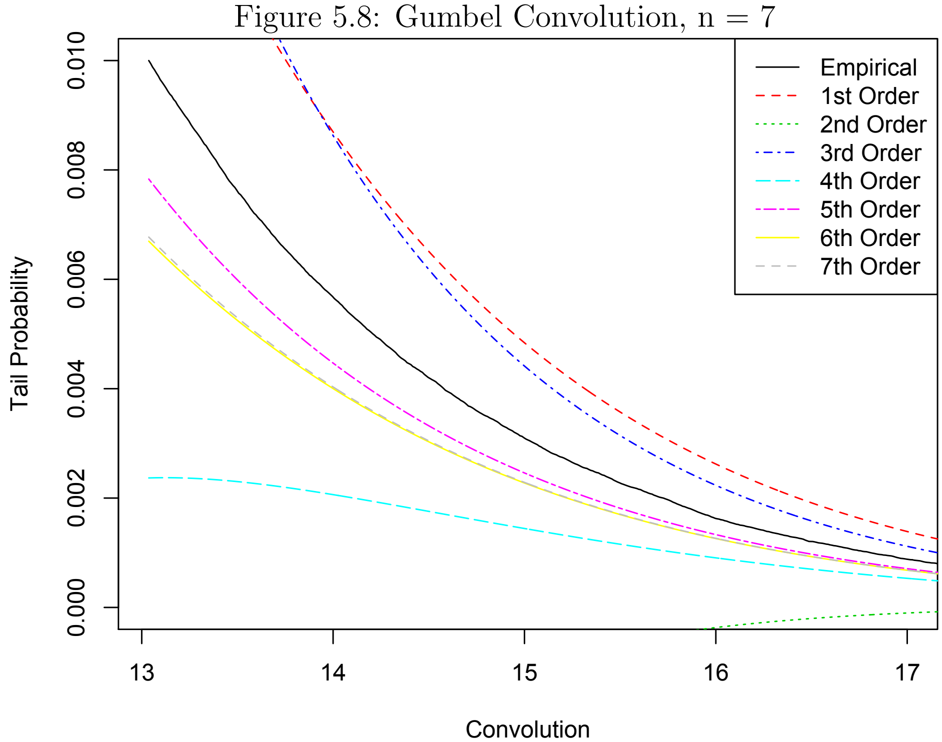


Figure 5.8: Gumbel Convolution, n = 7



Tables 5.4 and 5.5 contain the errors and relative errors in the probability approximations, given n and each approximation. We highlight the smallest error and also the optimal expansion in both tables. Examining these two tables plus the five figures, we make several observations. First, for $n = 3, \cdots, 6$ the third-order approximation is the most accurate, and adding any further terms actually is detrimental to the prediction. It is tempting, therefore, to conclude that Corollary 5.3 is the best approximation to use for $n \geq 3$. Unfortunately the $n = 7$ case suggests otherwise, for here a quinary expansion is best. These results show that in practice, it may be difficult to recommend a specific number of terms to include in the approximation.

Second, we see that the best formula to use sometimes underestimates ($n = 3, 4, 5, 7$) and sometimes overestimates ($n = 6$). Also note that for $n \geq 4$ the second-order approximation performs poorly and gets worse as n increases. A similar observation applies for the fourth-order formula. We therefore conjecture that whatever the ideal number of terms to include, choosing an odd number is optimal. Unfortunately in general, it is not clear as to when the formula over or underestimates the true probability. What does seem apparent is that the third-order approximation is always better than the first and second-orders.

Table 5.4: Errors in Approximations of Probabilities

n	**Approx**	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
2	1st	-0.0015	-0.0007	-0.0005	-0.0003	-0.0001	-2.2e-5	2.3e-6
	2nd	0.0003	7.6e-5	3.3e-5	-6.8e-6	-3.0e-5	-4.6e-6	3.6e-6
3	1st	-0.0034	-0.0021	-0.0017	-0.0009	-0.0005	-8.8e-5	-2.4e-6
	2nd	0.0103	0.0038	0.0028	0.0011	0.0004	5.8e-5	8.5e-6
	3rd	0.0076	0.0028	0.0021	0.0008	0.0003	4.3e-5	7.6e-6
4	1st	-0.0056	-0.0038	-0.0032	-0.0018	-0.0009	-0.0002	-1.7e-5
	2nd	0.0261	0.0104	0.0077	0.0032	0.0013	0.0002	1.2e-5
	3rd	0.0104	0.0043	0.0033	0.0013	0.0006	7.8e-5	6.2e-6
	4th	0.0122	0.0049	0.0037	0.0015	0.0006	8.4e-5	6.6e-6
5	1st	-0.0081	-0.0057	-0.0049	-0.0027	-0.0014	-0.0003	-3.4e-5
	2nd	0.0469	0.0196	0.0148	0.0063	0.0027	0.0004	2.5e-5
	3rd	0.0033	0.0022	0.0018	0.0009	0.0005	7.9e-5	3.5e-6
	4th	0.0178	0.0072	0.0054	0.0023	0.0010	0.0001	6.9e-6
	5th	0.0163	0.0068	0.0051	0.0022	0.0009	0.0001	6.8e-6
6	1st	-0.0107	-0.0077	-0.0066	-0.0038	-0.0021	-0.0004	-4.7e-5
	2nd	0.0724	0.0315	0.0241	0.0105	0.0046	0.0007	5.0e-5
	3rd	-0.0183	-0.0060	-0.0042	-0.0014	-0.0005	-2.0e-5	-8.9e-7
	4th	0.0328	0.0125	0.0093	0.0037	0.0015	0.0002	1.3e-5
	5th	0.0171	0.0075	0.0058	0.0025	0.0011	0.0002	1.1e-5
	6th	0.0186	0.0080	0.0061	0.0026	0.0011	0.0002	1.1e-5
7	1st	-0.0131	-0.0010	-0.0084	-0.0050	-0.0027	-0.0006	-6.2e-5
	2nd	0.1021	0.0461	0.0356	0.0160	0.0072	0.0011	8.4e-5
	3rd	-0.0583	-0.0223	-0.0164	-0.0065	-0.0025	-0.0003	-1.7e-5
	4th	0.0726	0.0268	0.0196	0.0076	0.0031	0.0004	2.4e-5
	5th	0.0054	0.0046	0.0039	0.0021	0.0010	0.0002	1.4e-5
	6th	0.0237	0.0099	0.0075	0.0032	0.0014	0.0002	1.5e-5
	7th	0.0221	0.0095	0.0073	0.0032	0.0014	0.0002	1.5e-5

Table 5.5: Relative Errors in Approximations of Probabilities

n	Approx	95%	97.5%	98%	99%	99.5%	99.9%	99.99%
2	1st	-0.0285	-0.0260	-0.0259	-0.0246	-0.0271	-0.0215	0.0234
	2nd	0.0064	0.0030	0.0017	-0.0007	-0.0059	-0.0046	0.0371
3	1st	-0.0632	-0.0785	-0.0796	-0.0834	-0.0856	-0.0807	-0.0234
	2nd	0.2583	0.1777	0.1619	0.1224	0.0947	0.0617	0.0931
	3rd	0.1793	0.1257	0.1154	0.0882	0.0682	0.0451	0.0826
4	1st	-0.1012	-0.1317	-0.1375	-0.1527	-0.1570	-0.1633	-0.1446
	2nd	1.0915	0.7071	0.6311	0.4601	0.3587	0.2191	0.1461
	3rd	0.2626	0.2091	0.1955	0.1557	0.1309	0.0843	0.0663
	4th	0.3229	0.2451	0.2266	0.1765	0.1456	0.0919	0.0701
5	1st	-0.1398	-0.1847	-0.1954	-0.2150	-0.2231	-0.2354	-0.2517
	2nd	15.1175	3.5973	2.8501	1.6946	1.1842	0.6547	0.3392
	3rd	0.0697	0.0953	0.0974	0.1024	0.1043	0.0853	0.0366
	4th	0.5543	0.4062	0.3709	0.2935	0.2442	0.1611	0.0744
	5th	0.4851	0.3714	0.3421	0.2764	0.2332	0.1564	0.0725
6	1st	-0.1763	-0.2357	-0.2490	-0.2770	-0.2970	-0.3039	-0.3185
	2nd	-3.2363	-4.8745	-5.9013	-19.7230	12.8988	2.4372	1.0015
	3rd	-0.2682	-0.1924	-0.1728	-0.1205	-0.0868	-0.0195	-0.0088
	4th	1.9096	1.0003	0.8608	0.5937	0.4395	0.2815	0.1488
	5th	0.5204	0.4321	0.4062	0.3401	0.2821	0.2148	0.1224
	6th	0.5935	0.4671	0.4345	0.3558	0.2915	0.2183	0.1236
7	1st	-0.2080	-0.2810	-0.2968	-0.3335	-0.3538	-0.3711	-0.3841
	2nd	-1.9603	-2.1868	-2.2811	-2.6607	-3.2717	-8.2609	5.4399
	3rd	-0.5383	-0.4718	-0.4511	-0.3922	-0.3366	-0.2275	-0.1440
	4th	-3.2119	-14.9082	52.1054	3.2271	1.5744	0.6712	0.3184
	5th	0.1222	0.2263	0.2436	0.2631	0.2617	0.2329	0.1592
	6th	0.8979	0.6564	0.6042	0.4791	0.3977	0.2874	0.1783
	7th	0.7909	0.6134	0.5708	0.4625	0.3886	0.2845	0.1775

5.8 Application to the Peachtree Creek Data

We close this chapter with an application of the Gumbel expansion to the Peachtree Creek data set from Chapter 2, followed by some open questions. Recall that a Gumbel distribution was the most appropriate fit for the seasons summer and winter, and furthermore we concluded that winter was stationary in the location and scale. For that reason, we focus on winter. We want to study the distribution of the sum of the maximum observed streamflow (in cubic feet per second) over, say, $n = 5$ years. This question would be of interest to hydrologists since it may help track trends in the river. First note how Theorem 5.6 will be used. If $X_i = \sigma Z_i + \mu$ where $Z_i \sim \Lambda$, then for large enough y

$$P\left(\sum_{t=1}^{n} X_i > y\right) = P\left(\sum_{t=1}^{n}(\sigma Z_i + \mu) > y\right) = P\left(\sum_{t=1}^{n} Z_i > \frac{y - n\mu}{\sigma}\right)$$
$$\approx \left\{\sum_{i=0}^{n-1} A_{n,n-1-i}\left(\frac{y-n\mu}{\sigma}\right)^{n-1-i}\right\}\exp\left[-\left(\frac{y-n\mu}{\sigma}\right)\right].$$

For the purposes of this example, we first refit the Gumbel distribution to the 52 winter observations without time trend in the location and scale parameters. That is, assuming that $X_1, \cdots, X_{52} \overset{iid}{\sim} \text{GEV}(0, \mu, \sigma)$, we find the new maximum likelihood estimates for μ and σ. The original values with time trend were 1687.4669 and 1106.9248, respectively, from Table 2.4, so we choose these estimates as starting values. The log likelihood to be maximized is

$$\text{lnL} = -\frac{52(\overline{x} - \mu)}{\sigma} - \sum_{t=1}^{52} \exp\left[-\left(\frac{x_t - \mu}{\sigma}\right)\right].$$

Table 5.6 summarizes the maximum likelihood estimates, both of which are not too far away from their nonstationary counterparts in Table 2.4.

Table 5.6: Estimated Stationary Parameters for Winter

Param.	**Estimate**	**SE**	**90% Confidence**
μ	1438.8241	118.2219	(1240.6955,1636.9527)
σ	810.8379	91.4888	(657.5114, 964.1644)

Using the fact that a tertiary expansion is optimal for $n = 5$ variables,

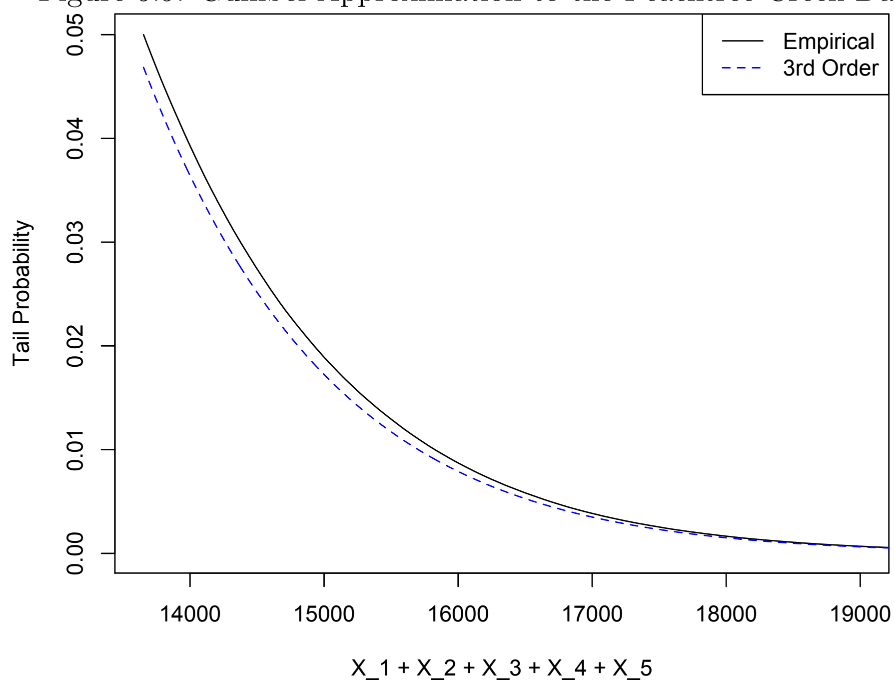
$$\overline{\Lambda^{*5}}\left(\frac{y-5\hat{\mu}}{\hat{\sigma}}\right) \approx \left\{A_{5,4}\left(\frac{y-5\hat{\mu}}{\hat{\sigma}}\right)^4 + A_{5,3}\left(\frac{y-5\hat{\mu}}{\hat{\sigma}}\right)^3 + A_{5,2}\left(\frac{y-5\hat{\mu}}{\hat{\sigma}}\right)^2\right\} e^{5\hat{\mu}/\hat{\sigma}} e^{-y/\hat{\sigma}}$$
$$= \left\{297.1979\left(\frac{y-5\hat{\mu}}{\hat{\sigma}}\right)^4 - 2242.121\left(\frac{y-5\hat{\mu}}{\hat{\sigma}}\right)^3 + 14172.78\left(\frac{y-5\hat{\mu}}{\hat{\sigma}}\right)^2\right\} e^{-y/\hat{\sigma}},$$

and when carefully foiled out and simplified the approximation is

$$\overline{\Lambda^{*5}}\left(\frac{y-5\hat{\mu}}{\hat{\sigma}}\right) \approx \left[6.9 \times 10^{-10} y^4 - 2.4 \times 10^{-5} y^3 + 0.3258 y^2 \right.$$
$$\left. -1987.2030y + 4523387\right] e^{-y/\hat{\sigma}}.$$

In Figure 5.9, we compare our approximation to the empirical distribution of the sum of 5 random variables from the GEV$(0, \hat{\mu}, \hat{\sigma})$. Although it underestimates, we can see that our approximation is quite accurate even for the 95th percentile, and certainly for the 99th and above. As an example, consider the 99th percentile of 15826.38. The correct probability is 0.01, and our formula predicts it to be 0.0091.

Figure 5.9: Gumbel Approximation to the Peachtree Creek Data



5.9 Open Questions

Theorem 5.6 is a significant contribution to extreme value theory. There are already established results on the convolution of random variables of common distributions, such as normal, exponential, gamma, Poisson, etc. Rootzén (1986) derived the first-order expansion of a more general class of distributions, from which one could specialize to the Gumbel. Our result establishes the n-term expansion for n Gumbel random variables. We also have a second term for a general class of distributions from Theorem 5.3.

There are, however, three questions that would be useful to investigate for future research, one theoretical and two computational topics. The first is to find the n-term expansion for a broader class of variables, of which the Gumbel would be a special case. That approximation would ideally have similar conditions as to those in Section 5.2 and in Rootzén (1986), plus additional assumptions as needed. As an example, we conjecture that the nth moment must be finite, among other stipulations.

The second is to find the optimal number of terms to include in the expansion, which would be useful for application purposes. We have developed a conjecture for the answer based on the simulation results, and it is easier to explain using an example. Consider the full expansion for 7 variables:

$$\overline{\Lambda^{*7}}(y) \approx \left(\frac{1}{6!}y^6 - 0.0253y^5 + 0.3526y^4 - 2.8773y^3 + 14.7598y^2 \right.$$
$$\left. -40.0139y + 34.5721\right) e^{-y}.$$

It has been shown that choosing the quinary expansion provides the best approximation. Observe that the coefficients, in absolute value, are increasing up to the sixth term, after which they begin to get smaller. Next, we look at the full 6-variable expansion:

$$\overline{\Lambda^{*6}}(y) \approx \left(\frac{1}{5!}y^5 - 0.1026y^4 + 1.0086y^3 - 5.1955y^2 + 14.1488y - 12.2325\right) e^{-y}$$

Here we should choose the tertiary expansion. In absolute value, the coefficients are increasing up to the fifth term. We therefore suspect a connection between the optimal number of terms and the turning point of the $|A_{n,n-r}|$. Plus recall that the optimal number was always odd. The conjecture therefore takes the following form for the convolution of $n \geq 5$ Gumbel random variables. (It does not work when $n < 5$, but those cases are easier to derive and inspect.)

1. The optimal choice of approximation should include an odd number of terms, starting with y^{n-1}.
2. Locate the r that provides $\sup\limits_{1 \leq r \leq n} |A_{n,n-r}|$.

3. If r is even, then the optimal approximation contains the first $(r-1)$ terms.

4. If r is odd, then the optimal approximation contains the first $(r-2)$ terms.

The third open topic concerns ways of computing the $A_{n,i}$ for higher values of n, namely $n \geq 8$. One motivation for this exploration is that a greater number of random variables may need to be added together for application purposes. We have covered cases such as $n = 4$ (weekly for a month, or quarterly) and $n = 7$ (daily for a week). However, it would be useful to have the expansions for, say, $n = 12$ (monthly for a year) or $n = 52$ (weekly for a year).

The main challenge is deriving the necessary θ_k and ζ_k constants, and then later the $\mu_{i,m}$. While we have the formulas, when k grows large these quantities become computationally intense to find. In fact, it was this handicap that motivated truncating the full expansion to just three or five terms. However, we now have an easier way of computing θ_k, and it involves a Taylor series expansion:

$$\begin{aligned}\theta_k &= -\int_0^\infty x^k \left[1 - \exp\left(-e^{-x}\right)\right] dx = -\int_0^\infty x^k \sum_{j=1}^{\infty} \frac{(-1)^{j-1} e^{-jx}}{j!} \\ &= \sum_{j=1}^{\infty} \frac{(-1)^j}{j!} \int_0^\infty x^k e^{-jx} dx = k! \sum_{j=1}^{\infty} \frac{(-1)^j}{j! j^{k+1}}.\end{aligned}$$

This alternative representation of θ_k is much easier to implement, which also explains why the values of θ_k in Table C.1 are fairly close to $-k!$. A similar series expansion could be performed on ζ_k. Thus, the third open question is to seek computational shortcuts like these to make implementation easier for higher k.

Chapter 6 Dealing With Ties In The AR(1) Process

In Chapters 3 and 4 we established the groundwork to build a two-term expansion for the AR(1) process with Gumbel innovations. The formula was constructed under the assumption that there were no ties in the coefficients, and especially in the highest and second highest. Chapter 5 discussed the approximation if all variables had exactly the same weights. The goal of this chapter is to investigate the interesting twist of having ties in the top two largest weights. There are multiple possible combinations in which this may occur, but we give the proof of only one of these. The techniques we employ may be used to carry out expansions under the other possibilities. This chapter constitutes the fourth project in the dissertation.

For ease of reference, we restate Theorem 4.3 here:

Theorem 6.1. *Let $d_k, k = 1, \cdots, n, n \geq 3$ be positive constants, and define $c_k = d_{(k)}$, the order statistics arranged from largest to smallest. That is, $c_1 > c_2 > c_3 \geq c_4 \geq \cdots \geq c_n$, and in particular $c_1 = \max(d_k)$. Assume that c_1 and c_2 have multiplicities of 1. If $0 < 2c_2 < c_1$, then as $y \to \infty$ a two-term expansion is given by*

$$P\left(\sum_{k=1}^{n} d_k Z_k > y\right) = \prod_{k=2}^{n} \Gamma\left(1 - \frac{c_k}{c_1}\right) e^{-y/c_1} - \frac{1}{2}\prod_{k=2}^{n} \Gamma\left(1 - \frac{2c_k}{c_1}\right) e^{-2y/c_1} + O\left(e^{-y/c_2}\right).$$

In the particular case where $0 < 2c_2 = c_1$, the expansion is

$$P\left(\sum_{k=1}^{n} d_k Z_k > y\right) = \prod_{k=2}^{n} \Gamma\left(1 - \frac{c_k}{c_1}\right) e^{-y/c_1} - \prod_{k=3}^{n} \Gamma\left(1 - \frac{2c_k}{c_1}\right) \frac{y}{c_1} e^{-2y/c_1} + O\left(e^{-y/c_2}\right).$$

Lastly, when $0 < c_1 < 2c_2$*, the expansion is*

$$P\left(\sum_{k=1}^{n} d_k Z_k > y\right) = -\frac{c_2}{c_1 - c_2}\Gamma\left(2 - \frac{c_1}{c_2}\right)\prod_{k=3}^{n}\Gamma\left(1 - \frac{c_k}{c_2}\right)e^{-y/c_2} \\ + \prod_{k=2}^{n}\Gamma\left(1 - \frac{c_k}{c_1}\right)e^{-y/c_1} + o\left(e^{-y/c_2}\right).$$

In this chapter, for simplicity we assume all the weights are nonnegative. Suppose the two highest constants c_1 and c_2 have multiplicities m_1 and m_2, respectively. Then for $n > m_1 + m_2$, define the series

$$Y_n \stackrel{d}{=} c_1 \sum_{k=1}^{m_1} Z_k + c_2 \sum_{k=1}^{m_2} Z_{m_1+k} + \sum_{k=3}^{n-m_1-m_2} c_k Z_{m_1+m_2+k}. \tag{6.1}$$

6.1 Necessary Lemmas for the Case where $m_1 \geq 2$ and $m_2 = 1$

There are no less than five possible scenarios we may consider for multiplicites among the two highest weights:

1. The case where $m_1 \geq 2$ and $m_2 = 1$.
2. The case where $m_1 = 1$ and $m_2 \geq 2$.
3. The case where $m_1 = m_2 \geq 2$.
4. The case where $m_1 > m_2 \geq 2$.
5. The case where $m_2 > m_1 \geq 2$.

For the purposes of illustrating how such a proof would be implemented, we choose only the first scenario as an example setting. This setting assumes that the largest weight c_1 occurs multiple times, but the second largest c_2 occurs only once. We conjecture that the remaining four cases may be worked out using similar techniques, and these are left as open questions. Therefore for the remainder of this chapter, unless otherwise noted we take $m_1 \geq 2$ and $m_2 = 1$, and (6.1) reduces to

$$Y_n \stackrel{d}{=} c_1 \sum_{k=1}^{m_1} Z_k + c_2 Z_{m_1+1} + \sum_{k=3}^{n-m_1+1} c_k Z_{m_1-1+k}. \tag{6.2}$$

Consider the probability

$$P\left(\sum_{k=1}^{m_1} Z_k + \frac{c_2}{c_1} Z_0 > \frac{y}{c_1}\right) = P(S + T > y^*), \tag{6.3}$$

where $S = \sum_{k=1}^{m_1} Z_k$, $T = \frac{c_2}{c_1} Z_0$, and $y^* = \frac{y}{c_1}$. Notice that the distribution of S is approximated from Theorem 5.6. We now examine some preliminary lemmas.

Lemma 6.1. *Recall from (5.80) that*

$$\begin{aligned} A_{m_1,m_1-r} &= \sum_{h=1}^{r-1} \binom{m_1}{h} \sum_{i=m_1-r}^{m_1-h-1} a_{m_1-h,i} \binom{i}{m_1-r} (-1)^{i-(m_1-r)} \mu_{i-(m_1-r),h} \\ &\qquad + a_{m_1,m_1-r}, \qquad r = 2, \cdots, m_1, \qquad \textit{and} \\ A_{m_1,m_1-1} &= \frac{1}{(m_1-1)!}. \end{aligned} \tag{6.4}$$

For all constants defined in Section 5.4, as $y \to \infty$

$$\overline{F}_S(y) = P(S > y) = \left\{ \sum_{i=0}^{m_1-1} A_{m_1,i} y^i \right\} e^{-y} + o(e^{-y}).$$

Proof. This is the main result from Chapter 5. Although the original had error term $o(e^{-\beta y})$ for some $\beta > 1$, for simplicity in this chapter we shall take $o(e^{-y})$ instead. □

Lemma 6.2. *For any* $\frac{c_2}{c_1} < a < 1$, *we have that* $1 - \exp\left(-e^{-ay/c_2}\right) = o\left(e^{-y/c_1}\right)$ *as* $y \to \infty$.

Proof. Observe that for y large enough,

$$1 - \exp\left(-e^{-ay/c_2}\right) = \sum_{j=1}^{\infty} \frac{(-1)^{j-1} e^{-jay/c_2}}{j!} \leq 2e^{-ay/c_2}.$$

Then we see that

$$\frac{1 - \exp\left(-e^{-ay/c_2}\right)}{e^{-y/c_1}} \leq \frac{2e^{-ay/c_2}}{e^{-y/c_1}} = 2\exp\left[-y\left(\frac{a}{c_2} - \frac{1}{c_1}\right)\right] \to 0,$$

because $\frac{a}{c_2} - \frac{1}{c_1} > 0$. □

Lemma 6.3. *For* $L = 0, 1, ..., m_1 - 1$, $\left|E\left(T^L e^T\right)\right| < \infty$. *In particular, it can be shown that*

$$E(e^T) = \Gamma\left(1 - \frac{c_2}{c_1}\right) \qquad \textit{and} \qquad E(Te^T) = \psi\left(1 - \frac{c_2}{c_1}\right).$$

Proof. First note that

$$\begin{aligned} E\left(T^L e^T\right) &= \int_{-\infty}^{\infty} t^L e^t \times \frac{c_1}{c_2} e^{-c_1 t/c_2} \exp\left(-e^{-c_1 t/c_2}\right) dt \\ &= \left(-\frac{c_2}{c_1}\right)^L \int_0^{\infty} (\log t)^L t^{-c_2/c_1} e^{-t} dt. \end{aligned} \tag{6.5}$$

Now write

$$|E\left(T^L e^T\right)| \leq \left(\frac{c_2}{c_1}\right)^L \left[\left|\int_1^{\infty} (\log t)^L t^{-c_2/c_1} e^{-t} dt\right| + \left|\int_0^1 (\log t)^L t^{-c_2/c_1} e^{-t} dt\right|\right]. \tag{6.6}$$

Consider the series $\sum_{k=1}^{\infty}(\log k)^L k^{-c_2/c_1}e^{-k}$. By the ratio test,

$$\frac{(\log(k+1))^L(k+1)^{-c_2/c_1}e^{-(k+1)}}{(\log k)^L k^{-c_2/c_1}e^{-k}} = \left(\frac{\log(k+1)}{\log k}\right)^L \left(\frac{k}{k+1}\right)^{c_2/c_1} e^{-1} \to e^{-1} < 1,$$

and therefore

$$\Big|\int_1^{\infty}(\log t)^L t^{-c_2/c_1}e^{-t}dt\Big| = K < \infty. \tag{6.7}$$

Next,

$$\Big|\int_0^{1}(\log t)^L t^{-c_2/c_1}e^{-t}dt\Big| \le \Big|\int_0^{1}(\log t)^L t^{-c_2/c_1}dt\Big|.$$

A change of variables results in

$$\left|(-1)^L\int_0^{\infty} t^L e^{-\left(1-\frac{c_2}{c_1}\right)t}dt\right| \le \int_0^{\infty} t^L e^{-\left(1-\frac{c_2}{c_1}\right)t}dt = L!\left(1-\frac{c_2}{c_1}\right)^{-(L+1)}. \tag{6.8}$$

Therefore putting (6.7) and (6.8) into (6.6),

$$|E\left(T^L e^T\right)| \le \left(\frac{c_2}{c_1}\right)^L\left[K + L!\left(1-\frac{c_2}{c_1}\right)^{-(L+1)}\right] < \infty.$$

Lastly, $E(e^T)$ follows from the moment generating function of T, while $E\left(Te^T\right)$ can be derived from the log-gamma density. □

Now we turn to the problem of solving $P\left(\sum_{i=1}^{m_1} Z_i + \frac{c_2}{c_1}Z_0 > y^*\right)$. Split into two integrals as

$$P(S+T>y^*) = \int_{ay^*}^{\infty} P(S>y^*-t)dF_T(t) + \int_{-\infty}^{ay^*} P(S>y^*-t)dF_T(t), \tag{6.9}$$

where a is chosen to satisfy $\frac{c_2}{c_1} < a < 1$. Observe that

$$\int_{ay^*}^{\infty} P(S > y^* - y) dF_T(t) \leq P(T > ay^*) = P\left(Z_0 > \frac{ay}{c_2}\right) = 1 - \exp\left(-e^{-ay/c_2}\right).$$

Therefore by Lemma 6.2,

$$\int_{ay^*}^{\infty} P(S > y^* - t) dF_T(t) = o\left(e^{-y/c_1}\right). \tag{6.10}$$

Now consider the second integral in (6.9). By Lemma 6.1, for y large enough

$$P(S > y^* - t) = \left[\sum_{q=0}^{m_1-1} A_{m_1,q(y^*-t)^q}\right] e^{-(y^*-t)} + R(y^* - t),$$

where for any $\epsilon > 0$

$$\sup_{t<ay^*} \left|R(y^* - t)e^{y^*-t}\right| < \epsilon.$$

Write the integral as

$$\begin{aligned} \int_{-\infty}^{ay^*} P(S > y^* - t) dF_T(t) &= \sum_{q=1}^{m_1-1} A_{m_1,q} \int_{-\infty}^{ay^*} (y^* - t)^q e^{-(y^*-t)} dF_T(t) \\ + \int_{-\infty}^{ay^*} &\left[\frac{\overline{F}_S(y^* - t) - \sum_{q=1}^{m_1-1} A_{m_1,q}(y^* - t)^q e^{-(y^*-t)}}{A_{m_1,0}e^{-(y^*-t)}}\right] A_{m_1,0}e^{-(y^*-t)} dF_T(t). \end{aligned} \tag{6.11}$$

