Predictive Modeling of Professional Figure Skating Tournament Data

by

Marlow Quintez Lemons

(Under the direction of Jaxk Reeves)

Abstract

This thesis examines various stochastic models for modeling the performance scores earned by competitors in figure skating competitions. Using the results from 107 professional figure skating tournaments over a 5-year period, a reasonably parsimonious model relating performance on the initial (short program) segment to performance on the final (long program) segment is derived. Tests of fit and checks for model validity and accuracy are also conducted to show that the model is adequate.

Index words: Figure Skating, Prediction, Stochastic Modeling

Predictive Modeling of Professional Figure Skating Tournament Data

by

Marlow Quintez Lemons

B.S., Georgia Southern University, 2001

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

Master of Science

Athens, Georgia

2003

Predictive Modeling of Professional Figure Skating Tournament Data

by

Marlow Quintez Lemons

Approved:

Major Professor: Jaxk Reeves

Committee: Lynne Seymour
Robert Lund

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2003

Acknowledgments

I would like to first acknowledge my major professor, Dr. Jaxk Reeves, who has been there to guide me through this research. Thank you so much for allowing me to be your student. I couldn't have picked a better professor to work with. Without your patience, knowledge, support, and advice, this wouldn't been possible. To the advising committee, Dr. Lynne Seymour and Dr. Robert Lund, I thank you for reviewing my thesis and providing the necessary comments to make this paper possible. Finally, I would like to thank Ms. Connie Durden for providing the final format check before submission to the graduate school.

Preface

My first dedication goes to the One above who gives me life each day. To my parents, I thank them for being there when I needed them most. My final dedication goes to Ms. Clara Taylor for being there to talk to during my roughest periods while writing this thesis.

Table of Contents

List of Tables

Chapter 1

Introduction

It was February 2002 in Salt Lake City, the site of the Winter Olympics. Sara Hughes, Olympic figure skater, trailed Michelle Kwan, Irina Slutskaya, and Sasha Cohen after the short program in the women's figure skating competition. Intense moments arose for both the Olympic skaters and the spectators when the final round, the long program, commenced. Kwan, known world-wide as a legendary skater, fell short in her long program routine by "two-footing a triple toe loop in combination and falling on a triple flip". Slutskaya's dry routine following a sloppy landing on a triple flip was rewarded with modest marks by the judges. Cohen's routine left the judges rendering bad marks as she lacked great footwork, and also fell on a triple toe loop. Hughes' presentation, on the other hand, was full of fire and the flawless techniques that the judges were expecting. A captivating long program score suddenly sent her from fourth place to a surprising first place finish. Hughes, followed by her coach, immediately fell to the ground in tears after reading the final scores from the scoreboard above. Rated by Sports Illustrated as 'The Most Unexpected Skater to Win the Olympics', Sarah Hughes left the ice that evening as a gold medalist.

In figure skating, switching adjacent positions from initial ranking to final rankings (such as third to second, fourth to fifth, or even second to first) is not uncommon. However, because of its scoring procedures, moving from third to first in figure skating is a somewhat rare occurrence, and moving from fourth to first is very rare. But exactly how rare was Ms. Hughes' accomplishment? Is it something that would occur once in one hundred competitions? Once in one thousand? Once in one million? A goal of this thesis is to formulate a model for figure skating scoring so that questions such as the above can be analyzed accurately.

Referring back to the 2002 Winter Olympics with Sarah Hughes, how likely was it for her to improve from fourth to first? One purpose of this project is to create a statistical model for predicting the conditional probability that a skater ranked ith after the first round (short program) will end up in jth place overall at the conclusion of the tournament. This will enable fans or analysts to make mid-tournament assessments of the expected placements of various competitors. It would also lend some objectivity to statements concerning how rare an event such as Sarah Hughes' comeback win was.

Chapter 2

Literature Review

2.1 History of Figure Skating

Ice skating, in particular figure skating, combines sport and art together in a way that sports such as football and baseball can't. The craft of interpreting a song through skating has now become, in our culture, a competition. The origins of ice skating are still debated. Some historians trace the start of ice skating back to Scotland in the mid 1600s, when the Scots used ice skates to cross over frozen fens and marshes in an attempt to invade enemy territory (Jonland and Fitzgerald). During this same period, it was said that Norway and Finland introduced ice skating as a means of transportation (Jonland and Fitzgerald). Northern Norwegians used ice skates and skis to move over frozen ice and slick roads where regular transportation couldn't. At that time, ice skates were made of bone rather than steel, and were attached to the bottom of skaters' boots. Eventually, ice skating was introduced to the United States, but not until two hundred years later. Here, it was not used as a means of transportation or war, but for art, recreation, and (beginning in 1932) competition.

The first skating club, called the Edinburgh Skating Club, dates back to 1742 in Edinburgh (Copley-Graves). By 1778, this organization had documented membership requirements by means of a skating test which required demonstration of artistic circles and high jumps (Copley-Graves). Most organizations didn't become established until the late nineteenth and early twentieth centuries. In the late nineteenth century (1898), the Cambridge Skating Club was founded by an American named

George Browne, and was internationally recognized as the first American skating club (Copley-Graves). It was then that Browne made popular what is now called the "ten-step", a technique for beginning figure skaters, which was introduced two years before the formation of the Figure Skating Club of London and the Princes' Skating Club in 1900. The market for figure skating competitions grew in the United States as small competitions blossomed in the northern part of the country. Because of this, clubs and organizations in figure skating became popular outdoor activities to join. Figure skating was inaugurated in the Winter Olympics in 1908, after Olympic officials noted the strong popularity of international competitions. Such competitions included the International Skating Tournament in Vienna in 1882 and the World Championship Figure Skating Competition beginning in 1896. Towards the end of the nineteenth century, the world began to recognize popular figure skaters. One expert was Ulrich Salchow, a Swedish ten-time world championship gold medalist who invented the Salchow jump still used in competitions today. Later stars included Gillis Grafstrom, a Swedish skater who won three Olympic titles (1920, 1924, and 1928), and Karl Schafer of Austria who won seven world championships and two gold medals. Today, his name remains known as it marks the name of a popular international skating competition in Austria. With so many figure skating organizations and competitions, the Dutch felt that there was a need to establish international standards to govern competitions. This would insure that international standards were followed when different countries competed against one another. In July 1892, the Dutch took the lead in calling a meeting of representatives from all countries interested in international ice skating competitions.

Thus, in that same year, the International Skating Union (ISU) was formed. Its purpose was to merge nationally recognized competitions and organizations together, and to host tournaments with rules compatible with every nation's skating guidelines. Currently, the ISU still operates with these same purposes. Now one hundred and eleven years old, the ISU is recognized as the oldest governing international winter sports federation. Although the ISU was formed in 1892, the United States felt that there should be a union of American clubs and organizations. Thus, almost thirty years later (in 1921), the United States Figure Skating Association (USFSA) was formed. The USFSA felt that state and local competitions in the US should be governed by principles complying with ISU's universal skating rules, but should be held under the jurisdiction of the USFSA. To achieve this, the same sanctioning rights provided by the ISU were granted by the USFSA, since it, too, was a member of ISU. Today, the purpose of the USFSA is to provide sanctioning and financial support to figure skating organization and competitions in the US. Although figure skating became popular at the start of the twentieth century, early competitions did not contain the musical choreography and dancing that we witness today. Originally, figure skating was more intense, as skaters were forced to perform a required set of skating techniques. In those days, dancing in Olympic competitions would have resulted in disqualification. But matters changed slowly. The judges at the World Championship allowed dance routines into figure skating in 1952, but dancing wasn't allowed in the Winter Olympics until 1976, more than two decades later. Today, a presentation without dancing is not pleasing to the judges, and would have no hope of winning.

2.2 Figure Skating Scoring Systems

Most figure skating competitions, in particular the professional level competitions, are composed of two rounds: a short program (SP) and a long program (LP). A competition begins with the SP, where n skaters are given two to three minutes to perform for the judges a required set of routines such as loops, lutzes, spirals, and axels. For each skater, the judges then render two types of marks, technical and presentation marks, each of which ranges from 0 to 6. The technical marks measure the difficulty of the program and the clean execution of the elements required. The presentation marks, on the other hand, reflect the skater's choreographic prowess and his/her ability to interpret the chosen music. The two marks from each judge are summed and used to assign each skater a placement, using the tournament's system of ranking the skaters. This will be explained later in this thesis, but the key concept is that each judge's relative rank of each skater, rather than his/her absolute rating of a skater, is the information from which the overall placements are obtained. These placements range from 1 to n (where the lowest placement, 1, represents the best skater from that round) with ties almost always precluded by the scoring system. The same scoring procedure is used in the long program (also known as the freeskate program), where skaters are not only required to perform the same type of techniques, but are rigorously rated on their choreography, interpretation of music through skating, speed, assurance, and balance of program, as well as on their ability to use the entire skating rink. After the LP round, skaters again receive technical and presentation marks which are converted to placements in the same manner. Because the LP round lasts twice as long and is considered to be more intense than the SP, its awarded placement is weighted twice as highly as the SP placement. The placements awarded for each skater are used to calculate a total score (TS) using the following formula:

$$TS = SP + 2(LP), \tag{2.1}$$

where SP is the placement given to a skater from the short program round and LP is the placement awarded by the judges from the long program. The skater with the lowest total score is declared the winner of the competition and other places are awarded by ascending point value. Although the ranking procedure used in obtaining the SP and LP placements generally precludes tied ranks for either component, it

Table 2.1: Hypothetical Figure Skating Results Example ($n = 6$)

Skaters	SP	LP	$TS = [SP + (2 * LP)]$
A	1	3	7
B	2	6	14
C	3	2	7
D	4	5	14
E	5	1	7
F	6	4	14

could easily happen that two or more competitors could earn the same total score. For example, suppose the placements among six competitors were as shown in Table 2.1. While the scenario of Table 2.1, as we shall see later, is not at all likely to occur in practice, there must be a rule to break ties which do occur. Should two or more skaters have the same total score, the skater whose LP rank is lowest is declared the victor. Thus, in the example above, the final placing from best to worst would be {E, C, A, F, D, B}. On the other hand, if skaters C and E had reversed their placements in the long program, the final order would have been {C, A, E, F, D, B}.

