THE ROLE OF SIMULATION IN SECONDARY STUDENTS' REASONING ABOUT PROBABILITY DISTRIBUTIONS

by

MARY H. BRUCE

(Under the Direction of Patricia S. Wilson)

ABSTRACT

Technological advancements have added a new dimension to the teaching and learning of mathematics. Research has praised the use of simulations as a technological tool in probability instruction. Using a social constructivist perspective, this study addressed secondary students' reasoning about probability distributions using simulations. A probabilistic thinking framework developed by Jones, Langrall, Thornton, and Mogill (1999) and the GAISE curriculum framework endorsed by the American Statistical Association (2007) were used to trace the evolution of secondary students' probabilistic reasoning in this study.

Four classes of Advanced Placement Statistics students were randomly assigned to two groups, a control group using a formulaic, textbook-oriented approach to learning about probability distributions and a simulation group using physical and technological simulations to supplement their learning. Students were subjectively designated as low- and high-level students based on course histories and performance in the class. A mixed-methods design included a pre-, post-, and retention test for quantitative sources of data and written feedback, student interviews, and observation notes for qualitative sources of data.

Results from the quantitative analysis did not show significant group differences on the posttest but did show significant group differences on the retention test. Performance level differences were significant on the posttest but not on the retention test, and the interaction of group and level was significant for the posttest scores but not for the retention scores. Low-level students in the simulation group appeared to benefit more from the simulations than the high-level students.  Qualitative results revealed that students in the simulation group reasoned differently about probability distributions than students in the control group. The simulations initiated new representational structures to aid the students in their conceptual understanding. Interviews, written feedback, and observations indicated a connected understanding of such critical concepts as randomness, variation, central tendency, distribution, and the law of large numbers.

Although the study reinforced the persistence of various probabilistic misconceptions, the results fuel an optimism that simulations can possibly lead to conceptual change in students' understanding of probabilistic concepts. The results indicate that using such exploratory instructional designs in the teaching of probability and probability distributions can lead to the achievement of an equilibrium among students in the classroom.

INDEX WORDS: Simulation, Probability distribution, Stochastic reasoning, Probabilistic reasoning, Statistical reasoning, Social constructivism, Zone of proximal development

# THE ROLE OF SIMULATION IN SECONDARY STUDENTS' REASONING ABOUT PROBABILITY DISTRIBUTIONS

by

MARY H. BRUCE

B.A., The University of North Carolina at Chapel Hill, 1983

M.A.T., Georgia State University, 1994

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

THE ROLE OF SIMULATION IN SECONDARY STUDENTS' REASONING ABOUT

PROBABILITY DISTRIBUTIONS

by

Mary H. Bruce

Major Professor: Patricia S. Wilson
Committee: James W. Wilson
Lynne Billard

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2012

# DEDICATION

To God, for giving me the strength to complete this when it seemed nearly impossible at the time.

## ACKNOWLEDGEMENTS

I am grateful to my committee members, Dr. Pat Wilson, Dr. Jim Wilson, and Dr. Lynne Billard, for their tremendous patience, support, and wisdom they offered toward the completion of this dissertation. I want to especially thank Dr. Pat Wilson for sticking with me over the many years and always being willing to offer her assistance regardless of the circumstances. For the pressure and many lost hours of sleep she endured in the last few months. Thank you to Dr. Stephen Olejnik for his statistical advice early in the study. I want to extend a warm thanks and appreciation to my fellow teachers who participated in the assessment development and to Carol Cloud and Trish Lee who gave their limited time to assist me with observations and valuable feedback. A special thanks to Margaret Sloan, Kathy Daymude , Susan Teaster, Shelley Churchill-Boggs, Trish Lee, Donna Conner, Dora Bell, and Stephanie Cookson for being there when I needed you all at different stages of the process. Finally, thank you to all the students in my Advanced Placement Statistics classes that were so open and willing to provide the data possible for this study to take place.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

# CHAPTER 1

# BACKGROUND AND RATIONALE

## Introduction

In his push for democratic ideals in education, John Dewey proclaimed the aim of education to enable students to think, to exercise "freedom of observation and of judgment" in an ever-changing world (Dewey, 1938/1997, p. 69). Advancements in technology have resulted in an information, data-driven society. For one to understand the current world in which one lives, it is necessary to process and interpret this deluge of data. In keeping with the push for social progress, the field of education must be in a non-static, ever-changing pursuit of this goal. As a mathematics educator for the past twenty-four years, I have been a buoy on the water – I have seen technology burst the dam of traditional, textbook-dominated instruction in the field of mathematics, thus opening an entire new frontier of learning. This new frontier is dynamic, animated, and allows students to delve much deeper into their cognitive potential than previous instruction. As I have witnessed a dramatic shift in the way we teach mathematics, I have also noticed a more subtle change in the curricular offerings at the secondary level. In the late 1990s, I was asked to teach an Advanced Placement (AP) Statistics course at my high school which tripled in enrollment the following year. A few years later when teaching Algebra Two for the first time in over seven years, I noticed an increased emphasis on statistics, probability, and data analysis in the course objectives. In addition to several sections of AP Statistics currently offered at my high school, the curriculum now includes a Discrete

Mathematics/Statistics course for college preparatory students, and all levels of the integrated math courses contain standards that include a data analysis component.

It appears that school systems are finally heeding the call toward the pursuit of statistical literacy. In his historical development of statistics education, Truran (2001) notes it was not until the last two decades that probability and statistics found a secure place in the primary and secondary agenda. The International Association for Statistical Education (IASE) replaced the former International Statistical Institute (ISI) Education Committee in 1991 which is when Vere-Jones claims "statistical education can be said to have come of age" (Vere-Jones, 1995, p. 3). The formation of this association was prompted by the International Conference on Teaching Statistics (ICOTS) Round Table conferences which began in the early 1980s and continue today. These gatherings of educators, researchers, government officials, and statisticians further the social and educative necessity of incorporating statistics and probability in the regular school curriculum. In 1994, in response to the growing concerns regarding world numeracy, the ISI Executive Committee formed the World Numeracy Project (WNP) with the goal "to spread quantitative skills around the world in areas and populations (especially in developing countries and among the young) that could benefit from increased knowledge of numbers and their applications, with particular regard to statistics." (International Statistical Literacy Project [ISLP], 2012, para. 1). IASE eventually reared this project thus increasing the statistical flavor of the initiative, and eventually in late 2001, the project was renamed the International Statistical Literacy Project (ISLP).

## Curriculum Standards

Concurrent with the increased interest in statistical literacy from the IASE was the realignment of the National Council of Teachers of Mathematics (NCTM) Principles and Standards (2000) with an increased emphasis on data analysis and probability. An entire content standard titled Data Analysis and Probability appeared at all grade levels from pre-kindergarten through grade twelve. Objectives in this content standard for grades 9-12 call for students to be able to:

- Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them
- Select and use appropriate statistical methods to analyze data
- Develop and evaluate inferences and predictions that are based on data
- Understand and apply basic concepts of probability (NCTM, 2000)

Sub-objectives for understanding concepts of probability call for students to be able to:

- Understand the concepts of sample space and probability distribution and construct sample spaces and distributions in simple cases
- Use simulations to construct empirical probability distributions
- Compute and interpret the expected value of random variables in simple cases
- Understand the concepts of conditional probability and independent events
- Understand how to compute the probability of a compound event (NCTM, 2000)

A decade later we see individual states encouraging the development of probabilistic and statistical reasoning with the formation of the Common Core State Standards Initiative (CCSSI) (CCSSI, 2011). In an effort to provide consistent and rigorous expectations across the state and to prepare students for both college and career goals, educators and standards experts developed the Common Core Georgia Performance Standards (CCGPS) which were adopted as part of the CCSSI in Georgia in July, 2010 (Georgia Department of Education, 2011). Statistics and probability is one of the six conceptual categories of the CCGPS that encourage modeling and the exploration of empirical situations to improve reasoning skills and to make better, informed decisions (Georgia Department of Education, 2011). With the emphasis on modeling, the CCSSI cites the value of technology with visualizations of consequences with changing parameters and the ability to compare predictions with empirical data (CCSSI, 2011). The CCGPS encourages the use of simulations in model testing and statistical inference and the development of probability distributions using both theoretical and empirical probabilities (Georgia Department of Education, 2011).

Both the CCSSI and the NCTM Standards reflect the importance of developing a deep, conceptual, data-driven understanding of mathematics as opposed to a formalist, procedural, mechanistic approach and promote the role of teacher as guide in presenting students with “the opportunity to learn important content through their explorations of the problems and to learn and practice a wide range of heuristic strategies.” (NCTM, 2000, p. 341). The added challenge facing those teaching probabilistic concepts is the abundance of erroneous judgmental heuristics that students, and even teachers, often use to solve problems involving uncertainty (Piaget & Inhelder, 1975; Kahneman, Slovic, & Tversky, 1982; Konold, 1989). Research shows that formal instruction in probability and statistics does not necessarily promote a deeper, conceptual understanding of the concepts of randomness and chance (Wilensky, 1995; Batanero & Sanchez, 2005). NCTM encourages the use of simulations to build probability distributions from empirical data and to help students overcome common probabilistic misconceptions (NCTM, 2000). Though the focus of this publication with regard to secondary probabilistic instruction appears to be on computer simulations, the text does emphasize the necessity of understanding the underlying reality that the simulation models purport to represent. Exposure to actual experiments involving dice and coins is encouraged. With such directive publications as NCTM's *Principles and Standards* and CCSSI that provide the framework for many mathematics curricula adopted by school districts, there is an obligation to ensure these prescriptive expectations elicit the desired outcomes. My study seeks to illuminate the value of using simulations in the teaching of probability and probability distributions.