We examine the first piece in (6.11). Observe that

$$\begin{aligned} &\sum_{q=1}^{m_1-1} A_{m_1,q} \int_{-\infty}^{\infty} (y^* - t)^q e^{-(y^*-t)} dF_T(t) \\ &\quad = \sum_{q=1}^{m_1-1} A_{m_1,q} \int_{\infty}^{\infty} \sum_{j=0}^{q} \binom{q}{j} (y^*)^j (-t)^{q-j} e^{-(y^*-t)} dF_T(t) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{q=1}^{m_1-1} A_{m_1,q} \sum_{j=0}^{q} \binom{q}{j} (-1)^{q-j} (y^*)^j e^{-y^*} \int_{\infty}^{\infty} t^{q-j} e^t dF_T(t) \\
&= \sum_{q=1}^{m_1-1} A_{m_1,q} \sum_{j=0}^{q} \binom{q}{j} (-1)^{q-j} E\left(T^{q-j} e^T\right) (y^*)^j e^{-y^*} \\
&= \sum_{q=1}^{m_1-1} A_{m_1,q} \sum_{j=1}^{q} \binom{q}{j} (-1)^{q-j} E\left(T^{q-j} e^T\right) (y^*)^j e^{-y^*} + \sum_{q=1}^{m_1-1} A_{m_1,q} (-1)^q E\left(T^q e^T\right) e^{-y^*} \qquad (6.12) \\
&= \sum_{j=1}^{m_1-1} \left[\sum_{q=j}^{m_1-1} A_{m_1,q} \binom{q}{j} (-1)^{q-j} E\left(T^{q-j} e^T\right) \right] (y^*)^j e^{-y^*} \\
&\qquad\qquad + \left[\sum_{q=1}^{m_1-1} A_{m_1,q} (-1)^q E\left(T^q e^T\right) \right] e^{-y^*}.
\end{aligned}
$$

By Lemma 6.3, all the expected values in (6.12) are finite. Now

$$
\begin{aligned}
&\sum_{q=1}^{m_1-1} A_{m_1,q} \int_{-\infty}^{ay^*} (y^* - t)^q e^{-(y^*-t)} dF_T(t) \\
&\quad = \sum_{j=1}^{m_1-1} \left[\sum_{q=j}^{m_1-1} A_{m_1,q} \binom{q}{j} (-1)^{q-j} E\left(T^{q-j} e^T\right) \right] (y^*)^j e^{-y^*} \qquad (6.13) \\
&\qquad + \left[\sum_{q=1}^{m_1-1} A_{m_1,q} (-1)^q E\left(T^q e^T\right) \right] e^{-y^*} - \sum_{q=1}^{m_1-1} A_{m_1,q} \int_{ay^*}^{\infty} (y^* - t)^q e^{-(y^*-t)} dF_T(t).
\end{aligned}
$$

We now analyze the final integrals in (6.13). First note that

$$
\int_{ay^*}^{\infty} (y^* - t)^q e^{-(y^*-t)} dF_T(t) \leq \sum_{j=0}^{q} \binom{q}{j} (y^*)^j e^{-y^*} \int_{ay^*}^{\infty} t^{q-j} e^t dF_T(t). \qquad (6.14)
$$

By integration by parts,

$$\int_{ay^*}^{\infty} t^{q-j} e^t dF_T(t) = (ay^*)^{q-j} e^{ay^*} \overline{F}_T(ay^*) + \int_{ay^*}^{\infty} t^{q-1-j}(t+q-j) e^t \overline{F}_T(t) dt. \tag{6.15}$$

Note that for y^* large enough,

$$\overline{F}_T(ay^*) = P\left(Z_0 > \frac{c_1 ay^*}{c_2}\right) = e^{-c_1 ay^*/c_2} + o\left(e^{-c_1 ay^*/c_2}\right),$$

and therefore

$$\begin{aligned}(ay^*)^{q-j} e^{ay^*} \overline{F}_T(ay^*) &= (ay^*)^{q-j} e^{-ay^*\left(\frac{c_1}{c_2}-1\right)} + o\left((y^*)^{q-j} e^{-ay^*\left(\frac{c_1}{c_2}-1\right)}\right) \\ &= o\left(e^{-y^*}\right) = o\left(e^{-y/c_1}\right),\end{aligned} \tag{6.16}$$

because $a\left(\frac{c_1}{c_2}-1\right) > 0$. Next, setting $P = \frac{c_1}{c_2} - 1$,

$$\begin{aligned}\int_{ay^*}^{\infty} t^{q-1-j}(t+q-j) e^t \overline{F}_T(t) dt &\leq 2\int_{ay^*}^{\infty} t^{q-1-j}(t+q-j) e^t e^{-c_1 t/c_2} dt \\ &= 2\int_{ay^*}^{\infty} t^{q-j} e^{-Pt} dt + 2(q-j)\int_{ay^*}^{\infty} t^{q-1-j} e^{-Pt} dt.\end{aligned} \tag{6.17}$$

Now $\int_{ay^*}^{\infty} e^{-Pt} dt = \frac{1}{P} e^{-Pay^*} = o\left(e^{-y^*}\right)$, and by integration by parts, it can be shown that for $L = 1, 2, ...$,

$$\int_{ay^*}^{\infty} t^L e^{-Pt} dt = \frac{1}{P} e^{-Pay^*}\left[(ay^*)^L + \sum_{i=1}^{L} \frac{(ay^*)^{L-i}}{P^i} \prod_{k=0}^{i-1}(L-k)\right] = o\left(e^{-y^*}\right). \tag{6.18}$$

Combining (6.16) and (6.18), (6.14) $= o\left(e^{-y^*}\right) = o\left(e^{-y/c_1}\right)$. Now we turn to the second piece in (6.11). Write

$$\frac{\overline{F}_S(y^*-t)-\sum_{q=1}^{m_1-1}A_{m_1,q}(y^*-t)^q e^{-(y^*-t)}}{A_{m_1,0}e^{-(y^*-t)}} = 1+\frac{R(y^*-t)e^{(y^*-t)}}{A_{m_1,0}} = 1+\epsilon(y^*-t),$$

and observe that $\sup_{t<ay^*}|\epsilon(y^*-t)| < \frac{\epsilon}{|A_{m_1,0}|}$. Now write

$$\begin{aligned}\int_{-\infty}^{ay^*}[1+\epsilon(y^*-t)]A_{m_1,0}e^{-(y^*-t)}dF_T(t) &= \int_{-\infty}^{ay^*}A_{m_1,0}e^{-(y^*-t)}dF_T(t)\\ &\quad+\int_{-\infty}^{ay^*}\epsilon(y^*-t)A_{m_1,0}e^{-(y^*-t)}dF_T(t)\\ &= (I)+(II).\end{aligned} \tag{6.19}$$

By dominated convergence, $(I) = A_{m_1,0}E(e^T)e^{-y^*}$. As for (II),

$$\begin{aligned}|(II)| &= \left|A_{m_1,0}\int_{-\infty}^{ay^*}\epsilon(y^*-t)A_{m_1,0}e^{-(y^*-t)}dF_T(t)\right|\\ &= \left|\int_{-\infty}^{ay^*}R(y^*-t)e^{y^*-t}e^{-(y^*-t)}dF_T(t)\right| < \epsilon E(e^T)e^{-y^*},\end{aligned}$$

and therefore $(II) = o\left(e^{-y^*}\right) = o\left(e^{-y/c_1}\right)$. Putting all the pieces together and simplifying, we arrive at Lemma 6.4.

Lemma 6.4. *If $Z_k, k = 0,1,\cdots,m_1, m_1 \geq 2$ are i.i.d. Gumbel random variables and $c_1 > c_2 > 0$ and $y^* = \frac{y}{c_1}$, then as $y \to \infty$*

$$P\left(\sum_{k=1}^{m_1}Z_k+\frac{c_2}{c_1}Z_0 > y^*\right) = \sum_{j=0}^{m_1-1}\left[\sum_{q=j}^{m_1-1}A_{m_1,q}\binom{q}{j}(-1)^{q-j}E\left(T^{q-j}e^T\right)\right](y^*)^j e^{-y^*} + o\left(e^{-y/c_1}\right).$$

6.2 The Expansion for when $m_1 \geq 2$ and $m_2 = 1$

Now we want to consider the random variable

$$X = \sum_{k=1}^{m_1} Z_k + \frac{c_2}{c_1} Z_{m_1+1} + \sum_{k=3}^{n-m_1+1} \frac{c_k}{c_1} Z_{m_1-1+k},$$

or namely the probability $P(V + W > y^*)$, where

$$V = \sum_{k=1}^{m_1} Z_k + \frac{c_2}{c_1} Z_{m_1+1}, \qquad W = \sum_{k=3}^{n-m_1+1} \frac{c_k}{c_1} Z_{m_1-1+k}, \qquad \text{and} \qquad y^* = \frac{y}{c_1}.$$

Luckily in this multiplicity scenario, because the Gumbel tail probability is negligible in comparison to the Gumbel convolution expansion, there is just one expansion to focus on rather than three. That is, there is no pivot point to worry about. The steps follow similar steps taken in Chapter 4. Take $\frac{c_3}{c_2} < a < 1$ and $\frac{1}{a} < \lambda < \frac{c_2}{c_3}$, then split the probability into two integrals as

$$P(V + W > y^*) = \int_{ay^*}^{\infty} P(V > y^* - w) dF_W(w) + \int_{-\infty}^{ay^*} P(V > y^* - w) dF_W(w). \tag{6.20}$$

Regarding the first integral,

$$\overline{F}_W(ay^*) \leq E\left[\exp\left(\frac{c_1 \lambda W}{c_2}\right)\right] \exp\left(-\frac{c_1 \lambda a y^*}{c_2}\right) = o\left(e^{-y/c_2}\right),$$

because $\lambda a > 1$, and therefore

$$\int_{ay^*}^{\infty} P(V > y^* - w) dF_W(w) = o\left(e^{-y/c_2}\right). \tag{6.21}$$

Now for $T = \frac{c_2}{c_1} Z_0$, define the constants

$$K_{m_1,j} = \sum_{q=j}^{m_1-1} A_{m_1,q} \binom{q}{j} (-1)^{q-j} E\left(T^{q-j} e^{T}\right), \qquad 0 \le j \le m_1 - 1. \tag{6.22}$$

If F_V represents the distribution of V, then observe that

$$\overline{F}_V(y^* - w) = \sum_{j=0}^{m_1-1} K_{m_1,j} (y^* - w)^j e^{-(y^*-w)} + R(y^* - w),$$

where for any $\epsilon > 0$

$$\sup_{w<ay^*} \left| R(y^* - w) e^{y^*-w} \right| < \epsilon.$$

We now rewrite the integral as

$$\begin{aligned}\int_{-\infty}^{ay^*} \overline{F}_V(y^* - w) dF_W(w) &= \sum_{j=1}^{m_1-1} K_{m_1,j} \int_{-\infty}^{ay^*} (y^* - w)^j e^{-(y^*-w)} dF_W(w) \\ + \int_{-\infty}^{ay^*} &\left[\frac{\overline{F}_V(y^* - w) - \sum_{j=1}^{m_1-1} K_{m_1,j} (y^* - w)^j e^{-(y^*-w)}}{K_{m_1,0} e^{-(y^*-w)}} \right] K_{m_1,0} e^{-(y^*-w)} dF_W(w).\end{aligned} \tag{6.23}$$

We examine the first piece in (6.23). Note that

$$\begin{aligned}&\sum_{j=1}^{m_1-1} K_{m_1,j} \int_{-\infty}^{\infty} \sum_{b=0}^{j} \binom{j}{b} (-1)^{j-b} (y^*)^b e^{-y^*} w^{j-b} e^{w} dF_W(w) \\ &= \sum_{j=1}^{m_1-1} \sum_{b=0}^{j} K_{m_1,j} \binom{j}{b} (-1)^{j-b} E\left(W^{j-b} e^{W}\right) (y^*)^b e^{-y^*} \\ &= \sum_{j=1}^{m_1-1} K_{m_1,j} (-1)^j E\left(W^j e^{W}\right) e^{-y^*} + \sum_{b=1}^{m_1-1} \sum_{j=b}^{m_1-1} K_{m_1,j} \binom{j}{b} (-1)^{j-b} E\left(W^{j-b} e^{W}\right) (y^*)^b e^{-y^*}.\end{aligned}$$

Therefore

$$\sum_{j=1}^{m_1-1} K_{m_1,j} \int_{-\infty}^{ay^*} (y^*-w)^j e^{-(y^*-w)} dF_W(w)$$

$$= -\sum_{j=1}^{m_1-1} K_{m_1,j} \int_{ay^*}^{\infty} (y^*-w)^j e^{-(y^*-w)} dF_W(w) + \sum_{j=1}^{m_1-1} K_{m_1,j} (-1)^j E\left(W^j e^W\right) e^{-y^*} \tag{6.24}$$

$$+ \sum_{b=1}^{m_1-1} \sum_{j=b}^{m_1-1} K_{m_1,j} \binom{j}{b} (-1)^{j-b} E\left(W^{j-b} e^W\right) (y^*)^b e^{-y^*}.$$

Using similar steps as taken in (6.14) through (6.18), the first piece in (6.24) is $o\left(e^{-y^*}\right)$. Now we turn to the second integral in (6.23). Observe that for y^* large enough,

$$\frac{\overline{F}_V(y^*-w) - \sum_{j=1}^{m_1-1} K_{m_1,j}(y^*-w)^j e^{-(y^*-w)}}{K_{m_1,0} e^{-(y^*-w)}} = 1 + \frac{1}{K_{m_1,0}} R(y^*-w) e^{y^*-w}$$

$$= 1 + \epsilon(y^*-w),$$

with $\sup_{w<ay^*} |\epsilon(y^*-w)| < \frac{\epsilon}{K_{m_1,0}}$. Thus the second integral in (6.23) can be written as

$$K_{m_1,0} \int_{-\infty}^{ay^*} [1+\epsilon(y^*-w)] e^{-(y^*-w)} dF_W(w)$$

$$= K_{m_1,0} e^{-y^*} \int_{-\infty}^{ay^*} e^w dF_W(w) + K_{m_1,0} \int_{-\infty}^{ay^*} \epsilon(y^*-w) e^{-(y^*-w)} dF_W(w) \tag{6.25}$$

$$= (I) + (II).$$

By dominated convergence, for y^* large $(I) = K_{m_1,0} E(e^W) e^{-y^*}$. As for (II),

$$|(II)| = \left| \int_{\infty}^{ay^*} R(y^*-w) e^{y^*-w} e^{-(y^*-w)} dF_W(w) \right| < \epsilon E(e^W) e^{-y^*},$$

and therefore $(II) = o\left(e^{-y^*}\right)$. Collecting all the pieces and simplifying, we finally summarize our findings.

Theorem 6.2. *Suppose $K_{m_1,j}$ and $A_{m_1,q}$ are as defined in (6.22) and (6.4), respectively, and that θ_k, ζ_k, and $\mu_{i,j}$ are as defined in (5.48), (5.49), and (5.50), respectively. Consider the series*

$$X = c_1 \sum_{k=1}^{m_1} Z_k + c_2 Z_{m_1+1} + \sum_{k=3}^{n-m_1+1} c_k Z_{m_1-1+k}.$$

Then as $y \to \infty$

$$P(X > y) = \sum_{b=0}^{m_1-1} \left[\sum_{j=b}^{m_1-1} K_{m_1,j} \binom{j}{b} (-1)^{j-b} E\left(W^{j-b} e^{W}\right) \right] \left(\frac{y}{c_1}\right)^b e^{-y/c_1} + o\left(e^{-y/c_1}\right).$$

We make a couple of remarks about Theorem 6.2. First, if c_1 and c_2 are the only occurring weights (*i.e.* there is no c_3, c_4, etc.), then the result reduces to the conclusion of Lemma 6.4. And second, whereas Theorem 5.6 had as many terms as there were Gumbel random variables, Theorem 6.2 has only m_1 terms. The sample size n appears only through the expressions $E\left(W^{j-b} e^{W}\right)$.

6.3 Examples

We now provide a few examples of Theorem 6.2, along with a simulation study.

Example 6.1. Take $m_1 = 2$ in this first example, so the largest weight occurs twice. Then

Theorem 6.2 reduces to

$$
\begin{aligned}
P(X>y) &= \sum_{b=0}^{1}\left[\sum_{j=b}^{1} K_{m_1,j}(-1)^{j-b}E\left(W^{j-b}e^{W}\right)\right]\left(\frac{y}{c_1}\right)^{b} e^{-y/c_1} + o\left(e^{-y/c_1}\right) \\
&= \sum_{j=0}^{1} K_{m_1,j}(-1)^{j}E\left(W^{j}e^{W}\right)e^{-y/c_1} + K_{m_1,1}E\left(e^{W}\right)\left(\frac{y}{c_1}\right)e^{-y/c_1} + o\left(e^{-y/c_1}\right) \\
&= \left[K_{m_1,0}E\left(e^{W}\right) - K_{m_1,1}E\left(We^{W}\right)\right]e^{-y/c_1} + K_{m_1,1}E\left(e^{W}\right)\left(\frac{y}{c_1}\right)e^{-y/c_1} \\
&\quad + o\left(e^{-y/c_1}\right).
\end{aligned}
\tag{6.26}
$$

Using Lemma 6.3 in (6.26), it can be shown that

$$
K_{m_1,0} = (1-2\gamma)\Gamma\left(1-\frac{c_2}{c_1}\right) - \Psi\left(1-\frac{c_2}{c_1}\right) \qquad \textit{and} \qquad K_{m_1,1} = \Gamma\left(1-\frac{c_2}{c_1}\right), \tag{6.27}
$$

and further we have that

$$
E(e^{W}) = \prod_{k=3}^{n-1}\Gamma\left(1-\frac{c_k}{c_1}\right) \tag{6.28}
$$

and

$$
E(We^{W}) = -\sum_{j=3}^{n-1}\frac{c_j}{c_1}\Psi\left(1-\frac{c_k}{c_1}\right)\prod_{k=3}^{n-1}\Gamma\left(1-\frac{c_k}{c_1}\right). \tag{6.29}
$$

Putting (6.27) through (6.29) into (6.26), as $y \to \infty$

$$
\begin{aligned}
P(X>y) = \Bigg\{&\left[(1-2\gamma)\Gamma\left(1-\frac{c_2}{c_1}\right) - \Psi\left(1-\frac{c_2}{c_1}\right)\right]E(e^{W}) \\
&-\Gamma\left(1-\frac{c_2}{c_1}\right)E(We^{W}) + \Gamma\left(1-\frac{c_2}{c_1}\right)E(e^{W})\left(\frac{y}{c_1}\right)\Bigg\}e^{-y/c_1} \\
&+ o\left(e^{-y/c_1}\right).
\end{aligned}
$$

Example 6.2. We next state the general two-term expansion. First note the following:

$$\begin{aligned}
K_{m_1,m_1-2} &= \frac{1}{(m_1-2)!}\left[(1-m_1\gamma)\Gamma\left(1-\frac{c_2}{c_1}\right) - \Psi\left(1-\frac{c_2}{c_1}\right)\right] \\
K_{m_1,m_1-1} &= \frac{1}{(m_1-1)!}\Gamma\left(1-\frac{c_2}{c_1}\right) \\
E(e^W) &= \prod_{k=3}^{n-m_1+1} \Gamma\left(1-\frac{c_k}{c_1}\right) \\
E(We^W) &= -\sum_{j=3}^{n-m_1+1} \frac{c_j}{c_1}\Psi\left(1-\frac{c_k}{c_1}\right)\prod_{k=3}^{n-m_1+1}\Gamma\left(1-\frac{c_k}{c_1}\right).
\end{aligned}$$

Then the two-term expansion for $m_1 \geq 2$ as $y \to \infty$ is

$$\begin{aligned}
P(X > y) = &\left[K_{m_1,m_1-2}E\left(e^W\right) - K_{m_1,m_1-1}E\left(We^W\right)\right]\left(\frac{y}{c_1}\right)^{m_1-2} e^{-y/c_1} \\
&+ K_{m_1,m_1-1}E\left(e^W\right)\left(\frac{y}{c_1}\right)^{m_1-1} e^{-y/c_1} + o\left(e^{-y/c_1}\right).
\end{aligned}$$

When $m_1 = 2$, this two-term expansion reduces to the one in the previous example.

Example 6.3. Now take $m_1 = 3$, where the largest weight occurs three times. Then using Table C.4, the $K_{3,j}$ are

$$\begin{aligned}
&K_{3,2} = \frac{1}{2}\Gamma\left(1-\frac{c_2}{c_1}\right), \qquad K_{3,1} = (1-3\gamma)\Gamma\left(1-\frac{c_2}{c_1}\right) - \Psi\left(1-\frac{c_2}{c_1}\right), \qquad \text{and} \\
&K_{3,0} = 0.8181\Gamma\left(1-\frac{c_2}{c_1}\right) - (1-3\gamma)\Psi\left(1-\frac{c_2}{c_1}\right) + \frac{1}{2}E\left(T^2e^T\right).
\end{aligned}$$

As $y \to \infty$, the three-term expansion is

$$\begin{aligned}
P(X > y) = &\left[\frac{K_{3,2}E\left(e^W\right)}{c_1^2}\right] y^2 e^{-y/c_1} + \left[\frac{K_1 E\left(e^W\right) - 2E\left(We^W\right)}{c_1}\right] ye^{-y/c_1} \\
&+ \left[K_{3,0}E\left(e^W\right) - K_{3,1}E\left(We^W\right) + K_{3,2}E\left(W^2e^W\right)\right] e^{-y/c_1} + o\left(e^{-y/c_1}\right).
\end{aligned}$$

This example also illustrates that, unlike the constants in Table C.4, the $K_{m_1,j}$ cannot easily be summarized in a table because they are a function of c_1 and c_2.

Example 6.4. We now present a data analysis example, using the Peachtree Creek data results from Tables 2.4 and 5.6. Suppose it is of interest to find the distribution of the total seasonal maximum streamflow over two years for summer and winter. In other words, we want the distribution of the sum of the maximum streamflow for summer and winter over two consecutive years (so 4 observations). We choose these two seasons because recall that they both fit the Gumbel distribution appropriately, whereas fall and spring did not. Further we had $X_{\text{summer, t}} \sim \text{GEV}(0, 727.6026 + 8.4951t, 429.7833 + 5.9982t)$ and $X_{\text{winter}} \sim \text{GEV}(0, 1438.8241, 810.8379)$ where $t = 1$ represents 1958. In this example we use the years 2008 and 2009, and therefore $t = 51, 52$.

Table 6.1 shows the location and scale parameters for the season and year, from which we can read off the c_k values.

Table 6.1: Estimated Parameters for 2008 and 2009

Param.	**Summer 2008**	**Winter 2008**	**Summer 2009**	**Winter 2009**
μ	1160.8526	1438.8241	1169.3477	1438.8241
σ	735.6915	810.8379	741.6897	810.8379

Here $c_1 = 810.8379$, $c_2 = 741.6897$, $c_3 = 735.6915$, and $m_1 = 2$. If X_k represents the Gumbel observation where $k = 1$ is Summer 2008 and $k = 4$ is Winter 2009, then $X_k = \sigma_k Z_k + \mu_k$ where $Z_k \sim \Lambda$. In which case, the probability is then

$$P\left(\sum_{k=1}^{4}(\sigma_k Z_k + \mu_k) > y\right) \stackrel{d}{=} P\left(Z_1 + Z_2 + \frac{c_2}{c_1}Z_3 + \frac{c_3}{c_1}Z_k > \frac{y - 5207.8485}{810.8379}\right).$$

After simplifying, Theorem 6.2 then provides the approximation

$$P\left(\sum_{k=1}^{4}(\sigma_k Z_k + \mu_k) > y\right) \approx (87.8023y - 1115987)\, e^{-y/810.8379}, \qquad (6.30)$$

provided $y > 12710.23$, a necessary condition to ensure that the approximation is positive.

Example 6.5. Understandably, conducting a general simulation on this topic is not as straightforward as it was in Chapters 3 through 5, since we have not only choices of m_1 but also various combinations of the weights c_k. We therefore perform a sample simulation on the previous example only. The study works by defining $X = \sum_{k=1}^{4}(\sigma_k Z_k + \mu_k)$ for μ_k, σ_k defined in Table 6.1, and simulating 10 million values of X.

Tables 6.2 and 6.3 show the errors and relative errors in estimation, respectively. Figure 6.1 shows the empirical probability for the 99th percentiles and higher. As in earlier chapters, we study the first and second-order approximations in each table. The second-order is the equation in (6.30), while the first-order is just $87.8023ye^{-y/810.8379}$.

Table 6.2: Errors in Approximations of Probabilities

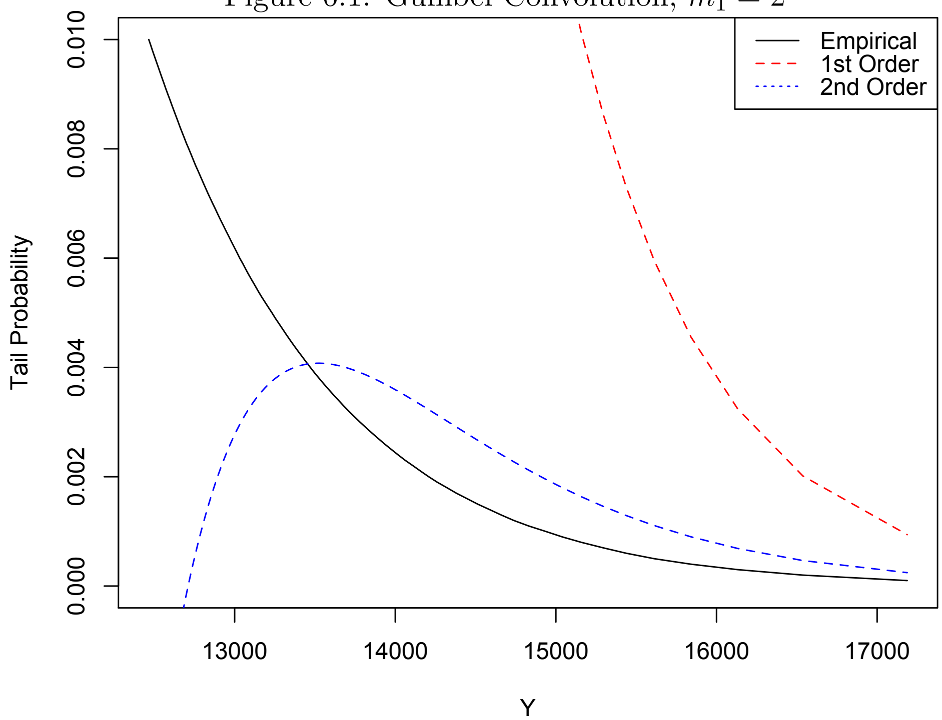
Approx	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
1st	-2.0306	-0.7565	-0.5547	-0.2195	-0.0895	-0.0124	-0.0008
2nd	0.4771	0.1146	0.0712	0.0144	0.0012	-0.0010	-0.0001

Table 6.3: Relative Errors in Approximations of Probabilities

Approx	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
1st	-0.9760	-0.9680	-0.9652	-0.9564	-0.9471	-0.9252	-0.8941
2nd	-1.1171	-1.2791	-1.3909	-3.2537	0.3300	-0.4950	-0.5930

Unfortunately the approximations are much worse than from earlier chapters. The first-order approximation is terrible, but the second-order at least gets close to the empirical probability

Figure 6.1: Gumbel Convolution, $m_1 = 2$



for the higher percentiles. However, this does not happen until around the 99.99th percentile. Also notice that the approximation cannot be used except for above the 99.2nd percentile, since below this mark the equation gives a negative probability. Note the smaller relative errors up to the 99th percentile, but these are clearly meaningless. We conclude that for this particular scenario, the second-order estimation is useful only for very high percentiles.

6.4 Open Questions

This fourth project begins to delve into the complicated topic of dealing with ties in the highest weights of Gumbel convolutions. We have answered the question for the first case, where $m_1 \geq 2$ and $m_2 = 1$. Several potential followup questions remain unanswered, and these would all be useful future topics to study.

First, we can obtain expansions for the other four cases listed at the beginning of the chapter. Each approach is conjectured to have similar steps as those taken in the preceding proofs, although we also expect subcases to emerge. For instance, pivot points may come into play, forcing us to divide into three further cases where $0 < \frac{c_2}{c_1} < \frac{1}{2}$, $\frac{c_2}{c_1} = \frac{1}{2}$, and $\frac{1}{2} < \frac{c_2}{c_1} < 1$.

Second, more involved simulation studies need to be investigated to see how accurate the approximation is, given choices of m_1 and m_2. Recall that in Chapter 5 we discovered that the full expansion should not always be used, and we derived a conjecture for the optimal number of terms to include in the expansion. A similar study could be conducted on this material, although it would have many cases to consider. Another difficulty would be the multitude of possible choices of weights c_k.

Third, we currently have expressions for $E\left(e^T\right)$ and $E\left(Te^T\right)$ as stated in Lemma 6.3. It would be of interest to derive closed-form expressions for higher moments, namely $E\left(T^L e^T\right)$, $L = 2, 3, \cdots$. This procedure would most likely involve characteristics of the gamma, digamma, and possibly polygamma functions. An integral representation is given in (6.5); however, expressions for each moment would be desirable for programming purposes. One temporary solution would be to note that $E\left(T^L e^T\right) = \left(-\frac{c_2}{c_1}\right)^L E\left[(\log G)^L\right]$ where $G \sim \text{Gamma}\left(1 - \frac{c_2}{c_1}, 1\right)$. Then one could simulate the value of $E\left(T^L e^T\right)$ through this alternative formula.

Chapter 7 Chain Dependent Linear Processes

7.1 Introduction

To close the dissertation, we now shift to studying an interesting topic on regularly varying variables, namely the Type II extreme value distribution. In the analysis of extremes for stochastic systems, it is often of interest to model the behavior of a chain dependent process. Such possible applications include the analysis of earthquake magnitudes, flood levels, insurance risk, and queueing theory. These topics, respectively, are discussed in the following references: Caers *et al.* (1999), Bruun and Tawn (1998), Asmussen (2001), and Borokov (1976). In a monograph series, McCormick and Seymour (2001) study the distribution of the maximum of a shot-noise process based on chain-dependent amplitudes. Finally, in another paper McCormick and Seymour (2001) analyze the maximum of a chain-dependent sequence as well as its rate of convergence.

We mention a few more sources that investigated similar work. Rootzén (1988) studied maxima of Markov chains and distributions of exceedances. Results from Barbe and McCormick (2005) that appeared in Section 3.3 will be used again in this chapter. The theorems in that paper helped extract the second term in the convolution of two regularly varying independent variables, and we will use them to derive the second term in the approximations in Section 7.2. Finally, Barbe and McCormick (2009) discuss how to derive three or higher terms in the independent and identically distributed case, and under what conditions further terms exist.

In this final chapter of the dissertation, and the fifth project, we consider the distribution of a linear process formed by taking a linear combination of Markov chain-dependent regularly varying random variables. We derive a first-order tail area approximation before turning our attention to the interesting problem of deriving a second term.