It is frequently the case that a figure skater placing first in the short program also places first in the long program (of course making that skater the winner). In other words, the rankings given from the SP round are frequently similar to those awarded in the LP round. However, the relationship isn't perfect. It is certainly possible for a skater to drop drastically from SP ranking to final ranking because of a poor LP performance. In statistical terms, one would say that SP and LP rankings are positively correlated, but the extent of this correlation must be quantified.

As figure skating became popular over this period, the scoring systems used became more consistent. The two most popular scoring systems currently used to obtain "placements" in figure skating competitions are the Ordinal system and the One-By-One system. This research considers only the total scores, so the ranking system used to convert the technical and presentation marks into placements in these tournaments is of secondary importance. Nonetheless, understanding how the placements are derived will help motivate the statistical models of this thesis.

As previously discussed, each judge renders a technical and presentation mark for every skater. Notationally, let $TMS(i, j)$ and $PMS(i, j)$ be the technical and presentation marks given to the ith skater by the jth judge in the short program, and, similarly, let $TML(i, j)$ and $PML(i, j)$ be the technical and presentation marks given in the long program. Note that these marks yield no intrinsic description of a skater's ability (Loosemore). Rather, the sums of these two marks are used only to assign each of the n skaters a relative rank for that judge, ranging from 1 (the best skater) to n (being the worst skater). That is $RS(i, j)$ is the relative rank [on a 1 to n scale] assigned to the ith skater by the jth judge. Following the judges' ranking, the tournament's chosen scoring system is used to assign placements, a process which is explained in more detail below. Finally, the placements awarded are used to calculate the total score, which is used to determine the winner. Table 2.2 outlines this process.

The process by which judges' marks are converted to judges ranks (step 1 $\rightarrow$ step 2 above) and by which total score is obtained from placements (step 3 $\rightarrow$ step 4) are easy to understand. However, the process of converting judges' individual ranks (step 2) to placements (step 3) was and is one of the thorniest issues in figure skating evaluation. In this thesis, we will examine two of the more commonly used methods for doing this, the Ordinal system and the One-by-One system.

Table 2.2: Steps Leading to the Total Score

		SP	**LP**
Step 1	**Judges' Marks**	$XS(i,j) = TMS(i,j) + PMS(i,j)$ ↓	$XL(i,j) = TML(i,j) + PML(i,j)$ ↓
Step 2 2	**Judges' Ranks**	$RS(i,j)$ ↓	$RL(i,j)$ ↓
Step 3	**Placements**	$SP(i)$ ↓	$LP(i)$ ↓
Step 4	**Total Score**	↘ $TS(i) = SP(i) + 2LP(i)$	↙

Still used in most skating competitions, the Ordinal system allows judges to offer placements based on the "majority vote". It is one method used to convert judges' ranks to placements. In order to explain how the Ordinal scoring system works, let's look at an example. Table 2.3 displays the results of the mens' short program (SP) from the 2001 Goodwill Games in Brisbane, Australia, starting from the judges' ranks (which we will call an R-Table). There were eleven skaters each evaluated by seven judges in this competition. The first judge's column (J1) shows that the first judge ranked Evengi Plushenko as his top ranked skater, followed by Johnny Weir in second, Ilia Klimkin in third, Anthony Liu in fourth, etc. If one were to consider the list to be the order in which the skaters presented, then ninth skater Takeshi Honda, with rank of 7 from judge number six could be written as $R(9,6) = 7$. With respect to the rows, Evengi Plushenko was ranked the top skater by all seven judges while Michael Weiss was ranked third by five judges, second by one judge, and seventh by one judge. In general, in order for a skater to have the majority vote, he/she must capture at least half of the judges' votes at that rank or lower. In this example, a

Table 2.3: R-Table of 2001 Goodwill Games

Skaters	**Judges**							**Ordinal**	**MJ**	**SMJ**	**Place.**
	J1	**J2**	**J3**	**J4**	**J5**	**J6**	**J7**				
E. Plushenko	1	1	1	1	1	1	1	1	7	7	1
M. Weiss	7	3	3	2	3	3	3	3	6	17	2
A. Yagudin	5	8	2	4	5	2	2	4	4	10	3
A. Liu	4	2	5	3	2	8	8	4	4	11	4
E. Stojko	10	7	4	5	6	4	5	5	4	18	5
C. Li	8	6	6	11	4	5	6	6	5	27	6
I. Klimkin	3	9	10	7	7	6	4	7	5	27	7
E. Sandhu	6	5	11	8	8	10	10	8	4	27	8
T. Honda	9	4	9	9	10	7	7	9	6	45	9
J. Weir	2	10	7	10	11	9	9	9	4	27	10
Y. Li	11	11	8	6	9	11	1	11	7	67	11

skater needs to capture four judges to receive the majority vote. To assign the first ordinal, the algorithm searches for a skater with four or more ranks of one. In this case, Evengi Plushenko satisfies this condition, therefore receiving the ordinal value of one. The algorithm then searches for the skater to receive the second ordinal using the same methodology. Here, no one has the majority vote, so no one is assigned the second ordinal and the algorithm searches for a skater to whom to award the third ordinal. In this case, Michael Weiss now satisfies this condition because he has at least four judges rendering him a rank of three or lower. For the fourth ordinal, there were two skaters with the majority vote: Alexei Yagudin and Anthony Liu. The process of distributing the ordinals is continued until all n skaters receive an ordinal. These ordinals are then converted to placements needed to calculate the TS, as explained next.

It could easily happen, as it did in the example above, that no skaters are assigned to a certain ordinal ('2' in the example above) or that more than one skater is assigned the same ordinal ('4' and '9' in the above example). In case of ties, the algorithm first examines the number of judges that determined the skater's majority vote (MJ column). The skater with the larger majority wins. In the case of Takeshi Honda and Johnny Weir, both with an ordinal of 9, Honda wins because he has 6 judges at 9 or below, whereas Weir has only 4. If two skaters remained tied after considering both ordinal and number of majority judges (as is the case with Yagudin and Liu at ordinal equaling 4), a further tie-breaker is needed. In this instance, Yagudin has four judges' ranks: a 2, 4, 2, and 2 (which came from judges three, four, six, and seven respectively). Anthony Liu also has four judges' ranks: 4, 2, 3, and 2 (from judges one, two, four and five). To break this tie, the sum of the judges' ranks that determined the majority vote (SMJ column) are calculated, and the skater holding the smallest total wins. So, Yagudin beats Liu 10-11. If a tie had still existed in ordinal, MJ, and SMJ values, the algorithm would refer back to the judges' original ranks, searching for the skater with the lowest sum over all n judges. This very common scoring method, known as the "rank-sum", is used only as a last resort tie-breaker. The reason that it is generally avoided is that it is much more susceptible to manipulation by biased judges than are the other methods. This entire system, in which the marks are converted to ranks, which are converted to ordinals, and ultimately to placements is repeated for the LP round. The SP and LP placements are finally used in (2.1) to calculate the total score for each skater, thus determining the overall results of the competition.

After the Ordinal scoring system had been in use for many years, ISU's president, Ottavio Cinquanta, during a meeting after the 1997 European Championship, asked the rules committee to adopt a new scoring system so that relative placing of two or more competitors who have already skated would not change when another skater's ranks were awarded (termed by Cinquanta as a "flip-flop"). Thus, in 1997, a new scoring system called the One-By-One system (OBO) was created. The objective was to compare the skaters pairwise rather than simultaneously as in the Ordinal system (Loosemore). However, in comparing two skaters, this system also considers a type of "majority vote" found in the Ordinal scoring system. An $n \times n$ OBO table is constructed where the cells of this table contain a two-variable response: a dichotomous decision variable, and an integer-valued "Judges-in-Favor" variable (JIF). The diagonals of this table would be blank since the skaters being compared must be distinct. Beginning with the ranks given from each judge (the R-table), each pair of skaters are compared by examining their judges' ranks to see which skater is most preferred in a head-to-head comparison. The most preferred skater would be declared the winner, and this entry of the table would have the value '1', followed by the number of judges who favored that skater over the other compared skater. At the same time, the losing skater's first value would be a '0' followed by his/her count of favored judges. This process is continued to complete all $(n-1)^2$ cells. Afterwards, two columns showing the total number of wins and the number of judges-in-favor is made from the OBO table. The distribution of the placements is dependent upon the skater's number of wins. In case of ties, the "Judges-in Favor" column is used. Note that, as far as the judges are concerned, there is no difference between this method and the Ordinal method. The algorithm uses the same judges' rankings, but combines them, in a perhaps slightly different way, to obtain the placement scores from which the total score is derived.