## Simulations

Published literature reviews indicate a growing interest in the use of simulations in the teaching of probabilistic and statistical reasoning, though many of these articles express merely anecdotal praises of the technology, affective benefits for the students, or small case studies (Becker, 1996; Mills, 2002). As Shaughnessy and Bergman (1993) pointed out in their directions of future research in probability and statistics, teacher partnerships in research can provide invaluable insight into the effects of instruction on student understanding of probability constructs, and simulations can offer a significant tool for exploratory situations involving data and chance . This study seeks to connect teaching and research and enrich the empirical literature by looking further into the use of simulations as a tool for reasoning about probability distributions, overcoming probabilistic misconceptions associated with distributions, and developing appropriate conceptions in a classroom setting.

Statistics education is peculiar in that the content permeates the boundaries of many disciplines, and advancements in technology have widened this span. Many academic fields have conducted research relative to probabilistic reasoning as uncertainty is an ubiquitous state. Though many adhere to the distinction among mathematics and statistics, and subsequently mathematical reasoning and statistical reasoning, the reality is that more statistics courses are taught by instructors from mathematics departments (Moore & Cobb, 2000). Most certainly, this rings true in the secondary setting as typically no statistics departments even exist. Yet, as Truran (2001, p. 10) states in his cursory distinction, “the ‘Mathematics’ of stochastics seems to be less well defined than for many other mathematical topics.” Indeed, it may be this ambiguity and ambivalence of meaning that causes frustration for so many when learning stochastic concepts as this starkly contrasts the deterministic nature of the pure mathematical content with which students are accustomed. Shaughnessy and Bergman (1993) clarify the use of the word "stochastics" as referring to both probability and statistics. Truran notes the insecurities many mathematics teachers experience with regard to teaching statistics and probability and proclaims the impending pedagogical challenge with his statement “Stochastics may have found a place within mathematics education, but this place is not yet secure” (2001, p. 10). This study seeks to inform secondary mathematics educators as they are being asked to implement probability and statistics into their Common Core curricula.

## Probabilistic Misconceptions

As a veteran teacher, I have always sought to understand how some mathematics topics evade students and create confusion. I seek to link concepts for children at their level of understanding and then attempt to find novel ways to present the mathematical topics to help the students have less difficulty. Repetitively, without fail, the topic that has most bewildered, alienated, and frustrated my Advanced Placement Statistics students each year has been the topic of probability. In their limited exposure to probability in such classes as Algebra Two, I have heard some of my brightest students say they were baffled by probability, and it contributed to their only "non-A" grades on mathematics tests. When I query the source of the difficulty, these students have been unable to put their finger on anything in particular, though they refer to "not understanding what the question is asking, not being able to gauge whether the answer seems correct or not, and not understanding what I am really finding." Most likely these students were taught a relatively quick unit on probability and were instructed primarily with the use of axiomatic formulas for probability. I believe that this formalistic approach to teaching probability leads to conceptual difficulties which promote feelings of alienation and frustration.

In an effort to help my students gain a better understanding of probability, I wanted to help my students place the confusing topic of probability in the bigger context of a probability distribution. I believe that some of the struggles with probability relate to errors from proportional reasoning most likely due to the prevalent use of proportions in rate calculations in prior mathematics courses. The students seem unable to distinguish between deterministic rate calculations and random outcome probabilities, both which make use of proportional reasoning but in different ways. I believe the evaluation of probabilistic tasks could be improved with more emphasis given to constructing and examining the sample space of outcomes prior to calculating the probability and thus an encouragement of distributional thinking (Prodromou, 2007b). Errors in both proportional reasoning and sample space have been noted in the literature (Jones & Thornton, 2005; Horvath & Lehrer, 1998).

A challenge lies in the design of instructional activities that can eradicate these probabilistic misconceptions. Time constraints prevent activities that require large spans of time, and the manner in which methods of instruction are enacted could hinder as well as help a student's understanding. Obviously, the abstract theoretical approach to teaching probability cannot eradicate these errors regarding proportional reasoning and sample space. The student has nothing concrete, visual, or real to support their conjectures making conceptual change and development unlikely. I believe the use of simulations in instruction can aid in improving these intuitive errors, though a critical consideration is *how* simulations can be used to improve them. A few years ago, I observed my student teacher using simulations to teach a unit on probability to my Algebra Two classes. The student teacher was not familiar with the use of simulations for teaching probability, and throughout the instruction of the unit, the concept of simulations and how they could be used to model random phenomena completely eluded the students. It reinforced the idea that simulations could possibly confound the already nebulous grasp of probability even more if not used in a constructive and meaningful way to promote understanding (Klein, 2005). I, too, had often sensed the reluctant use of simulations from students in earlier AP Statistics classes that I had taught as I assumed they were confused by the lack of understanding of the tool itself. Comments from students when reviewing simulations for the AP Statistics exam echo a familiar "Yeah, I don't really get those things."

By monitoring student reasoning during simulations oriented to the development of the notion of probability distribution, this study addresses the challenge of *how* teachers can juggle customary pedagogical constraints yet still make constructive use of simulations specifically in the development of probabilistic reasoning.

## Research Questions

This study seeks to examine the use of simulations as an instructional tool for secondary school instruction on probability. Recent studies have confirmed that secondary students are capable of developing conceptions of the simulation process (Zimmermann, 2002). The literature also confirms that despite this understanding of the construction of simulation models and belief in the model itself, students continue to be plagued by common misconceptions of probability (Zimmermann, 2002). Research incorporating the use of computer simulation tools in the understanding of sampling distributions with college students indicates a persistent, underlying, foundational struggle with the concept of distribution (Chance, delMas, & Garfield, 2004). I believe the struggle to understand the inferential applications of statistics is rooted in the lack of understanding of the concept of a probability distribution. My belief in the concept of distribution as the primary, consolidating structure between exploratory data analysis and inferential statistics echoes similar beliefs of other researchers (Cobb, 1999; Bakker & Gravemeijer, 2004; Scheaffer, Watkins, & Landwehr, 1998). The evolutionary development of the construct of distribution provides the scaffolding necessary to extend understanding to the abstract notions of sampling distributions. This study sought to narrow the analysis of the use of simulations in instruction specifically to monitor the use relative to the development of the construct of probability distribution. The following research questions propelled the study:

1. Does the use of simulations as an instructional tool aid in improving secondary students' understanding of probability distributions?
2. How does the use of simulations as an instructional tool help and/or hinder secondary students' understanding of probability distributions?

For purposes of this study, probability distributions will be restricted to discrete random variables and are defined as a list of all possible values of the random variable and their associated probabilities (Yates, Moore, & McCabe, 1999). Yates and colleagues (1999, p. 287) define a simulation as "the imitation of chance behavior, based on a model that accurately reflects the experiment under consideration." For the purposes of this study, simulations will include physical manipulatives such as dice as well as random number tables and technological simulations involving the TI-83/84 calculator and statistical software Minitab©. The majority of the simulations will use the Minitab© software.

# CHAPTER 2

# THEORETICAL PERSPECTIVES AND REVIEW OF THE LITERATURE

This chapter highlights the theoretical perspectives that frame this research. Both learning and cognitive philosophies are discussed followed by past and current research relevant to this study.

## Theoretical Perspectives

This study is framed from a social constructivist perspective. I believe that students learn through active construction of their own knowledge. Through facilitation and probing, a teacher can help a student progress cognitively from their current mental state to a progressively higher level. In accordance with Vygotsky's zone of proximal development (ZPD), a teacher can design a social setting of problem solving that fosters the development of a "budding" concept and guides a student towards the maturation of that cognitive development (Vygotsky, 1978). Cultural tools such as computers and calculators coupled with the social dynamics of the natural classroom setting can be used in a teacher's instructional design in such a way to encourage new conceptual models that help to promote understanding.

A theory that guided the design of my study is the theory of conceptual change. Although this theory of learning has its origin in science education, any field plagued with misconceptions could benefit from this theory which highlights testing prior conjectures and beliefs. The two primary components of conceptual change involve discovering students' misconceptions about a specific topic and then using a myriad of techniques to help students alter their erroneous conceptualizations (Davis, 2001). Posner, Strike, Hewson, and Gertzog (1982) initially proposed three conditions necessary to accommodate conceptual change. First, there must be a sense of recognized inadequacy and dissatisfaction with one's current conceptions. Subsequently, the new conception must be sensible and seem to fit cohesively into one's current schema. Finally, the new conception must be advantageous and show signs of productive and powerful ways of processing information (Posner, Strike, Hewson & Gertzog, 1982). This initial theory of conceptual change was challenged due to the absence of the emotional or affective and was revised in 1992 to include the role of social, institutional, and motivational factors as a fourth condition in the process of conceptual change. The theory was broadened to suggest the notion that alternative conceptions could be induced as a result of instructional decisions (Tyson, Venville, Harrison, & Treagust, 1997). Pintrich, Marx, and Boyle (1993) continued to push the motivational factor of learners as a necessary consideration for conceptual change and addressed the role that peer interaction, student-teacher interaction, learning communities, and classroom contexts may play in the actualization of that change (Pintrich, Marx, & Boyle, 1993). Teachers who promoted student engagement and inquiry-based tasks where students sensed control of their learning were more likely to see conceptual change in their students. Previous research supports that having students use simulations to confront and reason with contradictory predictions involving sampling distributions helps to foster possible cognitive change and improve the impact of the technological tools on student reasoning (Chance, delMas, & Garfield, 2004).

## Probabilistic Conceptions and Misconceptions

With the absence of data analysis and probability in Western academia until recent decades, much of the earlier research regarding probabilistic reasoning was conducted strictly with a developmental focus in the field of cognitive psychology. In their quest of explaining children's intuitive notions of chance, Piaget and Inhelder (1975) performed various random experiments with different age groups of children. With the primary goal to support his developmental theory, Piaget conducted various random tasks with children in three age groups: four to seven, seven to eleven, and eleven to twelve years. These groupings corresponded respectively with his proposed stages of cognitive development: the preoperational, concrete operational, and formal operational stages. Children in the preoperational stage predicted occurrences of random tasks with a sense of regularity and reversibility. They expected arrangements of colored balls to return to their previous state after being mixed. The awareness of possible outcomes did not follow an enumerative process, and intuitive judgments of prediction were highly subjective with reliance on such personal preferences as favorite color (Piaget & Inhelder, 1975). Children in the concrete operational stage recognized the pattern of random mixture though they were unable to formally operate within the mathematical structure of permutations. It was during this second stage of development that Piaget posited the recognition of uncertainty occurred. Subsequently, it was not until the age of eleven or twelve that Piaget noted the formal understanding of random behavior (Piaget & Inhelder, 1975). Piaget compared both random mixtures which could be observed physically as well as random tosses and drawings which he characterized as more abstract and requiring higher cognitive abilities than the physical observations.