Suppose $F_j, j = 1, ..., M$ are distribution functions. A chain dependent process X_n is such that

$$P\{X_{n+1} \leq x | X_i, J_i = j_i, i \leq n\} = F_{J_n}(x). \tag{7.1}$$

In the heavy tail setting, let F be a distribution and set $F_*(x) = \overline{F}(x) + F(-x)$ where $\overline{F} = 1 - F$ and $x > 0$. Assuming that F is continuous, F_* represents the tail distribution for $|X|$ where $X \sim F$. It is supposed that F_* is regularly varying at infinity with index of regular variation $-\alpha$, $\alpha > 0$, denoted as $F_* \in RV_{-\alpha}$. Furthermore F satisfies the tail balancing condition that $\overline{F}(x) \sim pF_*(x)$ and $F(-x) \sim qF_*(x)$ as $x \to \infty$ for some $0 \leq p \leq 1$ and $p + q = 1$.

For distributions F_j, the two-sided tail, denoted F_{j*}, is defined by $F_{j*}(x) = \overline{F_j}(x) + F_j(-x)$. In order to have nonnegligible components in the asymptotic analysis, it is assumed, for some $0 \leq p_j \leq 1$, $p_j + q_j = 1$, and positive constants $k_j, 1 \leq j \leq M$, that $\overline{F}_j \sim p_j F_{j*}$, $F_j(-x) \sim q_j F_{j*}(x)$, and $F_{j*} \sim k_j F_*$.

Consider the linear process

$$Y_n = \sum_{i=-\infty}^{\infty} c_i X_{n-i}, n \geq 1. \tag{7.2}$$

For the purposes of having a quick reference, here is the most common notation we shall use in this chapter. Set $c_i^+ = \max(c_i, 0)$, $c_i^- = \max(-c_i, 0)$, $d_1 = E_\pi(K_1(J_1))$, $d_2 = E_\pi(K_2(J_1))$, $K_1(a) = p_a k_a$, and $K_2(a) = q_a k_a$ for $1 \leq a \leq M$. For certain steps in the proof, we also rewrite $d_1 = \sum_{a=1}^{M} p_a k_a \pi_a$ and $d_2 = \sum_{a=1}^{M} q_a k_a \pi_a$ for $p_a + q_a = 1$. Further it is assumed that $\vec{\pi}$, the stationary distribution for the Markov chain, exists, and so the sequence Y_n is also stationary.

Finally, unless otherwise noted we shall use Z_i to denote a sequence of independent regularly varying random variables, and X_i if they are chain dependent.

7.2 The First-Order Tail Area Approximation

The first step is to derive a one-term expansion for the tail probability of Y_n. We shall ultimately prove that as $x \to \infty$

$$P(Y_1 > x) = \left(d_1 \sum_{i=-\infty}^{\infty} [c_i^+]^\alpha + d_2 \sum_{i=-\infty}^{\infty} [c_i^-]^\alpha \right) F_*(x) + o\left(F_*(x)\right). \tag{7.3}$$

The proof takes the following path. We first prove by induction that we can add up the finite sum with nonnegative constants. Then we prove that the same property holds when the constants are all negative. We then use these results to argue that the finite sum with any real choice of constants holds. Finally, we impose a suitable summability condition on the $\{c_i\}$ that allows us to move to the infinite sum.

Feller (1971) provides the following result.

Theorem 7.1. (Feller's Theorem) *Let Z_1 and Z_2 be independent random variables from F_1 and F_2, respectively, with the property that $\overline{F}_i(x) = x^{-\alpha} L_i(x)$ with $L_i(x)$ slowly varying, $i = 1, 2$. Then as $x \to \infty$*

$$P(Z_1 + Z_2 > x) = x^{-\alpha} \left[L_1(x) + L_2(x)\right] \left[1 + o(1)\right].$$

Recall the definition of slowly varying in Chapter 1. Let $c_1, c_2 > 0$, then it is easily seen that

$P(c_i Z_i > x) = P\left(Z_i > \frac{x}{c_i}\right) = c_i^\alpha x^{-\alpha} L_i\left(\frac{x}{c_i}\right), i = 1, 2$. Applying Theorem 7.1,

$$P(c_1 Z_1 + c_2 Z_2 > x) = x^{-\alpha}\left[c_1^\alpha L_1\left(\frac{x}{c_1}\right) + c_2^\alpha L_2\left(\frac{x}{c_2}\right)\right][1 + o(1)].$$

By induction,

$$P\left(\sum_{i=1}^{n} c_i Z_i > x\right) = x^{-\alpha}\left[\sum_{i=1}^{n} c_i^\alpha L_i\left(\frac{x}{c_i}\right)\right][1 + o(1)]. \tag{7.4}$$

Now we introduce the Markov chain described above. We first have that $P(|X_i| > x | J_{i-1} = j_{i-1}) \sim k_{j_{i-1}} x^{-\alpha} L(x)$. Because of the right tail balancing condition,

$$\begin{aligned} P(c_i X_i > x | J_{i-1} = j_{i-1}) &\sim p_{j_{i-1}} P(c_i |X_i| > x | J_{i-1} = j_{i-1}) \\ &= p_{j_{i-1}} k_{j_{i-1}} c_i^\alpha x^{-\alpha} L_i\left(\frac{x}{c_i}\right) + o\left(x^{-\alpha} L_i\left(\frac{x}{c_i}\right)\right). \end{aligned}$$

From this point forward, we avoid the $J_{i-1} = j_{i-1}$ notation and simply say J_{i-1} in the probabilities for notational convenience. Thus, conditioning on states J_0 and J_1, the random variables X_1 and X_2 become independent. We obtain

$$P(c_1 X_1 + c_2 X_2 > x | J_0, J_1) = x^{-\alpha}\left[p_{j_0} k_{j_0} c_1^\alpha L\left(\frac{x}{c_1}\right) + p_{j_1} k_{j_1} c_2^\alpha L\left(\frac{x}{c_2}\right)\right][1 + o(1)].$$

Because $L(\cdot)$ is slowly varying, $L\left(\frac{x}{c_i}\right) \Big/ L(x) \to 1$ as $x \to \infty$ for any $c_i > 0$. Using this fact, we see that

$$\begin{aligned}
&P(c_1X_1 + c_2X_2 > x|J_0, J_1) \\
&\quad = x^{-\alpha}L(x)\left(p_{j_0}k_{j_0}c_1^\alpha\left[\frac{L\left(\frac{x}{c_1}\right)}{L(x)}\right] + p_{j_1}k_{j_1}c_2^\alpha\left[\frac{L\left(\frac{x}{c_2}\right)}{L(x)}\right]\right)[1+o(1)] \\
&\quad = (p_{j_0}k_{j_0}c_1^\alpha + p_{j_1}k_{j_1}c_2^\alpha)F_*(x) + o\left(F_*(x)\right).
\end{aligned}$$

We now present the proof with positive constants.

Lemma 7.1. *Let* $\{c_i\}, i = 1, ..., n$ *be positive constants. Then as* $x \to \infty$

$$P\left(\sum_{i=1}^n c_iX_i > x\right) = \left(d_1\sum_{i=1}^n c_i^\alpha\right)F_*(x) + o\left(F_*(x)\right). \tag{7.5}$$

Proof. Our first goal is to establish that

$$P(c_1X_1 + c_2X_2 > x) = d_1(c_1^\alpha + c_2^\alpha)F_*(x) + o(F_*(x)).$$

Let $P_{j,j'}$ denote the probability of going from state j to state j' on the next move. Then

$$\begin{aligned}
P(c_1X_1 + c_2X_2 > x) &= \sum_{j_0}\sum_{j_1} P(c_1X_1 + c_2X_2 > x|J_0, J_1)P(J_0)P(J_1|J_0) \\
&= \left(\sum_{j_0}\sum_{j_1}(p_{j_0}k_{j_0}c_1^\alpha + p_{j_1}k_{j_1}c_2^\alpha)\pi_{j_0}P_{j_0,j_1}\right)F_*(x) + o(F_*(x)) \\
&= \left(\sum_{j_0}\sum_{j_1}(p_{j_0}k_{j_0}c_1^\alpha)\pi_{j_0}P_{j_0,j_1} + \sum_{j_0}\sum_{j_1}(p_{j_1}k_{j_1}c_2^\alpha)\pi_{j_0}P_{j_0,j_1}\right)F_*(x) + o(F_*(x)) \\
&= \left(c_1^\alpha\sum_{j_0}p_{j_0}k_{j_0}\pi_{j_0}\sum_{j_1}P_{j_0,j_1} + c_2^\alpha\sum_{j_1}p_{j_1}k_{j_1}\sum_{j_0}\pi_{j_0}P_{j_0,j_1}\right)F_*(x) + o(F_*(x)).
\end{aligned}$$

The first double sum is simple, since the sum over any row of a Markov chain is 1:

$$c_1^\alpha \sum_{j_0} p_{j_0} k_{j_0} \pi_{j_0} \sum_{j_1} P_{j_0,j_1} = c_1^\alpha \sum_{j_0} p_{j_0} k_{j_0} \pi_{j_0} = c_1^\alpha d_1. \tag{7.6}$$

To compute the second double sum, recall that $\vec{\pi}$ is the solution to $\vec{\pi}^T = \vec{\pi}^T P$ where P denotes the probability matrix of the Markov chain. In other words, for any $1 \leq j, a \leq M$, $\pi_j = \sum_{a=1}^M \pi_a P_{a,j}$. Thus

$$c_2^\alpha \sum_{j_1} p_{j_1} k_{j_1} \sum_{j_0} \pi_{j_0} P_{j_0,j_1} = c_2^\alpha \sum_{j_1} p_{j_1} k_{j_1} \pi_{j_1} = c_2^\alpha d_1. \tag{7.7}$$

Putting (7.6) and (7.7) together,

$$P(c_1 X_1 + c_2 X_2 > x) = d_1(c_1^\alpha + c_2^\alpha) F_*(x) + o(F_*(x)).$$

Next, assuming (7.5) is true, we show that the statement holds for $n+1$ and $c_{n+1} > 0$. By assumption,

$$\sum_{j_0} \cdots \sum_{j_{n-1}} \left(\sum_{i=1}^n c_i^\alpha p_{j_{i-1}} k_{j_{i-1}} \right) \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h, j_{h+1}} = d_1 \sum_{i=1}^n c_i^\alpha.$$

To begin,

$$\begin{aligned} P\left(\sum_{i=1}^{n+1} c_i X_i > x \right) &= \sum_{j_0} \cdots \sum_{j_n} P\left(\sum_{i=1}^{n+1} c_i X_i > x | J_0, \ldots, J_n \right) P(J_0) \prod_{h=0}^{n-1} P(J_{h+1}|J_h) \\ &= \left(\sum_{j_0} \cdots \sum_{j_n} \left(\sum_{i=1}^{n+1} c_i^\alpha p_{j_{i-1}} k_{ji-1} \right) \pi_{j_0} \prod_{h=0}^{n-1} P_{j_h, j_{h+1}} \right) F_*(x) + o(F_*(x)) \end{aligned}$$

$$
\begin{aligned}
&= \left(\sum_{j_0} \cdots \sum_{j_n} \left(\sum_{i=1}^{n} c_i^{\alpha} p_{j_{i-1}} k_{ji-1} \right) \pi_{j_0} \prod_{h=0}^{n-1} P_{j_h, j_{h+1}} \right) F_*(x) \\
&\quad + \left(\sum_{j_0} \cdots \sum_{j_n} \left(\sum_{i=1}^{n} c_{n+1}^{\alpha} p_{j_n} k_{j_n} \right) \pi_{j_0} \prod_{h=0}^{n-1} P_{j_h, j_{h+1}} \right) F_*(x) + o(F_*(x)) \\
&= (I) + (II) + o(F_*(x)).
\end{aligned}
$$

Computing (I) first,

$$
(I) = \left(\sum_{j_0} \cdots \sum_{j_{n-1}} \left(\sum_{i=1}^{n} c_i^{\alpha} p_{j_{i-1}} k_{ji-1} \right) \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h, j_{h+1}} \sum_{j_n} P_{j_{n-1}, j_n} \right) F_*(x),
$$

and therefore

$$
(I) = \left(\sum_{j_0} \cdots \sum_{j_{n-1}} \left(\sum_{i=1}^{n} c_i^{\alpha} p_{j_{i-1}} k_{j_{i-1}} \right) \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h, j_{h+1}} \right) F_*(x) = \left(d_1 \sum_{i=1}^{n} c_i^{\alpha} \right) F_*(x).
$$

Next,

$$
\begin{aligned}
(II) &= c_{n+1}^{\alpha} \left(\sum_{j_0} \cdots \sum_{j_{n-1}} \sum_{j_n} p_{j_n} k_{j_n} \pi_{j_0} \prod_{h=0}^{n-1} P_{j_h, j_{h+1}} \right) F_*(x) \\
&= c_{n+1}^{\alpha} \left(\sum_{j_0} \sum_{j_n} p_{j_n} k_{j_n} \pi_{j_0} P_{j_0, j_n}(n) \right) F_*(x).
\end{aligned}
$$

Here $P_{j,j'}(n)$ represents the probability of moving from state j to state j' in n steps. We have therefore applied the Chapman-Kolmogorov equation in (1.14) $n-1$ times. Continuing,

$$(II) = c_{n+1}^{\alpha}\left(\sum_{j_n} p_{j_n}k_{j_n}\sum_{j_0}\pi_{j_0}P_{j_0,j_n}(n)\right)F_*(x) = c_{n+1}^{\alpha}\left(\sum_{j_n} p_{j_n}k_{j_n}\pi_{j_n}\right)F_*(x)$$
$$= d_1 c_{n+1}^{\alpha}F_*(x).$$

The induction step follows. □

The next question to address is what happens when a constant c_i is zero. Clearly the corresponding X_i contributes nothing to the probability, so the model reduces to the case where all the remaining constants are positive. We can therefore include the case where constants are equal to zero.

Now we use the left tail-balancing condition to address the case when the constants take on negative values. In Lemma 7.2, all the constants are considered negative. Recall that $F_j(-x) \sim q_j F_{j*}(x)$, and let $c_1, c_2 < 0$. The left tail-balancing condition gives us

$$\begin{aligned}
P(c_iX_i > x|J_{i-1}) = P(-c_iX_i < -x|J_{i-1}) &= P\left(X_i < -\left(\frac{x}{-c_i}\right)\Big|J_{i-1}\right)\\
&\sim q_{i-1}P\left(|X_i| > \left(\frac{x}{-c_i}\right)\Big|J_{i-1}\right)\\
&= q_{i-1}[-c_i]^{\alpha}x^{-\alpha}L_i\left(\frac{x}{c_i}\right) + o\left(x^{-\alpha}L_i\left(\frac{x}{c_i}\right)\right).
\end{aligned}$$

Conditioning on states J_0 and J_1, the random variables X_1 and X_2 become independent:

$$\begin{aligned}
P(c_1X_1 + c_2X_2 > x|J_0,J_1) &= P(-(c_1X_1 + c_2X_2) < -x|J_0,J_1)\\
&= x^{-\alpha}\left[q_{j_0}k_{j_0}[-c_1]^{\alpha}L\left(\frac{x}{c_1}\right) + q_{j_1}k_{j_1}[-c_2]^{\alpha}L\left(\frac{x}{c_2}\right)\right][1+o(1)]\\
&= (q_{j_0}k_{j_0}[-c_1]^{\alpha} + q_{j_1}k_{j_1}[-c_2]^{\alpha})F_*(x) + o(F_*(x)).
\end{aligned}$$

We now present the result with negative constants.

Lemma 7.2. *Let* $\{c_i\}, i = 1, ..., n$ *be negative constants. Then as* $x \to \infty$

$$P\left(\sum_{i=1}^{n} c_i X_i > x\right) = \left(d_2 \sum_{i=1}^{n} [-c_i]^\alpha\right) F_*(x) + o\left(F_*(x)\right).$$

Proof. The proof follows the same steps of Lemma 7.1, only replacing the c_i with $[-c_i]$ and the p_i with q_i, $i = 1, ..., n$ in the conclusion. □

We have proven that for a finite sum, the first-order approximation holds provided all the constants are chosen to be nonnegative, or if they are all negative. The next step establishes that we can mix these results and choose a combination of any constants on the real line.

Without loss of generality, we may rewrite the order of $\sum_{i=1}^{n} c_i X_i$ so that the nonnegative constants come first, followed by all negative constants at the end. Define $U_n = \sum_{i=1}^{n} c_i^+ X_i$ and $V_n = -\sum_{i=1}^{n} c_i^- X_i$. Then we have reduced the sum to just two variables, and we have already proven that

$$P(U_n > x) = \left(d_1 \sum_{i=1}^{n} [c_i^+]^\alpha\right) F_*(x) + o(F_*(x)) \quad \text{and}$$
$$P(V_n > x) = \left(d_2 \sum_{i=1}^{n} [c_i^-]^\alpha\right) F_*(x) + o(F_*(x)).$$

Lemma 7.3. *Let* $\{c_i\}, i = 1, ..., n$ *be any real constants. Then as* $x \to \infty$

$$P\left(\sum_{i=1}^{n} c_i X_i > x\right) = \left(d_1 \sum_{i=1}^{n} [c_i^+]^\alpha + d_2 \sum_{i=1}^{n} [c_i^-]^\alpha\right) F_*(x) + o(F_*(x)).$$

Proof. Computing the probability,

$$
\begin{aligned}
P\left(\sum_{i=1}^{n} c_i X_i > x\right) &= \sum_{j_0} \cdots \sum_{j_{n-1}} P\left(\sum_{i=1}^{n} c_i X_i > x \middle| J_0, \ldots, J_{n-1}\right) P(J_0) \prod_{h=0}^{n-2} P(J_{h+1}|J_h) \\
&= \sum_{j_0} \cdots \sum_{j_{n-1}} \left(\left(d_1 \sum_{i=1}^{n} [c_i^+]^\alpha + d_2 \sum_{i=1}^{n} [c_i^-]^\alpha\right) \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h, j_{h+1}}\right) F_*(x) + o(F_*(x)) \\
&= \left(d_1 \sum_{i=1}^{n} [c_i^+]^\alpha + d_2 \sum_{i=1}^{n} [c_i^-]^\alpha\right) \left(\sum_{j_0} \cdots \sum_{j_{n-1}} \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h, j_{h+1}}\right) F_*(x) + o(F_*(x)) \\
&= \left(d_1 \sum_{i=1}^{n} [c_i^+]^\alpha + d_2 \sum_{i=1}^{n} [c_i^-]^\alpha\right) \left(\sum_{j_0} \pi_{j_0} \sum_{j_1} P_{j_0, j_1} \cdots \sum_{j_{n-1}} P_{j_{n-2}, j_{n-1}}\right) F_*(x) \\
&\qquad + o(F_*(x)) \\
&= \left(d_1 \sum_{i=1}^{n} [c_i^+]^\alpha + d_2 \sum_{i=1}^{n} [c_i^-]^\alpha\right) F_*(x) + o(F_*(x)).
\end{aligned}
$$

The last step follows because all the sums, evaluated from right to left, are 1. □

Up to this point we assumed that the sum is finite. This was necessary to establish the induction part of the proof. We are now going to leap from a finite sum to an infinite one, which will require establishing a suitable summability condition on the $\{c_i\}$. Following the technique used by Resnick (1987), we now derive that condition. The goal is to establish under what conditions the series $\sum_{i=-\infty}^{\infty} c_i X_{n-i}$ converges.

Given state J_{i-1} and the asymptotic tail distribution $k_{j_{i-1}}\overline{F}$, let $X_i|J_{i-1}$ be a random variable from this distribution. Also let X^* be a random variable from the distribution whose tail is $\overline{F}$. Because $X^* \in RV_{-\alpha}$, there exists $0 < \lambda < \alpha$ such that $E|X^*|^\lambda < \infty$. Choose λ such that $0 < \lambda < \min(\alpha, 1)$ and $E|X^*|^\lambda < \infty$.

We first need to establish an upper bound for $E(|X_1|^\lambda|J_0)$. We may use this particular X since the distribution of Y_1 is stationary. Choose $0 < x_0 < \infty$ such that $\overline{F}_{j_0} \leq 2k_{j_0}\overline{F}(x)$ for all $x > x_0$. Then

$$E(|X_1|^\lambda | J_0 = j_0) = \int_0^\infty \lambda x^{\lambda-1} \overline{F}_{j_0}(x)dx < \int_0^{x_0} \lambda x^{\lambda-1} dx + \int_{x_0}^\infty \lambda x^{\lambda-1} (2k_{j_0}) \overline{F}(x)dx$$
$$< (x_0)^\lambda + \int_0^\infty \lambda x^{\lambda-1} (2k_{j_0}) \overline{F}(x)dx$$
$$= (x_0)^\lambda + 2k_{j_0} E|X^*|^\lambda < \infty.$$

Therefore

$$E\left(E(|X_1|^\lambda | J_0 = j_0)\right) < (x_0)^\lambda + 2E|X^*|^\lambda E_\pi(k_{j_0}) = (x_0)^\lambda + 2E|X^*|^\lambda \sum_{j_0=1}^{M} k_{j_0}\pi_{j_0} < \infty.$$

By the triangle inequality and the stationarity of the series,

$$E\left|\sum_{i=-\infty}^{\infty} c_i X_{n-i}\right|^\lambda \leq \sum_{i=-\infty}^{\infty} |c_i|^\lambda E\left(E(|X_{n-i}|^\lambda)\right) = \sum_{i=-\infty}^{\infty} \sum_{j=1}^{M} |c_i|^\lambda E\left(E(|X_1|^\lambda | J_0 = j_0)\right)$$
$$< \sum_{i=-\infty}^{\infty} |c_i|^\lambda \left((x_0)^\lambda + 2E|X^*|^\lambda \sum_{j_0=1}^{M} k_{j_0}\pi_{j_0}\right).$$

This expectation will be finite provided $\sum_{i=-\infty}^{\infty} |c_i|^\lambda < \infty$, so $\sum_{i=-\infty}^{\infty} c_i X_{n-i}$ converges almost surely. Before continuing, we state two theorems from Resnick (1987).

Theorem 7.2. (Karamata's Theorem, page 17)
If $\rho \geq -1$, then $\overline{F} \in RV_\rho$ implies

$$\int_0^x \overline{F}(t)dt \in RV_{\rho+1} \quad \textit{and} \quad \lim_{x\to\infty} \frac{x\overline{F}(x)}{\int_0^x \overline{F}(t)dt} = \rho + 1.$$

Theorem 7.3. (Potter's Theorem, page 23)

Suppose $\overline{F} \in RV_\rho, \rho \in \mathbb{R}$. *Take* $\epsilon > 0$. *Then there exists* t_0 *such that for* $x \geq 1$ *and* $t \geq t_0$

$$(1-\epsilon)x^{\rho-\epsilon} < \frac{\overline{F}(tx)}{\overline{F}(t)} < (1+\epsilon)x^{\rho+\epsilon}.$$

Following the technique in Resnick (1987), we use Boole's Inequality and Markov's Inequality. First assume that all the constants are nonnegative:

$$\begin{aligned}
P\left(\sum_i c_i^+|X_i| > x\right) &= P\left(\sum_i c_i^+|X_i| > x, \bigvee_i c_i^+|X_i| > x\right) \\
&\quad + P\left(\sum_i c_i^+|X_i| > x, \bigvee_i c_i^+|X_i| \leq x\right) \\
&\leq P\left(\bigcup_i \left[c_i^+|X_i| > x\right]\right) + P\left(\sum_i c_i^+|X_i| I_{[c_i^+|X_i| \leq x]} > x, \bigvee_i c_i^+|X_i| \leq x\right) \\
&\leq \sum_i P\left(c_i^+|X_i| > x\right) + P\left(\sum_i c_i^+|X_i| I_{[c_i^+|X_i| \leq x]} > x\right) \\
&\leq \sum_i P\left(c_i^+|X_i| > x\right) + x^{-1}E\left(\sum_i c_i^+|X_i| I_{[|X_i| \leq x[c_i^+]^{-1}]}\right).
\end{aligned}$$

Therefore

$$\begin{aligned}
\frac{P\left(\sum_i c_i^+|X_i| > x\right)}{F_*(x)} &\leq \frac{\sum_i P\left(c_i^+|X_i| > x\right)}{F_*(x)} + \frac{x^{-1}E\left(\sum_i c_i^+|X_i| I_{[|X_i| \leq x[c_i^+]^{-1}]}\right)}{F_*(x)} \\
&= (I) + (II).
\end{aligned}$$

We can rewrite (I) as

$$(I) = \sum_i \frac{P(|X_i| > [c_i^+]^{-1}x)}{P(|X_i| > x)} \times \frac{P(|X_i| > x)}{F_*(x)}.$$

Next, we have

$$\frac{P(|X_i| > [c_i^+]^{-1}x)}{P(|X_i| > x)} = \frac{\sum_j P(|X_i| > [c_i^+]^{-1}x|J_{i-1} = j)P(J_{i-1} = j)}{\sum_j P(|X_i| > x|J_{i-1} = j)P(J_{i-1} = j)}$$
$$\sim \frac{\sum_j p_j k_j \pi_j F_*([c_i^+]^{-1}x)}{\sum_j p_j k_j \pi_j F_*(x)} = \frac{F_*([c_i^+]^{-1}x)}{F_*(x)}$$

and

$$\frac{P(|X_i| > x)}{F_*(x)} = \frac{\sum_j P(|X_i| > x|J_{i-1} = j)P(J_{i-1} = j)}{F_*(x)} \sim \sum_j p_j k_j \pi_j = d_1.$$

By Theorem 7.3, for all i such that $c_i^+ < 1$ (all but a finite number of i), there exists x_0 large enough such that for $x > x_0$,

$$d_1(1-\epsilon)[c_i^+]^{\alpha+\epsilon} < (I) < d_1(1+\epsilon)[c_i^+]^{\alpha-\epsilon}.$$

Both the lower and upper bounds are summable, and so by dominated convergence

$$\lim_{x\to\infty}(I) = d_1 \sum_i [c_i^+]^\alpha.$$

Next, we consider (II), first assuming that $0 < \alpha < 1$. From integration by parts,

$$\frac{E\left(\sum_i |X_i| I_{[|X_i| \le x]}\right)}{xF_*(x)} = \frac{\int_0^x P(|X_i| > u)du}{xP(|X_i| > x)} - 1,$$

and manipulating this, we get

$$\begin{aligned}
&\left\{ \frac{\sum_j \int_0^x P(|X_i| > u | J_{i-1} = j) P(J_{i-1} = j) du}{\sum_j x P(|X_i| > x | J_{i-1} = j) P(J_{i-1} = j)} \right. \\
&\qquad \left. \times \frac{\sum_j x P(|X_i| > x | J_{i-1} = j) P(J_{i-1} = j)}{F_*(x)} \right\} - 1 \\
&\qquad = \left\{ \frac{\sum_j \int_0^x p_j k_j \pi_j F_*(u) du}{\sum_j p_j k_j \pi_j x F_*(x)} \cdot \frac{\sum_j p_j k_j \pi_j F_*(x)}{F_*(x)} \right\} - 1 \\
&\qquad = \frac{d_1 \int_0^x F_*(u) du}{x F_*(x)} - 1.
\end{aligned} \tag{7.8}$$

Applying Theorem 7.2, as $x \to \infty$ (7.8) converges to

$$\frac{d_1}{1-\alpha} - 1 = \frac{d_1 + \alpha - 1}{1-\alpha}.$$

Therefore $E|X_i|I_{[|X_i| \le x]} \in RV_{1-\alpha}$. Once again using Theorem 7.3, we have, for all but a finite number of i and x_0 large enough, that for $x > x_0$ and some $k > 0$

$$\begin{aligned}
\frac{c_i^+ E(|X_i| I_{[|X_i| \le x[c_i^+]^{-1}]})}{x F_*(x)} &= c_i^+ \left(\frac{E(|X_i| I_{[|X_i| \le x[c_i^+]^{-1}]})}{E(|X_i| I_{[|X_i| \le x]})} \right) \left(\frac{E(|X_i| I_{[|X_i| \le x]})}{x F_*(x)} \right) \\
&\le k c_i^+ ([c_i^+]^{-1})^{1-\alpha+\alpha-\delta} = k [c_i^+]^{\delta}.
\end{aligned}$$

This upper bound is summable, so therefore

$$\limsup_{x \to \infty} (II) \le k \sum_i c_i^+ [c_i^+]^{\alpha - 1} = k \sum_i [c_i^+]^{\alpha}.$$

To conclude, for $0 < \alpha < 1$ and some constant $k' > 0$

$$\limsup_{x \to \infty} \frac{P\left(\sum_i c_i^+ |X_i| > x\right)}{F_*(x)} \le d_1 \sum_i [c_i^+]^{\alpha} + k \sum_i [c_i^+]^{\alpha} = k' \sum_i [c_i^+]^{\alpha}. \tag{7.9}$$

We now derive a similar inequality when $\alpha \geq 1$ by reducing to the previous case $0 < \alpha < 1$. Choose $\lambda \in (\alpha, \alpha\delta^{-1})$, and define $c = \sum_i c_i^+$ and $r_i^+ = c_i^+/c$. Then by Jensen's inequality,

$$\left(\sum_i c_i^+|X_i|\right)^\lambda = c^\lambda \left(\sum_i r_i^+|X_i|\right)^\lambda \leq c^\lambda \sum_i r_i^+|X_i|^\lambda = c^{\lambda-1}\sum_i c_i^+|X_i|^\lambda.$$

Let $|X_*| \in RV_{-\alpha}$ represent a random variable from the underlying tail distribution $F_*(x)$. Then

$$\frac{P\left(\sum_i c_i^+|X_i| > x\right)}{F_*(x)} \leq \frac{P\left(\sum_i c_i^+|X_i|^\lambda > c^{1-\lambda}x^\lambda\right)}{P(|X^*|^\lambda > x^\lambda)}. \tag{7.10}$$

Using the fact that $P(|X_*|^\lambda > x) \in RV_{-\alpha\lambda^{-1}}, \delta < \alpha\lambda^{-1} < 1$,

$$\begin{aligned}\frac{P\left(\sum_i c_i^+|X_i|^\lambda > c^{1-\lambda}x^\lambda\right)}{P(|X^*|^\lambda > x^\lambda)} &= \frac{\sum_j P\left(\sum_i c_i^+|X_i|^\lambda > c^{1-\lambda}x^\lambda | J_{i-1} = j\right) P(J_{i-1} = j)}{P(|X^*|^\lambda > x^\lambda)} \\ &\sim \frac{\sum_j p_j k_j \pi_j \sum_i [c_i^+]^{\alpha\lambda^{-1}} c^{-(1-\lambda)\alpha\lambda^{-1}} x^{-\alpha}}{x^{-(\alpha\lambda^{-1})\lambda}} = d_1 c^{\alpha(1-\lambda^{-1})} \sum_i [c_i^+]^{\alpha\lambda^{-1}} < \infty,\end{aligned}$$

which when combined with (7.10) gives

$$\limsup_{x\to\infty} \frac{P\left(\sum_i c_i^+|X_i| > x\right)}{F_*(x)} \leq d_1 c^{\alpha(1-\lambda^{-1})} \sum_i [c_i^+]^{\alpha\lambda^{-1}}. \tag{7.11}$$

We may finally establish the infinite limit. Choose any integer $m > 0$, then

$$\frac{P\left(\sum_i c_i^+|X_i| > x\right)}{F_*(x)} \geq \frac{P\left(\sum_{|i|\leq m} c_i^+|X_i| > x\right)}{F_*(x)} \to d_1 \sum_{|i|\leq m} [c_i^+]^\alpha,$$

using the already proven result for a finite sum. The constant m is arbitrary, so

$$\liminf_{x\to\infty} \frac{P\left(\sum_i c_i^+|X_i| > x\right)}{F_*(x)} \geq d_1 \sum_j [c_i^+]^\alpha. \tag{7.12}$$

Next, for any $\epsilon > 0$

$$\frac{P\left(\sum_i c_i^+|X_i| > x\right)}{F_*(x)} \leq \frac{P\left(\sum_{|i|\leq m} c_i^+|X_i| > (1-\epsilon)x\right)}{F_*(x)} + \frac{P\left(\sum_{|i|>m} c_i^+|X_i| > \epsilon x\right)}{F_*(x)}. \quad (7.13)$$

If $0 < \alpha < 1$, combine (7.9) and (7.13) to obtain

$$\limsup_{x\to\infty} \frac{P\left(\sum_i c_i^+|X_i| > x\right)}{F_*(x)} \leq d_1(1-\epsilon)^{-\alpha} \sum_{|i|\leq m} [c_i^+]^\alpha + k'\epsilon^{-\alpha} \sum_{|i|>m} [c_i^+]^\alpha. \quad (7.14)$$

On the other hand, if $\alpha \geq 1$ then (7.11) and (7.13) give

$$\limsup_{x\to\infty} \frac{P\left(\sum_i c_i^+|X_i| > x\right)}{F_*(x)} \leq d_1(1-\epsilon)^{-\alpha} \sum_{|i|\leq m} [c_i^+]^\alpha + d_1 c^{\alpha(1-\lambda^{-1})}\epsilon^{-\alpha} \sum_{|i|>m} [c_i^+]^{\alpha\lambda^{-1}}. \quad (7.15)$$

In both cases, first let $m \to \infty$, then

$$\limsup_{x\to\infty} \frac{P\left(\sum_i c_i^+|X_i| > x\right)}{F_*(x)} \leq d_1(1-\epsilon)^{-\alpha} \sum_i [c_i^+]^\alpha.$$

Now send $\epsilon \to 0$ to obtain

$$\limsup_{x\to\infty} \frac{P\left(\sum_i c_i^+|X_i| > x\right)}{F_*(x)} \leq d_1 \sum_i [c_i^+]^\alpha.$$

Combine this with (7.12) to conclude that

$$\lim_{x\to\infty} \frac{P\left(\sum_i c_i^+|X_i| > x\right)}{F_*(x)} = d_1 \sum_i [c_i^+]^\alpha. \quad (7.16)$$

Next, using similar steps with appropriate changes, it can be shown that when all the constants are negative,

$$\lim_{x\to\infty}\frac{P\left(\sum_i c_i^-|X_i|>x\right)}{F_*(x)}=d_2\sum_i[c_i^-]^\alpha. \tag{7.17}$$

Lastly, to leap to the general statement using any real constants, use the definitions $U_n = \sum_i c_i^+|X_i|$, $V_n = -\sum_i c_i^-|X_i|$, Theorem 7.1, (7.16), and (7.17) to obtain

$$\lim_{x\to\infty}\frac{P\left(\sum_i c_i|X_i|>x\right)}{F_*(x)}=d_1\sum_i[c_i^+]^\alpha+d_2\sum_i[c_i^-]^\alpha.$$

We have now proven the result for the infinite sum case and summarize our results from above and Lemmas 7.1 through 7.3.