Let's reconsider the data in Table 2.3 pertaining to the eleven skaters competing in the Goodwill Games. The cells contain individual ranks given by the judges. Table 2.4 is the (11 $\times$ 11) OBO table which would be created from Table 2.3, with the skaters' initials in the margins. From Table 2.4, if the One-by-One method were used rather than Ordinal scoring system, Evengi Plushenko would receive the first placement due to having the most number of wins. As can be noted, Michael Weiss and Alexis Yagudin both have same number of total wins. However, Michael Weiss has more "Judges-In-Favor" votes than Yagudin, therefore making him the recipient of the second placement. Therefore, the short program placement order from 1 to 11 would be Plushenko, Weiss, Yagudin, Liu, Stojko, Cheng Jiang Li, Klimkin, Honda, Weir, Sandhu, and Yunfei Li. Note that this order is the same as that of the Ordinal method for the first 7 skaters, but that ranks 8-10 are rearranged. Although a tie was found between Cheng Jian Li and Ilia Klimkin with their "Judges-In-Favor" votes, that is irrelevant if their number of wins differ. Although rarely happening, as with Ordinal scoring, should two skaters tie in both total wins and "Judges-In-Favor" votes, the tie is broken by calculating the sum of all judges' ranks. This same system is used again to award the placements for the LP round. After both the SP and LP placements are awarded, the total score is calculated to determine the winner of the competition. For eleven skaters, calculating the results manually is moderately easy, but in championship competitions, computer software is used in order to save time and prevent errors in calculations.

After many complaints concerning the flaws in the Ordinal system, the president and board members of ISU believed that a rapid conversion from the Ordinal to the OBO scoring system was necessary. When the OBO scoring system premiered in 1997, television and skating audiences did not understand it, nor did most understand the Ordinal system to begin with. The officials of the Nebelhorn Trophy were instructed to use the OBO system in scoring their competition to determine whether a significant difference in the results of the two systems would be found. The OBO system indeed was found to prevent "flip-flops" between rounds, but there were sometimes considerable differences in the skaters' placements between the two methods, which was something the ISU had not expected.

Table 2.4: One-By-One Judges' Ranks Table for 2001 Goodwill Games

	EP	**MW**	**AY**	**AL**	**ESt**	**CL**	**IK**	**ESa**	**TH**	**JW**	**YL**	**Total Wins**	**JIF**
EP	X	1_7	1_7	1_7	1_7	1_7	1_7	1_7	1_7	1_7	1_7	**10**	**70**
MW	0_0	X	0_3	1_4	1_7	1_7	1_6	1_6	1_7	1_6	1_7	**8**	**53**
AY	0_0	1_4	X	0_3	1_6	1_5	1_6	1_6	1_6	1_6	1_7	**8**	**49**
AL	0_0	0_3	0_4	X	1_4	1_5	1_4	1_7	1_5	1_6	1_7	**7**	**45**
ESt	0_0	0_0	0_1	0_3	X	1_4	1_5	1_5	1_5	1_6	1_7	**6**	**36**
CL	0_0	0_0	0_2	0_2	0_3	X	1_4	1_4	1_5	1_5	1_6	**5**	**31**
IK	0_0	0_1	0_1	0_3	0_2	0_3	X	1_6	1_5	1_5	1_5	**4**	**31**
ESa	0_0	0_1	0_1	0_0	0_2	0_3	0_2	X	0_3	0_3	1_5	**1**	**20**
TH	0_0	0_0	0_1	0_2	0_2	0_2	0_2	1_4	X	1_5	1_4	**3**	**23**
JW	0_0	0_1	0_1	0_1	0_1	0_2	0_2	1_4	0_2	X	1_5	**2**	**19**
YL	0_0	0_0	0_0	0_0	0_0	0_1	0_2	0_2	0_3	0_2	X	**0**	**10**

In concluding this discussion, it should be noted that President Cinquanta, in 2002, proposed that another scoring system should be adopted. Under this new method, as yet unnamed, a sample of judges' ranks is randomly chosen to produce a skater's rank and placing. The rationale for the proposed new scoring system is that it will be harder for unscrupulous judges to manipulate. However, most statisticians who have examined this issue don't feel that the new method will be an improvement, since the variability introduced by using a different set of judges for each competitor would be greater than typical victory margins.

As noted above, the actual scoring method (Ordinal or One-By-One) used to determine SP and LP placements is of only slight relevance to us. In our models, we are assuming that placements (without ties) for the SP and LP will be produced (and from these, via equation (2.1)), a total score will be produced. We will model the relationship between SP and LP scores, assuming that the process to produce these is stable. The dataset used primarily spans the five year period from January 1998 to December 2002 (a few tournaments outside this range), and includes data from both the Ordinal and One-By-One eras.

Chapter 3

Methodology

3.1 Data Collection

Data collection took place from Summer 2002 to Spring 2003. Most of the data were taken directly from websites dealing with figure skating. These websites included Icecalc.com (an online database containing figure skating results), the United States Figure Skating Association, the US World Championships, and International Olympics sites. The goal was to find results that contained placements (in both SP and LP) of all the skaters who competed in each tournament examined. The dataset was limited to national and international professional skating tournaments, with men's, women's, and pair divisions containing at least six skaters. The tournaments used in this research ranged from the years 1992 to 2003 (see Table A.1 of the Appendix). Those with fewer than seven skaters (or six teams for 'pairs' competitions) were discarded from the dataset. In all, this yielded 107 competitions, each having a distinctive tournament name, year, and division. There were 38 men's, 40 women's, and 29 pairs competitions, (see Table A.2 of the Appendix).

3.2 Data Cleaning and Summarization

The information collected at this stage consisted of about 1289 lines corresponding to 107 competitions, each with an average of 12 competitors. For each competitor in a competition, the SP and LP 'placements', along with final rank, were displayed. The data were checked for consistency by means of a SAS program, and inconsistencies

Table 3.1: Contingency Table (SP Placement vs. LP Placement)

LP **SP**	1st	2nd	3rd	4th	5th	6th	7th +	Total
1st	67	29	4	2	2	1	2	107
2nd	19	37	29	11	3	4	4	107
3rd	12	20	35	22	7	3	8	107
4th	7	12	21	29	15	7	16	107
5th	1	5	10	14	36	18	23	107
6th	0	3	5	13	15	34	37	107
7th +	1	1	3	16	29	40	552	642
Total	107	107	107	107	107	107	642	1284

were matched against original records for resolution until the dataset appeared to be 'clean'. From this cleaned dataset, I was able to create a 7×7 contingency table showing the joint (SP, LP) distribution for all places (1-7), where '7+' represents "rank 7 or higher". As a final verification that the dataset was free of errors, I checked that the marginal frequencies summed to the total number of tournaments in the dataset, 107. The results of these computations are shown in Table 3.1.

If a tournament had less than 12 entrants, fictitious entries for skaters who finished in jth place in both LP and SP were added to the table (i.e. the '552' shown in the (7+, 7+) cell of Table 3.1 is composed of about 410 actual observations and 142 fictitious observations corresponding to skaters who would have expected to fare poorly on both SP and LP if they had entered). This addition has minimal effect on the parameters fit by the model, but simplifies computations so that all tournaments may be considered to be of the same size ($n = 12$ skaters) in the final analyses.

Another SAS program was written to find the concordant/discordant matrix. The $A(i, j)$ entry of this 6×6 matrix (shown as Table 3.2) displays the number of

Table 3.2: Concordance-Discordance Matrix

j / i	1st	2nd	3rd	4th	5th	6th
1st	0	82	91	93	101	104
2nd	25	0	68	81	92	95
3rd	16	39	0	69	87	94
4th	14	26	38	0	74	80
5th	6	15	20	33	0	72
6th	3	12	13	27	35	0

times that the skater in ith place on the short program finished ahead (on the long program) of the skater who was in jth place in the short program. In comparing a pair of skaters from a tournament, it is typical to observe that the skater with the lower SP placement will also have the lower LP placement. This is called a concordant result. However, it is possible for the opposite to occur: a pair of skaters are said to be discordant if their relative orderings on SP and LP are reversed. The matrix in Table 3.2 is constrained-symmetric in that $A(i, j) + A(j, i) = N$, where N is the total number of tournaments (which equals 107). For example, from Table 3.2, $A(2, 3) = 68$ and $A(3, 2) = 39$, so the 2nd and 3rd ranked skaters on the short program finished in a concordant order (original #2 ahead of original #3) 68 times and in a discordant manner 39 times among the 107 competitions analyzed. From examining this matrix, as one might expect, concordant orderings are much more common than discordant (i.e. there is a positive association between SP and LP placements), and the discordant chances are higher for two skaters with adjacent ranks than for two who are widely separated in SP rank. For example, as can be seen from Table 3.2, a (3,2) discordance is much more likely to happen [39 occurrences] than a (6,2) discordance [12 occurrences].