Conclusions from both types of activities led to Piaget's claim that it is not until the formal operational stage that children exhibit a probabilistic intuition founded on progressive quantification.

Fischbein (1975) challenged Piaget's claims from a pedagogical standpoint. With a developmental rationale, Piaget had purposefully not considered the influence of instruction on children's understanding of chance. On the contrary, Fischbein assumed an interest in the notion of intuition and the interactive role of the instructional setting. He differentiated what he termed primary and secondary intuitions as intrinsic cognitions based on one's individual life experience in contrast to those based on task-specific instruction. Through Vygotsky's influence with the zone of proximal development, Fischbein believed that appropriate pedagogical choices could alter one's primary intuitions, though his extensive research that spanned several decades confirmed that this accomplishment would not be a simple endeavor. Similar to Piaget's cognitive developmental stages, Fischbein developed a framework for the stages of probabilistic intuitions but one that was founded on instructional intervention. For children under 7 years of age, the instructional effect was at a minimum though Fischbein did believe that children at this age exhibited some sense of the indeterminate nature of events. From 7 to 12 years of age, children did seem to develop a mental schemata relative to prediction and to respond to instructional strategies for comparing odds, though this was also the stage where probabilistic misconceptions were believed to materialize. Finally, in the formal operational stage, the quantification of probability became more developed, receptivity to instruction improved, and children seemed responsive to either reinforcement or refutation of their predictions (Fischbein, 1975). Fischbein's contributions fuel the optimism to continue to develop and refine instructional methods to help students overcome common misconceptions with probability. The responsiveness to the refutation of the children's conjectures suggests the possibility of conceptual change. Through his work with children and probabilistic misconceptions, Fischbein and Schnarch hailed "probability does not consist of mere technical information and procedures leading to solutions. Rather, it requires a way of thinking that is genuinely different from that required by most school mathematics" (1997, p. 104). The author supports Fischbein's notion of the overemphasis on formalistic instruction in the school system as an impediment to probabilistic reasoning.

Around the same time that Piaget, and subsequently Fischbein, were developing frameworks regarding the evolution of probabilistic reasoning, the emerging paradigm of cognitive psychology began to permeate the realm of decision-making. Much of the founding research on probabilistic misconceptions is credited to the cognitive psychologists, Kahneman, Slovic, and Tversky (1982). These psychologists noted specific judgmental heuristics that people tend to use when making decisions regarding uncertain events. Through studies of primarily undergraduate and graduate students as well as panels of experts, they identified several erroneous, judgmental heuristics in the context of uncertainty. The *representativeness* heuristic (Kahneman, Slovic, & Tversky, 1982) occurs when people assume random behavior in the short run should resemble the long-term behavior of the population. For example, students may think that the sequence of boys and girls in BGGB is more likely to represent children in a family than BBBB. In a similar vein, people tend to neglect the *effect of sample size* on the possible outcomes. Given, for example, a population of 50% green and 50% red marbles, many may believe the chances of getting 70% green would be the same regardless if you took a sample of 10 or a sample of 1000 marbles. The representative heuristic also signals lack of understanding or acceptance of the concept of independence often termed "*gambler's fallacy*", that what occurs on one trial will not affect the occurrence on the next trial. If someone flips a coin and gets four heads in a row, they are likely to believe that the chance of getting a tail on the next toss is higher due to the four previous tosses (Kahneman et al., 1982).

The propensity to base probabilities on one's unique experiences leads to the *availability* heuristic (Kahneman et al., 1982). One who experiences much sickness in their family may project the probability of various diseases to be higher than the actual figure. This error indicates the egocentric failure of one to consider the bigger population of probabilistic events.

Another debated heuristic is the neglect of base rates, often termed *base-rate fallacy*. When given prior probabilities, people tend to deem these as insignificant and only calculate probabilities on the limited information that they identify as important. Thus this heuristic produces faulty projections primarily with the calculation of conditional probabilities as people ignore the condition. Prior research on base-rate fallacy has often included examples that warrant personal stereotypical reactions. For example, subjects were informed of a 7:3 composition of engineers and lawyers. Yet when given descriptions of these individuals, the subjects assigned equal probabilities to the two professions (Kahneman et al., 1982). There is some debate regarding the validity of the base rate fallacy with incongruencies noted in laboratory settings versus real world situations. These conflicts urge studies to delve deeper into the subject's own mental processes and representation of the task at hand (Koehler, 1996).

The *conjunction fallacy* occurs when the chance of a compound event, A and B, is deemed more likely than that of one of the individual events and thus contradicts the rule of probability: $P(A \text{ and } B) = P(A \mid B) * P(B) = P(B \mid A) * P(A)$. After presenting subjects with descriptions such as " 'John is 27 years old, with an outgoing personality. At college he was an outstanding athlete but did not show much ability or interest in intellectual matters,'" Kahneman, Slovic, and Tversky (1982, p. 95) found that subjects claimed John was more likely to be a "gym teacher" than just a "teacher."

Konold (1989) furthered the misconception literature base initiated by Kahneman, Slovic, and Tversky with interviews that shed light on students' reasoning in uncertain situations. From the verbal responses, it was deduced that students held a conflicting, non-normative view of the goal of the probabilistic task. The "outcome approach" was the term given to this reasoning that rested on the prediction of just a single outcome. Students who use this reasoning fail to see the task as one among many thus ignoring the relative frequency component of probabilistic calculations. The decisions tend to produce qualitative predictions with certainty such as "Yes, this event will happen" or "No, this event will not happen" and due to the emphasis on the single trial, students would also subsequently judge their predictions with extremist evaluations of "right" or "wrong" once the event had occurred. The interpretation given to "50% chance" was admission that one was not sure what would happen, and chances between 0% and 100% were evaluated based on their proximities to the anchor values of 0%, 50%, and 100%. Another characteristic of those exhibiting the outcome approach is the causal justification of the resulting outcomes. For example, in a bone tossing experiment (Konold, 1989), students felt that inspection of the bone or consideration of the thrower rather than inspection of repeated trials would yield a more accurate probability estimate. In a similar vein, these students rejected the use of an urn model with weighted stones as a representation of the sides of the bone. This result has significant implications for the use and acceptance of simulation models to imitate real-world random behavior. In the design of my study, I was interested in students' acceptance of both physical and technological simulations as models of random behavior due to this discrepancy noted in the literature.

Green (1990) conducted a four-year longitudinal study with a large sample of 7-11 year olds on tasks involving random patterns and comparison of odds. He divided the students into groups according to gender and ability based on subjective evaluations of general reasoning ability. Over the four-year period, students showed similar improvement in random pattern tasks according to both gender and ability. There was, however, a significant difference in improvement according to ability on the more difficult comparison of odds questions with the higher ability students showing a dramatically higher percentage correct. Green speculated the role of school mathematics in this discrepancy as ratio and proportion were typically well-developed concepts in the school mathematics curriculum and random patterns were not.

The equiprobability bias has been observed in both younger and older students primarily evidenced with the previously mentioned misuse of the phrase “50-50.” Researchers have found that this penchant for equal assignment of probabilities extends even beyond events with more than two possible outcomes (Lecoutre, 1992). By changing tasks involving outcomes of a pair of dice to outcomes of geometric figures on cards, Lecoutre suggested an improvement in the equiprobability bias for situations deemed less random by the students, though this improvement did not seem to extend to new situations. I believe this tendency comes partly from the type of proportional and part-whole reasoning the students have been exposed to in previous mathematics instruction and partly from their neglect of combinatorial reasoning with regard to order and construction of sample space. A greater emphasis on the visualization of sample space relative to the assigned probabilities may allow the students to modify this tendency.

As defined by Yates, Moore, & McCabe (1999), conditional probability of event B occurring given that A occurs is $P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$ when P(A) > 0, and the concept of independence of events is presented as satisfaction of the conditional statement P(B|A) = P(B). Not surprisingly, Fischbein and Gazit (1984) discovered that middle school students showed misconceptions related to the changing sample space with conditional probability of without replacement tasks. Students showed proportional errors when comparing part to part without considering the changing total number of outcomes.

Based on their research of middle school students, Tarr and Jones (1997) developed a framework for students' thinking in conditional probability and independence that corresponded with other comparable frameworks showing cognitive development. In summary, level one is characterized by subjective evaluation of probabilities with a preponderance in the belief that one can control outcomes. Level Two shows an attempt to employ numerical reasoning in the calculation of probabilities, though representativeness and equiprobability misconceptions dominate their calculations. Level Three shows recognition of changing probabilities of outcomes depending on preceding events though still shows a tendency to revert to representativeness, whereas Level Four distinguishes with and without replacement probabilities with minimal evidence of misconceptions. This framework was the validation study in conjunction with the development of a later probabilistic thinking framework developed from a socioconstructivist case study of third-grade students to address the instructional impact on probabilistic reasoning (Jones, Langrall, Thornton, & Mogill, 1999). With the same levels one through four, this study specifically addressed the constructs of sample space, probability of an event, probability comparisons, and conditional probability for both one-outcome and two-outcome events. Though this framework was used in the evaluation of an instructional program for third-graders, the four-level design adapted from Biggs and Collis' levels of cognitive thinking (Biggs & Collis, 1991) and the socioconstructivist orientation make it appropriate to use as a framework for probabilistic reasoning in my study as well. Instructional tasks were used in a case study to monitor students' maturation across the four constructs. Probabilistic reasoning was tracked relative to the four levels of development - subjective, transitional between subjective and quantitative, informal quantitative and numerical. Though the researchers were able to conclude that instruction helped to promote significant movement from one level to the next, results showed that progression was not necessarily uniform as some students reverted back to previous levels on later assessments. Students showed persistent struggles with listing outcomes of the sample space and incorporating part-part and part-whole reasoning to accurately calculate probabilities.