Theorem 7.4. *Let Y_1 be as defined earlier, and choose $0<\lambda<\min(\alpha,1)$ such that $\int_0^\infty \lambda x^{\lambda-1}\overline{F}(x)dx<\infty$ and $\sum_i|c_i|^\lambda<\infty$. Then as $x\to\infty$*

$$P(Y_1>x)=\left(d_1\sum_{i=-\infty}^{\infty}[c_i^+]^\alpha+d_2\sum_{i=-\infty}^{\infty}[c_i^-]^\alpha\right)F_*(x)+o(F_*(x)).$$

7.3 The Second-Order Tail Area Approximation

In this section we derive a second-order approximation formula. This result will hold for finite n; the infinite case remains an open question. In what follows, we first assume that the $\{c_i\}, i=1,...,n$ are positive and that $n<\infty$. Further suppose that $\alpha\geq 1$ and that the mean of F_* is finite, and consequently so are the means μ_j of the underlying distributions F_j, $j=1,...,M$. This situation will be dealt with in Section 7.3.1. The cases where $\alpha=1$ and $0<\alpha<1$ will be handled later in Sections 7.3.2 and 7.3.3. We shall make the following additional assumptions:

1. F is asymptotically smooth and right-tail dominant as defined in Section 3.4.

2. For each $i = 1, ..., n$, as $x \to \infty$

$$\overline{F}_i\left(\frac{x}{c_i}\right) = c_i^\alpha \overline{F}_i(x) + o\left(\frac{\overline{F}_i(x)}{x}\right) \leftrightarrow x\left(\frac{\overline{F}_i(x/c_i)}{\overline{F}_i(x)} - c_i^\alpha\right) \to 0. \tag{7.18}$$

We begin by extending the formula for two independent variables to the sum of an n-variable process as described earlier. Recall that for Z_1, Z_2 independent and $Z_i \sim F_i, i = 1, 2$, we had from Theorem 3.3

$$P(Z_1 + Z_2 > x) = \overline{F}_1(x) + \overline{F}_2(x) + \frac{\alpha}{x}[\overline{F}_1(x)\mu_{F_2} + \overline{F}_2(x)\mu_{F_1}][1 + o(1)].$$

Introducing constants $c_i > 0, i = 1, 2$,

$$P(c_1Z_1+c_2Z_2 > x) = c_1^\alpha\overline{F}_1(x)+c_2^\alpha\overline{F}_2(x)+\frac{\alpha}{x}[c_1^\alpha c_2\mu_{F_2}\overline{F}_1(x)+c_2^\alpha c_1\mu_{F_1}\overline{F}_2(x)][1+o(1)]. \tag{7.19}$$

The first goal is to use induction to establish the same formula for the sum of n independent variables $Z_i \sim \overline{F}_i \in RV_{-\alpha}$.

Lemma 7.4. *Let $Z_i \sim F_i$ be independent regularly varying random variables and $c_i > 0$, $i = 1, ..., n$. Assuming that the F_i are asymptotically smooth and right-tail dominant, and further that condition (7.18) is satisfied, a two-term expansion for the weighted convolution is*

$$P\left(\sum_{i=1}^n c_iZ_i > x\right) = \sum_{i=1}^n c_i^\alpha\overline{F}_i(x) + \frac{\alpha}{x}\left\{\sum_{i=1}^n\sum_{g\neq i} c_i^\alpha c_g\mu_{F_g}\overline{F}_i(x)\right\}[1 + o(1)].$$

Proof. (7.19) provides the $n = 2$ case. Now define $\overline{H}(x) = P\left(\sum_{i=1}^n c_iZ_i > x\right)$, and let $c_{n+1} > 0$ and $Z_{n+1} \sim F_{n+1}$ with $\overline{F}_{n+1} \in RV_{-\alpha}$, independent of the previous random variables.

Notice that the mean of the process up to the nth variable is $\sum_{i=1}^{n} c_i \mu_{F_i}$. Then (7.19) provides

$$P\left(\sum_{i=1}^{n} c_i Z_i + c_{n+1} Z_{n+1} > x\right) = \overline{H}(x) + c_{n+1}^{\alpha} \overline{F}_{n+1}(x)$$
$$+ \frac{\alpha}{x} \left\{ \overline{H}(x) c_{n+1} \mu_{F_{n+1}} + c_{n+1}^{\alpha} \overline{F}_{n+1}(x) \sum_{g=1}^{n} c_g \mu_{F_g} \right\} [1 + o(1)]$$

$$= \sum_{i=1}^{n+1} c_i^{\alpha} \overline{F}_i(x) + \frac{\alpha}{x} \left\{ \sum_{i=1}^{n} \sum_{g=1}^{n} c_i^{\alpha} c_g \mu_{F_g} \overline{F}_i(x) + c_{n+1}^{\alpha} \overline{F}_{n+1}(x) \sum_{g \neq n+1} c_g \mu_{F_g} \right.$$
$$\left. + c_{n+1} \mu_{F_{n+1}} \sum_{i=1}^{n} c_i^{\alpha} \overline{F}_i(x) + \frac{\alpha \mu_{F_{n+1}}}{x} \sum_{i=1}^{n} \sum_{g \neq i} c_i^{\alpha} c_g \mu_{F_g} \overline{F}_i(x) \right\} [1 + o(1)]$$
$$= \sum_{i=1}^{n+1} c_i^{\alpha} \overline{F}_i(x) + \frac{\alpha}{x} \left\{ \sum_{i=1}^{n} \sum_{g \neq i} c_i^{\alpha} c_g \mu_{F_g} \overline{F}_i(x) + c_{n+1}^{\alpha} \overline{F}_{n+1}(x) \sum_{g \neq n+1} c_g \mu_{F_g} \right.$$
$$\left. + c_{n+1} \mu_{F_{n+1}} \sum_{i=1}^{n} c_i^{\alpha} \overline{F}_i(x) \right\} [1 + o(1)].$$

The last line above results from the fact that one of the terms involves $\frac{\alpha^2}{x^2}$, so it is negligible. Continuing,

$$P\left(\sum_{i=1}^{n+1} c_i Z_i > x\right) = \sum_{i=1}^{n+1} c_i^{\alpha} \overline{F}_i(x) + \frac{\alpha}{x} \left\{ \sum_{i=1}^{n} c_i^{\alpha} \overline{F}_i(x) \left[\sum_{g \neq i} c_g \mu_{F_g} + c_{n+1} \mu_{F_{n+1}} \right] \right.$$
$$\left. + c_{n+1}^{\alpha} \overline{F}_{n+1}(x) \sum_{g \neq n+1} c_g \mu_{F_g} \right\} [1 + o(1)].$$

The term $\sum_{i=1}^{n} \sum_{g \neq i} c_g \mu_{F_g} + c_{n+1} \mu_{F_{n+1}}$ collapses into $\sum_{i=1}^{n} \sum_{g \neq i} c_g \mu_{F_g}$ because the index in the first sum is from 1 to n. Therefore g never equals $n+1$ in the finite sum anyway. Finally, the probability is

$$\sum_{i=1}^{n+1} c_i^\alpha \overline{F}_i(x) + \frac{\alpha}{x}\left\{\sum_{i=1}^{n}\sum_{g\neq i} c_i^\alpha c_g \mu_{F_g}\overline{F}_i(x) + c_{n+1}^\alpha \overline{F}_{n+1}(x)\sum_{g\neq n+1} c_g\mu_{F_g}\right\}[1+o(1)]$$
$$= \sum_{i=1}^{n+1} c_i^\alpha \overline{F}_i(x) + \frac{\alpha}{x}\left\{\sum_{i=1}^{n+1}\sum_{g\neq i} c_i^\alpha c_g \mu_{F_g}\overline{F}_i(x)\right\}[1+o(1)].$$

This completes the induction step. □

7.3.1 The Case With $\alpha \geq 1$ and Finite Means

Let $X_1, ..., X_n$ be Markov chain dependent random variables where the M-state chain has stationary distribution $\vec{\pi}$. For now, take $\alpha \geq 1$ and the underlying distributions to have finite means, and assume that $c_i > 0, i = 1, ..., n$.

Before going further, we need to establish some stronger tail balance properties. In the previous section we assumed that $\overline{F}_j = p_j k_j \overline{F}_* + o(F_*)$, but that order term was appropriate since we were only concerned with a first-order term. In this section we need a stronger assumption, namely that for $j = 1, ..., M$,

$$x\left\{\left(\frac{\overline{F}_j(x)}{F_{*j}(x)} - p_j\right)\left(\frac{F_{*j}}{F_*}\right) + p_j\left(\frac{F_{*j}}{F_*} - k_j\right)\right\} \to 0 \quad \text{as} \quad x \to \infty. \tag{7.20}$$

As a consequence, as $x \to \infty$

$$P(c_i X_i > x | J_{i-1}) = p_{j_{i-1}} k_{j_{i-1}} c_i^\alpha F_*(x) + o\left(\frac{F_*(x)}{x}\right).$$

In addition, we also have that $E[c_i X_i | J_{i-1} = j_{i-1}] = c_i \mu_{F_{J_{i-1}}} = c_i \mu_{j_{i-1}}$. Now we can find the conditional probability given all the previous states of the chain. Lemma 7.4 provides

$$
\begin{aligned}
P\left(\sum_{i=1}^{n} c_i X_i > x \middle| J_l, 0 \le l \le n-1\right) &= \sum_{i=1}^{n} p_{j_{i-1}} k_{j_{i-1}} c_i^{\alpha} F_*(x) \\
&\quad + \frac{\alpha F_*(x)}{x} \left\{ \sum_{i=1}^{n} \sum_{g \neq i} p_{j_{i-1}} k_{j_{i-1}} c_i^{\alpha} E\left[c_g X_g | J_{g-1} = j_{g-1}\right] \right\} + o\left(\frac{F_*(x)}{x}\right) \\
&= \left(\sum_{i=1}^{n} p_{j_{i-1}} k_{j_{i-1}} c_i^{\alpha}\right) F_*(x) + \frac{\alpha F_*(x)}{x} \left\{ \sum_{i=1}^{n} \sum_{g \neq i} p_{j_{i-1}} k_{j_{i-1}} c_i^{\alpha} c_g \mu_{j_{g-1}} \right\} + o\left(\frac{F_*(x)}{x}\right).
\end{aligned} \tag{7.21}
$$

We now uncondition (7.21):

$$
\begin{aligned}
P\left(\sum_{i=1}^{n} c_i X_i > x\right) &= \sum_{j_0} \cdots \sum_{j_{n-1}} P\left(\sum_{i=1}^{n} c_i X_i > x \middle| J_l, 0 \le l \le n-1\right) \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h j_{h+1}} \\
&= \sum_{j_0} \cdots \sum_{j_{n-1}} \left(\sum_{i=1}^{n} p_{j_{i-1}} k_{j_{i-1}} c_i^{\alpha}\right) F_*(x) \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h j_{h+1}} \\
&\quad + \frac{\alpha F_*(x)}{x} \sum_{j_0} \cdots \sum_{j_{n-1}} \left\{ \sum_{i=1}^{n} \sum_{g \neq i} p_{j_{i-1}} k_{j_{i-1}} c_i^{\alpha} c_g \mu_{j_{g-1}} \right\} \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h j_{h+1}} + o\left(\frac{F_*(x)}{x}\right) \\
&= (I) + (II) + o\left(\frac{F_*(x)}{x}\right).
\end{aligned}
$$

(I) is equal to $\left(d_1 \sum_{i=1}^{n} c_i^{\alpha}\right) F_*(x)$. To simplify (II), define $h_g^+(x,y) = c_g p(x) k(x) \mu(y)$ where $p(x)k(x) = p_{j_x} k_{j_x}$ and $\mu(y) = \mu_{j_y}$. Then (II) can be expressed as the sum of expected values of the $h_g^+(\cdot,\cdot)$:

$$
(II) = \frac{\alpha F_*(x)}{x} \left\{ \sum_{i=1}^{n} \sum_{g \neq i} c_i^{\alpha} E\left[h_g^+(J_{i-1}, J_{g-1})\right] \right\}.
$$

We therefore have

Lemma 7.5. *Suppose $\alpha \geq 1$, and choose $c_i \geq 0, i = 1, \cdots, n$. Then as $x \to \infty$*

$$P\left(\sum_{i=1}^{n} c_i X_i > x\right) = \left\{d_1 \sum_{i=1}^{n} c_i^{\alpha} + \frac{\alpha}{x} \sum_{i=1}^{n} \sum_{g \neq i} c_i^{\alpha} E\left[h_g^{+}(J_{i-1}, J_{g-1})\right]\right\} F_*(x) + o\left(\frac{F_*(x)}{x}\right).$$

Unlike before, this approximation formula requires the transition matrix probabilities of the Markov chain, which in practice will be unknown. Therefore in order to use this formula, the transition matrix must first be estimated from the data set.

We now redo the proof of a finite sum, this time assuming the $\{c_i\}$ are all negative. First consider the random variable $c_i X_i$ conditional on J_{i-1}:

$$\begin{aligned} P(c_i X_i > x | J_{i-1}) = P(-c_i X_i < -x | J_{i-1}) &= P\left(X_i < \frac{x}{[-c_i]} \middle| J_{i-1}\right) \\ &= q_{j_{i-1}} k_{j_{i-1}} P\left(|X_i| > \frac{x}{[-c_i]} \middle| J_{i-1}\right) + o\left(\frac{F_{*j_{i-1}}(x)}{x}\right) \\ &= q_{j_{i-1}} k_{j_{i-1}} [c_i]^{\alpha} F_{*j_{i-1}}(x) + o\left(\frac{F_{*j_{i-1}}(x)}{x}\right). \end{aligned}$$

As before, we condition on all previous states of the chain, making the variables independent, and therefore by Lemma 7.4

$$\begin{aligned} P\left(\sum_{i=1}^{n} c_i X_i > x \middle| J_l, 0 \leq l \leq n-1\right) &= \sum_{i=1}^{n} q_{j_{i-1}} k_{j_{i-1}} [-c_i]^{\alpha} F_*(x) \\ &+ \frac{\alpha F_*(x)}{x} \left\{\sum_{i=1}^{n} \sum_{g \neq i} q_{j_{i-1}} k_{j_{i-1}} [-c_i]^{\alpha} E[c_g X_g | J_g = j_g]\right\} + o\left(\frac{F_*(x)}{x}\right) \qquad (7.22) \\ &= \left\{\sum_{i=1}^{n} q_{j_{i-1}} k_{j_{i-1}} [-c_i]^{\alpha} + \frac{\alpha}{x} \sum_{i=1}^{n} \sum_{g \neq i} q_{j_{i-1}} k_{j_{i-1}} [-c_i]^{\alpha} c_g \mu_{j_{g-1}}\right\} F_*(x) + o\left(\frac{F_*(x)}{x}\right). \end{aligned}$$

Unconditioning (7.22), for $0 \le l \le n-1$

$$
\begin{aligned}
P\left(\sum_{i=1}^{n} c_i X_i > x\right) &= \sum_{j_0} \cdots \sum_{j_{n-1}} P\left(\sum_{i=1}^{n} c_i X_i > x \Big| J_l = j_l\right) \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h j_{h+1}} \\
&= F_*(x) \Bigg\{ \sum_{j_0} \cdots \sum_{j_{n-1}} \left(\sum_{i=1}^{n} q_{j_{i-1}} k_{j_{i-1}} [-c_i]^{\alpha}\right) \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h j_{h+1}} \\
&\quad + \frac{\alpha}{x} \sum_{j_0} \cdots \sum_{j_{n-1}} \left[\sum_{i=1}^{n} \sum_{g \neq i} q_{j_{i-1}} k_{j_{i-1}} [-c_i]^{\alpha} c_g \mu_{j_{g-1}}\right] \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h j_{h+1}} \Bigg\} + o\left(\frac{F_*(x)}{x}\right) \\
&= (I) + (II) + o\left(\frac{F_*(x)}{x}\right).
\end{aligned}
$$

The first piece (I) is equal to $\left(d_2 \sum_{i=1}^{n} [-c_i]^{\alpha}\right) F_*(x)$. Define $h_g^-(x,y) = c_g q(x) k(x) \mu(y)$ where $q(x)k(x) = q_{j_x} k_{j_x}$ and $\mu(y) = \mu_{j_y}$, then (II) can be written as

$$
(II) = \frac{\alpha F_*(x)}{x} \left\{ \sum_{i=1}^{n} \sum_{g \neq i} [-c_i]^{\alpha} E\left[h_g^-(J_{i-1}], J_{g-1})\right\}.
$$

We therefore have

Lemma 7.6. *Suppose $\alpha \ge 1$, and choose negative constants $\{c_i\}$. Then as $x \to \infty$*

$$
\begin{aligned}
P\left(\sum_{i=1}^{n} c_i X_i > x\right) = \left\{ d_2 \sum_{i=1}^{n} [-c_i]^{\alpha} + \frac{\alpha}{x} \sum_{i=1}^{n} \sum_{g \neq i} [-c_i]^{\alpha} E\left[h_g^-(J_{i-1}, J_{g-1})\right] \right\} F_*(x) \\
+ o\left(\frac{F_*(x)}{x}\right).
\end{aligned}
$$

Now that we have proven the cases where the constants are either all nonnegative or negative, we can mix the two results. Define $U_n = \sum_{i=1}^{n} [c_i^+] X_i$ and $V_n = -\sum_{i=1}^{n} [c_i^-] X_i$. Observe that conditional on the $\{J_i\}$, U_n and V_n are independent, so (7.19) can be used. To recap, we have

$$P\left(U_n > x\right) = \left\{ d_1 \sum_{i=1}^{n} [c_i^+]^\alpha + \frac{\alpha}{x} \sum_{i=1}^{n} \sum_{g \neq i} [c_i^+]^\alpha E\left[h_g^+(J_{i-1}, J_{g-1})\right] \right\} F_*(x) + o\left(\frac{F_*(x)}{x}\right)$$

$$P\left(V_n > x\right) = \left\{ d_2 \sum_{i=1}^{n} [c_i^-]^\alpha + \frac{\alpha}{x} \sum_{i=1}^{n} \sum_{g \neq i} [c_i^-]^\alpha E\left[h_g^-(J_{i-1}, J_{g-1})\right] \right\} F_*(x) + o\left(\frac{F_*(x)}{x}\right).$$

Theorem 7.5. *Let* $\{c_i\}, i = 1, \cdots, n$ *be a sequence of real constants. Suppose* $\alpha \geq 1$ *so that each of the underlying distributions has finite mean* μ_j. *Define* $\overline{\mu} = \sum_{j=1}^{M} \mu_j \pi_j$. *Then as* $x \to \infty$

$$\begin{aligned} P\left(Y_1 > x\right) = &\left(d_1 \sum_{i=1}^{n} [c_i^+]^\alpha + d_2 \sum_{i=1}^{n} [c_i^-]^\alpha \right) F_*(x) \\ &+ \frac{\alpha F_*(x)}{x} \left\{ \sum_{i=1}^{n} \sum_{g \neq i} [c_i^+]^\alpha E\left[h_g^+(J_{i-1}, J_{g-1})\right] + \sum_{i=1}^{n} \sum_{g \neq i} [c_i^-]^\alpha E\left[h_g^-(J_{i-1}, J_{g-1})\right] \right. \\ &\left. - d_1 \overline{\mu} \sum_{i=1}^{n} \sum_{g=1}^{n} [c_i^+]^\alpha [c_g^-] + d_2 \overline{\mu} \sum_{i=1}^{n} \sum_{g=1}^{n} [c_i^-]^\alpha [c_g^+] \right\} + o\left(\frac{F_*(x)}{x}\right). \end{aligned}$$

Proof. Observe that $P(U_n + V_n > x)$ is equal to

$$P(U_n > x) + P(V_n > x) + \frac{\alpha F_*(x)}{x}[P(U_n > x)E(V_n) + P(V_n > x)E(U_n)][1 + o(1)].$$

Further

$$E(U_n) = E\left[E\left(\sum_{i=1}^{n} [c_i^+] X_i | J_{i-1}\right)\right] = \sum_{i=1}^{n} [c_i^+] E(\mu_{J_{i-1}}) = \sum_{i=1}^{n} [c_i^+] \sum_{g=1}^{M} \mu_j \pi_j = \overline{\mu} \sum_{i=1}^{n} [c_i^+].$$

Similarly, $E(V_n) = -\overline{\mu}\sum_{i=1}^{n}[c_i^-]$. Finally,

$$
\begin{aligned}
P\left(\sum_{i=1}^{n} c_i X_i > x\right) &= \left(d_1 \sum_{i=1}^{n}[c_i^+]^\alpha + d_2 \sum_{i=1}^{n}[c_i^-]^\alpha\right) F_*(x) \\
&\quad + \frac{\alpha F_*(x)}{x}\left\{\sum_{i=1}^{n}\sum_{g\neq i}[c_i^+]^\alpha E\left[h_g^+(J_{i-1}, J_{g-1})\right] + \sum_{i=1}^{n}\sum_{g\neq i}[c_i^-]^\alpha E\left[h_g^-(J_{i-1}, J_{g-1})\right]\right\} \\
&\quad + \frac{\alpha F_*(x)}{x}\left\{d_1 \sum_{i=1}^{n}[c_i^+]^\alpha E(V_n) + d_2 \sum_{i=1}^{n}[c_i^-]^\alpha E(U_n)\right\} + o\left(\frac{F_*(x)}{x}\right) \\
&= \left(d_1 \sum_{i=1}^{n}[c_i^+]^\alpha + d_2 \sum_{i=1}^{n}[c_i^-]^\alpha\right) F_*(x) \\
&\quad + \frac{\alpha F_*(x)}{x}\left\{\sum_{i=1}^{n}\sum_{g\neq i}[c_i^+]^\alpha E\left[h_g^+(J_{i-1}, J_{g-1})\right] + \sum_{i=1}^{n}\sum_{g\neq i}[c_i^-]^\alpha E\left[h_g^-(J_{i-1}, J_{g-1})\right]\right. \\
&\quad \left. - d_1\overline{\mu}\sum_{i=1}^{n}\sum_{g=1}^{n}[c_i^+]^\alpha[c_g^-] + d_2\overline{\mu}\sum_{i=1}^{n}\sum_{g=1}^{n}[c_i^-]^\alpha[c_g^+]\right\} + o\left(\frac{F_*(x)}{x}\right).
\end{aligned}
$$

□

The major issue with this result is that we need another, simpler form for computational purposes. We shall derive computational results for the two sums in Theorem 7.5 that contain the expected values. Ultimately we will write the transition probabilities in terms of the number of steps m needed to get from state j to state j'. Recall that this is denoted as $P_{j,j'}(m)$. The first goal is to simplify

$$
\begin{aligned}
\sum_{i=1}^{n}\sum_{g\neq i}[c_i^+]^\alpha E\left[h_g^+(J_{i-1}, J_{g-1})\right] = \sum_{i=1}^{n}\sum_{g\neq i}[c_i^+]^\alpha c_g \sum_{j_0}\cdots\sum_{j_{n-1}} p_{j_{i-1}} k_{j_{i-1}} \mu_{j_{g-1}} \\
\times \pi_{j_0}\prod_{h=0}^{n-2} P_{j_h, j_{h+1}}.
\end{aligned}
\tag{7.23}
$$

We proceed by first fixing $i = 1$, then letting $i = t, 1 < t < n$, and then finally ending with

$i = n$. The contribution from the $i = 1$ term provides

$$[c_1^+]^\alpha \sum_{g \neq 1} E\left[h_g^+(J_0, J_{g-1})\right] = [c_1^+]^\alpha \sum_{g=2}^{n} c_g \sum_{j_0} \cdots \sum_{j_{n-1}} p_{j_0} k_{j_0} \mu_{j_{g-1}} \pi_{j_0} \prod_{h=0}^{n-2} P_{j_h, j_{h+1}}.$$

Simplifying,

$$\begin{aligned}
&[c_1^+]^\alpha \sum_{g=2}^{n} c_g \sum_{j_0} \cdots \sum_{j_{n-1}} p_{j_0} k_{j_0} \mu_{j_{g-1}} \pi_{j_0} P_{j_0, j_1} \cdots P_{j_{n-2}, j_{n-1}} \\
&\qquad = [c_1^+]^\alpha \sum_{g=2}^{n} c_g \sum_{j_0} \cdots \sum_{j_{g-1}} p_{j_0} k_{j_0} \mu_{j_{g-1}} \pi_{j_0} P_{j_0, j_1} \cdots P_{j_{g-2}, j_{g-1}} \\
&\qquad = [c_1^+]^\alpha \sum_{g=2}^{n} c_g \sum_{j_0} \sum_{j_{g-1}} p_{j_0} k_{j_0} \mu_{j_{g-1}} \pi_{j_0} P_{j_0, j_{g-1}}(g-1).
\end{aligned}$$

Rewriting this sum using the new indices a and b for notational convenience, we have that the $i = 1$ term contributes

$$[c_1^+]^\alpha \sum_{g \neq 1} E\left[h_g^+(J_0, J_{g-1})\right] = [c_1^+]^\alpha \sum_{g=2}^{n} c_g \sum_{a=1}^{M} \sum_{b=1}^{M} p_a k_a \mu_b \pi_a P_{a,b}(g-1). \tag{7.24}$$

Now we examine the $i = t$ term, $1 < t < n$. Notice we have two cases to check here: when $1 \leq g < t$ and when $t < g \leq n$. First let $1 \leq g < t$:

$$\begin{aligned}
&[c_t^+]^\alpha \sum_{g=1}^{t-1} c_g \sum_{j_0} \cdots \sum_{j_{n-1}} p_{j_{t-1}} k_{j_{t-1}} \mu_{j_{g-1}} \pi_{j_0} P_{j_0, j_1} \cdots P_{j_{n-2}, j_{n-1}} \\
&\quad = [c_t^+]^\alpha \sum_{g=1}^{t-1} c_g \sum_{j_0} \cdots \sum_{j_{g-1}} \cdots \sum_{j_{t-1}} p_{j_{t-1}} k_{j_{t-1}} \mu_{j_{g-1}} \pi_{j_0} P_{j_0, j_1} \cdots P_{j_{g-1}, j_g} \cdots P_{j_{t-2}, j_{t-1}} \\
&\quad = [c_t^+]^\alpha \sum_{g=1}^{t-1} c_g \sum_{j_{g-1}} \cdots \sum_{j_{t-1}} p_{j_{t-1}} k_{j_{t-1}} \mu_{j_{g-1}} \pi_{j_{g-1}} P_{j_{g-1}, j_g} \cdots P_{j_{t-2}, j_{t-1}}
\end{aligned}$$

$$= [c_t^+]^\alpha \sum_{g=1}^{t-1} c_g \sum_{j_{g-1}} \sum_{j_{t-1}} p_{j_{t-1}} k_{j_{t-1}} \mu_{j_{g-1}} \pi_{j_{g-1}} P_{j_{g-1},j_{t-1}}(t-g).$$