3.3 Research Objectives

The summarized data shown in Tables 3.1 and 3.2 are the results against which one must check any proposed models. The relevant statistical question would be: "If 107 tournaments of n=12 skaters were generated from a proposed model, how likely is it for the results generated to be more extreme [from what is expected] than those observed in Tables 3.1 and 3.2?" Of course, for this question to make sense, we must define what is meant by both "expected" and "extreme", but assuming we can do so, we can search through a hierarchy of models for the simplest model which adequately explains the observed data. The information contained in Tables 3.1 and 3.2 is non-hierarchical (i.e. neither table can be derived from the other), so fits to both should be used to assess a model's adequacy. However, although the tables are non-hierarchical, a little thought will reveal that Table 3.1 contains much more information about Table 3.2 than vice-verse, since Table 3.2 merely records concordance/discordance, whereas Table 3.1 yields a type of joint (SP,LP) distribution. Indeed, as we shall see in the next sections, it is relatively easy to postulate a class of models which yield results consistent with Table 3.2, but to find the subset of this class of models which is also simultaneously consistent with the data in Table 3.1 is much more difficult.

3.4 Creating The Statistical Model

In formulating a model for the ranking of the skaters, one should remember the following:

(a) We don't know the skaters' true relative orderings.

(b) We don't observe actual performance scores, but, rather, relative rankings.

(c) The long program lasts approximately twice as long as the short program and, hence, should be a more precise measure of a skater's ability.

Considering this, one would like to obtain some parameters $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and a distribution $\boldsymbol{f}$, such that $f(\lambda_i)$, the probability distribution of the performance scores of the ith best skater, is stochastically less than $f(\lambda_j)$ for $i < j$, and yields results that are consistent with the data collected. (We are assuming here that a low score is better, congruent with the ranking system.) The major problem here is to pick an appropriate function, $\boldsymbol{f}$, and to obtain the parameters $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. The most standard simple model might be:

$$f_{SP}(\lambda_i) \sim N(\lambda_i, 1) \tag{3.1}$$

and

$$f_{LP}(\lambda_i) \sim N(\lambda_i, 1/2), \tag{3.2}$$

where $\lambda_1 = 0$, $\lambda_1 < \lambda_2 < \ldots < \lambda_n$, and where $N(a, b)$ indicates a normally distributed random variable with mean, a, and variance b. This models states that the true abilities of the ith best skater are normally distributed with increasing mean (but variance assumed to equal one) as rank increases, and that the long program score is half as variable as the short program scores. The trick would be in finding $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ such that the data are well fitted. This problem is complicated by the fact that the actual performance scores are not released, and that the true abilities aren't known (the third place skater may not really be the third best, etc.). Setting $\lambda_1 = 0$ and setting the variance of the SP normal distribution equal to 1 have no effect on the problem; this is done for scaling convenience.

Deciding on the correct type of distribution, $\boldsymbol{f}$, is a challenge. One must consider the shape of the distribution, the number of parameters to estimate, and, more importantly, the distribution's tail. Originally, an exponential distribution for $f(\lambda_i)$ was hypothesized. During the simulation process, it became apparent that the exponential model's fit was very poor due to the lack of control of the distribution's tail. Since the exponential is a one-parameter family, once the mean is set, there is no further control on the tail's behavior. Much attention and control was needed to the tail of the distribution in order to account for the small, but not too rare, chance that a skater would lose or gain many places between SP and LP due to a possible fall in either of the presentations.

A normal distribution's model, after simulations, was concluded to yield a more sufficient fit than the exponential distribution. This isn't too surprising, since the normal is a two-parameter family and would thus yield more flexibility in modeling both mean and variation. With the decision to use the normal distribution came the extra challenge of deciding on the variability to associate with the short and long programs. Since the long program is twice as long as the short program, it seems reasonable to set the variance of the LP's raw scores to be one-half as large as that of the short program, as noted in equation 3.2. The variance for the short program can arbitrarily be set to one (as the expected performance score of the 'best' contestant can arbitrarily be set equal to zero), since location-scale families are invariant to shifts in location and scale. Unfortunately, as simulations showed, no matter how the λ's were chosen, the fit as given by the normal models of equations (3.1) and (3.2) was not adequate. This occurred primarily because there is a small but non-negligible probability that a skater could fall during either the SP or LP and thus perform much worse than s/he typically performs. This led to the consideration of the normal mixture model:

$$f_{SP}(i) = \begin{cases} N(\lambda_i, 1) & \text{with prob.} \quad (1 - p/2) \\ N(\lambda_i + \delta, 1) & \text{with prob.} \quad p/2 \end{cases}, \tag{3.3}$$

and

$$f_{LP}(i) = \begin{cases} N(\lambda_i, 1/2) & \text{with prob.} \quad (1-p) \\ N(\lambda_i + \delta, 1/2) & \text{with prob.} \quad p \end{cases}, \tag{3.4}$$

where $\{\lambda_1, \lambda_2, \ldots, \lambda_{12}\}$, are ability parameters, $\delta > 0$ is a 'screw-up' parameter, and p is the probability of a major screw-up. This model is similar to the model given in (3.1) and (3.2), but acknowledges that there is a small probability, p, that skater i could do much worse than usual $(\lambda_i + \delta)$ on either round because of a major screw-up, such as fall or poor skating technique required in the performance. If this happens, the skater typically obtains a much worse (i.e. 'higher' in the scale used) ranking than would be expected without a fall or mistake in technique. There is no symmetric "un-screw-up" event - a skater may have a very good day, but that would be expressed by the variability inherent in the original normal distribution, which doesn't require the skew necessitated by a fall. The 'screw-up' probability, p, is one-half as large for the short program as for the long program, since the SP is one-half as long as the LP and, thus, should offer approximately one-half the opportunity for a major screw-up. Of course, the paramaterization above could be made even more general by allowing either p or δ to depend on i, but this would increase the numbers of parameters tremendously. As shown later, such generality doesn't appear to be necessary to obtain an adequate fit; constant values for p and δ appear to work well.

FORTRAN programming was used to estimate the parameters of the statistical model. In addition to requiring values from the concordant-discordant matrix (Table 3.2) and the SP-LP contingency table (Table 3.1), the program also requires initial estimates of the 13 free parameters $(\lambda_2, \lambda_3, \ldots, \lambda_{12}, p, \delta)$. Values $\lambda_1 = 0$ and $\sigma = 1$ don't count as free parameters, since they are arbitrary location and scale parameters.

Recall from the previous section that there are two tables of data (Table 3.1 and Table 3.2) which we may use to assess the fit of the model given by Equations

(3.3) and (3.4). Table 3.2 is a 6×6 matrix with diagonal values equal to zero and constrained symmetry ($A(i,j)+A(j,i) = 107$), so it has 15 independent cells. Table 3.1 is a 7×7 matrix with row sums constrained to 107, so it has 42 independent cells. If one use a χ^2 test to assess the fit of a table with T independent cells by a model with K independent parameters, the degrees of freedom (df) for the resulting test statistic is

$$df = T - K, \tag{3.5}$$

which would be df $= 15 - 13 = 2$ for Table 3.2 and df $= 42 - 13 = 29$ for Table 3.1, assuming all $K = 13$ model parameters are estimated, as noted in the previous paragraph. For the concordant-discordant table (Table 3.2), it is relatively easy to find sets of parameters that will fit the data quite well, even with only 2 df available. For Table 3.1, on the other hand, even with 29 df, it is difficult to find a set of parameters which will yield a chi-squared statistic which is near what would be considered feasible ($\chi^2_{29,.05} = 42.56$, for example). The remainder of this section describes the process by which the sample space of 13 parameters $(\lambda_2, \lambda_3, \ldots, \lambda_{12}, p, \delta)$ is searched to find the approximate MLE of the parameters.

This problem is considerably more difficult than most MLE/goodness-of-fit calculations for two reasons, one common and one not. The 'common' reason, very relevant here, is that it is difficult to find the true MLE $(\hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4, \ldots, \hat{\lambda}_{12}, \hat{p}, \hat{\delta})$ in a high dimensional space. There are many nearly equivalent parameterizations, and local maxima can obscure the truth. The more unusual aspect, not nearly as common in contingency table goodness-of-fit tests, is that this likelihood can't be immediately evaluated, even for a known set of parameters. In this case, we evaluated a model by running 1000 times as many simulations as in the original experiment ($1000 \times 107 = 107,000$), and using the average test statistic over these simulations as an estimate of the true chi-squared value for that configuration of parameters.

Based on the fit, the parameters were tweaked in an MCMC (Monte Carlo Markov Chain) way such that a better expected fit could be found.