## Technology and Simulation

Kaput and Roschelle's view of technology in the field of mathematics education was one of liberation. As more and more students were expected to learn higher and higher levels of mathematics, they posited that computational media in the evolution of mathematical representations could result in a "democratisation of access to mathematical reasoning" (Kaput and Roschelle, 1997, p. 1). Through their work with the SimCalc Project and *MathWorlds* software, Kaput and Roschelle saw the tremendous value and potential of using technology to link multiple representations and add motion to a child's understanding of otherwise static conceptualizations. Though most of their work was with algebra students in learning the concepts of rates of change in preparation for advanced concepts in calculus, their promotion of the mathematics of change and variation was catalytic in the encouragement of simulations in probability and statistics instruction (Kaput and Roschelle, 1998). I agree with Kaput and Roschelle regarding the necessity to incorporate fluidity in our pedagogical decisions regarding probability distributions to avoid the shallow understandings of variation resulting from deterministic, formalistic methods of instruction. Though simulations have shown mixed results as an instructional tool in probability, as technological tools continue to advance, researchers continue to push for its use in developing students' empirical understanding of probability and to allow for connections to be made between the various camps of probability theory. To date, the amount of formal research conducted on simulations in a realistic setting continues to be sparse (Batanero & Sanchez, 2005). Existing research is dominated by smaller case studies and qualitative analyses. The most successful examples of students' conceptual change regarding probabilistic reasoning have been instructional settings that encourage students to make predictions and then use empirical results to compare, confront, and reflect on their predictions (Tarr & Lannin, 2005).

Whole-class studies on simulations with a focus on probability distributions are sparse especially in the secondary setting. With the increased attention to probability and data analysis in the school curriculum, more studies are now being done to further Fischbein's call and address the instructional impact on probabilistic reasoning. In her study with a comparable sample of AP Statistics students, Zimmermann (2002) conducted a whole-class teaching experiment with twenty-three participants to examine students' reasoning during probability simulations. Over a 12-day experiment, students were taught how to use probability simulations with their calculator to solve contextual problems involving probability. Students showed through instruction they were capable of understanding the underlying assumptions of a probability simulation mainly with regard to evaluating the validity of a probability generator, constructing a valid probability generator to simulate various contextual, probabilistic situations and use the results of the simulation to calculate empirical probabilities. Zimmermann noted that invalid reasoning typically resulted from student errors in proportional reasoning and determining appropriate sample spaces. Initial difficulties with two-dimensional trials tended to improve with instruction, and most of the students accepted the "randomness" of the output from the simulation devices. The primary tool used in Zimmermann's study was the TI-83 calculator.

delMas, Garfield and Chance (1999) have been conducting evolutionary studies on the effect of computer simulations with college students over a seven-year period looking for continuous improvement in the instructional tool. These researchers specifically looked at undergraduate students' statistical reasoning about sampling distributions using a computer program called *Sampling Distributions* programmed by delMas. The program was a student-controlled program that allowed for manipulations of shapes and changes in sample size. While results indicated improvement in students' statistical reasoning, most of the improvement was relative to visual assessment items indicating a persistent struggle to develop a deeper, conceptual understanding of sampling distributions and their connectedness with the concepts of sample, population, distribution, sampling, and sampling variability. From my own experience with sampling distribution applets that allow for student manipulation, I believe some of the difficulties resulting from the *Sampling Distributions* software stem from the student's disconnect with the technological generation of the distributions. A possible scaffolding with students' own generation of sample observations and sample means could possibly help foster a deeper connection between the resulting graphical representations and the underlying conceptual structures.

## Probability Distributions

Pratt and Noss (2002) questioned the scarcity of coverage of randomness and distributions in the middle and secondary curriculums contesting the implication that only high achievers could achieve this level of conceptual understanding. Like Fischbein, Pratt and Noss believed that instruction dominated by deterministic models hindered children's probabilistic intuitions. Their studies with 10 and 11-year old children interacting with computational devices used to simulate random devices called the *Chance-Maker* microworld showed that children began to see that manipulation of the workings box and repetition of trials led to fairness in the simulated random generators.

Pratt reminded us of the need to question the transference of knowledge from the virtual microworld setting to conventional settings and to pedagogically strive to merge the two closer together.

Prodromou and Pratt (2006) conducted design experiments in an attempt to design microworlds that encourage an assimilation between two perspectives of distribution, the data-centric perspective and the modeling perspective. The data-centric perspective refers to the data-driven empirical distribution resulting from the collection of actual data, whereas the modeling perspective refers to the theoretical distribution of all possible outcomes and their probabilities. Through iterations of simulations involving a Basketball microworld with small paired groupings of 14-15 year olds, Prodromou and Pratt sought to study the students' conceptions of the duality of causal and stochastic factors in the development of the construct of distribution. Both Prodromou and Pratt believe the key to statistical inference lies in the understanding of the merging of the two perspectives of distribution. Through their iterations of the Basketball microworld, Prodromou and Pratt concluded that the students were able to detect causal agents of variation as they manipulated the speed and release angle buttons. As the researchers began to introduce random error arrows, the students then began to discern another uncontrollable source of variation. In conclusion, the researchers deemed the students were able to perceive the difference between the data-centric model as the actual outcome versus the modeling distribution as the intended outcome with a growing sense of how the Law of Large Numbers helped to merge these two perspectives. On the other hand, their discourse did not illustrate an understanding of the underlying probabilistic source of both types of distributions. With an emerging emphasis on exploratory data analysis (EDA) in statistics instruction, it is significant to remember to address the theoretical distribution as a source of the data-driven distribution. Again, however, with a strong reliance on animated computer simulations, one has to question the transference of the situated learning to a different setting and context.

## Curriculum Frameworks

The mission of the American Statistical Association (ASA) includes, among other aspects of promoting statistical science, the improvement of statistics education for all students (American Statistical Association, 2007). As part of this strategic mission, the association provided grants for the creation of the GAISE report (Guidelines for Assessment and Instruction in Statistics Education) in August 2005 which detailed a curriculum framework for statistics instruction and assessment at both the Pre K-12 level as well as the undergraduate introductory level. The development of the framework is justified as the need for statistical literacy abounds in decision making inherent in both our personal and professional lives (ASA, 2007). The formal report was published in 2007 and underlines the importance of the understanding of the concept of variation as the salient and striking distinguishing aspect between statistics and mathematics and thus includes progression in the understanding of the nature of variability as the driving factor in the development of the four components of statistical problem solving (Metz, 2010). The framework presents these four components descriptively in terms of the application of variability at each stage:

- Formulate Questions - Anticipating Variability - Making the Statistics Question Distinction
- Collect Data - Acknowledging Variability - Designing for Differences

- Analyze Data - Accounting of Variability - Using Distributions
- Interpret Results - Allowing for Variability - Looking beyond the Data (ASA, 2007, pp. 11-12)

Although the Analyze Data component is the only one that specifies the use of distributions in the framework, all components essentially require the understanding and application of distributions as variation and distributions go hand in hand. The GAISE report recognizes that statistical literacy is a developmental process and thus specifies three levels of development within each of the four components. Levels A, B, and C assume a sequential developmental process, although these levels are not necessarily age-specific or grade-level specific. From an instructional standpoint, the framework posits that Level A is a more teacher-centered goal, whereas B and C are more student-centered and aligned with a constructivist approach. For the Analyze Data component, the development in using distributions across levels is specified as:

- Level A - Use particular properties of distributions in the context of a specific example
- Level B - Learn to use particular properties of distributions as tools of analysis
- Level C - Understand and use distributions in analysis as a global concept (ASA, 2007, p. 14)

Though this framework was published after the design of my study, the results can be analyzed within this framework, and implications from my study can be made with regard to future instructional decisions regarding probability distributions and the GAISE framework.

This research is well-situated at a time when exploratory data analysis is emerging as an integral part of teaching statistics. Modeling activities are encouraged as a way to make learning relevant and realistic to today's society. Researchers agree that a conceptual understanding of distributions is a foundational component of understanding variation and making the connection to statistical inference. Computer simulations have shown promise in helping students formulate the relationship between a distribution formed by the collection of data and the conceptual abstraction of a probability distribution. With animated computer simulations, however, one must question the dependency on the machine itself and whether profound understanding of the underlying structures occurs. My study seeks to embellish the developing literature with regard to the role that random number simulations play in the development of students' reasoning of probability distributions in a whole-class setting. My study will specifically look at the connections made between sample space, probabilistic misconceptions, data distributions and probability distributions.

# CHAPTER 3

# RESEARCH METHODOLOGY

This chapter outlines the action research methodology incorporated in this teacher-researcher study along with a description of the methods, procedures, instrumentation and data sources used in the study. The chapter is introduced with a description of a pilot study conducted two years prior to the actual study which was used to inform and develop the larger body of research.

## Pilot Study

The embryonic stages of this research took place with a brief pilot study I conducted with my Advanced Placement statistics students two years prior to the actual study discussed in this dissertation. At that time my rudimentary focus was on the role of simulations in students' understanding of randomness. The primary purpose of the pilot study was to provide a framework for developing the simulation tasks and classroom procedures that would eventually become the protocol for my larger study. The pilot study consisted of a two-day activity that allowed students to test their conjectures on various questions regarding randomness and probability by using dice manipulatives and group comparisons. The activity worksheet can be found in Appendix A. Twenty-seven students participated in the study that provided myriad data sources including students' written artifacts, researcher observations, witness observations, and audiotapes. The pilot study took place in February so the students had already covered the units on probability and were currently learning concepts regarding sampling distributions and statistical inference.