Therefore when $1 \leq g < t$, we obtain

$$[c_t^+]^\alpha \sum_{g=1}^{t-1} c_g \sum_{a=1}^{M} \sum_{b=1}^{M} p_b k_b \mu_a \pi_a P_{a,b}(t-g). \tag{7.25}$$

Now let $t < g \leq n$ and repeat:

$$\begin{aligned}
&[c_t^+]^\alpha \sum_{g=t+1}^{n} c_g \sum_{j_0} \cdots \sum_{j_{n-1}} p_{j_{t-1}} k_{j_{t-1}} \mu_{j_{g-1}} \pi_{j_0} P_{j_0,j_1} \cdots P_{j_{n-2},j_{n-1}} \\
&\quad = [c_t^+]^\alpha \sum_{g=t+1}^{n} c_g \sum_{j_0} \cdots \sum_{j_{t-1}} \cdots \sum_{j_{g-1}} p_{j_{t-1}} k_{j_{t-1}} \mu_{j_{g-1}} \pi_{j_0} P_{j_0,j_1} \cdots P_{j_{t-1},j_t} \cdots P_{j_{g-2},j_{g-1}} \\
&\quad = [c_t^+]^\alpha \sum_{g=t+1}^{n} c_g \sum_{j_{t-1}} \cdots \sum_{j_{g-1}} p_{j_{t-1}} k_{j_{t-1}} \mu_{j_{g-1}} \pi_{j_{t-1}} P_{j_{t-1},j_t} \cdots P_{j_{g-2},j_{g-1}} \\
&\quad = [c_t^+]^\alpha \sum_{g=t+1}^{n} c_g \sum_{j_{t-1}} \sum_{j_{g-1}} p_{j_{t-1}} k_{j_{t-1}} \mu_{j_{g-1}} \pi_{j_{t-1}} P_{j_{t-1},j_{g-1}}(g-t).
\end{aligned}$$

Therefore when $t < g \leq n$, we get

$$[c_t^+]^\alpha \sum_{g=t+1}^{n} c_g \sum_{a=1}^{M} \sum_{b=1}^{M} p_a k_a \mu_b \pi_a P_{a,b}(g-t). \tag{7.26}$$

We combine (7.25) and (7.26) to get the $i = t$ contribution

$$[c_t^+]^\alpha \sum_{g \neq t} E h_g^+(J_{t-1}, J_{g-1}) = [c_t^+]^\alpha \sum_{g=1}^{t-1} c_g \sum_{a=1}^{M} \sum_{b=1}^{M} p_b k_b \mu_a \pi_a P_{a,b}(t-g) \\ + [c_t^+]^\alpha \sum_{g=t+1}^{n} c_g \sum_{a=1}^{M} \sum_{b=1}^{M} p_a k_a \mu_b \pi_a P_{a,b}(g-t). \tag{7.27}$$

Finally, we compute the $i = n$ case:

$$\begin{aligned}
&[c_n^+]^\alpha \sum_{g \neq n} E\left[h_g^+(J_{n-1}, J_{g-1})\right] \\
&\quad = [c_n^+]^\alpha \sum_{g=1}^{n-1} c_g \sum_{j_0} \cdots \sum_{j_{n-1}} p_{j_{n-1}} k_{j_{n-1}} \mu_{j_{g-1}} \pi_{j_0} P_{j_0,j_1} \cdots P_{j_{n-2},j_{n-1}} \\
&\quad = [c_n^+]^\alpha \sum_{g=1}^{n-1} c_g \sum_{j_{g-1}} \cdots \sum_{j_{n-1}} p_{j_{n-1}} k_{j_{n-1}} \mu_{j_{g-1}} \pi_{j_{g-1}} P_{j_{g-1},j_g} \cdots P_{j_{n-2},j_{n-1}} \\
&\quad = [c_n^+]^\alpha \sum_{g=1}^{n-1} c_g \sum_{j_{g-1}} \sum_{j_{n-1}} p_{j_{n-1}} k_{j_{n-1}} \mu_{j_{g-1}} \pi_{j_{g-1}} P_{j_{g-1},j_{n-1}}(n-g).
\end{aligned}$$

The contribution from $i = n$ is therefore

$$[c_n^+]^\alpha \sum_{g \neq n} E\left[h_g^+(J_{n-1}, J_{g-1})\right] = [c_n^+]^\alpha \sum_{g=1}^{n-1} c_g \sum_{a=1}^{M} \sum_{b=1}^{M} p_b k_b \mu_a \pi_a P_{a,b}(n-g). \tag{7.28}$$

Put (7.24), (7.27), and (7.28) together to get

$$\sum_{i=1}^{n} \sum_{g \neq i} [c_i^+]^\alpha E h_g^+(J_{i-1}, J_{g-1}) = \sum_{i=1}^{n-1} \sum_{g=i+1}^{n} [c_i^+]^\alpha c_g \sum_{a=1}^{M} \sum_{b=1}^{M} p_a k_a \mu_b \pi_a P_{a,b}(|i-g|) \\ + \sum_{i=2}^{n} \sum_{g=1}^{i-1} [c_i^+]^\alpha c_g \sum_{a=1}^{M} \sum_{b=1}^{M} p_b k_b \mu_a \pi_a P_{a,b}(|i-g|). \tag{7.29}$$

Using similar derivations,

$$\sum_{i=1}^{n}\sum_{g\neq i}[c_i^-]^\alpha E\left[h_g^-(J_{i-1},J_{g-1})\right] = \sum_{i=1}^{n-1}\sum_{g=i+1}^{n}[c_i^-]^\alpha c_g \sum_{a=1}^{M}\sum_{b=1}^{M} q_a k_a \mu_b \pi_a P_{a,b}(|i-g|) + \sum_{i=2}^{n}\sum_{g=1}^{i-1}[c_i^-]^\alpha c_g \sum_{a=1}^{M}\sum_{b=1}^{M} q_b k_b \mu_a \pi_a P_{a,b}(|i-g|). \tag{7.30}$$

The computational version of Theorem 7.5 follows, using (7.29) and (7.30).

Theorem 7.6. *Suppose $\alpha \geq 1$ so that each of the underlying distributions has finite mean μ_j. Then as $x \to \infty$*

$$\begin{aligned} P\left(\sum_{i=1}^{n} c_i X_i > x\right) &= \left\{d_1 \sum_{i=1}^{n}[c_i^+]^\alpha + d_2 \sum_{i=1}^{n}[c_i^-]^\alpha\right\} F_*(x) \\ &+ \frac{\alpha F_*(x)}{x}\left\{\sum_{i=1}^{n-1}\sum_{g=i+1}^{n}\sum_{a=1}^{M}\sum_{b=1}^{M} c_g\left[[c_i^+]^\alpha p_a + [c_i^-]^\alpha q_a\right] k_a \mu_b \pi_a P_{a,b}(|i-g|)\right. \\ &+ \sum_{i=2}^{n}\sum_{g=1}^{i-1}\sum_{a=1}^{M}\sum_{b=1}^{M} c_g\left[[c_i^+]^\alpha p_b + [c_i^-]^\alpha q_b\right] k_b \mu_a \pi_a P_{a,b}(|i-g|) \\ &\left. -\overline{\mu} d_1 \sum_{i=1}^{n}\sum_{g=1}^{n}[c_i^+]^\alpha c_g^- + \overline{\mu} d_2 \sum_{i=1}^{n}\sum_{g=1}^{n}[c_i^-]^\alpha c_g^+\right\} + o\left(\frac{F_*(x)}{x}\right). \end{aligned}$$

7.3.2 The Case With $\alpha = 1$ and Infinite Means

In this section, we state the proven results from the previous section under the assumptions that $\alpha = 1$ and the means μ_j are infinite. This happens when the underlying distribution $F_*(x)$ has an infinite mean. As stated in Theorem 3.3, the means μ_j are replaced by truncated means $\mu_j(x) = \int_{-x}^{x} tF_{j*}(t)dt$. Using this same idea, define the truncated average mean $\overline{\mu}(x) = \sum_{j=1}^{M}\mu_j(x)\pi_j$. Since our results are generalizations of results from Barbe and McCormick (2005), it suffices to replace the μ_j with $\mu_j(x)$ and $\overline{\mu}$ with $\overline{\mu}(x)$.

Let the expected values be defined as $h_g^+(x, y, t) = c_g p(x) k(x) \mu(y, t)$ where $p(x)k(x) = p_{j_x} k_{j_x}$ and $\mu(y) = \mu_{j_y}(t)$. Similarly, $h_g^-(x, y, t) = c_g q(x) k(x) \mu(y, t)$ where $q(x)k(x) = q_{j_x} k_{j_x}$.

Theorem 7.7. *For $\alpha = 1$ and x large enough, the theoretical two-term expansion is*

$$
\begin{aligned}
P\left(\sum_{i=1}^n c_i X_i > x\right) &= \left(d_1 \sum_{i=1}^n c_i^+ + d_2 \sum_{i=1}^n c_i^-\right) F_*(x) \\
&+ \frac{F_*(x)}{x} \left\{ \sum_{i=1}^n \sum_{g \neq i} c_i^+ E\left[h_g^+(J_{i-1}, J_{g-1})\right] + \sum_{i=1}^n \sum_{g \neq i} c_i^- E\left[h_g^-(J_{i-1}, J_{g-1})\right] \right. \\
&\left. - d_1 \overline{\mu}(x) \sum_{i=1}^n \sum_{g=1}^n c_i^+ c_g^- + d_2 \overline{\mu}(x) \sum_{i=1}^n \sum_{g=1}^n c_i^- c_g^+ \right\} + o\left(\frac{F_*(x)}{x}\right).
\end{aligned}
$$

The computational result is

$$
\begin{aligned}
P\left(\sum_{i=1}^n c_i X_i > x\right) &= \left(d_1 \sum_{i=1}^n c_i^+ + d_2 \sum_{i=1}^n c_i^-\right) F_*(x) \\
&+ \frac{F_*(x)}{x} \left\{ \sum_{i=1}^{n-1} \sum_{g=i+1}^{n} \sum_{a=1}^{M} \sum_{b=1}^{M} c_g \left[c_i^+ p_a + c_i^- q_a\right] k_a \mu_b(x) \pi_a P_{a,b}(|i-g|) \right. \\
&+ \sum_{i=2}^{n-1} \sum_{g=1}^{i-1} \sum_{a=1}^{M} \sum_{b=1}^{M} c_g \left[c_i^+ p_b + c_i^- q_b\right] k_b \mu_a(x) \pi_a P_{a,b}(|i-g|) \\
&\left. - d_1 \overline{\mu}(x) \sum_{i=1}^n \sum_{g=1}^n c_i^+ c_g^- + d_2 \overline{\mu}(x) \sum_{i=1}^n \sum_{g=1}^n c_i^- c_g^+ \right\} + o\left(\frac{F_*(x)}{x}\right).
\end{aligned}
$$

7.3.3 The Case With $0 < \alpha < 1$

Up until now we have assumed that $\alpha \geq 1$, adjusting to the particular case when $\alpha = 1$ but the means are infinite. It is now time to consider the case when $0 < \alpha < 1$. First, we reference Theorem 3.4, restated here.

Theorem 7.8. *Define the quantity*

$$I(\alpha) = \int_0^{1/2} \left((1-y)^{-\alpha} - 1\right) \alpha y^{-\alpha-1} dy. \tag{7.31}$$

Now let F_1 and F_2 be regularly varying, asymptotically smooth distribution functions supported on the nonnegative real line. Then

$$\lim_{x\to\infty} \frac{1 - F_1 * F_2(x) - \overline{F}_1(x) - \overline{F}_2(x)}{\overline{F}_1(x)\overline{F}_2(x)} = 2I(\alpha) + 2^{2\alpha} - 2^{\alpha+1}.$$

For our work, all the underlying distributions have the same index $-\alpha$. Define

$$\Theta = 2I(\alpha) + 2^{2\alpha} - 2^{\alpha+1}. \tag{7.32}$$

If $Z_i \sim F_i, i = 1, 2$ are independent, the result of Theorem 7.8 can be rearranged to

$$P(Z_1 + Z_2 > x) = \overline{F}_1(x) + \overline{F}_2(x) + \Theta\overline{F}_1(x)\overline{F}_2(x)[1 + o(1)].$$

For $c_i > 0, i = 1, 2,$

$$P(c_1Z_1 + c_2Z_2 > x) = c_1^\alpha\overline{F}_1(x) + c_2^\alpha\overline{F}_2(x) + \Theta c_1^\alpha c_2^\alpha\overline{F}_1(x)\overline{F}_2(x)[1 + o(1)]. \tag{7.33}$$

We now leap to n variables.

Lemma 7.7. *For $c_i \geq 0$ and independent random variables $Z_i \sim F_i, i = 1, ..., n,$ all regularly* *varying of index $-\alpha$ where $0 < \alpha < 1$ and satisfying the assumptions in Theorem 7.8,*

$$P\left(\sum_{i=1}^{n} c_i Z_i > x\right) = \sum_{i=1}^{n} c_i^\alpha \overline{F}_i(x) + \Theta \sum_{i=1}^{n-1} \sum_{g=i+1}^{n} c_i^\alpha c_g^\alpha \overline{F}_i(x) \overline{F}_g(x)[1 + o(1)].$$

Proof. The case where $n = 2$ is (7.33). To establish the induction step, consider $c_{n+1} \geq 0$ and random variable $Z_{n+1} \sim F_{n+1}$, independent of the previous variables. Then, neglecting the terms where more than two distribution tails are multiplied together,

$$\begin{aligned}
P\left(\sum_{i=1}^{n+1} c_i Z_i > x\right) &= P\left(\sum_{i=1}^{n} c_i Z_i > x\right) + P\left(c_{n+1} Z_{n+1} > x\right) \\
&\quad + \Theta P\left(\sum_{i=1}^{n} c_i Z_i > x\right) P\left(c_{n+1} Z_{n+1} > x\right)[1 + o(1)] \\
= \sum_{i=1}^{n+1} c_i^\alpha \overline{F}_i(x) &+ \Theta \Bigg\{ \sum_{i=1}^{n-1} \sum_{g=i+1}^{n} c_i^\alpha c_g^\alpha \overline{F}_i(x) \overline{F}_g(x) + \sum_{i=1}^{n} c_i^\alpha c_{n+1}^\alpha \overline{F}_i(x) \overline{F}_{n+1}(x) \\
&\quad + c_{n+1}^\alpha \overline{F}_{n+1}(x) \sum_{i=1}^{n-1} \sum_{g=i+1}^{n} c_i^\alpha c_g^\alpha \overline{F}_i(x) \overline{F}_g(x) \Bigg\} [1 + o(1)] \\
= \sum_{i=1}^{n+1} c_i^\alpha \overline{F}_i(x) &+ \Theta \Bigg\{ \sum_{i=1}^{n-1} \sum_{g=i+1}^{n} c_i^\alpha c_g^\alpha \overline{F}_i(x) \overline{F}_g(x) \\
&\quad + \sum_{i=1}^{n} c_i^\alpha c_{n+1}^\alpha \overline{F}_i(x) \overline{F}_{n+1}(x) \Bigg\} [1 + o(1)] \\
= \sum_{i=1}^{n+1} c_i^\alpha \overline{F}_i(x) &+ \Theta \sum_{i=1}^{n} \sum_{g=i+1}^{n+1} c_i^\alpha c_g^\alpha \overline{F}_i(x) \overline{F}_g(x)[1 + o(1)].
\end{aligned}$$

□

Adapting to our work, and under the assumptions (7.18) and (7.20), Lemma 7.7 provides

$$P\left(\sum_{i=1}^{n} c_i X_i > x \middle| J_l, 0 \le l \le i-1\right) = \sum_{i=1}^{n} p_{j_{i-1}} k_{j_{i-1}} [c_i^+]^\alpha F_*(x) + \Theta \sum_{i=1}^{n-1} \sum_{g=i+1}^{n} [c_i^+ c_g^+]^\alpha p_{j_{i-1}} p_{j_{g-1}} k_{j_{i-1}} k_{j_{g-1}} [F_*(x)]^2 + o\left([F_*(x)]^2\right). \tag{7.34}$$

To uncondition (7.34) define $h^+(x,y) = p_x p_y k_x k_y$, then

$$P\left(\sum_{i=1}^{n} c_i X_i > x\right) = \left(d_1 \sum_{i=1}^{n} [c_i^+]^\alpha\right) F_*(x) + \Theta [F_*(x)]^2 \sum_{i=1}^{n-1} \sum_{g=i+1}^{n} [c_i^+ c_g^+]^\alpha E[h^+(J_{i-1}, J_{g-1})] + o\left([F_*(x)]^2\right). \tag{7.35}$$

Similarly, letting the constants $\{c_i\}$ be all negative, it can be shown that

$$P\left(\sum_{i=1}^{n} c_i X_i > x \middle| J_l = j_l, 0 \le l \le i-1\right) = \sum_{i=1}^{n} q_{j_{i-1}} k_{j_{i-1}} [c_i^-]^\alpha F_*(x) + \Theta \sum_{i=1}^{n-1} \sum_{g=i+1}^{n} [c_i^- c_g^-]^\alpha q_{j_{i-1}} q_{j_{g-1}} k_{j_{i-1}} k_{j_{g-1}} [F_*(x)]^2 + o\left([F_*(x)]^2\right),$$

and consequently, letting $h^-(x,y) = q_x q_y k_x k_y$,

$$P\left(\sum_{i=1}^{n} c_i X_i > x\right) = \left(d_2 \sum_{i=1}^{n} [c_i^-]^\alpha\right) F_*(x) + \Theta [F_*(x)]^2 \sum_{i=1}^{n-1} \sum_{g=i+1}^{n} [c_i^- c_g^-]^\alpha E[h^-(J_{i-1}, J_{g-1})] + o\left([F_*(x)]^2\right). \tag{7.36}$$

Moving to any real constants, define $U_n = \sum_{i=1}^{n} [c_i^+] X_i$ and $V_n = -\sum_{i=1}^{n} [c_i^-] X_i$. Then (7.35) and (7.36) provide

$$
\begin{aligned}
P(U_n+V_n>x) &= P(U_n>x)+P(V_n>x)+\Theta P(U_n>x)P(V_n>x)[1+o(1)] \\
&= \left(d_1\sum_{i=1}^{n}[c_i^+]^\alpha + d_2\sum_{i=1}^{n}[c_i^-]^\alpha\right)F_*(x) \\
&+ \Theta[F_*(x)]^2\left\{\sum_{i=1}^{n-1}\sum_{g=i+1}^{n}\left([c_i^+c_g^+]^\alpha E[h^+(J_{i-1},J_{g-1})]] + [c_i^-c_g^-]^\alpha E[h^-(J_{i-1},J_{g-1})]\right)\right. \\
&\left. + d_1d_2\sum_{i=1}^{n}\sum_{g=1}^{n}[c_i^+c_g^-]^\alpha\right\} + o\left([F_*(x)]^2\right).
\end{aligned}
\tag{7.37}
$$

Lastly, if one desires a computational version of (7.37), one can use simplifying techniques as in the proof of Theorem 7.6 to obtain

$$
\begin{aligned}
P(U_n+V_n>x) &= \left(d_1\sum_{i=1}^{n}[c_i^+]^\alpha + d_2\sum_{i=1}^{n}[c_i^-]^\alpha\right)F_*(x) \\
&+ \Theta[F_*(x)]^2\left\{\sum_{i=1}^{n-1}\sum_{g=i+1}^{n}\sum_{a=1}^{M}\sum_{b=1}^{M}\left([c_i^+c_g^+]^\alpha p_ap_b + [c_i^-c_g^-]^\alpha q_aq_b\right)k_ak_b\pi_aP_{a,b}(g-i)\right. \\
&\left. + d_1d_2\sum_{i=1}^{n}\sum_{g=1}^{n}[c_i^+c_g^-]^\alpha\right\} + o\left([F_*(x)]^2\right).
\end{aligned}
\tag{7.38}
$$

We have now established the following result.

Theorem 7.9. *For $0<\alpha<1$ and x large enough, the theoretical two-term expansion and computational result are as given in (7.36) and (7.37), respectively.*

7.4 Open Questions

There are a couple of additional conjectures we have about the chain dependent regularly varying random variables, both of which currently remain unproven. Recall that the first-order approximation was completely derived, even for an infinite sum. The second-order formula, however, has only been worked out for the sum of n variables. It is strongly suspected that the infinite sum holds, but at this time it is not clear specifically what additional assumptions need to be made. As an example, taking $\alpha \geq 1$ and the underlying means to be finite, it is conjectured that under suitable conditions, as $x \to \infty$

$$\begin{aligned} P(Y_1 > x) &= \left(d_1 \sum_{i=-\infty}^{\infty} [c_i^+]^\alpha + d_2 \sum_{i=-\infty}^{\infty} [c_i^-]^\alpha \right) F_*(x) \\ &+ \frac{\alpha F_*(x)}{x} \Bigg\{ \sum_{i=-\infty}^{\infty} \sum_{g=i+1}^{\infty} \sum_{a=1}^{M} \sum_{b=1}^{M} c_g \left[[c_i^+]^\alpha p_a + [c_i^-]^\alpha q_a \right] k_a \mu_b \pi_a P_{a,b}(|i-g|) \\ &+ \sum_{i=-\infty}^{\infty} \sum_{g=-\infty}^{i-1} \sum_{a=1}^{M} \sum_{b=1}^{M} c_g \left[[c_i^+]^\alpha p_b + [c_i^-]^\alpha q_b \right] k_b \mu_a \pi_a P_{a,b}(|i-g|) \\ &- \overline{\mu} d_1 \sum_{i=-\infty}^{\infty} \sum_{g=1}^{n} [c_i^+]^\alpha c_g^- + \overline{\mu} d_2 \sum_{i=-\infty}^{\infty} \sum_{g=1}^{n} [c_i^-]^\alpha c_g^+ \Bigg\} + o\left(\frac{F_*(x)}{x} \right). \end{aligned}$$

The outcome of Theorem 7.4 required $\sum_{i=-\infty}^{\infty} |c_i|^\lambda < \infty$ for some $0 < \lambda < \min(\alpha, 1)$. It is believed that a similar summability condition is needed in the three possible cases. Further assumptions on asymptotic smoothness and right-tail dominance may come into play as well.

The second conjecture concerns an interesting application of the three-case second order approximations. We derived computational results, but they all depend on the individual entries of the underlying Markov chain. Even if these probabilities are known, we still have to store possibly massive amounts of matrices to get the specific Markov chain for moving in k steps. That is, we would need to store in computer memory $P^k, k = 1, 2, \cdots$. We propose that all three approximations can be further approximated by simply replacing the transition probability $P_{a,b}(k)$ with π_b. Of course, some amount of error is to be expected from such an operation, but we believe that the amount is negligible for large enough x.

To illustrate how the conjecture works, let $Q_n(x)$ denote $P\left(\sum_{i=1}^n c_i X_i > x\right)$ but with the transition probabilities replaced with the stationary probabilities. When $\alpha \geq 1$ and the means are finite, we have

$$\begin{aligned} Q_n(x) &= \left(d_1 \sum_{i=1}^n [c_i^+]^\alpha + d_2 \sum_{i=1}^n [c_i^-]^\alpha\right) F_*(x) \\ &+ \alpha\overline{\mu}\left\{d_1\left(\sum_{i=1}^n \sum_{g\neq i} [c_i^+]^\alpha c_g - \sum_{i=1}^n \sum_{g=1}^n [c_i^+]^\alpha c_g^-\right)\right. \\ &\left. + d_2\left(\sum_{i=1}^n \sum_{g\neq i} [c_i^-]^\alpha c_g + \sum_{i=1}^n \sum_{g=1}^n [c_i^-]^\alpha c_g^+\right)\right\} \frac{F_*(x)}{x} + o\left(\frac{F_*(x)}{x}\right). \end{aligned} \tag{7.39}$$

When $\alpha = 1$ and the means are infinite,

$$\begin{aligned} Q_n(x) = &\left(d_1 \sum_{i=1}^n c_i^+ + d_2 \sum_{i=1}^n c_i^-\right) F_*(x) + \left\{d_1\left(\sum_{i=1}^n \sum_{g\neq i} c_i^+ c_g - \sum_{i=1}^n \sum_{g=1}^n c_i^+ c_g^-\right)\right. \\ &\left. + d_2\left(\sum_{i=1}^n \sum_{g\neq i} c_i^- c_g + \sum_{i=1}^n \sum_{g=1}^n c_i^- c_g^+\right)\right\} \frac{\overline{\mu}(x) F_*(x)}{x} + o\left(\frac{F_*(x)}{x}\right). \end{aligned} \tag{7.40}$$

When $0 < \alpha < 1$,

$$\begin{aligned} Q_n(x) = &\left(d_1 \sum_{i=1}^n [c_i^+]^\alpha + d_2 \sum_{i=1}^n [c_i^-]^\alpha\right) F_*(x) + \Theta\left\{\sum_{i=1}^{n-1} \sum_{g=i+1}^n \left(d_1^2 [c_i^+ c_g^+]^\alpha + d_2^2 [c_i^- c_g^-]^\alpha\right)\right. \\ &\left. + d_1 d_2 \sum_{i=1}^n \sum_{g=1}^n [c_i^+ c_g^-]^\alpha\right\} [F_*(x)]^2 + o\left([F_*(x)]^2\right). \end{aligned} \tag{7.41}$$

One obvious question that arises from using an approximation to the approximation, as it were, is what kind of error results in doing so. Clearly using the $Q_n(x)$ should result in some sort of error, but at this time it is not clear as to that discrepancy's exact behavior. The error appears through $|P_{i,j}(n) - \pi_j|$. There are numerous published results on this topic, and we suspect that the answer lies in at least one of these. For instance, Roberts and Tweedie (1999) discussed the problem in the geometrically ergodic case. They proved that for some function $M(x)$ and some $\rho < 1$, $|P_{i,j}(n) - \pi_j| \leq M(x)\rho^n$. For further discussion of this matter, see Meyn and Tweedie (1994) and Lund and Tweedie (1996).

Lastly, there are instances where one may want to pursue further terms in the expansion. First, a third term may provide a more accurate approximation if both the first and second-orders are a little inaccurate. And second, there are instances where the second term may vanish, necessitating the need for an extra term. One example is when the underlying distributions are T with degrees of freedom at least 2, in which case the means are all 0. Barbe and McCormick (2009) discuss this problem and provide several examples.

7.5 Examples of Chain-Dependent Processes

We now present some examples of distributions that satisfy the requirements of the regularly varying setting.

Example 7.1. Suppose $Z_1, \cdots, Z_n \overset{iid}{\sim} F_*$ with $F_* \sim kx^{-\alpha}$. Then the Markov chain has only one state, and the first-order approximation is

$$P\left(\sum_{i=1}^{n} c_i Z_i > x\right) = k\left(p\sum_{i=1}^{n}[c_i^+]^\alpha + q\sum_{i=1}^{n}[c_i^-]^\alpha\right)x^{-\alpha} + o(x^{-\alpha}).$$

If F is defined only on the positive half line, then the approximation reduces to

$$P\left(\sum_{i=1}^{n} c_i Z_i > x\right) = \left(k\sum_{i=1}^{n}[c_i^+]^\alpha\right)x^{-\alpha} + o(x^{-\alpha}),$$

which is the result from Resnick (1987). Because of the independence, the second-order approximations reduce to the results given in Barbe and McCormick (2009) provided $\alpha > 1$. If $n = 2$, then for any $\alpha > 0$ we recover the results in Barbe and McCormick (2005).

Example 7.2. The standard Pareto distribution is $\overline{F}(x) = x^{-\alpha}, \alpha > 0, x \geq 1$. Clearly $\overline{F} \in RV_{-\alpha}$ with $L(x) = 1$, and the mean is $\frac{\alpha}{\alpha-1}$ for $\alpha > 1$. If $\alpha = 1$, then the truncated mean is $\log(x)$.

Example 7.3. Define the Cauchy distribution as

$$\overline{F}(x) = \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{x-\mu}{\sigma}\right), \sigma > 0, -\infty < x < \infty.$$

It can be shown that $\overline{F}(x) \in RV_{-1}$, and further that $\overline{F}(x) \sim \frac{\sigma}{\pi} x^{-1}$. The truncated mean is $\frac{1}{\pi} \log(1 + x^2)$.

Example 7.4. The T distribution on 2 degrees of freedom is defined by

$$\overline{F}(x) = \frac{1}{2}\left[1 - \frac{x}{\sqrt{2 + x^2}}\right], -\infty < x < \infty.$$

It can be shown that $\overline{F}(x) \in RV_{-2}$ and $\overline{F}(x) \sim \frac{1}{2} x^{-2}$. This example illustrates a potential problem with using the second-order expansion, which depends on the means of the underlying distributions. Because the mean of the T distribution is 0, the second term in the expansion is also 0. This is an instance of where further terms in the expansion might be appropriate, as mentioned in the previous section.