The object of the FORTRAN program is to conduct 100 Monte Carlo (MC) trials each containing 107,000 simulations. Before running the first MC trial, the FORTRAN program read in the total number of simulations, a set of random seeds, the screw-up probability, the concordant-discordant and SP-LP matrices (Tables 3.2 and 3.1), and initial estimates of the thirteen parameters of the model. Beginning with the first trial, the program used the random seeds and the scale parameter, $\sigma = 1$, to randomly position the means of the performance parameters. The twelve given parameters served as a beginning location or 'center', for where the best estimates could be located. Thus, a tolerance bound was made around these estimates, and the algorithm randomly generated $\{\underline{\lambda}, p, \delta\}$ centered at the input values. Once the program has chosen a set of $\{\underline{\lambda}, p, \delta\}$ parameters, it generates a random short and long program score for each of the twelve skaters, according to the models given in equations (3.3) and (3.4). On average, it would be expected by this algorithm that the SP scores for skaters one through twelve would end up being in increasing order (best to worst). In generating each skater's SP score, it is quite possible for the first skater to have a higher SP score than the second, third, or even higher positioned skater due simply to random variation. Therefore, once each skater has his/her SP and LP score, they're placed in increasing order by the SP score. A combinatorial comparison, starting with the first skater, is then made with the other eleven skaters to determine if a concordance or discordance happened amongst the SP and LP acores. A concordance, discussed in Chapter 3, means that a skater placed ith in the SP round again finished ahead of a skater placed jth in the SP round (with the restriction that $(i < j)$). Therefore, for any one simulation, we could represent the concordance and discordances through a 12×12 matrix, similar to Table 3.3, where each element, r_{ij}, would be either 0 or 1. Elements below the

Table 3.3: Concordance/Discordance Matrix of One Simulation

$i\backslash j$	1	2	3	4	5	6	7	8	9	10	11	12
1	0	$r_{1,2}$	$r_{1,3}$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$r_{1,11}$	$r_{1,12}$
2	$r_{2,1}$	0	$r_{2,3}$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$r_{2,12}$
3	$r_{3,1}$	$r_{3,2}$	0	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\vdots$
4	$\vdots$	$\ldots$	$\ldots$	0	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\vdots$
5	$\vdots$	$\ldots$	$\ldots$	$\ldots$	0	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\vdots$
6	$\vdots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	0	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\vdots$
7	$\vdots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	0	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\vdots$
8	$\vdots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	0	$\ldots$	$\ldots$	$\ldots$	$\vdots$
9	$\vdots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	0	$\ldots$	$\ldots$	$\vdots$
10	$\vdots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	0	$r_{10,11}$	$r_{10,12}$
11	$r_{11,1}$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$r_{11,10}$	0	$r_{11,12}$
12	$r_{12,1}$	$r_{12,2}$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$r_{12,10}$	$r_{12,11}$	0

diagonal would represent discordances while those above would be concordances. Because any particular comparison in one simulation will be one or the other, $r_{ij} + r_{ji} = 1$.

At the end of an MC trial, a 12×12 matrix has been created containing the results of 107,000 simulations that should 'averagely' represent the results of 107,000 figure skating tournaments (call this matrix **E**). The FORTRAN program then operates on this **E** matrix in two ways. It first converts the values of the **E** matrix in terms of 107 tournaments by dividing each cell in the matrix by 1000. It then truncates this 12×12 matrix to a 6×6 matrix, keeping only the first six rows and columns. This is necessary in order to compute a Chi-Squared test statistic representing the degree of difference between the expected matrix, **E**, and the observed matrix (Table 3.2).

This Chi-Squared test statistic, denoted as $\chi^2(C/D)$, is defined as

$$\chi^2(C/D) = \sum_{i=1}^{6}\sum_{j=1}^{6}\frac{(O_{ij} - E_{ij})^2}{E_{ij}}, \tag{3.6}$$

with O_{ij} representing the observed concordant/discordant matrix from Table 3.2, and E_{ij} representing our best estimate (based on 107,000 simulations) of the expected values for cell (ij) under the $\{\underline{\lambda}, p, \delta\}$ parameterization used in this MCMC trial.

Before ending each MC trial, the FORTRAN program performs another calculation, using the results from the twelve skaters to create an SP/LP "expected" matrix. Similar to the calculations used to make the C/D "expected" matrix, **E**, each simulation created a 12×12 matrix filled with binary elements (values of 0 or 1) similar to Table 3.3. An element from such a table would be in the form, p_{ijk}, which equals 1 if from the kth simulation a skater placed ith in the SP and jth in the LP (otherwise, $p_{ijk} = 0$).

After 107,000 simulations for each MC trial, an "expected" matrix, **P**, counting the number of skaters placed ith in the SP and jth in the LP would be evaluated as

$$p_{ij.} = \sum_{k=1}^{10,700} p_{ijk} \quad . \tag{3.7}$$

Similar to what was done to matrix **E** for the concordant/discordant matrix, the elements of this **P** matrix would be divided by 1000 and truncated to a 7×7 matrix so that it could be used to compute a Chi-Squared value measuring the degree of difference between Table 3.1 and this new SP/LP expected matrix, P^*, defined as

$$P^* = \begin{bmatrix} P^*_{11.} & P^*_{12.} & P^*_{13.} & P^*_{14.} & P^*_{15.} & P^*_{16.} & P^*_{17.} \\ P^*_{21.} & \cdots & \cdots & \cdots & \cdots & \cdots & P^*_{27.} \\ \vdots & & \ddots & & & & \vdots \\ \vdots & & & \ddots & & & \vdots \\ \vdots & & & & \ddots & & \vdots \\ P^*_{61.} & \cdots & \cdots & \cdots & \cdots & \cdots & P^*_{67.} \\ P^*_{71.} & P^*_{72.} & P^*_{73.} & P^*_{74.} & P^*_{75.} & P^*_{76.} & P^*_{77.} \end{bmatrix},$$

$$\text{where} \quad P^*_{ij} = \left\{ \begin{array}{cc} P_{ij.}, & (i < 7) \text{ and } (j < 7) \\ \sum_{m=7}^{12} p_{im.}, & (i < 7) \text{ and } (j = 7) \\ \sum_{m=7}^{12} p_{mj.}, & (i = 7) \text{ and } (j < 7) \\ \sum_{m=7}^{12} \sum_{n=7}^{12} p_{mn.}, & (i = j = 7) \end{array} \right\}. \tag{3.8}$$

The Chi-Squared test statistic measuring the degree of difference between Table 3.1 and this new SP/LP expected matrix, P^*, is defined as

$$\chi^2(SP/LP) = \sum_{i=1}^{7} \sum_{j=1}^{7} \frac{(O_{ij} - P_{ij})^2}{P_{ij}}, \tag{3.9}$$

where matrix O_{ij} is the matrix from Table 3.1, and the P_{ij} has an interpretation similar to that of E_{ij} on the previous page.

Upon the conclusion of an MCMC trial (for both the C/D and SP/LP matrices) and before entering the next MCMC trial, the FORTRAN program generates a new set of $\{\underline{\lambda}, \delta, p\}$ to use as the parameters for the next set of 107,000 simulations. If the χ^2 (SP/LP) statistic found in the previous simulation is not the best thus far found, the new values are generated using the original centers and tolerances. If the χ^2 (SP/LP) statistic for the previous trial is the best (lowest) thus far found, the process is recalibrated with centers at the $\{\underline{\lambda}, \delta, p\}$ used in the previous trial, but with tighter tolerances. In general, the tolerances decreased inversely proportional to the square root of the number of new 'best fits' found. Of course, as in any MCMC or Metroplis-Hasting type algorithm, one needs to adjust these tolerances ('tuning parameters') a bit before finding values that work well.

For each MCMC trial, the FORTRAN program yielded the twelve λ parameters, the screw-up probability (p), the screw-up parameter (δ), and the Chi-Squared test values from both the C/D matrix (Table 3.2) and the SP/LP matrix (Table 3.1). After 100 MC trials, the program reports the lowest Chi-Squared test value found from the 100 SP/LP trials, along with its MCMC trial number. We're more interested in the Chi-Squared test statistic from the SP/LP matrix because it doesn't fit as well as χ^2 (C/D), which is easy to fit. As can be noted from the FORTRAN output (Appendix C), the χ^2 (C/D) fits extremely well. One will also observe, however, that the bigger concern is the possibility of the Chi-Squared test statistic for the SP/LP table not fitting well, since large values were reported.

This entire process was repeated a number of times. In the beginning, it was very easy to find new 'best parameterizations', and the χ^2-value decreased often as the search progressed. However, as the χ^2-value approached the region of acceptability (that of a $\chi^2_{df=29}$ random variable), new minima became harder to find. In addition, many apparent new minima weren't due to a better parameterization, but due to randomness in the P_{ij} estimates, since they are based on 107,000 (rather than an infinite number) of simulations. Ultimately, the parameterization shown in Table 3.4 was chosen to be the 'best' parameterization. This is surely not the true MLE, but it is in a close neighborhood of the MLE. More importantly, it is acceptable by both χ^2 criterion (at $\alpha = .05$), since the critical values are $\chi^2(df = 2, \alpha = .05) = 5.99$ and $\chi^2(df = 29, \alpha = .05) = 42.56$, and the values which were obtained were $\chi^2(C/D) = 2.733$ and $\chi^2(SP/LP) = 41.393$, respectively.

With these 'best' parameter estimates, the final phase of this project was to run a large number of simulations to analyze the concordances/discordances in the short

Table 3.4: Best Parameter Estimates

Parameters:	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7
Estimates:	**0.000**	**1.050**	**1.690**	**2.370**	**3.580**	**3.900**	**4.348**
Parameters:	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	p	δ
Estimates:	**5.750**	**5.850**	**6.700**	**6.850**	**8.360**	**0.070**	**3.080**

Average χ^2 for SP/LP Table = 41.393
Average χ^2 for Concord./Discord. Table = 2.733

versus long program scores. These simulations will be useful in answering examples of the research questions mentioned in the Introduction. For example, what is the chance of a skater finishing in first place given that he/she ended the short program round in fourth place?