On the first day students were given four multiple-choice questions to answer individually with written explanations justifying their cognitive processes. In addition, they were asked to rate their confidence level on each question from 1 (not confident at all) to 5(very confident). On the second day the students used a fair die with five faces marked black and one face marked white to simulate the following question from the previous day:

> Five faces of a fair die are painted black, and one face is painted white. The die is rolled six times. Which of the following results is more likely?
> (a) Black side up on five of the rolls; white side up on the other roll
> (b) Black side up on all six rolls
> (c) Choices a and b are equally likely

Students were asked to compare their individual simulation results with their original conjectures and to discuss any new discoveries. Belief in the simulation was tested as the students were asked whether they would change their original answer to the question and why. Subsequently, students were then asked to compare their results with other group members and pool their results with their group and eventually the entire class.

Quantitative analysis was used to calculate percentages of correct responses, and qualitative analysis was used to code responses according to appropriate reasoning, common misconceptions cited in the literature, and new emerging patterns or nuances. The following table summarizes the quantitative results:

Table 3.1
*Pilot Study Percentage of Correct Responses and Confidence Rating Results*

| | Question #1 | Question #2 | Question #3 | Question #4 |
|---|---|---|---|---|
| Percentage of correct responses | 44% | 63% | 22% | 100% |
| Confidence Rating Mean | 3.5 | 3.6 | 3.7 | 4.3 |
| Confidence Rating Standard Deviation | 0.8 | 1.1 | 1.1 | 0.8 |

The low scores on three of the four questions substantiated the persistent difficulties cited in the literature despite prior instruction in probability (delMas, Garfield & Chance, 1999). Based on the confidence ratings, the students were fairly accurate in discerning their abilities to answer the questions appropriately, but interestingly, the qualitative analysis revealed that correct responses were not always justified with appropriate probabilistic reasoning. The question that presented the greatest struggle for students accentuated the documented difficulties with compound events and deficiencies in combinatorial thinking (Jones, Langrall, Thornton, & Mogill, 1999; Zimmermann, 2002). Qualitatively, students seemed to have a secure acceptance of the concept of independence as it related to Bernoulli trials and did not fall victim to the representativeness heuristic and positive or negative recency effects as shown by the results of Question #4. Linguistic analysis indicated a secure understanding of the construct as students were able to support the terminology with descriptive definitions. In addition, many students understood the Law of Large Numbers at some level and could describe why a hospital with fewer births would yield a gender proportion significantly different from 0.50. The most significant result to come from the pilot study was the affirmation that students could, indeed, construct their own knowledge from doing the simulations with minimal intervention from my teacher directives. Despite the simplicity of the study, the results were rich in confident promises of teachable moments. The students were not afraid to admit their mathematical errors, and, in fact, welcomed the ability of the simulations to correct them in their thinking. Many students expressed initial confusion and frustration with the simulations, but the group and whole class comparisons proved beneficial in transforming the perturbations to an enlightened awareness of variation differences in small and large samples. Curiously, the student perturbations seemed to elegantly mimic the noise and high variation inherent in random behavior - initial confusion and frustration with an eventual calming as stability increased, variation lessened, and patterns emerged with the increasing number of trials. Karol was a strong mathematics student who favored using formulas over doing the simulations, but she proclaimed at the end, "It appears that the law of large numbers is very important because the beginning simulations did not match the math but the more simulations we took, the more the results resembled the math. I am a firm believer in the math and this reinforced that feeling because my individual results were so different from my group, the class, and what the math said." Several of the stronger students failed to validate the simulation results that conflicted with their formulaic answers as they steadfastly supported and believed only “the math.”

In summary, pedagogical decisions that arose from the pilot study were the omission of audiotapes as they were too muffled, contained too much empty space, and did not provide value-added data to the analysis, inclusion of tasks that included peer-to-peer and group comparisons to model short-term and long-term behavior, and a focus primarily on technological simulations over physical simulations due to a significant mistrust in the randomness of the die. Surprisingly, 67% of the students believed that people have some form of control over the outcomes that result from rolling the die. The stronger students' stubbornness to validate the simulation was curious to me and influenced my decision to address performance level as a factor in my study.

## Research Methods

This study incorporated a mixed methods analysis with both qualitative and quantitative methods used to analyze the data collected. The structure of the analysis and frame for the research questions was a comparative design experiment to address differences in probabilistic understanding for students using simulations to guide their learning with students who learned the probability concepts in a more formalistic lecture and textbook setting. Quantitative methods consisted of a multivariate MANCOVA design to analyze posttest and retention scores from assessments designed to measure statistical reasoning. Using pretest scores as the covariate, mean scores on the posttest and retention tests were compared between the treatment group using simulations in the learning of probability concepts and the control group (traditional learning) to seek evidence of the effect of the simulations on student learning. Qualitative analysis was conducted to probe deeper into the conceptual differences between the two groups of students. Interview responses and written responses from student worksheets were mined for these differences in student cognitive processing.

As discussed by Greene, Caracelli, and Graham (1989), the mixed method design for this particular study serves the purposes of triangulation, complementarity, and expansion in an effort to give an extensive look into the common construct of students' probabilistic reasoning. Description of subjects involved in the study, the sources of data, the instruments used in the analysis, and the methods of analysis are all described in this chapter.

Much of the research on misconceptions of probability has been done in artificial, non-instructional settings (Piaget & Inhelder, 1975; Kahneman, Slovic, & Tversky, 1982). Shaughnessy and Bergman (1993) encouraged the inclusion of teachers in future research regarding probability and statistics in the schools to help fuse the gap between the teaching and research communities and to allow for longitudinal record-keeping of changes in student conceptions and misconceptions in a realistic classroom setting. As both the principal investigator and teacher of the students in the sample, the research methodology for this study falls under the paradigm of action research with the primary goal to improve both teaching and learning in the specific mathematical context of probability. As described by Doerr and Tinto, the action research methodology evokes images of a "cyclic process of problem identification, action, and reflection aimed at changes in practice." (Doerr & Tinto, 2000, p. 403). With action research, the teacher assumes a more dominant role in the generation of knowledge to add to the research base, and the immersion of research in one's practice helps to fuse the perceived rift in research and practice. The effect of this type of research results in the reciprocity of benefits: the existing research is enhanced by the applicability and relevance of knowledge gained in a practical setting, and the practitioner is informed and enlightened so as to improve his or her professional expertise. The cyclical process whereby reflection helps to solve current problems regarding the teaching of a concept as well as to generate new problems is represented by Schön's nodes of inquiry for the reflective practitioner (Doerr & Tinto, 2000). This process consists of bidirectional links between knowing in action and reflection in action to produce a continuum of discovery and new learning. In essence, through this reflective inquiry that characterizes action research, the teacher begins to develop their dual role as teacher and researcher (Feldman & Minstrell, 2000). From their action research studies with sampling distributions, delMas, Garfield and Chance (1999) developed their own model of classroom research specifically adapted to statistics education. Four foundational questions are suggested to guide the instructors in their classroom research:

- What is the problem? What is not working in the class? What difficulties are students having learning a particular topic or learning from a particular type of instructional activity? The identification of the problem emerges from experience in the classroom, as the teacher observes students, reviews student work, and reflects on this information. As a clearer understanding of the problem emerges, the teacher may also refer to published research to better understand the problem, to see what has already been learned and what is suggested regarding this situation, and to understand what might be causing the difficulty.
- What technique can be used to address the learning problem? A new instructional technique may be designed and implemented in class, a modification may be made to an existing technique, or alternative materials may be used, to help eliminate the learning problem.
- What type of evidence can be gathered to show whether the implementation is effective? How will the teacher know if the new

technique or materials are successful? What type of assessment data will be gathered? How will it be used and evaluated?

- What should be done next, based on what was learned? Once a change has been made, and data have been gathered and used to evaluate the impact of the change, the situation is again appraised. Is there still a problem? Is there a need for further change? How might the technique or materials be further modified to improve student learning? How should new data be gathered and evaluated? (delMas, Garfield, & Chance, 1999, para. 1).

My study followed this sequential model of classroom research and focused on the topics that, as a teacher, I had observed to produce the greatest difficulty.

As a veteran teacher who has seen teacher autonomy, respect and empowerment dwindle over the span of decades in her career, this exercise in study, reflection and change became a liberating experience renewing my faith in my expertise as a professional, my ability to come to know my students and intimately explore their understanding, and that through my direction and research efforts, I could come to better know my students and could responsibly make improved, informed decisions about the betterment of their learning.

## Participants

The setting for my study was a large, public high school in the southeastern United States in a suburban, predominantly middle to upper-class neighborhood. Four classes for a total of fifty-five students taking Advanced Placement (AP) Statistics comprised my sample. The subjects were predominantly seniors who had completed three years of mathematics prior to taking AP Statistics including Geometry, Algebra Two, and

Precalculus or Trigonometry respectively in that order. Most of the students were on the honors or gifted track and, therefore, had taken Precalculus instead of Trigonometry. Differences in the curriculum for these two courses were minimal with rigor being the primary distinction. Five of the students were gifted sophomores who were concurrently taking gifted Algebra Two, and five of the students were juniors who were concurrently taking Precalculus. There were eight AP Statistics classes offered at the school, and the subjects of this study had been placed in four of these classes based on their scheduling needs. This study was conducted prior to Common Core curricular revisions so exposure to data analysis and probability was limited. The study was conducted from November until early January of the 2007-08 school year. The subjects had completed units on exploratory analysis, normal distributions, regression, and experimental design prior to the start of the study. For all of these students, this course was their first extensive exposure to probability and statistics. The only prior instruction in probability was either in review courses for the Scholastic Aptitude Test or a small unit taught two years earlier in an Algebra Two course. Previous exposure to probability consisted primarily of the fundamental counting principle, compound events, and beginning combinatorics in the context of simple probabilistic tasks describing such things as marbles, coins, spinners, and dice. Students had very little exposure to technological simulations, and none of the students had prior experience with Minitab© software. An extensive unit on probability fits into the course outline for AP Statistics objectives with simulations of random behavior and probability distributions as required sub-topics under the probability strand. On the end-of-course AP examination, the students are often asked to analyze a string of randomly generated digits and to calculate and compare experimental and theoretical probabilities. The dominant method of simulation in the AP Statistics curriculum is the random number generator on the Texas Instrument TI-83 graphing calculator. Students are expected to know and apply theoretical probability formulas, but simulations are recognized by readers of the exam as legitimate means for arriving at solutions to the free response and multiple choice questions on the exam.