Example 7.5. Now we introduce the AR(1) process to the Markov chain scenario. If $\alpha > 1$, then set $c_i = \theta^{i-1}$ for $0 < \theta < 1$. Then Theorem 7.6 reduces to

$$\begin{aligned} P\left(\sum_{i=1}^{n} c_i X_i > x\right) &= d_1 \left(\frac{1-\theta^{\alpha n}}{1-\theta^{\alpha}}\right) F_*(x) \\ &+ \frac{\alpha F_*(x)}{x} \left\{ \sum_{i=1}^{n-1} \sum_{g=i+1}^{n} \sum_{a=1}^{M} \sum_{b=1}^{M} \theta^{\alpha(i-1)+g-1} p_a k_a \mu_b \pi_a P_{a,b}(|i-g|) \right. \\ &\quad \left. + \sum_{i=2}^{n} \sum_{g=1}^{i-1} \sum_{a=1}^{M} \sum_{b=1}^{M} \theta^{\alpha(i-1)+g-1} p_b k_b \mu_a \pi_a P_{a,b}(|i-g|) \right\} + o\left(\frac{F_*(x)}{x}\right). \end{aligned} \tag{7.42}$$

Assuming that (7.39) holds, a more succinct way of writing (7.42) is

$$Q_n(x) = d_1 \left(\frac{1-\theta^{\alpha n}}{1-\theta^{\alpha}}\right) F_*(x) + \frac{\alpha d_1 \overline{\mu} F_*(x)}{x} \left(\sum_{i=1}^{n} \sum_{g \neq i} \theta^{\alpha(i-1)+g-1}\right) + o\left(\frac{F_*(x)}{x}\right).$$

Simplifying the double sum,

$$\begin{aligned} Q_n(x) &= \alpha d_1 \overline{\mu} \theta^{\alpha} \left(\frac{1-\theta^{\alpha(n-1)} - \theta^{n-1} + \theta^{(\alpha-1)n-\alpha-1}}{(1-\theta)(1-\theta^{\alpha})} + \frac{1}{1-\theta^{\alpha+1}}\right) \frac{F_*(x)}{x} \\ &\quad + d_1 \left(\frac{1-\theta^{\alpha n}}{1-\theta^{\alpha}}\right) F_*(x) + o\left(\frac{F_*(x)}{x}\right). \end{aligned} \tag{7.43}$$

Finally, if the infinity conjecture holds, then as $n \to \infty$

$$Q_n(x) \to \frac{d_1 F_*(x)}{1-\theta^{\alpha}} + \left(\frac{\alpha d_1 \overline{\mu} \theta^{\alpha}(2-\theta^{\alpha+1}-\theta+\theta^{\alpha+1}-\theta^{\alpha})}{(1-\theta)(1-\theta^{\alpha})(1-\theta^{\alpha+1})}\right) \frac{F_*(x)}{x} + o\left(\frac{F_*(x)}{x}\right). \tag{7.44}$$

Example 7.6. Next, we consider the AR(1) process where $-1 < \theta < 0$, and thus $c_i = \theta^{i-1}$. This time we illustrate Theorem 7.9 and take $0 < \alpha < 1$. We also assume that the $Q_n(x)$ conjecture holds and, for simplicity, take infinite sums. One has to be careful defining the constants as the $\{c_i\}$ alternate signs. Observe that $c_i^+ = \theta^{i-1} I_{[\text{i odd}]}$ and $c_i^- = \theta^{i-1} I_{[\text{i even}]}$

where I is the indicator function. Then as $n \to \infty$

$$Q_n(x) \to \left(\frac{d_1 + \theta^\alpha d_2}{1-\theta^{2\alpha}}\right) F_*(x) + \Theta\theta^\alpha \left(\frac{(d_1^2+d_2^2)\theta^\alpha + d_1 d_2}{1-\theta^{2\alpha}}\right) [F_*(x)]^2 + o\left([F_*(x)]^2\right). \quad (7.45)$$

Example 7.7. This final example illustrates a specific Markov chain with state space 3, and also provides motivation for obtaining a future proof of the $Q_n(x)$. Because of the multitude of variables needed for the approximation, we do not use the Peachtree Creek data set like we did in earlier chapters, but instead we perform a simulation study on this specific example. Define the transition probability matrix and corresponding stationary distribution by

$$P = \begin{pmatrix} 0.2 & 0.5 & 0.3 \\ 0.5 & 0 & 0.5 \\ 0.4 & 0.4 & 0.2 \end{pmatrix} \quad \text{and} \quad \vec{\pi} = \begin{pmatrix} 0.35928 & 0.31138 & 0.32934 \end{pmatrix}.$$

Take $F_*(x) = x^{-\alpha}, x \geq 1$, the Pareto distribution. Now define the underlying distributions to be

$$\overline{F}_j(x) = \left(\frac{2j+1}{x}\right)^\alpha, x \geq 2j+1, j = 1, 2, 3. \quad (7.46)$$

Thus, the Markov chain distributions are related to the parent Pareto distribution via $\overline{F}_j(x) \sim k_j \overline{F}(x)$ where $k_j = (2j+1)^\alpha$ for x large enough. Next, define the 20-dependent moving average process $Y_n = \sum_{i=0}^{19} \left(\frac{1}{3}\right)^i X_{n-i}$. Finally, take $\alpha = 1.5$. Then $p_j = 1$, $q_j = 0$, $[c_i^+]^\alpha = \left(\frac{1}{3}\right)^{1.5(i-1)}$, $[c_i^-]^\alpha = 0$, $d_1 = 11.44767$, and $d_2 = 0$. Further $\mu_j = 3(2j+1)$. The trickiest part of the formula is computing the $P_{a,b}(|i-g|)$, for this requires utilizing the spectral decomposition on $P^{|i-g|}$. It can be shown that

$$P_{1,1} = 0.35928 + 0.43680(-0.15858)^{|i-g|} + 0.20392(-0.44142)^{|i-g|}$$

$$P_{2,1} = 0.35928 - 0.06320(-0.15858)^{|i-g|} - 0.29608(-0.44142)^{|i-g|}$$

$$P_{3,1} = 0.35928 - 0.41675(-0.15858)^{|i-g|} + 0.05747(-0.44142)^{|i-g|}$$

$$P_{1,2} = 0.31138 + 0.18093(-0.15858)^{|i-g|} - 0.49231(-0.44142)^{|i-g|}$$

$$P_{2,2} = 0.31138 - 0.02618(-0.15858)^{|i-g|} + 0.71480(-0.44142)^{|i-g|}$$

$$P_{3,2} = 0.31138 - 0.17263(-0.15858)^{|i-g|} - 0.13875(-0.44142)^{|i-g|}$$

$$P_{1,3} = 0.32934 - 0.61773(-0.15858)^{|i-g|} + 0.28839(-0.44142)^{|i-g|}$$

$$P_{2,3} = 0.32934 + 0.08938(-0.15858)^{|i-g|} - 0.41872(-0.44142)^{|i-g|}$$

$$P_{3,3} = 0.32934 + 0.58938(-0.15858)^{|i-g|} + 0.08128(-0.44142)^{|i-g|}.$$

The two-term expansion is

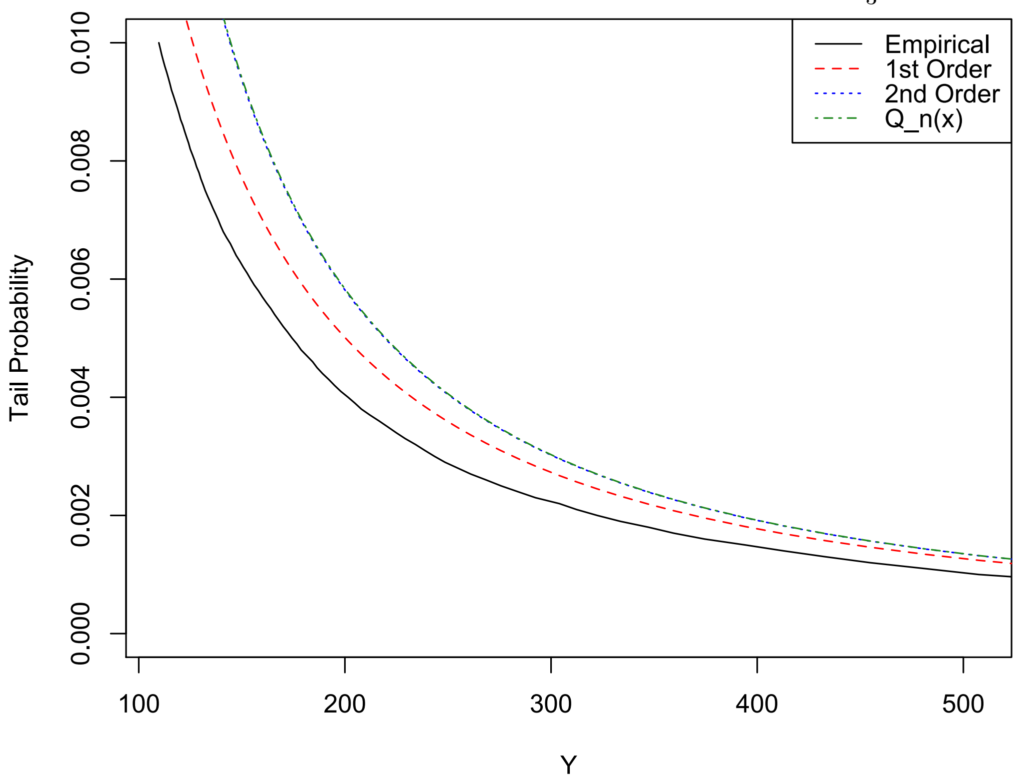
$$P\left(\sum_{i=1}^{n} c_i X_i > x\right) \approx 14.1758x^{-1.5} + 454.3261x^{-2.5}. \tag{7.47}$$

In this case, we have a three-state Markov chain, so storing the $P^{|i-g|}$ is not unreasonable. However, the calculations quickly grow out of control for larger states, and using the spectral theorem would be computationally intense. For this reason, we also compute the probability using the $Q_n(x)$ conjecture:

$$Q_n(x) \approx 14.1758x^{-1.5} + 465.2446x^{-2.5}. \tag{7.48}$$

While the second coefficients in (7.47) and (7.48) are slightly different, for x large enough the difference is going to be negligible. This example illustrates that using the $Q_n(x)$ approximation may be beneficial in practice provided the percentiles are high enough. Now we examine the three approximations in Figure 7.1.

Figure 7.1: Chain Dependent Convolution, $\beta = \frac{1}{3}$



Curiously, the first-order approximation is the closest to the truth, and even then the estimated probabilities are not very good. In this particular example the second-order is worse. The good news is that, as we conjectured, the $Q_n(x)$ formula is very close to the actual second-order approximation at the high percentiles. This observation provides hope that in general, we may be able to use the $Q_n(x)$ instead in order to make computation much easier. Tables 7.1 and 7.2, as usual, show the errors and relative errors in estimation. Note the very similar numbers for the second-order and the $Q_n(x)$ approximations.

Table 7.1: Errors in Approximations of Chain Dependent Convolution

Approx	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
1st	-0.0122	-0.0061	-0.0049	-0.0023	-0.0012	-0.0002	-1.6e-5
2nd	-0.0657	-0.0229	-0.0165	-0.0059	-0.0023	-0.0003	-1.7e-5
$Q_n(x)$	-0.0670	-0.0233	-0.0167	-0.0060	-0.0023	-0.0003	-1.8E-5

Table 7.2: Relative Errors in Approximations of Chain Dependent Convolution

Approx	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.9%**	**99.99%**
1st	-0.1967	-0.1955	-0.1957	-0.1890	-0.1878	-0.1936	-0.1412
2nd	-0.5679	-0.4779	-0.4515	-0.3723	-0.3139	-0.2415	-0.1523
$Q_n(x)$	-0.5727	-0.4823	-0.4556	-0.3757	-0.3164	-0.2426	-0.1525

We should question why the approximations in this example are a little off. There are unfortunately a number of possible reasons - the Markov chain, the choice of $\{c_i\}$, the choice of k_j, the α, even the underlying distribution. The multitude of possibilities explains why doing a general simulation on this matter is difficult. However, we are hopeful that in other situations the approximations will be better behaved. One reason why they did not work as well here is because all the constants happened to be positive, and the first-order approximation was already an overestimate. In an instance where the $\{c_i\}$ alternate signs, or where the underlying distribution has mass on both halves of the real line, we may see the desired results. For instance, this may happen when $c_i = \rho^{i-1}, i \geq 1, -1 < \rho < 0$. This topic of exactly when the formulas are reliable will be left as an open question.

Chapter 8 Conclusions

The five projects in the dissertation together have established some very significant results in extreme value theory, both theoretical and computational. As much as possible, the theory has been well balanced with illustrative examples and computational results. We started with a thorough data analysis of the Peachtree Creek, establishing that the creek's median height is increasing significantly in summer and fall. In doing so, we have arrived at a significant result that will be of interest to insurance companies and hydrologists. This project also enabled us to tie extreme value theory with maximum likelihood estimation.

The second project derived theoretical expansions of the AR(1) process where the random variables satisfied the Gumbel distribution. This type of process is ubiquitous in a variety of disciplines, and now we have results for its upper tail probabilities. Furthermore, we also established a realistic range of β values over which each approximation should or should not be used.

The third project established a general two-term expansion for the convolution of a particular class of random variables, of which the Gumbel is a special case. This theoretical result is analogous to the one in Rootzén (1986), only with a second term. However, we discovered by simulation that further terms were needed, so at this point we focused on getting n terms for the Gumbel situation. We concluded that at least some more terms were beneficial, but an interesting open question is precisely how many.

The fourth project gave an example of how to handle ties in the largest weights of the AR(1) process. There are plenty of open questions from this topic, not the least of which are working out more cases and carrying out a series of simulation studies to test for overall effectiveness.

Finally, the fifth project provided a fully worked-out first-order expansion for a process of Markov chain dependent regularly varying random variables. The second term in the expansion has also been derived, taking one of three different forms. We established theoretical and computational results, and we suspect that under certain additional assumptions the second term will hold for the infinite process. That extra proof would make the second-order analysis complete for theoretical purposes. In addition, that chapter provided some potentially useful alternative computational formulas, but the error expended in using these has yet to be studied.

All five projects contribute significantly to extreme value theory in one way or another, especially with results on the Gumbel distribution. However, plenty of open questions have arisen from these studies, and these would all be useful ideas for future study. We recap the most interesting questions below.

1. A followup study could be conducted with the Peachtree Creek, or any other creek or river, with more covariates. We suspect that variables like population density, along with some measure of the amount of concrete in the surrounding city, should be included.

2. One could study the AR(1) processes in Chapter 4, but with negative weights. We proved the result for standard Gumbels, but the nonstandard case remains open.

3. Precisely how many terms are needed in the expansion for the convolution of n Gumbel random variables should be investigated. The conjecture provided at the end of Chapter 5 should get this future study started.

4. The remaining four cases outlined at the beginning of Chapter 6 would give a more complete analysis of the AR(1) process with various combinations of ties in the weights.

5. The $Q_n(x)$ formulas should be further studied, with emphasis on the error resulting in going from the second-order approximation to the conjectured equivalents. We suspect that various results from papers by Tweedie will help start this project.

On top of all these suggested future studies, there is the opportunity to do more involved simulational results. The fourth and fifth projects in particular could benefit from more simulations, although as mentioned the combinations of parameters to keep track of are plentiful. Results from such future work would help us work out how and when each approximation should be used in practice.

Chapter 9 Bibliography

1. Asmussen, S. (2001). *Ruin Probabilities.* World Scientific, Singapore.

2. Balkema, A. and de Haan, L. (1974). Residual life time at great age. *Annals of Probability* **2**, 792804.

3. Barbe, P. and McCormick, W. P. (2005). Asymptotic expansions of convolutions of regularly varying distributions. *Australian Mathematical Society* **78**, 339-371.

4. Barbe, P. and McCormick, W. P. (2009). Asymptotic expansions for infinite weighted convolutions of heavy tail distributions and applications. *Memoirs American Mathematical Society* **197**, No. 922.

5. Berman, A. and Plemmons, R. J. (1979). *Nonnegative Matrices in the Mathematical Sciences.* Academic Press, New York.

6. Bingham, N. H., Goldie, C.M., and Teugels, J. L. (1989). *Regular Variation*, 2nd ed. Cambridge University Press, New York.

7. Borokov, A. A. (1976). *Stochastic process in Queueing Theory.* Springer, Berlin.

8. Broniatowski, M, and Fuchs, A. (1995). Tauberian theorems, Chernoff inequality, and the tail behavior of finite convolutions of distribution functions. *Advances in Mathematics* **116**, 12-33.

9. Bruun, J. T. and Tawn, J. A. (1998). Comparisons of approaches for estimating the probability of coastal flooding. *Applied Statistics* **47**, 405-423.

10. Butler, A., Heffernan, J. E., Tawn. J. A. and Flather, R. A. (2007). Trend estimation in extremes of synthetic North Sea surges. *Applied Statistics* **56**, 395-414.

11. Caers, J., Beirlant, J. and Maes, M. A. (1999). Statistics for modelling heavy tailed distributions in geology: Part I. Methodology. *Mathematical Geology* **31**, 391-410.

12. Chandra, M., Singpurwalla, N. D. and Stephens, M. A. (1981). Kolmogorov statistics for test of fit for the extreme-value and Weibull distributions. *Journal of the American Statistical Association* **76**, 729-731.

13. Choi, J. and Seo, T. Y. (1998). Integral formulas for Euler's constant. *Comm. Korean Math. Soc.* **13**, No 3, 683-689.

14. Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values.* Springer-Verlag London Limited, London.

15. Coles S and Davison A. (2008). Statistical modeling of extreme values [PDF document]. Retrieved from http://www.cces.ethz.ch/projects/hazri /EXTREMES/reports.

16. Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Volume II, 2nd ed. Wiley, New York.

17. Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society* **24**, 180-190.

18. Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématique* **6**, 93116.

19. Geluk, J. L., de Hann, L., Resnick, S. I., and Stărică, C. (1997). Second-order regular variation, convolution and the central limit theorem. *Stochastic Processes and their Applications* **69**, 139-159.

20. Gradshteyn, I. S. and Ryzhik, I. M. (1980). *Tables of Integrals, Series, and Products* (Corrected and Enlarged Ed. by A. Jeffrey). Academic Press, New York.

21. Gumbel, E. J. (1958). *Statistics of Extremes.* Columbia University Press, New York.

22. Hollis GE. 1975. The Effect of Urbanization on Floods of Different Recurrence Interval. *Water Resources Research* **11**: 431-435.

23. Hosking, J. R. M. (1984). Testing whether the shape parameter is zero in the generalized extreme-value distribution. *Biometrika* **72**, 367-374.

24. Katz, R. W, Parlange, M. B., and Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources* **25**: 1287-1304.

25. Kinnison, R. R. (1985). *Applied Extreme Value Statistics*. Macmillan, New York.

26. Kinnison, R. R. (1989). Correlation coefficient goodness-of-fit test for the extreme value distribution. *Journal of the American Statistical Association* **Vol 43, No. 2**, 98-100.

27. Lancaster, P. and Tismenetsky, M. (1985). *The Theory of Matrices*, 2nd ed. Academic Press, San Diego.

28. Leadbetter, M. R., Lindgren, G. and Rootzen, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, New York.

29. Lund, R. B. and Tweedie, R. L. (1996). Geometric convergence rates for stochastically ordered Markov chains. *Mathematics of Operations Research* **21**, 182-194.

30. McCormick, W. P. and Seymour, L. (2001). Extreme values for a class of shot-noise processes. *Selected Proceedings of the Symposium on Inference for Stochastic Processes*, Institute of Mathematical Statistics Lecture Notes-Monograph Series v. 37, 33-46.

31. McCormick, W. P. and Seymour, L. (2001). Rates of convergence and approximations to the distribution of the maximum of chain-dependent sequences. *Extremes* **4**, 23-52.

32. Meyn, S. P. and Tweedie, R. L. (1994). Computable bounds for convergence rates of Markov chains. *Annals of Applied Probability*, **4**, 981-1011.

33. Norros, I. (2003). Large deviations of queues with long-range dependent input. *Theory and Applications of Long-Range Dependence.* Doukhan, P., Oppenheim, G., Taqqu, M. S. editors. Birkhäuser, Boston.

34. Omey, E. (1988). Asymptotic properties of convolution products of functions. *Publications de l'Institut Mathématique (Beograd) (N.S.)* **43**, 41-57.

35. Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics* **3**, 119131.

36. Pinder, J. E., Wiener, J. G. and Smith, M. H. (1978). The Weibull distribution: a new method of summarizing survivorship data. *Ecological Society of America* **59**, 175-179.

37. Pisarenko, V. F. and Sornette, D. (2008). Characterization of the frequency of extreme events by the generalized Pareto distribution. *Pure and Applied Geophysics* **160**, 2343-2364.

38. Resnick, S. I. (1986). Point processes, regular variation and weak convergence. *Advances in Applied Probability* **18**, 66-138.

39. Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes.* Springer-Verlag, New York.

40. Resnick, S. I. (1992). *Adeventures in Stochastic Processes.* Birkhäuser, Boston.

41. Resnick, S. I. (1999). *A Probability Path.* Birkhäuser, Boston.

42. Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer Science + Business Media, LLC, New York.

43. Roberts, G. O. and Tweedie, R. L. (1999). Bounds on regeneration times and convergence rates for Markov chains. *Stochastic Processes and their Applications* **80**, 211-229.

44. Roman, S. (1980). The Formula of Faà di Bruno. *American Mathematical Monthly* **87**, 805-809.

45. Rootzén, H. (1986). Extreme value theory for moving average processes. *The Annals of Probability* **14**, 612-652.

46. Rootzén, H. (1988). Maxima and exceedances of stationary Markov chains. *Advances in Probability Theory* **20**, 371-390.

47. Seo, T. Y., Kang, J. S. and Choi, J. (1997). On the Euler's constant. *Pusan Kyongnam Math. J.* **13**, 33-41.

48. Smith, R. (1999). Trends in Rainfall Extremes. Preprint, University of North Carolina, Chapel Hill.

49. Smith, R. and Shivey, T. S. (1995). Point process approach to modeling trends in tropospheric ozone based on exceedances of a high threshold. *Atmospheric Environment* **29**, 3489-3499.

50. USGS. "USGS 02336300 Peachtree Creek at Atlanta, GA." *http://waterdata.usgs.gov/ga/nwis.*

51. Villarini G *et al.*. 2009. Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. *Advances in Water Resources* **32**: 1255-1266.

52. Weibull, W. (1951). A statistical distribution function of wide applicability. *ASME Journal of Applied Mechanics* **18**, 293-297.

53. Whitt, W. (2002). *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and their Application to Queues*. Springer, Berlin.

Appendix A The Hessian Matrices for the Peachtree Creek Data

After numerically solving for the maximum likelihood estimators via an evolutionary algorithm, we next need to compute the standard errors, T-statistics, and p-values. None of these are possible without first computing the Hessian matrix. We chose to compute the exact second derivatives and evaluate each of them at the estimators. This appendix shows the equations we used.

If H_i denotes the individual Hessian matrix for season $i, i = 1, \cdots, 4$, and H is the overall 18×18 Hessian matrix, then

$$H = \begin{pmatrix} H_1 & 0 & \cdots & 0 \\ 0 & H_2 & \ddots & \vdots \\ \vdots & \ddots & H_3 & 0 \\ 0 & \cdots & 0 & H_4 \end{pmatrix}.$$

Since H_1 and H_3 are Gumbel fits and therefore $\xi_1 = \xi_3 = 0$, we derive their Hessian matrix forms separately from H_2 and H_4. The results are in Appendices A.1 and A.2, respectively.

A.1 The Hessian Matrices for Summer and Winter

For $i = 1, 3$, H_i has the form

$$H_i = \begin{pmatrix} \frac{\partial^2 \ln \mathrm{L}_i}{\partial \mu_i^2} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \mu_i \partial \sigma_i} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \mu_i \partial \eta_i} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \mu_i \partial \phi_i} \\ \frac{\partial^2 \ln \mathrm{L}_i}{\partial \sigma_i \partial \mu_i} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \sigma_i^2} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \sigma_i \partial \eta_i} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \sigma_i \partial \phi_i} \\ \frac{\partial^2 \ln \mathrm{L}_i}{\partial \eta_i \partial \mu_i} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \eta_i \partial \sigma_i} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \eta_i^2} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \eta_i \partial \phi_i} \\ \frac{\partial^2 \ln \mathrm{L}_i}{\partial \phi_i \partial \mu_i} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \phi_i \partial \sigma_i} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \phi_i \partial \eta_i} & \frac{\partial^2 \ln \mathrm{L}_i}{\partial \phi_i^2} \end{pmatrix}.$$

We now state all ten unique second derivatives. Define $f(x_{ti}) = \exp\left[-\left(\frac{x_{ti}-\mu_i-\eta_i t}{\sigma_i+\phi_i t}\right)\right]$.

$$\frac{\partial^2 \text{lnL}_i}{\partial \mu_i^2} = -\sum_{t=1}^{52}\left(\frac{1}{\sigma_i+\phi_i t}\right)^2 f(x_{t,i})$$

$$\frac{\partial^2 \text{lnL}_i}{\partial \mu_i \partial \sigma_i} = -\sum_{t=1}^{52}\left(\frac{1}{\sigma_i+\phi_i t}\right)^2 \left\{1 - f(x_{ti}) + \left(\frac{x_{ti}-\mu_i-a_i t}{\sigma_i+\phi_i t}\right) f(x_{t,i})\right\}$$

$$\frac{\partial^2 \text{lnL}_i}{\partial \mu_i \partial \eta_i} = -\sum_{t=1}^{52} t\left(\frac{1}{\sigma_i+\phi_i t}\right)^2 f(x_{t,i})$$

$$\frac{\partial^2 \text{lnL}_i}{\partial \mu_i \partial \phi_i} = -\sum_{t=1}^{52} t\left(\frac{1}{\sigma_i+\phi_i t}\right)^2 \left\{1 - f(x_{ti}) + \left(\frac{x_{ti}-\mu_i-a_i t}{\sigma_i+\phi_i t}\right) f(x_{t,i})\right\}$$

$$\frac{\partial^2 \text{lnL}_i}{\partial \sigma_i^2} = \sum_{t=1}^{52}\left(\frac{1}{\sigma_i+\phi_i t}\right)^2 \left\{1 - 2\left(\frac{x_{ti}-\mu_i-\eta_i t}{\sigma_i+\phi_i t}\right) + 2\left(\frac{x_{ti}-\mu_i-\eta_i t}{\sigma_i+\phi_i t}\right) f(x_{ti}) - \left(\frac{x_{ti}-\mu_i-\eta_i t}{\sigma_i+\phi_i t}\right)^2 f(x_{ti})\right\}$$

$$\frac{\partial^2 \text{lnL}_i}{\partial \sigma_i \partial \eta_i} = -\sum_{t=1}^{52} t\left(\frac{1}{\sigma_i+\phi_i t}\right)^2 \left\{1 - f(x_{ti}) + \left(\frac{x_{ti}-\mu_i-\eta_i t}{\sigma_i+\phi_i t}\right) f(x_{ti})\right\}$$

$$\frac{\partial^2 \text{lnL}_i}{\partial \sigma_i \partial \phi_i} = \sum_{t=1}^{52} t\left(\frac{1}{\sigma_i+\phi_i t}\right)^2 \left\{1 - 2\left(\frac{x_{ti}-\mu_i-\eta_i t}{\sigma_i+\phi_i t}\right) + 2\left(\frac{x_{ti}-\mu_i-\eta_i t}{\sigma_i+\phi_i t}\right) f(x_{ti}) - \left(\frac{x_{ti}-\mu_i-\eta_i t}{\sigma_i+\phi_i t}\right)^2 f(x_{ti})\right\}$$

$$\frac{\partial^2 \text{lnL}_i}{\partial \eta_i^2} = -\sum_{t=1}^{52}\left(\frac{t}{\sigma_i+\phi_i t}\right)^2 f(x_{ti})$$

$$\frac{\partial^2 \ln \mathrm{L}_i}{\partial \eta_i \partial \phi_i} = -\sum_{t=1}^{52} \left(\frac{t}{\sigma_i + \phi_i t} \right)^2 \left\{ 1 - f(x_{ti}) + \left(\frac{x_{ti} - \mu_i - \eta_i t}{\sigma_i + \phi_i t} \right) f(x_{ti}) \right\}$$

$$\begin{aligned} \frac{\partial^2 \ln \mathrm{L}_i}{\partial \phi_i^2} = \sum_{t=1}^{52} \left(\frac{t}{\sigma_i + \phi_i t} \right)^2 \Bigg\{ 1 &- 2 \left(\frac{x_{ti} - \mu_i - \eta_i t}{\sigma_i + \phi_i t} \right) + 2 \left(\frac{x_{ti} - \mu_i - \eta_i t}{\sigma_i + \phi_i t} \right) f(x_{ti}) \\ &- \left(\frac{x_{ti} - \mu_i - \eta_i t}{\sigma_i + \phi_i t} \right)^2 f(x_{ti}) \Bigg\} \end{aligned}$$

The Hessian matrices for summer and winter, respectively, are

$$H_1 = \begin{pmatrix} -0.00016 & 0.00007 & -0.00351 & 0.00132 \\ 0.00007 & -0.00029 & 0.00132 & -0.00574 \\ -0.00351 & 0.00132 & -0.11317 & 0.04575 \\ 0.00132 & -0.00574 & 0.04575 & -0.19106 \end{pmatrix}$$

and

$$H_3 = \begin{pmatrix} -0.00009 & 0.00004 & -0.00311 & 0.00130 \\ 0.00004 & -0.00015 & 0.00130 & -0.00479 \\ -0.00311 & 0.00130 & -0.12646 & 0.05841 \\ 0.00130 & -0.00479 & 0.05841 & -0.18238 \end{pmatrix}.$$

If H_i^{-1} is the inverse Hessian matrix for season i, then

$$H_1^{-1} = \begin{pmatrix} 22112.67287 & 5700.84739 & -693.87291 & -185.32091 \\ 5700.84739 & 10184.84795 & -184.01871 & -311.05733 \\ -693.87291 & -184.01871 & 31.55885 & 8.31256 \\ -185.32091 & -311.05733 & 8.31256 & 15.30152 \end{pmatrix}$$

and

$$H_3^{-1} = \begin{pmatrix} 79631.30101 & 31388.20592 & -2063.35865 & -918.57453 \\ 31388.20592 & 55656.84538 & -908.37433 & -1528.30638 \\ -2063.35865 & -908.37433 & 62.95349 & 29.33502 \\ -918.57453 & -1528.30638 & 29.33502 & 48.45324 \end{pmatrix}.$$

The standard errors are found by taking the square roots of the diagonal entries.