3.5 Model Validation

The last phase of the project was to determine whether the estimates obtained as the true parameters of $\{\underline{\lambda}, p, \delta\}$ produce a response similar to that found in the observed data. The idea was to simulate the $\{\underline{\lambda}, p, \delta\}$ process a large enough number of times to answer questions about the joint SP/LP and SP/FINAL distributions. As a check, of course, these simulated results should yield tables statistically close to what was observed (over 107 tournaments) in Tables 3.1 and 3.2. In order to keep conversions simple, the decision was made to simulate 1,070,000 times; 10,000 times larger than the observed sample. Using the $\{\underline{\lambda}, p, \delta\}$ parameter estimates from Table 3.4, SAS programming was used to simulate the results of 1,070,000 skating tournaments (twelve skaters per tournament). The SAS program is shown in Appendix D, with interim results being shown in Appendix E. These results were used to construct

Table 3.5: Conditional Probability of LP Rank Given SP Rank

LP / **SP**	1st	2nd	3rd	4th	5th	6th	7th +
1st	0.5751	0.2222	0.1074	0.0545	0.0196	0.0117	0.0096
2nd	0.2401	0.3176	0.2296	0.1233	0.0370	0.0246	0.0278
3rd	0.1093	0.2408	0.2724	0.1985	0.0730	0.0482	0.0577
4th	0.0459	0.1328	0.2100	0.2405	0.1428	0.1076	0.1204
5th	0.0162	0.0477	0.0925	0.1610	0.2288	0.2053	0.2485
6th	0.0076	0.0208	0.0459	0.1094	0.2193	0.2279	0.3691
7th +	0.0010	0.0030	0.0070	0.0193	0.0466	0.0624	0.8607

two truncated 7×7 contingency tables: an SP/LP contingency table similar to Table 3.1, and an SP/FINAL contingency table, which analyzes a skater's SP rank and standing at the conclusion of the tournament. Both tables were also rescaled by a factor of 1/10,000 in order to interpret each cell in terms of 107 figure skating tournaments. To answer the questions brought out in the introduction of the thesis, we examine the conditional probabilities given by these tables.

The first table, Table 3.5, lists the conditional probabilities, $P(LP = j|SP = i)$, between the SP and LP ranks. For example, cell (3,4) contains ".1985", which means that we estimate the probability that a third place (after SP) skater will earn 4th rank in the LP is about 19.85%. Similarly, from the table, a skater has only a 13.93% chance of finishing in the top six in the long program if he/she was seventh place or lower after the SP round. Of course, not all of the 13.93% who score in the top six on the LP after scoring in [7+] during the SP will actually score in the top six overall. That type of conditional probability information is what is displayed in Table 3.6.

Table 3.6 gives the conditional probabilities, $P[Final = j|SP = i]$. This table is useful in answering conditional probability questions of the type proposed earlier. For example, cell $(6, 3) = .0372$ says that there is a 3.72% chance of finishing a tournament in third, given that one completed the SP round in sixth place. The question of interest throughout this thesis has been the conditional probability, $P[Final = 1|SP = 4]$. From Table 3.6, this probability is estimated to be .0241. Hence, from Table 3.5, the probability of a fourth-ranked (on SP) skater performing best on the LP is about 4.6%, but only slightly more than one-half the time $(.0241/.0459 = .5251)$ will this jump in LP performance be sufficient enough for the overall score to become first. In terms of the initial discussion of this thesis, what Sarah Hughes did was rare, but not extremely rare, since a fourth- place-to-first place jump would be expected to occur in about 2.4% of all tournaments.

The final row appears to be a bit unorthodox because the highest probability is not along the table's diagonal. According to Table 3.6, it appears that a skater with a '7+' SP rank has a better chance of finishing in sixth place overall rather than in seventh (or higher). It should be remembered, however, that row '7+' contains the results for all skaters ranked 7-12 on SP. So, what these results mean is that the probability that some (unspecified) skater in the '7+' group finishes in 6th place overall is higher than this chance is for the individual skater in 6th place after SP. However, if these skaters' probabilities were displayed individually, one would, of course, see that the 6th place (SP) skater has better final chances than the 7th place (SP) skater, who, in turn, would be better than 8th, etc. A final discussion of these tables is in the conclusion, Chapter 4.

Table 3.6: Conditional Probability of Final Rank Given SP Rank

FINAL **SP**	1st	2nd	3rd	4th	5th	6th	7th +
1st	0.6120	0.2259	0.1042	0.0410	0.0117	0.0042	0.0011
2nd	0.2484	0.3432	0.2338	0.1174	0.0329	0.0169	0.0074
3rd	0.1112	0.2498	0.2886	0.2049	0.0760	0.0428	0.0267
4th	0.0241	0.1341	0.2226	0.2713	0.1614	0.1056	0.0809
5th	0.0040	0.0350	0.0949	0.1811	0.2613	0.2131	0.2107
6th	0.0004	0.0095	0.0372	0.1109	0.2357	0.2481	0.3584
7th+	0.0001	0.0025	0.0188	0.0734	0.2210	0.3693	0.3149

Chapter 4

Conclusion

In a large figure skating competition, it occasionally happens that a competitor decides to forfeit after the SP round. In some cases, this forfeiture is a result of withdrawal due to a major injury. Often, however, a competitor forfeits due to a belief that s/he has low chances of winning or placing in high standings. Conditional contingency tables, such as Table 3.6, make it easier, now, to see why.

If a skater finishes first, second, or third in the SP and wins in the LP, s/he will automatically be the overall winner as can be seen from the scoring rules of section 2.2. Skaters ranked lower than third on SP are not guaranteed a first-place finish even if they win the LP round. As discussed in the previous chapter, Sarah Hughes' accomplishment (moving from fourth to first) was a rare one, but not extremely rare. Could she have managed a victory if she had been in fifth, sixth, or seventh (or worse) place after the SP round? Table 4.1 is a contingency table analyzing the chances of winning a tournament given that a skater finished the SP round in the ith position, but ranked 1st in the LP round. As can be noted, these conditional chances drop steeply as SP rank increases beyond third. Also, it should be reiterated that these are the chances conditional on the fact that the ith rated SP skater can win the LP. In fact, the probabilities of that happening decrease at the rate of a Geometric ($p = .41$) random variable, as can be seen from the first column of Table 3.5 in the previous section. Hence, the conditional probability of winning given SP rank is as shown in the first column of Table 3.6.

Table 4.1: Conditional Probability of Wining (given $LP = 1$)

LP SP	1st
1st	1.0000
2nd	1.0000
3rd	1.0000
4th	0.5251
5th	0.2469
6th	0.0526
7+	0.0100

Most figure skaters compete to win, and would forfeit only if the chances of winning were impossible, so even the [7+] skaters have a theoretical chance. If one wanted to forfeit if one's chances were below 1% of winning, then all skaters ranked 5th or lower on the SP would be advised to quit. If one cared about earning a medal (1st, 2nd, or 3rd place), fifth place or higher on the SP have greater than 10% chance of medalling, and 6th place has about a 5% chance. The '7+' ranks are pooled together in Tables 3.5 and 3.6, but even the best of these (#7 on SP) has less than a 1% chance of attaining a medal. Thus, the decision used in the Olympics (and many other large competitions) to include exactly the top 6 SP skaters in the final LP competition is very sound.

Of the many tables given in this thesis, the most useful would probably be Table 3.6, giving conditional probabilities of final rank as a function of SP rank. The model used to derive these appears to fit the results of the 107 observed tournaments quite well, but affords much smoother estimates of the conditional probabilities in Table 3.6 than the raw tournament data could. One still might feel that using 13 parameters to fit this model is excessive, and a more parsimonious model for the 12 performance parameters ($\underline{\lambda}$) could probably be found. The only restriction used in the model of Chapter 3 was monotonicity ($\lambda_1 = 0 < \lambda_2 < \lambda_3 < \ldots < \lambda_{12}$) , but some sort of three-parameter curve would probably work almost as well, with many fewer degrees of freedom utilized. Finding such a model parameterization is left as a topic for future researchers.

Bibliography

[1] Copley-Graves, Lynn (1992). *Figure Skating History: The Evolution of Dance on Ice*, Columbus: Platoro Press.

[2] DeLeeuw, D., and Steve Leherman (1978). *Figure Skating*, Hanover: Halliday Lithograph Corporation

[3] Heller, Mark (1979). *The Illustrated Encyclopedia of Ice Skating*, New York: Paddington Press.

[4] Jonland, E., and Jim Fitzgerald (1978). *Inside Ice Skating*, Chicago: Contemporary Books Incorporated.

[5] Loosemore, Sandra (1997). *The Truth about OBO*, "Frogs on Ice" website, **http://www.frogsonice.com/skateweb/obo/score-obo.shtml**.

[6] Loosemore, Sandra (1997). If It Ain't Broke, Don't Fix It: An Analysis of the Figure Skating Scoring System, "Frogs on Ice" website, http://www.frogsonice.com/skateweb/obo/score-tech.shtml.

[7] Morris, Jim (2003). *Two New Judging Systems To Be Sent at Skate Canada in Quebec City*, "Yahoo! Canada Sports" website, **http://ca.sports.yahoo.com/021031/6/pxxh.html**.

[8] Ottavio, Cinquanta (2003). *Some Key Dates in ISU History*, "ISU" website, **http://www.isu.org/vsite/vcontent**.

[9] Sullivan, Robert (2002). *Person of the Week: Sarah Hughes*, "Time Magazine" website, **http://www.time.com/time/olympics2002.html**.