I was the primary teacher for all four of the classes used in the study thus controlling for teacher variation. I have been a mathematics teacher in the public school system for the past twenty-four years and have taught statistics at the high school and college level for the past fourteen years. Although I have always believed in teaching mathematics for conceptual understanding, I had relied primarily on traditional teaching and lecture methods in the past. As a statistics instructor, I believe that technological advancements provide an opportunity for a greater population of students to access the learning of this subject. The technology provides a tool with the ability to quickly process large amounts of data and offer multiple representations for various statistical and probabilistic constructs. Students who struggle to understand the abstractions inherent in the formality of mathematics have a renewed hope in grasping a conceptual understanding of data analysis, probability, and statistics with the use of technology as a pedagogical tool. The AP Statistics curriculum calls for the use of the Texas Instruments TI-83 or TI-84 calculator as the technological tool of choice. The students who take the Advanced Placement test in May are allowed to use the calculators throughout the test. As a result, most of my prior instruction with my AP classes centered around using the calculator as opposed to doing computer simulations. I had some familiarity with using Minitab© from my college-level statistics courses so I chose this software to support the majority of the simulation tasks in my study as I felt the simulations would be easier to program, easier to understand, easier to process large amounts of data, and had better visual graphics than the calculator to help the students form their conceptions of a distribution.

Two other AP Statistics teachers from my high school were used as witness teachers during some of the lab sessions. They observed the students and took notes during the simulation exercises. Although these teachers did not observe all lab sessions, their intermittent observations served to give third-party, unbiased feedback as well as to validate my own qualitative conclusions from the labs.

## Procedures

At the time of the study, I taught five sections of AP Statistics. Two weeks prior to the onset of the instructional unit on probability, I randomly selected one of the classes to serve as a comparison group in the development of the assessment instrument. A group of ten mathematics teachers from my high school and the students in the comparison group were given a pretest assessment. Results from this comparison were analyzed both quantitatively and qualitatively and were then used to adjust the questions on the pretest to its final form. The remaining four classes of AP Statistics were then randomly assigned to the treatment and control groups - two of the classes were used in each of the groups. One week prior to the onset of the instructional unit on probability, these four classes were given the revised pretest assessment. Two days before the start of the study, the two simulation classes were taken to the lab to be introduced to the Minitab© software to control for confounding effects from lack of understanding of the technology. The students were taught how to generate random data using the software, and assumptions regarding randomness and independence were discussed. All of the students in the study were designated as low- or high-performing level based on previous course histories, prior grades in mathematics courses, and performance in the AP class thus far in the school year. Students were asked to volunteer to participate in one-on-one interviews throughout the instructional unit, and from this pool I selected three from each of the two comparative groups with a mixture of both low- and high-level students from each group. The decision to address performance level in my study was prompted by differences noted in my pilot study regarding the acceptance of the simulations and evidence in the literature that technological tools could aid in the teaching and learning of those students not as experienced in formal mathematics (Kaput & Schorr, 2002). Unstructured, open-ended interviews were given to the volunteer students at various times throughout the study based on my observations during each teaching lesson and based on both the students' verbal and written comments following each lesson. The instructional unit was organized according to the text used in the class titled The Practice of Statistics (Yates, Moore, & McCabe, 1999). I chose and developed the simulation tasks based on my own research, similar tasks cited in the literature, my own beliefs regarding the use of sample space in developing the concept of a distribution, and the sequence of topics presented in the text. Both the treatment and control groups learned the same topics and were given the same notes on matched days throughout the instructional unit, although practice for the control group was conducted with textbook exercises during class and practice for the treatment group was conducted with simulations in the computer lab and in the classroom. At the end of the instructional unit, all four classes were given a posttest and subsequently a retention test three weeks later following winter break.

## Instruments

The instruments used in the study consisted of a pretest given one week prior to the start of the instructional unit, a posttest given one week following the conclusion of the instructional unit, a retention test given three weeks following the conclusion of the instructional unit, and daily task worksheets given throughout the instructional unit. I developed all the instruments used in the study though they were adapted from research tools used in related studies and cited in the literature. One instrument, in particular, was used extensively in the development of the pretest, posttest, and retention test. Garfield (1998) initially developed *The Statistical Reasoning Assessment* (SRA) as an instrument to be used in the ChancePlus Project to meet the increasing need of assessing secondary students' understanding and application of appropriate statistical reasoning as the inclusion of these concepts in the secondary curriculum was on the rise. Early in my research on students' struggles with probabilistic reasoning, Garfield's research on sampling distributions sparked my interest which led to my pursuit in developing students' conceptual understanding of distributions. The SRA is an objective assessment with 20 multiple-choice items heavily embedded with probabilistic computation and reasoning. The construction of the assessment was built on various types of reasoning associated with data analysis and statistical reasoning: reasoning about data, reasoning about representations of data, reasoning about statistical measures, reasoning about uncertainty, reasoning about samples, and reasoning about association (Garfield, 1998). Due to the prevalence of erroneous probabilistic intuitions evident from the research, the SRA also included items based on these common misconceptions cited in the literature – misconceptions involving averages, the outcome approach, sample percentages of the population, law of small numbers, representativeness, and the equiprobability bias. Test retest reliability results for the SRA were 0.70 for the correct total score and 0.75 for the incorrect reasoning scores (Liu, 1998). From the call for improved assessments in statistics education, a NSF-funded project called ARTIST (Assessment Resource Tools for Improving Statistical Thinking) spearheaded by Garfield and colleagues produced a refined and broader assessment called the Comprehensive Assessment of Outcomes in Statistics (CAOS). Many of the 40 multiple-choice items on this assessment were selected from a database of questions developed from the ARTIST project. The relevant scales on this assessment were broadened to include such topics as bivariate data, sampling distributions, confidence intervals and significance testing (delMas, Garfield, Ooms, & Chance, 2007). Many of the items included in my pretest, posttest, and retention tests were pulled from this database.

My adapted assessment included questions primarily from the ARTIST database as well as some used in other studies on probabilistic reasoning (delMas, Garfield, & Chance, 1999; Konold, 1995; Fischbein & Schnarch, 1997; Konold, 1989). In addition, as my course was a preparatory course for the Advanced Placement test in May, I included some sample questions on probability from *Barron's How to Prepare for the AP Statistics Examination* (Sternstein, 2000). My original pretest consisted of 25 questions including multiple-choice and true-false items. These items were primarily selected from the data representation and probability scales from the ARTIST database. Whereas the ARTIST scale identification labeled these items as probability literacy, probability computation reasoning, probability reasoning, and data representation literacy, I specifically chose items to correspond more specifically to the topics of sample space, shape of distribution, proportional reasoning, probabilistic reasoning, representativeness, independence, compound events, conditional probability, law of large numbers, binomial distribution, and geometric distribution. To test the validity of my own assessment, the original pretest was distributed to a sample of ten mathematics instructors at my high school. Only four of these teachers had taught AP Statistics, three had tutored AP Statistics, and the remainder of the sample had very little exposure or experience in teaching probability and statistics concepts. The sample of "experts" was used to calculate a KR-20 internal consistency reliability coefficient, to support content validity with the appropriateness of the items for a secondary course on the introduction of probability and to make suggestions with regard to wording, bias or omission of topics. The contrasted groups approach was used to establish construct validity. One of five of my AP Statistics classes was randomly selected for instrument development. A two-sample t-test comparing teacher scores on the pretest with student scores showed that teachers fared significantly higher ($t = 4.71$, p value $= 0.000$, $df = 14$). The questions that proved the most difficult for both groups were those involving compound events. A KR20 coefficient for the pretest was calculated to be 0.77 showing a favorable measure of internal consistency. Based on a qualitative analysis of teacher comments and suggestions, five more questions were added, and some of the questions were reworded for clarity and gender neutrality reducing the confounding effect from question ambiguity.

## Interviews

Interviews were given to selected students from each of the two groups intermittently throughout the study. The interview questionnaires were typically unstructured, open-ended probability questions either identical to exercises covered in class discussion or items analogous to these exercises. These interviews were informal and were often catapulted from a quick, cursory glance at the responses from the daily task worksheets or from interesting developments noted during the classroom tasks that warranted clarification or that generated some special interest in further pursuit. Interviews were conducted either during a student's advisement period during the day or after school. Following a similar "interview teaching approach" to other studies on students' probabilistic reasoning (Rubel, 2007), I typically watched a student perform a given exercise and asked the student to think out loud so I could more closely monitor their thinking. From their initial response to the question, I was able to further probe into the cognitive processing of the interviewed students. Journal notes and audiotapes were taken during each of the interviews to add to the data sources.

## Worksheets and Lab Activities

Worksheets for both groups and lab activities for the simulation group were used for each of the designated tasks in the study. The worksheets that were given each day prior to the simulation activity were the same for both groups to allow for me to monitor developmental differences throughout the study.  Theories of social constructivism and conceptual change guided my design of the labs. From a social constructivist perspective, I strategically designed the labs to build on knowledge developed in previous labs so the tasks would fall within the zone of proximal development (ZPD) (Vygotsky, 1978). Questions that I asked on the lab worksheets and ones I asked aloud as I monitored classroom activity were designed to provide scaffolding  so connections and learning could take place. I designed each lab activity with a goal to promote conceptual change through the simulation activity (Posner, Strike, Hewson, & Gertzog, 1982). Typically the lab would begin with a scenario and question prompting the student to come up with an answer or make a conjecture or both. The student would then be guided on how to conduct a simulation using the technology and would subsequently be asked if they would change their answer or conjecture after conducting the simulation. Common misconceptions noted in the literature were used to choose activities that were likely to generate conflict with students' conjectures and predictions (Fischbein & Schnarch, 1997; Kahneman, Slovic, & Tversky, 1982; Konold, 1989; Lecoutre, 1992). Students were given the opportunity to comment on discoveries made through the simulation so as to monitor conceptual change and cognitive processing during the activity.