A.2 The Hessian Matrices for Fall and Spring

For $i = 2, 4$, H_i has the form

$$H_i = \begin{pmatrix} \frac{\partial^2 \ln L_i}{\partial \xi_i^2} & \frac{\partial^2 \ln L_i}{\partial \xi_i \partial \mu_i} & \frac{\partial^2 \ln L_i}{\partial \xi_i \partial \sigma_i} & \frac{\partial^2 \ln L_i}{\partial \xi_i \partial \eta_i} & \frac{\partial^2 \ln L_i}{\partial \xi_i \partial \phi_i} \\ \frac{\partial^2 \ln L_i}{\partial \mu_i \partial \xi_i} & \frac{\partial^2 \ln L_i}{\partial \mu_i^2} & \frac{\partial^2 \ln L_i}{\partial \mu_i \partial \sigma_i} & \frac{\partial^2 \ln L_i}{\partial \mu_i \partial \eta_i} & \frac{\partial^2 \ln L_i}{\partial \mu_i \partial \phi_i} \\ \frac{\partial^2 \ln L_i}{\partial \sigma_i \partial \xi_i} & \frac{\partial^2 \ln L_i}{\partial \sigma_i \partial \mu_i} & \frac{\partial^2 \ln L_i}{\partial \sigma_i^2} & \frac{\partial^2 \ln L_i}{\partial \sigma_i \partial \eta_i} & \frac{\partial^2 \ln L_i}{\partial \sigma_i \partial \phi_i} \\ \frac{\partial^2 \ln L_i}{\partial \eta_i \partial \xi_i} & \frac{\partial^2 \ln L_i}{\partial \eta_i \partial \mu_i} & \frac{\partial^2 \ln L_i}{\partial \eta_i \partial \sigma_i} & \frac{\partial^2 \ln L_i}{\partial \eta_i^2} & \frac{\partial^2 \ln L_i}{\partial \eta_i \partial \phi_i} \\ \frac{\partial^2 \ln L_i}{\partial \phi_i \partial \xi_i} & \frac{\partial^2 \ln L_i}{\partial \phi_i \partial \mu_i} & \frac{\partial^2 \ln L_i}{\partial \phi_i \partial \sigma_i} & \frac{\partial^2 \ln L_i}{\partial \phi_i \partial \eta_i} & \frac{\partial^2 \ln L_i}{\partial \phi_i^2} \end{pmatrix}.$$

We now state all fifteen unique second derivatives. Define $f(x_{ti}) = 1 + \frac{\xi_i(x_{ti}-\mu_i-\eta_i t)}{\sigma_i+\phi_i t}$.

$$\begin{aligned} \frac{\partial^2 \ln L_i}{\partial \xi_i^2} &= -\frac{2}{\xi_i^3}\sum_{t=1}^{52}\log f(x_{ti}) + \frac{2}{\xi_i^2}\sum_{t=1}^{52}\left(\frac{x_{ti}-\mu_i-\eta_i t}{\sigma_i+\phi_i t}\right)[f(x_{ti})]^{-1} \\ &\quad + \frac{1}{\xi_i^3}\sum_{t=1}^{52}[f(x_{ti})]^{-1/\xi_i}\log f(x_{ti})\left\{2 - \frac{1}{\xi_i}\log f(x_{ti})\right\} \\ &\quad - \frac{2}{\xi_i^2}\sum_{t=1}^{52}\left(\frac{x_{ti}-\mu_i-\eta_i t}{\sigma_i+\phi_i t}\right)[f(x_{ti})]^{-1/\xi_i-1}\left\{1 - \frac{1}{\xi_i}\log f(x_{ti})\right\} \\ &\quad + \frac{\xi_i+1}{\xi_i}\sum_{t=1}^{52}\left(\frac{x_{ti}-\mu_i-\eta_i t}{\sigma_i+\phi_i t}\right)^2[f(x_{ti})]^{-2}\left\{1 - \frac{1}{\xi_i}[f(x_{ti})]^{-1/\xi_i}\right\} \\ \frac{\partial^2 \ln L_i}{\partial \xi_i \partial \mu_i} &= \frac{1}{\sigma_i}\sum_{t=1}^{52}[f(x_{ti})]^{-1}\left\{1 - \frac{1}{\xi_i^2}[f(x_{ti})]^{-1/\xi_i}\log f(x_{ti})\right\} \\ &\quad - \left(\frac{\xi_i+1}{\sigma_i^2}\right)\sum_{t=1}^{52}(x_{ti}-\mu_i-\eta_i t)[f(x_{ti})]^{-2}\left\{1 - \frac{1}{\xi_i}[f(x_{ti})]^{-1/\xi_i}\right\} \\ \frac{\partial^2 \ln L_i}{\partial \xi_i \partial \sigma_i} &= \frac{1}{\sigma_i^2}\sum_{t=1}^{52}(x_{ti}-\mu_i-\eta_i t)[f(x_{ti})]^{-1}\left\{1 - \frac{1}{\xi_i^2}[f(x_{ti})]^{-1/\xi_i}\log f(x_{ti})\right\} \\ &\quad - \left(\frac{\xi_i+1}{\sigma_i^3}\right)\sum_{t=1}^{52}(x_{ti}-\mu_i-\eta_i t)^2[f(x_{ti})]^{-2}\left\{1 - \frac{1}{\xi_i}[f(x_{ti})]^{-1/\xi_i}\right\} \end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2 \ln \mathrm{L}_i}{\partial \xi_i \partial \eta_i} &= \sum_{t=1}^{52} \left(\frac{t}{\sigma_i + \phi_i t} \right) [f(x_{ti})]^{-1} \left\{ 1 - \frac{1}{\xi_i^2} [f(x_{ti})]^{-1/\xi_i} \log f(x_{ti}) \right\} \\
&\quad - (\xi_i + 1) \sum_{t=1}^{52} t \left(\frac{1}{\sigma_i + \phi_i t} \right)^2 (x_{ti} - \mu_i - \eta_i t) [f(x_{ti})]^{-2} \left\{ 1 - \frac{1}{\xi_i} [f(x_{ti})]^{-1/\xi_i} \right\} \\
\frac{\partial^2 \ln \mathrm{L}_i}{\partial \xi_i \partial \phi_i} &= \sum_{t=1}^{52} t \left(\frac{1}{\sigma_i + \phi_i t} \right)^2 (x_{ti} - \mu_i - \eta_i t) [f(x_{ti})]^{-1} \left\{ 1 - \frac{1}{\xi_i^2} [f(x_{ti})]^{-1/\xi_i} \log f(x_{ti}) \right\} \\
&\quad - (\xi_i + 1) \sum_{t=1}^{52} t \left(\frac{1}{\sigma_i + \phi_i t} \right)^3 (x_{ti} - \mu_i - \eta_i t)^2 [f(x_{ti})]^{-2} \left\{ 1 - \frac{1}{\xi_i} [f(x_{ti})]^{-1/\xi_i} \right\}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 \ln \mathrm{L}_i}{\partial \mu_i^2} &= (\xi_i + 1) \sum_{t=1}^{52} \left(\frac{1}{\sigma_i + \phi_i t} \right)^2 [f(x_{ti})]^{-2} \left\{ \xi_i - [f(x_{ti})]^{-1/\xi_i} \right\} \\
\frac{\partial^2 \ln \mathrm{L}_i}{\partial \mu_i \partial \sigma_i} &= -\sum_{t=1}^{52} \left(\frac{1}{\sigma_i + \phi_i t} \right)^2 [f(x_{ti})]^{-1} \left\{ (\xi_i + 1) - [f(x_{ti})]^{-1/\xi_i} \right\} \\
&\quad + (\xi_i + 1) \sum_{t=1}^{52} \left(\frac{1}{\sigma_i + \phi_i t} \right)^3 (x_{ti} - \mu_i - \eta_i t) [f(x_{ti})]^{-2} \left\{ \xi_i - [f(x_{ti})]^{-1/\xi_i} \right\} \\
\frac{\partial^2 \ln \mathrm{L}_i}{\partial \mu_i \partial \eta_i} &= (\xi_i + 1) \sum_{t=1}^{52} t \left(\frac{1}{\sigma_i + \phi_i t} \right)^2 [f(x_{ti})]^{-2} \left\{ \xi_i - [f(x_{ti})]^{-1/\xi_i} \right\} \\
\frac{\partial^2 \ln \mathrm{L}_i}{\partial \mu_i \partial \phi_i} &= -\sum_{t=1}^{52} t \left(\frac{1}{\sigma_i + \phi_i t} \right)^2 [f(x_{ti})]^{-1} \left\{ (\xi_i + 1) - [f(x_{ti})]^{-1/\xi_i} \right\} \\
&\quad + (\xi_i + 1) \sum_{t=1}^{52} t \left(\frac{1}{\sigma_i + \phi_i t} \right)^3 (x_{ti} - \mu_i - \eta_i t) [f(x_{ti})]^{-2} \left\{ \xi_i - [f(x_{ti})]^{-1/\xi_i} \right\}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 \ln \mathrm{L}_i}{\partial \sigma_i^2} &= -2 \sum_{t=1}^{52} \left(\frac{1}{\sigma_i + \phi_i t} \right)^3 (x_{ti} - \mu_i - \eta_i t) [f(x_{ti})]^{-1} \left\{ \xi_i + 1 - [f(x_{ti})]^{-1/\xi_i} \right\} \\
&\quad + (\xi_i + 1) \sum_{t=1}^{52} \left(\frac{1}{\sigma_i + \phi_i t} \right)^4 (x_{ti} - \mu_i - \eta_i t)^2 [f(x_{ti})]^{-2} \left\{ \xi_i - [f(x_{ti})]^{-1/\xi_i} \right\} \\
&\quad + \sum_{t=1}^{52} \left(\frac{1}{\sigma_i + \phi_i t} \right)^2
\end{aligned}
$$

$$\frac{\partial^2 \ln \mathrm{L}_i}{\partial \sigma_i \partial \eta_i} = -\sum_{t=1}^{52} t \left(\frac{1}{\sigma_i + \phi_i t}\right)^2 [f(x_{ti})]^{-1} \left\{\xi_i + 1 - [f(x_{ti})]^{-1/\xi_i}\right\}$$
$$+ (\xi_i + 1) \sum_{t=1}^{52} t \left(\frac{1}{\sigma_i + \phi_i t}\right)^3 (x_{ti} - \mu_i - \eta_i t)[f(x_{ti})]^{-2} \left\{\xi_i - [f(x_{ti})]^{-1/\xi_i}\right\}$$

$$\frac{\partial^2 \ln \mathrm{L}_i}{\partial \sigma_i \partial \phi_i} = -2\sum_{t=1}^{52} t \left(\frac{1}{\sigma_i + \phi_i t}\right)^3 (x_{ti} - \mu_i - \eta_i t)[f(x_{ti})]^{-1} \left\{\xi_i + 1 - [f(x_{ti})]^{-1/\xi_i}\right\}$$
$$+ (\xi_i + 1) \sum_{t=1}^{52} t \left(\frac{1}{\sigma_i + \phi_i t}\right)^4 (x_{ti} - \mu_i - \eta_i t)^2[f(x_{ti})]^{-2} \left\{\xi_i - [f(x_{ti})]^{-1/\xi_i}\right\}$$
$$+ \sum_{t=1}^{52} t \left(\frac{1}{\sigma_i + \phi_i t}\right)^2$$

$$\frac{\partial^2 \ln \mathrm{L}_i}{\partial \eta_i^2} = (\xi_i + 1) \sum_{t=1}^{52} t^2 \left(\frac{1}{\sigma_i + \phi_i t}\right)^2 [f(x_{ti})]^{-2} \left\{\xi_i - [f(x_{ti})]^{-1/\xi_i}\right\}$$

$$\frac{\partial^2 \ln \mathrm{L}_i}{\partial \eta_i \partial \phi_i} = -\sum_{t=1}^{52} t^2 \left(\frac{1}{\sigma_i + \phi_i t}\right)^2 [f(x_{ti})]^{-1} \left\{\xi_i + 1 - [f(x_{ti})]^{-1/\xi_i}\right\}$$
$$+ (\xi_i + 1) \sum_{t=1}^{52} t^2 \left(\frac{1}{\sigma_i + \phi_i t}\right)^3 (x_{ti} - \mu_i - \eta_i t)[f(x_{ti})]^{-2} \left\{\xi_i - [f(x_{ti})]^{-1/\xi_i}\right\}$$

$$\frac{\partial^2 \ln \mathrm{L}_i}{\partial \phi_i^2} = -2\sum_{t=1}^{52} t^2 \left(\frac{1}{\sigma_i + \phi_i t}\right)^3 (x_{ti} - \mu_i - \eta_i t)[f(x_{ti})]^{-1} \left\{\xi_i + 1 - [f(x_{ti})]^{-1/\xi_i}\right\}$$
$$+ (\xi_i + 1) \sum_{t=1}^{52} t^2 \left(\frac{1}{\sigma_i + \phi_i t}\right)^4 (x_{ti} - \mu_i - \eta_i t)^2[f(x_{ti})]^{-2} \left\{\xi_i - [f(x_{ti})]^{-1/\xi_i}\right\}$$
$$+ \sum_{t=1}^{52} t^2 \left(\frac{1}{\sigma_i + \phi_i t}\right)^2$$

The Hessian matrices for fall and spring, respectively, are

$$H_2 = \begin{pmatrix} -70.05764 & -0.04483 & 0.02668 & -0.77989 & 0.10155 \\ -0.04483 & -0.00023 & 0.00017 & -0.00355 & 0.00238 \\ 0.02668 & 0.00017 & -0.00033 & 0.00238 & -0.00477 \\ -0.77989 & -0.00355 & 0.00238 & -0.10428 & 0.07342 \\ 0.10155 & 0.00238 & -0.00477 & 0.07342 & -0.13407 \end{pmatrix}$$

and

$$H_4 = \begin{pmatrix} -39.83390 & -0.07984 & 0.05635 & -1.84513 & 1.06889 \\ -0.07984 & -0.00044 & 0.00032 & -0.01302 & 0.00931 \\ 0.05635 & 0.00032 & -0.00030 & 0.00931 & -0.00847 \\ -1.84513 & -0.01302 & 0.00931 & -0.47611 & 0.33372 \\ 1.06889 & 0.00931 & -0.00847 & 0.33372 & -0.30177 \end{pmatrix}.$$

If H_i^{-1} is the inverse Hessian matrix for season i, then

$$H_2^{-1} = \begin{pmatrix} 0.01734 & -0.72415 & 2.29779 & -0.17886 & -0.17937 \\ -0.72415 & 19478.62600 & 11627.35253 & -716.92017 & -460.62141 \\ 2.29779 & 11627.35253 & 13515.25007 & -482.34530 & -536.43491 \\ -0.17886 & -716.92017 & -482.34530 & 44.51125 & 28.65698 \\ -0.17937 & -460.62141 & -536.43491 & 28.65698 & 33.90823 \end{pmatrix}.$$

and

$$H_4^{-1} = \begin{pmatrix} 0.05057 & -9.25956 & 12.63677 & -0.08487 & -0.55511 \\ -9.25956 & 56281.33366 & 52124.69912 & -1404.46811 & -1313.39844 \\ 12.63677 & 52124.69912 & 73634.87371 & -1446.63307 & -2014.58926 \\ -0.08487 & -1404.46811 & -1446.63307 & 46.42803 & 48.33709 \\ -0.55511 & -1313.39844 & -2014.58926 & 48.33709 & 70.85107 \end{pmatrix}.$$

The standard errors are found by taking the square roots of the diagonal entries.

Appendix B Supplemental Proofs for the AR(1) Process

This collection of appendices gives supplemental proofs that were excluded from Chapter 4. They concern two parts from Theorem 4.1 and two more parts from Theorem 4.2. We also include two tables of numerical values for the infinite products.

B.1 Proof of Theorem 4.1 for $\beta = -\frac{\sqrt{2}}{2}$

If $\beta = -\frac{\sqrt{2}}{2}$ and Z_k are i.i.d. standard Gumbel random variables, then as $y \to \infty$

$$P\left(\sum_{k=0}^{\infty}(0.5)^k Z_k > y\right) = \prod_{k=1}^{\infty}\Gamma\left(1-(0.5)^k\right)\prod_{k=0}^{\infty}\Gamma\left(1+\sqrt{2}(0.5)^{k+1}\right)e^{-y}$$
$$-\prod_{k=1}^{\infty}\Gamma\left(1-(0.5)^k\right)\prod_{k=0}^{\infty}\Gamma\left(1+\sqrt{2}(0.5)^{k}\right)ye^{-2y} + O\left(e^{-2y}\right).$$

Proof. First notice that by Lemma 4.9,

$$P(S > y) = \prod_{k=1}^{\infty}\Gamma\left(1-(0.5)^k\right)\left[e^{-y} - ye^{-2y}\right] + O\left(e^{-2y}\right). \tag{2.1}$$

Let $X = S + T$, and for some $0 < a < 1$ write

$$P(X > y) = \int_{ay}^{\infty} P(S > y-t)dF_T(t) + \int_{-\infty}^{ay} P(S > y-t)dF_T(t). \tag{2.2}$$

First,

$$\int_{ay}^{\infty} P(S > y-t)dF_T(t) \le P(T > ay),$$

which by Lemma 4.5 is $o\left(e^{-y/\beta^2}\right)$ as $y \to \infty$. Now by (2.1) we have that for y large enough and any $t < ay$

$$P(S > y - t) = \prod_{k=1}^{\infty} \Gamma\left(1 - (0.5)^k\right) \left[e^{-(y-t)} - (y-t)e^{-2(y-t)}\right] + R(y-t),$$

where for some $c > 0$

$$\sup_{t<ay} \left|R(y-t)e^{2(y-t)}\right| < c.$$

Set $K_1 = \prod_{k=1}^{\infty} \Gamma(1 - (0.5)^k)$. Then observe that

$$\begin{aligned} \int_{-\infty}^{ay} \overline{F}_S(y-t)dF_T(t) = \int_{-\infty}^{ay} \left[\frac{\overline{F}_S(y-t) - K_1 e^{-(y-t)}}{-K_1(y-t)e^{-2(y-t)}}\right] \times -K_1(y-t)e^{-2(y-t)}dF_T(t) \\ + K_1 \int_{-\infty}^{ay} e^{-(y-t)}dF_T(t). \end{aligned} \tag{2.3}$$

Now

$$\frac{\overline{F}_S(y-t) - K_1 e^{-(y-t)}}{-K_1(y-t)e^{-2(y-t)}} = 1 - \frac{R(y-t)e^{2(y-t)}}{K_1(y-t)} = 1 + \epsilon(y-t).$$

We examine the $\epsilon(y-t)$:

$$\sup_{t<ay} |\epsilon(y-t)| = \frac{1}{K_1} \sup_{t<ay} \left|\frac{R(y-t)e^{2(y-t)}}{y-t}\right| < \frac{c}{K_1} \sup_{t<ay} \left|(y-t)^{-1}\right| = \frac{c}{K_1(1-a)y}.$$

Therefore $\epsilon(y-t) \to 0$ as $y \to \infty$ uniformly in $t < ay$. Consider the piece

$$\begin{aligned} -K_1 \int_{-\infty}^{ay} [1 + \epsilon(y-t)](y-t)e^{-2(y-t)}dF_T(t) = -K_1 y e^{-2y} \int_{-\infty}^{ay} e^{2t}dF_T(t) \\ + K_1 e^{-2y} \int_{-\infty}^{ay} te^{2t}dF_T(t) - K_1 \int_{-\infty}^{ay} \epsilon(y-t)(y-t)e^{-2(y-t)}dF_T(t) \\ = (I) + (II) + (III). \end{aligned} \tag{2.4}$$

Examining (I) in (2.4), as $y \to \infty$

$$\int_{-\infty}^{ay} e^{2t} dF_T(t) \to E\left(e^{2T}\right) = \prod_{k=0}^{\infty} E\left(e^{2\beta^{2k+1} Z_1}\right) = \prod_{k=0}^{\infty} \Gamma\left(1 + \sqrt{2}(0.5)^k\right).$$

Therefore by dominated convergence,

$$(I) = -K_1 \prod_{k=0}^{\infty} \Gamma\left(1 + \sqrt{2}(0.5)^k\right) y e^{-2y}. \tag{2.5}$$

Turning to (II), as $y \to \infty$

$$K_1 \int_{-\infty}^{ay} t e^{2t} dF_T(t) \to K_1 E\left(T e^{2T}\right),$$

and by an extension of Lemma 4.4, $(II) = O\left(e^{-2y}\right)$. Next,

$$|(III)| = \left| \int_{-\infty}^{ay} R(y-t) e^{2(y-t)} e^{-2(y-t)} dF_T(t) \right| < cE\left(e^{2T}\right) e^{-2y},$$

and therefore $(III) = O(e^{-2y})$. Now note that

$$K_1 \int_{-\infty}^{\infty} e^t dF_T(t) = K_1 E\left(e^T\right) = K_1 \prod_{k=0}^{\infty} \Gamma\left(1 + \sqrt{2}(0.5)^{k+1}\right). \tag{2.6}$$

We now consider the integral

$$K_1 e^{-y} \int_{-\infty}^{ay} e^t dF_T(t) = K_1 \prod_{k=0}^{\infty} \Gamma\left(1 + \sqrt{2}(0.5)^{k+1}\right) e^{-y} - K_1 e^{-y} \int_{ay}^{\infty} e^t dF_T(t). \tag{2.7}$$

By Lemma 4.7, the last integral in (2.7) is $o\left(e^{-y/\beta^2}\right)$, and the result follows. $\square$

B.2 Proof of Theorem 4.1 for $-1 < \beta < -\frac{\sqrt{2}}{2}$

If $-1 < \beta < -\frac{\sqrt{2}}{2}$ and Z_k are i.i.d. standard Gumbel random variables, then we shall show that as $y \to \infty$

$$P\left(\sum_{k=0}^{\infty} \beta^k Z_k > y\right) = \prod_{k=1}^{\infty} \Gamma\left(1-\beta^k\right) e^{-y} + o\left(e^{-y/\beta^2}\right) - \frac{\beta^2}{1-\beta^2}\Gamma(1-\beta^{-1})\Gamma(2-\beta^{-2}) \prod_{k=1}^{\infty} \Gamma\left(1-\beta^k\right) e^{-y/\beta^2}.$$

Proof. First notice that by Lemma 4.10,

$$P(S > y) = \prod_{k=1}^{\infty} \Gamma\left(1-\beta^{2k}\right) \left[e^{-y} - \frac{\beta^2}{1-\beta^2}\Gamma(2-\beta^{-2})e^{-y/\beta^2}\right] + o\left(e^{-y/\beta^2}\right). \tag{2.8}$$

Let $X = S + T$, and choose some $0 < a < 1$. By earlier work,

$$P(X > y) = \int_{-\infty}^{ay} P(S > y-t)dF_T(t) + o\left(e^{-y/\beta^2}\right). \tag{2.9}$$

Now by (2.8) we have that for y large enough and any $t < ay$

$$P(S > y-t) = \prod_{k=1}^{\infty} \Gamma\left(1-\beta^{2k}\right) \left[e^{-(y-t)} - \frac{\beta^2}{1-\beta^2}\Gamma(2-\beta^{-2})e^{-(y-t)/\beta^2}\right] + R(y-t),$$

where for any $\epsilon > 0$

$$\sup_{t<ay} \left|R(y-t)e^{(y-t)/\beta^2}\right| < \epsilon.$$

Set $K_1 = \prod_{k=1}^{\infty} \Gamma(1-\beta^{2k})$ and $K_2 = -\frac{\beta^2}{1-\beta^2}\Gamma(2-\beta^{-2})K_1$. Then observe that

$$\int_{-\infty}^{ay} \overline{F}_S(y-t)dF_T(t)$$
$$= \int_{-\infty}^{ay} \left[\frac{\overline{F}_S(y-t) - K_1 e^{-(y-t)}}{K_2 e^{-(y-t)/\beta^2}}\right] K_2 e^{-(y-t)/\beta^2} dF_T(t) + K_1 \int_{-\infty}^{ay} e^{-(y-t)} dF_T(t). \tag{2.10}$$

Now

$$\frac{\overline{F}_S(y-t) - K_1 e^{-(y-t)}}{K_2 e^{-(y-t)/\beta^2}} = 1 + \frac{1}{K_2} R(y-t) e^{(y-t)/\beta^2} = 1 + \epsilon(y-t).$$

Note that $\sup_{t<ay} |\epsilon(y-t)| < \dfrac{\epsilon}{|K_2|}$. Consider the piece

$$\int_{-\infty}^{ay} [1 + \epsilon(y-t)] K_2 e^{-(y-t)/\beta^2} dF_T(t)$$
$$= K_2 e^{-y/\beta^2} \int_{-\infty}^{ay} e^{t/\beta^2} dF_T(t) + K_2 \int_{-\infty}^{ay} \epsilon(y-t) e^{-(y-t)/\beta^2} dF_T(t). \tag{2.11}$$

Examining the first integral in (2.11),

$$\lim_{y\to\infty} \int_{-\infty}^{ay} e^{t/\beta^2} dF_T(t) = E\left(e^{T/\beta^2}\right) = E\left[\exp\left(\frac{1}{\beta^2}\sum_{k=0}^{\infty} \beta^{2k+1} Z_k\right)\right]$$
$$= E\left[\exp\left(\sum_{k=0}^{\infty} \beta^{2k-1} Z_k\right)\right] = \Gamma\left(1-\beta^{-1}\right) \prod_{k=0}^{\infty} \Gamma\left(1-\beta^{2k+1}\right).$$

Thus, by dominated convergence the first integral in (2.11) is

$$-\frac{\beta^2}{1-\beta^2} \Gamma\left(1-\beta^{-1}\right) \Gamma\left(2-\beta^{-2}\right) \prod_{k=1}^{\infty} \Gamma\left(1-\beta^k\right) e^{-y/\beta^2}. \tag{2.12}$$

Turning to the second integral in (2.11),

$$
\begin{aligned}
&\left|K_2\int_{-\infty}^{ay}\epsilon(y-t)e^{-(y-t)/\beta^2}dF_T(t)\right| = \left|\int_{-\infty}^{ay}R(y-t)dF_T(t)\right| \\
&\quad = \left|\int_{-\infty}^{ay}R(y-t)e^{(y-t)/\beta^2}e^{-(y-t)/\beta^2}dF_T(t)\right| < \epsilon\int_{-\infty}^{ay}e^{-(y-t)/\beta^2}dF_T(t) \\
&\quad \leq \epsilon e^{-y/\beta^2}\int_{-\infty}^{\infty}e^{t/\beta^2}dF_T(t) = \epsilon E\left(e^{T/\beta^2}\right)e^{-y/\beta^2}.
\end{aligned}
$$

By Lemma 4.6, $E\left(e^{T/\beta^2}\right) < \infty$ and therefore

$$
K_2\int_{-\infty}^{ay}\epsilon(y-t)e^{-(y-t)/\beta^2}dF_T(t) = o\left(e^{-y/\beta^2}\right). \tag{2.13}
$$

Next, note that

$$
K_1\int_{-\infty}^{\infty}e^t dF_T(t) = K_1E\left(e^T\right) = K_1\prod_{k=0}^{\infty}\Gamma\left(1-\beta^{2k+1}\right) = \prod_{k=1}^{\infty}\Gamma\left(1-\beta^k\right). \tag{2.14}
$$

We now consider the integral

$$
K_1e^{-y}\int_{-\infty}^{ay}e^t dF_T(t) = \prod_{k=1}^{\infty}\Gamma(1-\beta^k)e^{-y} - K_1\int_{ay}^{\infty}e^{-(y-t)}dF_T(t). \tag{2.15}
$$

Collecting (2.11) through (2.15),

$$
\begin{aligned}
\int_{-\infty}^{ay}\overline{F}_S(y-t)dF_T(t) &= \prod_{k=1}^{\infty}\Gamma(1-\beta^k)e^{-y} + o\left(e^{-y/\beta^2}\right) - K_1e^{-y}\int_{ay}^{\infty}e^t dF_T(t) \\
&\quad - \frac{\beta^2}{1-\beta^2}\Gamma\left(1-\beta^{-1}\right)\Gamma\left(2-\beta^{-2}\right)\prod_{k=1}^{\infty}\Gamma\left(1-\beta^k\right)e^{-y/\beta^2}.
\end{aligned}
$$

By Lemma 4.7, the integral in the above expression is $o\left(e^{-y/\beta^2}\right)$, and the result follows. □

B.3 Proof of Theorem 4.2 for $\frac{\beta^{(1)}}{\beta^{(0)}} = \frac{1}{2}$

Here we present the proofs of the second part of Theorem 4.2. Recall that $V = Z_0 + \frac{\beta^{(1)}}{\beta^{(0)}} Z_1$, $W = \sum_{k=2}^{n-1} \frac{\beta^{(k)}}{\beta^{(0)}} Z_k$, and $Y = V + W$. Then we shall derive the probability $P(Y_n > y_n^*)$ where $y_n^* = \frac{y - \overline{\mu}_n}{\beta^{(0)}}$. When $\frac{\beta^{(1)}}{\beta^{(0)}} = \frac{1}{2}$, then as $y \to \infty$ the end result will be

$$P(Y_n > y_n^*) = \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right) e^{-y_n^*} - \prod_{k=2}^{n-1} \Gamma\left(1 - \frac{2\beta^{(k)}}{\beta^{(0)}}\right) y_n^* e^{-2y_n^*} + O\left(e^{-y/\beta^{(1)}}\right).$$

Proof. Let a be a positive real such that $\frac{\beta^{(2)}}{\beta^{(1)}} < a < 1$, and choose $1 < \lambda < \frac{\beta^{(1)}}{\beta^{(2)}}$ such that $\lambda > \frac{1}{a}$. Then observe that

$$P(V + W > y_n^*) = \int_{ay_n^*}^{\infty} P(V > y_n^* - w) dF_W(w) + \int_{-\infty}^{ay_n^*} P(V > y_n^* - w) dF_W(w).$$

Earlier in the proof of Theorem 4.2, we established that the first integral above is $o\left(e^{-y/\beta^{(1)}}\right)$. Now by Theorem 3.1 we have for y_n^* large enough and any $w < ay_n^*$ that

$$P(V > y_n^* - w) = \sqrt{\pi} e^{-(y_n^* - w)} - (y_n^* - w) e^{-2(y_n^* - w)} + R(y_n^* - w),$$

where for some $c > 0$

$$\sup_{w < ay_n^*} \left| R(y_n^* - w) e^{2(y_n^* - w)} \right| < c.$$

Observe that

$$\int_{-\infty}^{ay_n^*} \overline{F}_V(y_n^* - w) dF_W(w) = \sqrt{\pi} \int_{-\infty}^{ay_n^*} e^{-(y_n^* - w)} dF_W(w)$$
$$+ \int_{-\infty}^{ay_n^*} \left[\frac{\overline{F}_V(y_n^* - w) - \sqrt{\pi} e^{-(y_n^* - w)}}{-(y_n^* - w) e^{-2(y_n^* - w)}} \right] \times -(y_n^* - w) e^{-2(y_n^* - w)} dF_W(w).$$

Now

$$\frac{\overline{F}_V(y_n^*-w)-\sqrt{\pi}e^{-(y_n^*-w)}}{-(y_n^*-w)e^{-2(y_n^*-w)}} = 1-\frac{R(y_n^*-w)e^{2(y_n^*-w)}}{y_n^*-w} = 1+\epsilon(y_n^*-w).$$

We examine the $\epsilon(y_n^*-w)$:

$$\sup_{w<ay_n^*}|\epsilon(y_n^*-w)| = \sup_{w<ay_n^*}\left|\frac{R(y_n^*-w)e^{2(y_n^*-w)}}{y_n^*-w}\right| < c\sup_{w<ay_n^*}\left|(y_n^*-w)^{-1}\right| \le \frac{c}{(1-a)y_n^*}.$$

Therefore $\epsilon(y_n^*-w)\to 0$ as $y_n^*\to\infty$ uniformly in $w<ay_n^*$. Consider the piece

$$\begin{aligned}
&\int_{-\infty}^{ay_n^*}[1+\epsilon(y_n^*-w)]\times -(y_n^*-w)e^{-2(y_n^*-w)}dF_W(w)\\
&\quad = -e^{-2y_n^*}\int_{-\infty}^{ay_n^*}(y_n^*-w)e^{2w}dF_W(w) - \int_{-\infty}^{ay_n^*}\epsilon(y_n^*-w)(y_n^*-w)e^{-2(y_n^*-w)}dF_W(w)\\
&\quad = -y_n^*e^{-2y_n^*}\int_{-\infty}^{ay_n^*}e^{2w}dF_W(w) + e^{-2y_n^*}\int_{-\infty}^{ay_n^*}we^{2w}dF_W(w)\\
&\qquad\quad - \int_{-\infty}^{ay_n^*}\epsilon(y_n^*-w)(y_n^*-w)e^{-2(y_n^*-w)}dF_W(w)\\
&\quad = (I)+(II)+(III).
\end{aligned}$$

Examining (I),

$$-\lim_{y_n^*\to\infty}\int_{-\infty}^{ay_n^*}e^{2w}dF_W(w) = -E\left(e^{2W}\right) = -\prod_{k=2}^{n-1}\Gamma\left(1-\frac{2\beta^{(k)}}{\beta^{(0)}}\right).$$