Appendix A

Miscellaneous Contingency Tables

Table A.1: Tournaments Collected for Dataset (by year & division)

Tournament Name	**Abbr:**	**Years Available:**		
		Mens	**Womens**	**Pairs**
Winter Olympics	WO	98,02	92,98,02	92,02
Vienna Cup	VC	00	00	
Goodwill Games	GG	98,01	98,01	98,01
World Figure Skating Championship	WSC	01,02	01,02	01,02
Cup of Russia	COR	99,00,01,02	99,00,01,02	99,00,01,02
Nebelhorn Trophy	NBT	98,99,00,01,02	98,99,00,01,02	98,99,00,01,02
Karl Schaffer Memorial	KSM	98,99,00,01,02	98,99,00,01,02	02
Skate America	SAM	98,99,01,02	98,99,00,01,02	98,99,00,01,02
North American Skate	NAC	98,01,02	98,01,02	02
Challenge Finlandia Trophy	FIN	98,01,02	98,01,02	02
Skate Canada	SKC	98,01,02	98,01,02	98,01,02
Four Continents Figure Skating Championships	FCF	01,02	01,02	01,02
O. Nepela Memorial	NPM	01	01	
NHK Trophy	NHK	02	02	02

Table A.2: Distribution of Tournaments (by year and gender)

Gender Year	**Men's**	**Women's**	**Pairs**	**Total**
Before 1998	0	1	1	2
1998	8	8	4	20
1999	4	4	3	11
2000	4	5	3	12
2001	11	11	7	29
2002	11	11	11	33
Total	**38**	**40**	**29**	**107**

Appendix B

FORTRAN MLE Program

```
C PROGRAM skatemcmc2.f IS A FORTRAN PROGRAM TO USE MCMC METHODS TO FIND THE
C OPTIMAL PARAMETER CONFIGURATIONS FOR THE DISTRIBUTION OF
C of skaters for the Marlow Lemons thesis problem.
C 12 skaters
C NMC = # of MCMC runs to examine
C NSIM # of simulations. (N=107,000)
C SDE=SD OF N(ZU,SDE) DISTRIBUTION FOR SHORT PROGRAM
C SDF=SDE/SQRT(2)
C ZU(I)= MEAN OF TRUE Ith GROUP, I=1,2,...,12
C P = PROB OF SERIOUS SCREW-UP ON LONG PROGRAM (P/2 for SP)
C TOLZ, TOLP, TOLC = ORIGINAL TOLERANCES FOR PARAMETER AND CHI VARIATION.
C
C  PROGRAM CREATED 1/26/03 FROM SKATESIM.f . REVISED 3/09/03.
      DIMENSION IC(12,12),NSW(12,12),X(12,2),NR(12)
      DIMENSION ZU(12),ZC(12),CHI(1000),CHJ(1000)
      DIMENSION GCH(12,12),IR(12,12),IS(12,12),GSH(12,12)
      DATA IR/144*0/
      DATA IS/144*0/
      DATA GCH/144*0.0/
      DATA GSH/144*0.0/
C
      READ(5,30) NMC,TOLZ,TOLP,TOLC,CHIMIN
30    FORMAT(1X,I4,3(1X,F5.3),1X,F7.4)
      READ(5,1) NSIM,IXX,IYY,IZZ,P
1     FORMAT(I6,3(1X,I5),1X,F5.3)
      READ(5,2) SDE,(ZU(I),I=1,12)
2     FORMAT(13(1X,F5.3))
      DO 46 I=1,7
      READ(5,3) (IR(I,J),J=1,7)
3     FORMAT(7(I3))
46    CONTINUE
      DO 48 I=1,6
```

```
      READ(5,4) (IS(I,J),J=1,6)
4     FORMAT(6(I3))
48    CONTINUE
      IW=0
      NNM=0
      NDIV=(NSIM/107)
C
      DO 1000 MC=1,NMC
      TOLZC=TOLZ/SQRT(NNM+1.0)
      TOLPC=TOLP/SQRT(NNM+1.0)
C
      DO 686 I=1,12
      CALL UNIF(IXX,IYY,IZZ,WA,WB,WC)
      ZC(I)=ZU(I)+((WA-0.5)*2*TOLZC)
      DO 685 J=1,12
      IC(I,J)=0
      NSW(I,J)=0
685   CONTINUE
686   CONTINUE
C
      SDC=SDE+((WB-0.5)*2*TOLPC)
      SDF=SDC/SQRT(2.0)
      PC=P+((WC-0.5)*2*TOLPC)
      WRITE(6,99) MC,IXX,IYY,IZZ,TOLZC,TOLPC
99    FORMAT(1X,/,1X,'MC= ',I5,3(1X,I7),2(1X,F5.3))
C
      DO 100 M=1,NSIM
C
      DO  80 I=1,12
      NR(I)=I
      CALL UNIF(IXX,IYY,IZZ,WA,WB,WC)
      CALL NORM(WA,W)
      CALL NORM(WB,V)
      IZE=0
      IZF=0
      IF(WC.LT.(PC/2)) IZE=1
      IF(WC.GT.(1-PC)) IZF=1
      X(I,1)=ZC(I)+(W*SDC)+(IZE*3.080*SDC)
      X(I,2)=ZC(I)+(V*SDF)+(IZF*3.080*SDC)
80    CONTINUE
C
      DO 90 I=1,11
      DO 91 J=I+1,12
```

```
      H1=X(I,1)
      H2=X(I,2)
      IF(X(J,1).GT.X(I,1)) GO TO 91
      X(I,2)=X(J,2)
      X(I,1)=X(J,1)
      X(J,2)=H2
      X(J,1)=H1
91    CONTINUE
90    CONTINUE
C
      DO 93 I=1,11
      DO 94 J=I+1,12
      IF (X(J,2).GT.X(I,2)) IC(I,J)=IC(I,J)+1
94    CONTINUE
93    CONTINUE
C
      DO 95 I=1,11
      DO 96 J=I+1,12
      NH=NR(I)
      H1=X(I,1)
      H2=X(I,2)
      IF(X(J,2).GT.X(I,2)) GO TO 96
      NR(I)=NR(J)
      X(I,2)=X(J,2)
      X(I,1)=X(J,1)
      NR(J)=NH
      X(J,2)=H2
      X(J,1)=H1
96    CONTINUE
95    CONTINUE
C
C
      DO 98 I=1,12
      NSW(NR(I),I)=NSW(NR(I),I)+1
98    CONTINUE
100   CONTINUE
C
      CHJ(MC)=0.0
      DO 201 I=1,12
C     WRITE(6,203) I,(IC(I,J),J=1,12)
203   FORMAT(13(1X,I4))
      IF(I.GE.6) GO TO 201
      DO 207 J=I+1,6
```

```
      E=IC(I,J)/(NDIV*1.0)
      IF(E.LT.0.01) E=0.005
      GSH(I,J)=(IS(I,J)-E)/SQRT(E)
      YY=GSH(I,J)**2
      CHJ(MC)=CHJ(MC)+YY
207   CONTINUE
201   CONTINUE
C
      DO 704 I=1,7
      DO 706 J=8,12
      NSW(I,7)=NSW(I,7)+NSW(I,J)
706   CONTINUE
704   CONTINUE
C
      DO 904 J=1,7
      DO 906 I=8,12
      NSW(7,J)=NSW(7,J)+NSW(I,J)
906   CONTINUE
904   CONTINUE
C
      DO 804 I=8,12
      DO 806 J=8,12
      NSW(7,7)=NSW(7,7)+NSW(I,J)
806   CONTINUE
804   CONTINUE
C
      CHI(MC)=0.0
      DO 204 I=1,7
      DO 206 J=1,7
      E=NSW(I,J)/(NDIV*1.0)
      IF(E.LT.0.01) E=0.005
      GCH(I,J)=(IR(I,J)-E)/SQRT(E)
      YY=GCH(I,J)**2
      CHI(MC)=CHI(MC)+YY
206   CONTINUE
C     WRITE(6,205) I,(NSW(I,J),J=1,7)
205   FORMAT(8(1X,I4))
204   CONTINUE
C
      DO 304 I=1,7
C     WRITE(6,305) I,(GCH(I,J),J=1,7)
305   FORMAT(1X,I1,7(1X,F5.2))
304   CONTINUE
```

```
C
      DO 604 I=1,6
C     WRITE(6,605) I,(GSH(I,J),J=1,6)
605   FORMAT(1X,I1,6(1X,F5.2))
604   CONTINUE
C
      WRITE(6,72) CHI(MC),CHJ(MC),SDC,PC,(ZC(I),I=1,12)
72    FORMAT(1X,'CHI= ',F13.4,2X,'CHJ= ',F13.4,2X,'SDC= ',F5.3,
     * 1X,'PC= ',F5.3,1X,/,12(1X,F5.3))
C
      IF(CHI(MC).GE.CHIMIN) GO TO 1000
      NNM=NNM+1
      IW=MC
      CHIMIN=CHI(MC)
      P=PC
      SDE=SDC
      DO 998 I=1,12
      ZU(I)=ZC(I)
998   CONTINUE
1000  CONTINUE
C
      WRITE(6,999) NNM,IW,CHIMIN
999   FORMAT(1X,/,1X,'NNM= ',I2,'IW= ',I3,2X,'CHIMIN= ',F9.4)
C
      STOP
      END
C
      SUBROUTINE UNIF(IXX,IYY,IZZ,U,U3,U4)
C THIS SUBROUTINE GENERATES U(0,1) R.V.'s.
      U=RAND(IXX)
      U3=RAND(IYY)
      U4=RAND(IZZ)
      IXX=10000*((-3.46*ALOG(U))+(-72.59*ALOG(U4))+(-33.147*ALOG(U3)))
      IYY=10000*((-5.27*ALOG(U))+(-26.81*ALOG(U4))+(-13.580*ALOG(U3)))
      IZZ=10000*((-9.26*ALOG(U))+(-12.49*ALOG(U4))+(-53.247*ALOG(U3)))
      RETURN
      END
C
      SUBROUTINE NORM(U,Z)
C THIS SUBROUTINE GENERATES APPROXIMATE NORMAL(0,1) RANDOM DEVIATES, Z, FROM A
C UNIFORM(0,1) RANDOM VARIABLE, U.
      U2=U
      IF(U.LT.0.5) U2=1.0-U
```

```
T=SQRT(-2.*ALOG(1.-U2))
AT=2.30753 + (0.27061*T)
BT=1.0+(0.99229*T)+(0.04481*(T**2))
Z=T-(AT/BT)
IF(U.LE.0.5) Z=-Z
RETURN
END
```