## Research Journal

I kept a journal throughout the entire study recording my observations throughout the daily tasks for both the treatment group and control group as well as during the student interviews. Significant discourse was of particular interest as well as concepts that seemed particularly easy or difficult for the students. Two other statistics teachers observed the lab sessions on several occasions and recorded field notes of their observations as well. These notes were compared so triangulation could be achieved to strengthen validity of the researcher's observations and to supplement my own reflections of significant activity.

# CHAPTER 4

# RESULTS OF THE STUDY

The following research questions guided the analysis of the data:

1. Does the use of simulations as an instructional tool aid in improving secondary students' understanding of probability distributions?
2. How does the use of simulations as an instructional tool help and/or hinder secondary students' understanding of probability distributions?

The first question was addressed through quantitative analyses of the pretest, posttest, and retention tests given one week prior to the study, one week following the study, and three weeks following the study and winter break respectively. Qualitative analyses were conducted from comments on the pretest, responses on the daily task worksheets, journal notes from daily observations, student interviews, and comments on the posttest. The regularity of the qualitative sources of data allowed me to trace the developmental changes of the students from the beginning of the study through the end and to note transitions according to the probabilistic thinking framework and GAISE framework mentioned in the literature review (Jones, Langrall, Thornton, & Mogill, 1999; American Statistical Association, 2007).

## Quantitative Analysis of Results

To answer my first research question and confirm whether simulations as an instructional tool showed a significant treatment effect, four AP Statistics classes were randomly assigned, two each, to the treatment and control groups. The treatment group was taught a unit on probability distributions over a two-week period with daily usage of either the TI-83 graphing calculator or Minitab© software in the computer lab. The control group was taught the same content over the same period of time but through textbook-dominated instruction, formulaic solutions, and notes given in class. Kaput and Roschelle posited that technology used as a pedagogical tool could "democratize access to ideas that have historically required extensive algebraic prerequisites." (Kaput, & Roschelle, 1997, p. 1). In accordance with Kaput and Roschelle's proposition and my own observations of performance-level differences from my pilot study, I used student course histories and performance in the class up to the point of the study to designate students as either low- or high-level students. A low-level designation does not mean the student is a low-performing student, in general, as all participants were placed in Advanced Placement Statistics. These low and high designations are relative only to other participants.

The assignment to groups resulted in the following student counts:

Table 4.1
*Summary of Group Counts by Level*

| Number of Students | Low Level | High Level | Total |
|---|---|---|---|
| Group 0<br>(Control Group) | 17 | 12 | 29 |
| Group 1<br>(Treatment Group - Simulations) | 12 | 14 | 26 |
| Total | 29 | 26 | 55 |

Comparisons were made using a 2x2 factorial MANCOVA design using the pretest scores as the covariate. The two dependent variables were posttest and retention scores with two independent variables, Group (Treatment vs. Control) and Level (High vs. Low). Means and standard deviations of the three assessments in the study are presented in the following table. Scores on each assessment were calculated by the number of questions correct with maximum scores possible of 30.

Table 4.2
*Summary of Assessment Means and Standard Deviations*

| Assessment | Mean | Standard Deviation | N |
|---|---|---|---|
| Pretest | 16.73 | 3.24 | 55 |
| Posttest | 20.20 | 3.43 | 55 |
| Retention | 19.47 | 4.26 | 55 |

Descriptive statistics with consideration of the independent and dependent variables are summarized in the table below:

Table 4.3
*Summary of Posttest and Retention Test Means and Standard Deviations*

| | Group | Level | Mean | Std. Deviation | N |
|---|---|---|---|---|---|
| Post | Control | Low | 16.82 | 2.88 | 17 |
| | | High | 22.75 | 2.56 | 12 |
| | | Total | 19.28 | 4.02 | 29 |
| | Treatment | Low | 20.58 | 1.83 | 12 |
| | | High | 21.79 | 2.55 | 14 |
| | | Total | 21.23 | 2.29 | 26 |
| | Total | Low | 18.38 | 3.10 | 29 |
| | | High | 22.23 | 2.55 | 26 |
| | | Total | 20.20 | 3.43 | 55 |
| Retention | Control | Low | 15.94 | 4.04 | 17 |
| | | High | 20.92 | 3.90 | 12 |
| | | Total | 18.00 | 4.64 | 29 |
| | Treatment | Low | 21.08 | 3.26 | 12 |
| | | High | 21.14 | 3.16 | 14 |
| | | Total | 21.12 | 3.14 | 26 |
| | Total | Low | 18.07 | 4.49 | 29 |
| | | High | 21.04 | 3.45 | 26 |
| | | Total | 19.47 | 4.26 | 55 |

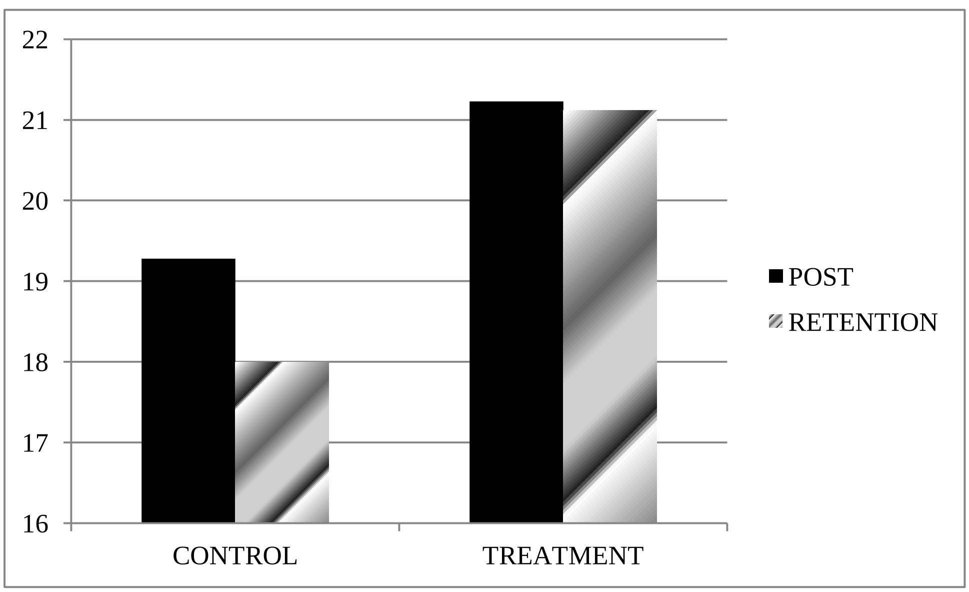



*Figure 4.1.* Group Means on Posttest and Retention Test

From Figure 4.1 above we see that the means for the simulation group were higher on both the posttest and retention test than the control group.

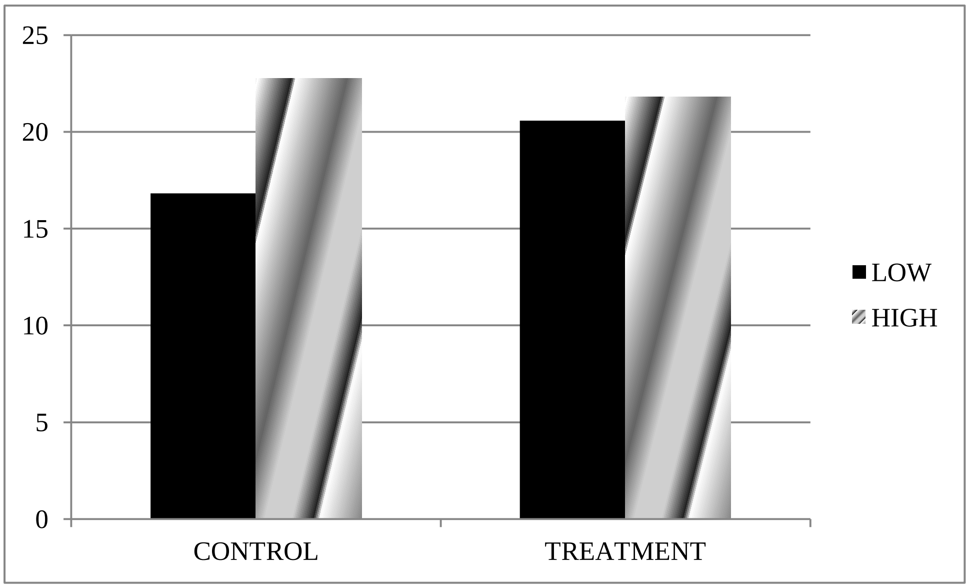



*Figure 4.2* Group Means on Posttest for Low- and High-level Students

From Figure 4.2 we see the posttest means for low-level students were higher in the simulation group than the control group, but means for the high-level students were slightly higher in the control group than the simulation group.

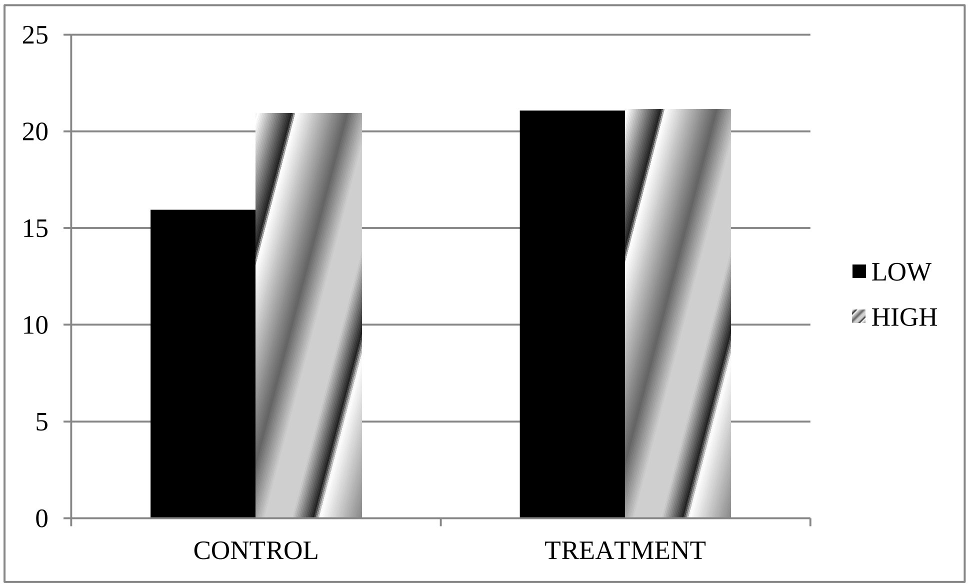



*Figure 4.3.* Group Means on Retention Test for Low- and High-level Students

From Figure 4.3 we see the retention test means for low-level students were higher in the simulation group than the control group, but these means for high-level students were nearly the same among both groups.