Next, $(II)=E\left(We^{2W}\right)e^{-2y_n^*} = O\left(e^{-2y_n^*}\right) = O\left(e^{-y/\beta^{(1)}}\right)$. Turning to (III),

$$\begin{aligned}
|(III)| &= \left|\int_{-\infty}^{ay_n^*}R(y_n^*-w)dF_W(w)\right| = \left|\int_{-\infty}^{ay_n^*}R(y_n^*-w)e^{2(y_n^*-w)}e^{-2(y_n^*-w)}dF_W(w)\right|\\
&< c\left|\int_{-\infty}^{ay_n^*}e^{-2(y_n^*-w)}dF_W(w)\right| \le cE\left(e^{2W}\right)e^{-2y_n^*}.
\end{aligned}$$

Therefore $(III) = O\left(e^{-2y_n^*}\right) = O\left(e^{-y/\beta^{(1)}}\right)$. Next, note that

$$\int_{-\infty}^{\infty} \sqrt{\pi} e^w dF_W(w) = \sqrt{\pi} \prod_{k=2}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right) = \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right).$$

We now consider the integral

$$\sqrt{\pi} e^{-y_n^*} \int_{-\infty}^{ay_n^*} e^w dF_W(w) = \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right) e^{-y_n^*} - \sqrt{\pi} e^{-y_n^*} \int_{ay_n^*}^{\infty} e^w dF_W(w).$$

Collecting all the terms, we arrive at

$$\begin{aligned}\int_{-\infty}^{ay_n^*} \overline{F}_V(y_n^* - w) dF_W(w) &= \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right) e^{-y_n^*} \\ &- \prod_{k=2}^{n-1} \Gamma\left(1 - \frac{2\beta^{(k)}}{\beta^{(0)}}\right) y_n^* e^{-2y_n^*} + O\left(e^{-y/\beta^{(1)}}\right) - \sqrt{\pi} e^{-y_n^*} \int_{ay_n^*}^{\infty} e^w dF_W(w).\end{aligned}$$

The last integral above was shown to be $o\left(e^{-y/\beta^{(1)}}\right)$ earlier in the proof of Theorem 4.2. □

B.4 Proof of Theorem 4.2 for $\frac{1}{2} < \frac{\beta^{(1)}}{\beta^{(0)}} < 1$

Here we present the proof of the third part of Theorem 4.2. When $\frac{1}{2} < \frac{\beta^{(1)}}{\beta^{(0)}} < 1$ and $K_2 = -\frac{\beta^{(1)}}{\beta^{(0)}-\beta^{(1)}}\Gamma\left(2 - \frac{\beta^{(0)}}{\beta^{(1)}}\right)$, then as $y \to \infty$ the end result will be

$$P(Y_n > y_n^*) = \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right) e^{-y_n^*} + K_2 \prod_{k=2}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(1)}}\right) e^{-\beta^{(0)} y_n^*/\beta^{(1)}} + o\left(e^{-y/\beta^{(1)}}\right).$$

Proof. Let a be a positive real such that $\frac{\beta^{(2)}}{\beta^{(1)}} < a < 1$. Then by earlier work

$$P(V + W > y_n^*) = \int_{-\infty}^{a y_n^*} P(V > y_n^* - w) dF_W(w) + o\left(e^{-y/\beta^{(1)}}\right).$$

By Theorem 3.1, we have for y_n^* large enough and any $w < a y_n^*$ that

$$P(V > y_n^* - w) = \Gamma\left(1 - \frac{\beta^{(1)}}{\beta^{(0)}}\right) e^{-(y_n^* - w)} + K_2 e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}} + R(y_n^* - w),$$

where for any $\epsilon > 0$

$$\sup_{w < a y_n^*} \left| R(y_n^* - w) e^{\beta^{(0)}(y_n^* - w)/\beta^{(1)}} \right| < \epsilon.$$

Further set $K_1 = \Gamma\left(1 - \frac{\beta^{(1)}}{\beta^{(0)}}\right)$. Then observe that

$$\begin{aligned} \int_{-\infty}^{a y_n^*} \overline{F}_V(y_n^* - w) dF_W(w) &= K_1 \int_{-\infty}^{a y_n^*} e^{-(y_n^* - w)} dF_W(w) \\ &+ \int_{-\infty}^{a y_n^*} \left[\frac{\overline{F}_V(y_n^* - w) - K_1 e^{-(y_n^* - w)}}{K_2 e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}}} \right] K_2 e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}} dF_W(w). \end{aligned}$$

Now

$$\frac{\overline{F}_V(y_n^* - w) - K_1 e^{-(y_n^* - w)}}{K_2 e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}}} = 1 + \frac{1}{K_2} R(y_n^* - w) e^{\beta^{(0)}(y_n^* - w)/\beta^{(1)}} = 1 + \epsilon(y_n^* - w).$$

We examine the $\epsilon(y_n^* - w)$:

$$\sup_{w < a y_n^*} |\epsilon(y_n^* - w)| = \frac{1}{|K_2|} \sup_{w < a y_n^*} \left| R(y_n^* - w) e^{\beta^{(0)}(y_n^* - w)/\beta^{(1)}} \right| < \frac{\epsilon}{|K_2|}.$$

Consider the piece

$$\begin{aligned}
&\int_{-\infty}^{a y_n^*} [1 + \epsilon(y_n^* - w)] K_2 e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}} dF_W(w) \\
&= K_2 e^{-\beta^{(0)} y_n^*/\beta^{(1)}} \int_{-\infty}^{a y_n^*} e^{\beta^{(0)} w/\beta^{(1)}} dF_W(w) \\
&\qquad + K_2 \int_{-\infty}^{a y_n^*} \epsilon(y_n^* - w) e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}} dF_W(w).
\end{aligned}$$

Examining the first integral,

$$\begin{aligned}
\lim_{y_n^* \to \infty} \int_{-\infty}^{a y_n^*} K_2 e^{\beta^{(0)} w/\beta^{(1)}} dF_W(w) &= K_2 E\left(e^{\beta^{(0)} W/\beta^{(1)}}\right) = K_2 E\left[\exp\left(\sum_{k=2}^{n-1} \frac{2\beta^{(k)}}{\beta^{(1)}} Z_k\right)\right] \\
&= -\frac{\beta^{(1)}}{\beta^{(0)} - \beta^{(1)}} \Gamma\left(2 - \frac{\beta^{(0)}}{\beta^{(1)}}\right) \prod_{k=2}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(1)}}\right).
\end{aligned}$$

Turning to the second integral,

$$\begin{aligned}
\left| K_2 \int_{-\infty}^{a y_n^*} \epsilon(y_n^* - w) e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}} dF_W(w) \right| &= \left| \int_{-\infty}^{a y_n^*} R(y_n^* - w) dF_W(w) \right| \\
< \epsilon \int_{-\infty}^{a y_n^*} e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}} dF_W(w) &\leq \epsilon e^{-\beta^{(0)} y_n^*/\beta^{(1)}} \int_{-\infty}^{\infty} e^{\beta^{(0)} w/\beta^{(1)}} dF_W(w) \\
= \epsilon E\left(e^{\beta^{(0)} W/\beta^{(1)}}\right) e^{-\beta^{(0)} y_n^*/\beta^{(1)}}.
\end{aligned}$$

Therefore

$$K_2 \int_{-\infty}^{ay_n^*} \epsilon(y_n^* - w) e^{-\beta^{(0)}(y_n^* - w)/\beta^{(1)}} dF_W(w) = o\left(e^{-\beta^{(0)} y_n^* / \beta^{(1)}}\right) = o\left(e^{-y/\beta^{(1)}}\right).$$

Next, note that

$$K_1 \int_{-\infty}^{\infty} e^w dF_W(w) = \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right).$$

We now consider the integral

$$K_1 e^{-y_n^*} \int_{-\infty}^{ay_n^*} e^w dF_W(w) = \prod_{k=1}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right) e^{-y_n^*} - K_1 e^{-y_n^*} \int_{ay_n^*}^{\infty} e^w dF_W(w).$$

Gathering all the terms,

$$\begin{aligned}\int_{-\infty}^{ay_n^*} \overline{F}_V(y_n^* - w) dF_W(w) = &\prod_{k=1}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(0)}}\right) e^{-y_n^*} + o\left(e^{-y/\beta^{(1)}}\right) \\ &- \frac{\beta^{(1)}}{\beta^{(0)} - \beta^{(1)}} \Gamma\left(2 - \frac{\beta^{(0)}}{\beta^{(1)}}\right) \prod_{k=2}^{n-1} \Gamma\left(1 - \frac{\beta^{(k)}}{\beta^{(1)}}\right) e^{-\beta^{(0)} y_n^* / \beta^{(1)}} \\ &- K_1 e^{-y_n^*} \int_{ay_n^*}^{\infty} e^w dF_W(w),\end{aligned}$$

and by previous work the latter integral is $o\left(e^{-y/\beta^{(1)}}\right)$. Collecting the pieces, the result follows. $\square$

B.5 Infinite Products 1

Table B.1 displays values of $\prod_{k=1}^{\infty}\Gamma(1-\beta^k)$ for values of β between -0.89 and 0.69, given to three decimal places. For $\beta < -0.89$ and $\beta > 0.69$, the products quickly grow large and are not worth reproducing in a table.

Table B.1: Infinite Products 1

β	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
-0.8	3.536	4.005	4.607	5.397	6.463	7.947	10.091	13.334	18.522	27.427
-0.7	1.687	1.769	1.862	1.970	2.095	2.240	2.412	2.617	2.863	3.164
-0.6	1.229	1.257	1.287	1.320	1.357	1.398	1.443	1.494	1.551	1.615
-0.5	1.052	1.064	1.077	1.091	1.106	1.122	1.140	1.160	1.181	1.204
-0.4	0.974	0.979	0.985	0.991	0.998	1.005	1.013	1.022	1.031	1.041
-0.3	0.941	0.943	0.945	0.947	0.950	0.953	0.956	0.960	0.964	0.969
-0.2	0.937	0.937	0.936	0.936	0.936	0.936	0.937	0.937	0.938	0.940
-0.1	0.956	0.953	0.951	0.948	0.946	0.944	0.942	0.941	0.939	0.938
-0.0	1.000	0.994	0.989	0.984	0.979	0.975	0.971	0.967	0.963	0.960
0.0	1.000	1.006	1.012	1.019	1.026	1.033	1.041	1.049	1.057	1.066
0.1	1.076	1.085	1.096	1.107	1.118	1.130	1.143	1.156	1.170	1.185
0.2	1.200	1.216	1.233	1.251	1.270	1.290	1.312	1.334	1.357	1.382
0.3	1.409	1.437	1.467	1.498	1.532	1.567	1.605	1.646	1.689	1.735
0.4	1.785	1.838	1.895	1.956	2.022	2.094	2.171	2.255	2.347	2.446
0.5	2.555	2.674	2.805	2.949	3.109	3.286	3.483	3.703	3.950	4.228
0.6	4.544	4.903	5.314	5.787	6.336	6.976	7.729	8.622	9.689	10.977

B.6 Infinite Products 2

Table B.2 shows values of $\prod_{k=1}^{\infty} \Gamma(1-2\beta^k)$ for values of β between -.70 and 0.49, given to three decimal places. Notice that since this product is only defined on $-\frac{1}{\sqrt{2}} < \beta < \frac{1}{2}$ as explained in Section 4.4, the table only goes down to -0.69.

Table B.2: Infinite Products 2

β	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
-0.6	3.787	4.303	4.962	5.827	6.998	8.648	11.099	15.032	22.145	38.138
-0.5	1.716	1.810	1.916	2.038	2.180	2.344	2.538	2.767	3.042	3.376
-0.4	1.194	1.225	1.260	1.298	1.340	1.386	1.438	1.496	1.561	1.634
-0.3	0.999	1.012	1.025	1.040	1.057	1.075	1.094	1.116	1.140	1.166
-0.2	0.927	0.931	0.935	0.940	0.946	0.953	0.960	0.968	0.978	0.988
-0.1	0.928	0.925	0.923	0.921	0.920	0.920	0.920	0.921	0.922	0.924
-0.0	1.000	0.989	0.979	0.970	0.961	0.954	0.947	0.941	0.936	0.932
0.0	1.000	1.012	1.025	1.040	1.055	1.072	1.090	1.110	1.131	1.155
0.1	1.180	1.207	1.236	1.268	1.303	1.341	1.382	1.426	1.475	1.528
0.2	1.587	1.651	1.721	1.799	1.885	1.980	2.086	2.205	2.338	2.487
0.3	2.657	2.850	3.070	3.324	3.619	3.963	4.369	4.853	5.435	6.147
0.4	7.029	8.144	9.585	11.499	14.135	17.944	23.831	33.916	54.577	117.738

Appendix C Constants for Gumbel Convolution

This collection of appendices establishes numerical values for θ_k, ζ_k, $\mu_{i,m}$, $a_{k,i}$, and $A_{k,i}$ in Chapter 5.

C.1 Values for θ_k and ζ_k

Recall that $\theta_k = -\int_0^\infty x^k \left[1 - \exp\left(-e^{-x}\right)\right] dx$ and $\zeta_k = \int_{-\infty}^0 x^k \exp\left(-e^{-x}\right) dx$. Table C.1 lists the numerical values for $k = 0, \cdots, 16$. Observe that in each case, the value of θ_k is fairly close to $-k!$, and for $k \geq 13$ there is little to no difference.

Table C.1: Values for θ_k and ζ_k

k	θ_k	ζ_k
0	-0.7966	0.2194
1	-0.8912	-0.0978
2	-1.8862	0.0712
3	-5.8239	-0.0664
4	-23.6405	0.0727
5	-119.0888	-0.0891
6	-717.2406	0.1193
7	-5030.2812	-0.1715
8	-40280.9602	0.2614
9	-362703.8226	-0.4187
10	-3627917.4412	0.7001
11	-39911939.7644	-1.2161
12	-478972413.8420	2.1846
13	-6226830982.4340	-4.0452
14	-87176961974.0852	7.6989
15	-1307664396286.7170	-15.0244
16	-20922710100842.7969	30.0029

C.2 Values for $\mu_{i,m}$

Recall that for $0 \le i \le m$,

$$\mu_{i,m} = \sum \left(\frac{i!}{k_1! \cdots k_i!} \right) \left(\frac{m!}{(m - [k_1 + \cdots + k_i])!} \right) \zeta_0^{m-(k_1+\cdots+k_i)} \prod_{L=1}^{i} \left(\frac{\zeta_L}{L!} \right)^{k_L},$$

where the sum is taken over all nonnegative integers $k_1, \cdots, k_i$ such that $k_1+2k_2+\cdots+ik_i = i$. Table C.2 lists values for $0 \le i \le m \le 7$.

Table C.2: Values for $\mu_{i,m}$

i \ m	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	1.0000	0.2194	0.0481	0.0106	0.0023	0.0005	0.0001	2.4e-5
1	—	-0.0978	-0.0429	-0.0141	-0.0041	-0.0011	-0.0003	-7.6e-5
2	—	—	0.0504	0.0229	0.0085	0.0029	0.0009	0.0003
3	—	—	—	-0.0427	-0.0198	-0.0079	-0.0028	-0.0010
4	—	—	—	—	0.0506	0.0238	0.0099	0.0037
5	—	—	—	—	—	-0.0770	-0.0365	-0.0156
6	—	—	—	—	—	—	0.1432	0.0683
7	—	—	—	—	—	—	—	-0.3148

It may also be handy to have explicit formulas for $\mu_{i,m}$ in terms of the ζ_k. They quickly get messy, so we only give them for the values listed in Table C.2. First observe that $\mu_{0,m} = \zeta_0^m, m = 0, 1, \cdots$, and $\mu_{1,m} = m\zeta_1^{m-1}, m = 1, 2, \cdots$. We now list formulas for the remaining constants.

$$\mu_{2,2} = 2\zeta_1^2 + 2\zeta_0\zeta_2 \qquad \mu_{2,5} = 20\zeta_0^3\zeta_1^2 + 5\zeta_0^4\zeta_2$$

$$\mu_{2,3} = 6\zeta_0\zeta_1^2 + 3\zeta_0^2\zeta_2 \qquad \mu_{2,6} = 30\zeta_0^4\zeta_1^2 + 6\zeta_0^5\zeta_2$$

$$\mu_{2,4} = 12\zeta_0^2\zeta_1^2 + 4\zeta_0^3\zeta_2 \qquad \mu_{2,7} = 42\zeta_0^5\zeta_1^2 + 7\zeta_0^6\zeta_2$$

$$\mu_{3,3} = 6\zeta_1^3 + 18\zeta_0\zeta_1\zeta_2 + 3\zeta_0^2\zeta_3 \qquad \mu_{3,6} = 120\zeta_0^3\zeta_1^3 + 90\zeta_0^4\zeta_1\zeta_2 + 6\zeta_0^5\zeta_3$$

$$\mu_{3,4} = 24\zeta_0\zeta_1^3 + 36\zeta_0^2\zeta_1\zeta_2 + 4\zeta_0^3\zeta_3 \qquad \mu_{3,7} = 210\zeta_0^4\zeta_1^3 + 126\zeta_0^5\zeta_1\zeta_2 + 7\zeta_0^6\zeta_3$$

$$\mu_{3,5} = 60\zeta_0^2\zeta_1^3 + 60\zeta_0^3\zeta_1\zeta_2 + 5\zeta_0^4\zeta_3$$

$$\mu_{4,4} = 24\zeta_1^4 + 144\zeta_0\zeta_1^2\zeta_2 + 48\zeta_0^2\zeta_1\zeta_3 + 36\zeta_0^2\zeta_2^2 + 4\zeta_0^3\zeta_4$$

$$\mu_{4,5} = 120\zeta_0\zeta_1^4 + 360\zeta_0^2\zeta_1^2\zeta_2 + 80\zeta_0^3\zeta_1\zeta_3 + 60\zeta_0^3\zeta_2^2 + 5\zeta_0^4\zeta_4$$

$$\mu_{4,6} = 360\zeta_0^2\zeta_1^4 + 720\zeta_0^3\zeta_1^2\zeta_2 + 120\zeta_0^4\zeta_1\zeta_3 + 90\zeta_0^4\zeta_2^2 + 6\zeta_0^5\zeta_4$$

$$\mu_{4,7} = 840\zeta_0^3\zeta_1^4 + 1260\zeta_0^4\zeta_1^2\zeta_2 + 168\zeta_0^5\zeta_1\zeta_3 + 126\zeta_0^5\zeta_2^2 + 7\zeta_0^6\zeta_4$$

$$\mu_{5,5} = 120\zeta_1^5 + 1200\zeta_0\zeta_1^3\zeta_2 + 600\zeta_0^2\zeta_1^2\zeta_3 + 900\zeta_0^2\zeta_1\zeta_2^2 + 100\zeta_0^3\zeta_1\zeta_4 + 200\zeta_0^3\zeta_2\zeta_3 + 5\zeta_0^4\zeta_5$$

$$\mu_{5,6} = 720\zeta_0\zeta_1^5 + 3600\zeta_0^2\zeta_1^3\zeta_2 + 1200\zeta_0^3\zeta_1^2\zeta_3 + 1800\zeta_0^3\zeta_1\zeta_2^2 + 150\zeta_0^4\zeta_1\zeta_4 + 300\zeta_0^4\zeta_2\zeta_3 + 6\zeta_0^5\zeta_5$$

$$\mu_{5,7} = 2520\zeta_0^2\zeta_1^5 + 8400\zeta_0^3\zeta_1^3\zeta_2 + 2100\zeta_0^4\zeta_1^2\zeta_3 + 3150\zeta_0^4\zeta_1\zeta_2^2 + 210\zeta_0^5\zeta_1\zeta_4 + 420\zeta_0^5\zeta_2\zeta_3 + 7\zeta_0^6\zeta_5$$

$$
\begin{aligned}
\mu_{6,6} = {} & 720\zeta_1^6 + 10800\zeta_0\zeta_1^4\zeta_2 + 7200\zeta_0^2\zeta_1^3\zeta_3 + 16200\zeta_0^2\zeta_1^2\zeta_2^2 + 1800\zeta_0^3\zeta_1^2\zeta_4 \\
& + 7200\zeta_0^3\zeta_1\zeta_2\zeta_3 + 180\zeta_0^4\zeta_1\zeta_5 + 450\zeta_0^4\zeta_2\zeta_4 + 300\zeta_0^4\zeta_3^2 + 1800\zeta_0^3\zeta_2^3 + 6\zeta_0^5\zeta_6 \\
\mu_{6,7} = {} & 5040\zeta_0\zeta_1^6 + 37800\zeta_0^2\zeta_1^4\zeta_2 + 16800 * \zeta_0^3\zeta_1^3\zeta_3 + 37800 * \zeta_0^3\zeta_1^2\zeta_2^2 + 3150\zeta_0^4\zeta_1^2\zeta_4 \\
& + 12600\zeta_0^4\zeta_1\zeta_2\zeta_3 + 252\zeta_0^5\zeta_1\zeta_5 + 630\zeta_0^5\zeta_2\zeta_4 + 420\zeta_0^5\zeta_3^2 + 3150\zeta_0^4\zeta_2^3 + 7\zeta_0^6\zeta_6
\end{aligned}
$$

$$
\begin{aligned}
\mu_{7,7} = {} & 5040\zeta_1^7 + 105840\zeta_0\zeta_1^5\zeta_2 + 88200\zeta_0^2\zeta_1^4\zeta_3 + 264600\zeta_0^2\zeta_1^3\zeta_2^2 + 29400\zeta_0^3\zeta_1^3\zeta_4 \\
& + 176400\zeta_0^3\zeta_1^2\zeta_2\zeta_3 + 4410\zeta_0^4\zeta_1^2\zeta_5 + 88200\zeta_0^3\zeta_1\zeta_2^3 + 14700\zeta_0^4\zeta_1\zeta_3^2 \\
& + 22050\zeta_0^4\zeta_1\zeta_2\zeta_4 + 294\zeta_0^5\zeta_1\zeta_6 + 22050\zeta_0^4\zeta_2^2\zeta_3 + 882\zeta_0^5\zeta_2\zeta_5 + 1470\zeta_0^5\zeta_3\zeta_4 + 7\zeta_0^6\zeta_7
\end{aligned}
$$

C.3 Values for $a_{n,i}$

Recall the formula for $a_{n,i}$ and $n \geq 3$:

$$a_{n,i} = \begin{cases} \displaystyle\sum_{j=0}^{n-2} a_{n,j}(-1)^j \theta_j + a_{n-1,0}, & i = 0 \\ \displaystyle\frac{a_{n-1,i-1}}{i} + \sum_{j=i}^{n-2} \binom{j}{i} (-1)^{j-i} a_{n-1,j} \theta_{j-i}, & i = 1, \cdots, n-2 \\ \displaystyle\frac{1}{(n-1)!}, & i = n-1. \end{cases}$$

Table C.3 displays the $a_{n,i}$ constants up to $n = 7$. For example, if we needed $a_{4,2}$, then we obtain -1.0932.

Table C.3: Values for $a_{n,i}$

n \ **i**	**0**	**1**	**2**	**3**	**4**	**5**	**6**
1	1.0000	—	—	—	—	—	—
2	-0.5932	1.0000	—	—	—	—	—
3	0.7706	-1.3898	0.5000	—	—	—	—
4	-2.0250	2.7689	-1.0932	0.1667	—	—	—
5	5.0884	-7.1223	2.7009	-0.4972	0.0417	—	—
6	-14.2873	19.3600	-7.5134	1.4449	-0.1575	0.0083	—
7	41.6497	-55.9312	21.7958	-4.3741	0.5238	-0.0381	0.0014

It may also be handy to have explicit formulas for $a_{n,i}$. They quickly get complicated, so we only give them for the values listed in Table C.3.

$$a_{1,0} = 1 \qquad a_{3,0} = 1 + 3\theta_0 + 2\theta_0^2 - \theta_1$$

$$a_{2,0} = 1 + 2\theta_0 \qquad a_{3,1} = 1 + 3\theta_0$$

$$a_{2,1} = 1 \qquad a_{3,2} = \frac{1}{2}$$

$$a_{4,0} = 1 + 4\theta_0 + 5\theta_0^2 + 2\theta_0^3 - 4\theta_0\theta_1 - 2\theta_1 + \frac{1}{2}\theta_2$$

$$a_{4,1} = 1 + 4\theta_0 + 5\theta_0^2 - 2\theta_1$$

$$a_{4,2} = \frac{1}{2}[1 + 4\theta_0]$$

$$a_{4,3} = \frac{1}{6}$$

$$a_{5,0} = 1 + 5\theta_0 + 9\theta_0^2 + 7\theta_0^3 + 2\theta_0^4 - 10\theta_0\theta_1 - 9\theta_0^2\theta_1 + \frac{5}{2}\theta_0\theta_2 - 3\theta_1 + 2\theta_1^2 + \theta_2 - \frac{1}{6}\theta_3$$

$$a_{5,1} = 1 + 5\theta_0 + 9\theta_0^2 + 7\theta_0^3 - 10\theta_0\theta_1 - 3\theta_1 + \theta_2$$

$$a_{5,2} = \frac{1}{2}[1 + 5\theta_0 + 9\theta_0^2 - 3\theta_1]$$

$$a_{5,3} = \frac{1}{6}[1 + 5\theta_0] \qquad a_{5,4} = \frac{1}{24}$$

$$a_{6,0} = 1 + 6\theta_0 + 14\theta_0^2 + 16\theta_0^3 + 9\theta_0^4 + 2\theta_0^5 - 18\theta_0\theta_1 - 28\theta_0^2\theta_1 - 16\theta_0^3\theta_1 + 12\theta_0\theta_1^2$$
$$+ 7\theta_0^2\theta_2 + 6\theta_0\theta_2 - \theta_0\theta_3 - 4\theta_1 + 5\theta_1^2 - \frac{5}{2}\theta_1\theta_2 + \frac{3}{2}\theta_2 - \frac{1}{3}\theta_3 + \frac{1}{24}\theta_4$$

$$a_{6,1} = 1 + 6\theta_0 + 14\theta_0^2 + 16\theta_0^3 + 9\theta_0^4 - 18\theta_0\theta_1 - 28\theta_0^2\theta_1 + 6\theta_0\theta_2 - 4\theta_1 + 5\theta_1^2$$
$$+ \frac{3}{2}\theta_2 - \frac{1}{3}\theta_3$$

$$a_{6,2} = \frac{1}{2}\left[1 + 6\theta_0 + 14\theta_0^2 + 16\theta_0^3 - 18\theta_0\theta_1 - 4\theta_1 + \frac{3}{2}\theta_2\right]$$

$$a_{6,3} = \frac{1}{6}\left[1 + 6\theta_0 + 14\theta_0^2 - 4\theta_1\right]$$

$$a_{6,4} = \frac{1}{24}[1 + 6\theta_0] \qquad a_{6,5} = \frac{1}{120}$$

$$\begin{aligned}
a_{7,0} = {} & 1 + 7\theta_0 + 20\theta_0^2 + 30\theta_0^3 + 25\theta_0^4 + 11\theta_0^5 + 2\theta_0^6 - 28\theta_0\theta_1 - 60\theta_0^2\theta_1 - 60\theta_0^3\theta_1 \\
& - 25\theta_0^4\theta_1 + 40\theta_0^2\theta_1^2 + 35\theta_0\theta_1^2 - \frac{35}{2}\theta_0\theta_1\theta_2 + \frac{21}{2}\theta_0\theta_2 + 20\theta_0^2\theta_2 + 15\theta_0^3\theta_2 \\
& - \frac{7}{3}\theta_0\theta_3 - \frac{10}{3}\theta_0^2\theta_3 + \frac{7}{24}\theta_0\theta_4 - 5\theta_1 + 9\theta_1^2 - 5\theta_1^3 - 6\theta_1\theta_2 + \theta_1\theta_3 + 2\theta_2 + \frac{3}{4}\theta_2^2 \\
& - \frac{1}{2}\theta_3 + \frac{1}{12}\theta_4 \\
a_{7,1} = {} & 1 + 7\theta_0 + 20\theta_0^2 + 30\theta_0^3 + 25\theta_0^4 + 11\theta_0^5 - 28\theta_0\theta_1 - 60\theta_0^2\theta_1 - 60\theta_0^3\theta_1 + 35\theta_0\theta_1^2 \\
& + 20\theta_0^2\theta_2 + \frac{21}{2}\theta_0\theta_2 - \frac{7}{3}\theta_0\theta_3 - 5\theta_1 + 9\theta_1^2 - 6\theta_1\theta_2 + 2\theta_2 - \frac{1}{2}\theta_3 + \frac{1}{12}\theta_4 \\
a_{7,2} = {} & \frac{1}{2}\left[1 + 7\theta_0 + 20\theta_0^2 + 30\theta_0^3 + 25\theta_0^4 - 28\theta_0\theta_1 - 60\theta_0^2\theta_1 + \frac{21}{2}\theta_0\theta_2 - 5\theta_1 + 9\theta_1^2\right. \\
& \left. +2\theta_2 - \frac{1}{2}\theta_3\right] \\
a_{7,3} = {} & \frac{1}{6}[1 + 7\theta_0 + 20\theta_0^2 + 30\theta_0^3 - 28\theta_0\theta_1 - 5\theta_1 + 2\theta_2] \\
a_{7,4} = {} & \frac{1}{24}[1 + 7\theta_0 + 20\theta_0^2 - 5\theta_1] \\
a_{7,5} = {} & \frac{1}{120}[1 + 7\theta_0] \qquad a_{7,6} = \frac{1}{720}
\end{aligned}$$

C.4 Values for $A_{n,i}$

Recall the formula for $A_{n,i}$ and $n \geq 3$:

$$A_{n,i} = \sum_{m=1}^{n-i-1} \binom{n}{m} \sum_{j=i}^{n-m-1} a_{n-m,j} \binom{j}{i} (-1)^{j-i} \mu_{j-i,m} + a_{n,i}, \qquad i = 0, \cdots, n-2, \qquad \text{and} \tag{3.1}$$

$$A_{n,n-1} = \frac{1}{(n-1)!}.$$

Table C.4 displays the $A_{n,i}$ constants up to $n = 7$. For example, if we needed $A_{4,2}$, then we obtain -0.6544.

Table C.4: Values for $A_{n,i}$

n \ i	0	1	2	3	4	5	6
1	1.0000	—	—	—	—	—	—
2	-0.1544	1.0000	—	—	—	—	—
3	0.8181	-0.7316	0.5000	—	—	—	—
4	-1.7642	2.2294	-0.6544	0.1667	—	—	—
5	4.3381	-5.2887	1.9870	-0.3143	0.0417	—	—
6	-12.2325	14.1488	-5.1955	1.0086	-0.1026	0.0083	—
7	34.5721	-40.0139	14.7598	-2.8773	0.3526	-0.0253	0.0014