Appendix C

FORTRAN MLE Output

```
MC=      1  483460  187210  348173 0.000 0.000
CHI=        40.8089  CHJ=         2.8128  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=      2 1094726  504650  880006 0.000 0.000
CHI=        41.9706  CHJ=         2.6427  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=      3 1620769  701351  826673 0.000 0.000
CHI=        41.9127  CHJ=         2.7463  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=      4  739071  310665  710605 0.000 0.000
CHI=        43.1643  CHJ=         3.1906  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=      5  723131  299237  292702 0.000 0.000
CHI=        41.1888  CHJ=         2.5547  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=      6  623350  258046  186820 0.000 0.000
CHI=        39.0172  CHJ=         2.5306  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=      7 1213862  506184 1375555 0.000 0.000
CHI=        39.6573  CHJ=         2.5215  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=      8 2355251  954856 2594883 0.000 0.000
CHI=        41.5025  CHJ=         3.1831  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360
```

```
MC=     9 3234593 1238046 1590361 0.000 0.000
CHI=       40.3797  CHJ=        2.4347  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=    10 1116394  421074  212261 0.000 0.000
CHI=       39.9466  CHJ=        2.6052  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=    11 1853004  831823 1721086 0.000 0.000
CHI=       43.2338  CHJ=        2.8822  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=    12  378055  202095  440207 0.000 0.000
CHI=       42.6395  CHJ=        2.8679  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=    13 1815358  721506  963920 0.000 0.000
CHI=       40.2082  CHJ=        2.7914  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=    14  577665  232436  751523 0.000 0.000
CHI=       41.4466  CHJ=        2.6820  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=    15  127288   63015  107853 0.000 0.000
CHI=       40.1436  CHJ=        2.4624  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=    16  457249  195075  365893 0.000 0.000
CHI=       42.3792  CHJ=        2.6552  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=    17  975399  432890 1403431 0.000 0.000
CHI=       44.5058  CHJ=        3.0479  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=    18 1329825  517769  945259 0.000 0.000
CHI=       39.5925  CHJ=        2.5168  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=    19  846920  328993  371044 0.000 0.000
CHI=       37.4769  CHJ=        2.3159  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360
```

```
MC=    20 1070074  483229  898345 0.000 0.000
CHI=       41.5693  CHJ=        2.7935  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

[MC 21-98 Were Deleted For Thesis Insertion]

MC=    99 1282932  555279  819550 0.000 0.000
CHI=       41.7570  CHJ=        2.6972  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

MC=   100 2126014  819678  813384 0.000 0.000
CHI=       43.5127  CHJ=        2.6594  SDC= 1.000 PC= 0.070
0.000 1.050 1.690 2.370 3.580 3.900 4.340 5.750 5.850 6.700 6.850 8.360

NNM=  3IW=  19  CHIMIN=   37.4769
```

Appendix D

SAS Simulation Code

```
* To write a program that simulates 1,070,000 tournaments of size n=12 to
        compare with the actual values from the table.*

proc iml;
sp_f=j(7,7,0);
sp_lp=j(7,7,0);
sp_lp7t=j(12,12,0);
sp_lp7=j(7,7,0);
numtrials=10700;
do trials = 1 to numtrials;
mu={0.000,1.050,1.690,2.370,3.580,3.900,4.340,5.750,5.850,6.700,6.850,8.360};
spscr=j(12,12,0);
lpscr=j(12,12,0);
sc_sp=j(12,12,0);
sc_lp=j(12,12,0);
skater=j(12,1,0);
spscore=j(12,1,0);
lpscore=j(12,1,0);
sds=j(12,1,1);
tfp=j(12,1,0);
final=j(12,1,1);
ft=j(12,1,0);
lpt=j(12,1,0);
sdl=j(12,1,.5);
sp=j(12,1,1);
spt=j(12,1,0);
lp=j(12,1,1);
z=j(12,1,0);
do i = 1 to 12;
skater[i,1]=i;
z[i,1]=rannor(0);
```

```
spscr[i,1]=uniform(0);
lpscr[i,1]=uniform(0);
if spscr[i,1]<=.07 then sc_sp[i,1]=3;
else sc_sp[i,1]=0;
if lpscr[i,1]<=.07 then sc_lp[i,1]=3;
else sc_lp[i,1]=0;
spscore[i,1]=(z[i,1]*sds[i,1])+mu[i,1]+sc_sp[i,1];
lpscore[i,1]=(z[i,1]*sdl[i,1])+mu[i,1]+sc_lp[i,1];
end;
do j = 1 to 12;
do k = 1 to 12;
if spscore[j,1] > spscore[k,1] then sp[j,1]=sp[j,1]+1;
else if spscore[j,1] <= spscore[k,1] then sp[j,1]=sp[j,1];
if lpscore[j,1] > lpscore[k,1] then lp[j,1]=lp[j,1]+1;
else if lpscore[j,1] <= lpscore[k,1] then lp[j,1]=lp[j,1];
end;
end;
do i = 1 to 12;
tfp[i]=(sp[i,1]*.5)+(lp[i,1]*1);
tfpcheck=(sp[i,1]*.5)+((lp[i,1]*1)+(lp[i,1]*.01));
end;
do j = 1 to 12;
do k= 1 to 12;
if tfp[j,1] > tfp[k,1] then final[j,1]=final[j,1]+1;
else if tfp[j,1] < tfp[k,1] then final[j,1]=final[j,1];
end;
end;
do i = 1 to 12;
if sp[i] > 7 then spt[i]=7;
else spt[i]=sp[i];
if lp[i] > 7 then lpt[i]=7;
else lpt[i]=lp[i];
if final[i] > 7 then ft[i]=7;
else ft[i]=final[i];
end;
do i = 1 to 12;
sp_f[(spt[i,1]),(ft[i,1])]=sp_f[(spt[i,1]),(ft[i,1])]+1;
sp_lp[(spt[i,1]),(lpt[i,1])]=sp_lp[(spt[i,1]),(lpt[i,1])]+1;
end;
end;
div=numtrials/107;
sp_lp_trun = (1/div)*sp_lp;
sp_f_trun = (1/div)*sp_f;
```

```
print sp_f sp_lp sp_f_trun sp_lp_trun;
/*
obs_splp={ 67 29 4 2 2 1 2,
19 37 29 11 3 4 4,
12 20 35 22 7 3 8,
7 12 21 29 15 7 16,
1 5 10 14 36 18 23,
0 3 5 13 15 34 27,
1 1 3 16 29 40 552};
diff= (obs_splp - sp_lp)**2/sp_lp;
chi={0};
do i = 1 to 7;
do j = 1 to 7;
chi=diff[i,j]+chi;
end;
end;
print diff chi;
*/
quit;
```

Appendix E

SAS Post-Simulation Results

SP_F_TRUN

65.4790	24.1701	11.1490	4.3841	1.2525	0.4453	0.1200
26.5768	36.7179	25.0159	12.5660	3.5170	1.8132	0.7932
11.8996	26.7328	30.8789	21.9218	8.1291	4.5787	2.8591
2.5752	14.3480	23.8151	29.0333	17.2733	11.2974	8.6577
0.4258	3.7500	10.1491	19.3750	27.9571	22.8016	22.5414
0.0384	1.0116	3.9782	11.8637	25.2192	26.5452	38.3437
0.0052	0.2696	2.0138	7.8561	23.6518	39.5186	568.6849

SP_LP_TRUN

61.5325	23.7769	11.4956	5.8265	2.0931	1.2478	1.0276
25.6892	33.9800	24.5696	13.1941	3.9569	2.6329	2.9773
11.6955	25.7682	29.1493	21.2405	7.8076	5.1607	6.1782
4.9108	14.2060	22.4734	25.4357	15.2831	11.5107	13.1803
1.7298	5.1046	9.8980	17.2272	24.4790	21.9720	26.5894
0.8122	2.2216	4.9130	11.7065	23.4602	24.3896	39.4969
0.6300	1.9427	4.5011	12.3695	29.9201	40.0863	552.5503

			DIFF				CHI
0.4858	1.1473	4.8874	2.5130	0.0041	0.0492	0.9201	40.7387
1.7417	0.2684	0.7988	0.3648	0.2314	0.7098	0.3512	
0.0079	1.2912	1.1743	0.0271	0.0835	0.9046	0.5372	
0.8888	0.3425	0.0965	0.4994	0.0052	1.7676	0.6032	
0.3079	0.0021	0.0010	0.6045	5.4223	0.7180	0.4845	
0.8122	0.2727	0.0015	0.1429	3.0509	3.7868	0.1578	
0.2173	0.4574	0.5006	1.0655	0.0282	0.0001	0.0005	