For the MANCOVA analysis, the Box's Test of Equality of Covariance Matrices was checked for the homogeneity of variance assumption and was not rejected ($p = .540$). In addition, Levene's Test of Equality met the assumption of equal error variance of the dependent variables across groups for both dependent variables, posttest ($p = .241$) and retention test ($p = .148$). The pretest covariate showed a strong relationship with both posttest scores ($p < .01$) and retention scores ($p < .001$) after controlling for main and interaction effects. Finally, the Kolmogorov-Smirnov test was initially used to satisfy the normality assumption of the dependent variables across all groupings of the independent variables.

Table 4.4 *Kolmogorov-Smirnov Results*

| Assessment | Group | Level | K-S Statistic | df | Significance (lower bound)* |
|---|---|---|---|---|---|
| Posttest | Control | Low | .104 | 17 | .200* |
| Retention | Control | Low | .165 | 17 | .200* |
| Posttest | Control | High | .187 | 12 | .200* |
| Retention | Control | High | .193 | 12 | .200* |
| Posttest | Treatment | Low | .208 | 12 | .159 |
| Retention | Treatment | Low | .111 | 12 | .200* |
| Posttest | Treatment | High | .193 | 14 | .169 |
| Retention | Treatment | High | .161 | 14 | .200* |

The multivariate tests of between-subjects effects did not show GROUP (treatment vs. control) as a significant main effect for the posttest scores ($F(1, 50) = 2.154$, $p = .148$) but did show GROUP as a significant main effect for the retention scores ($F(1, 50) = 4.309$, $p = .043$). The tests showed LEVEL (low vs. high) as a significant main effect for the posttest scores ($F(1, 50) = 21.456$, $p < .001$) but not for the retention scores ($F(1, 50) = 2.943$, $p = .092$). The GROUP*LEVEL interaction was significant for the posttest scores ($F(1, 50) = 6.934$, $p = .011$) but not for the retention scores ($F(1, 50) = 1.804$, $p = .185$). Main effect and interaction plots serve to illuminate these significant relationships:

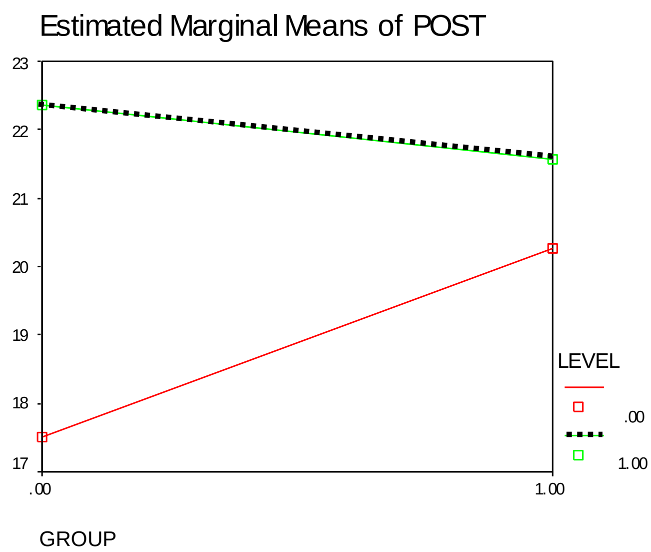



*Figure 4.4.* Marginal Means of Posttest by Group

The interaction plot of marginal means on the posttest show increased scores for the simulation group versus the control group for low-level students. A different pattern emerges for the high-level students as we see a slight decrease in the simulation group scores versus the control group scores.

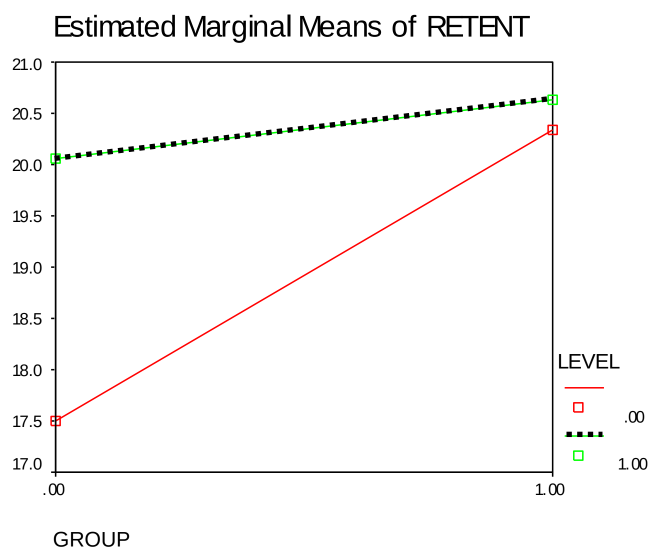



*Figure 4.5.* Marginal Means of Retention Test by Group

The interaction plot of marginal means on the retention test show increased scores for the simulation group versus the control group for both low-level and high-level students; however, as we see from the steeply-sloped line with the low-level students, the change is more prominent for the low-level students than the high-level students.

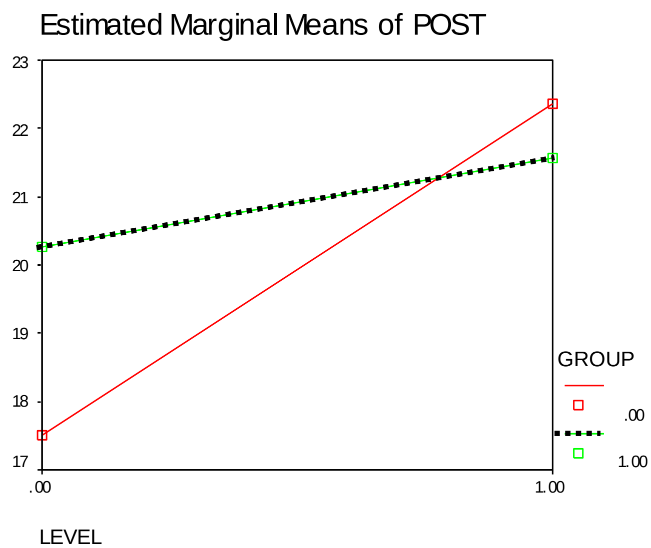



*Figure 4.6.* Marginal Means of Posttest by Level

Switching LEVEL to the horizontal axis more clearly shows the significant interaction effect on the posttest (p = .011). The control group shows a marked difference in performance between the low and high-level students on the posttest. The simulation group shows a greater balance in scores between the two different levels of students.

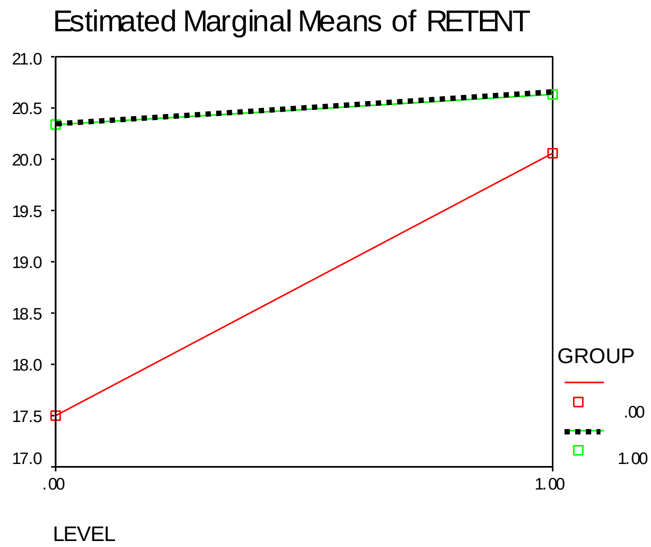



*Figure 4.7.* Marginal Means of Retention Test by Level

From the retention test interaction plot, we once again see a greater imbalance in scores within the control group compared to the nearly horizontal line for the simulation group.

In summary, the quantitative results are informative, significant, and exciting. The evidence supports Kaput and Roschelle's proposition of technology's ability to provide "democratization of access to mathematical reasoning" as we see substantially higher scores for the lower level students in the simulation group compared to their control group counterparts on both the posttest and the retention test (Kaput & Roschelle, 1997, p. 1). On both the posttest and retention tests, we see an equitable balance between the low- and high-level students' performance which is in stark contrast to the asymmetric inequity easily apparent in the control group results. In addition, the retention test discrepancy between the simulation group and the control group for both levels of student, but primarily the low-level students, hints at possibly a deeper, sustainable understanding of probability and probability distributions. For the low-level students who are often left bewildered and confused by formal probabilistic notions and formulas, the simulations appeared to offer a grounded, simpler approach to making sense of complex probabilistic ideas.

## Qualitative Analysis of Results

This next section of this chapter will take an evolutionary look at the qualitative differences in reasoning about probability and probability distributions from the beginning of the study with the pretest, through the daily labs and activities that ensued, and finally with the posttest and retention test. Recorded observations, audiotaped interviews with selected students of both levels, and student responses on daily worksheets provided the basis for these qualitative conclusions. The probabilistic thinking framework developed by Jones and colleagues provided a foundation for monitoring and organizing the growth and development in the students' thinking throughout the study (Jones et al., 1999).

### The Pretest

An item-by-item analysis of the pretest peformance for both the simulation group and the control group showed similar patterns. The pretest can be found in Appendix B.

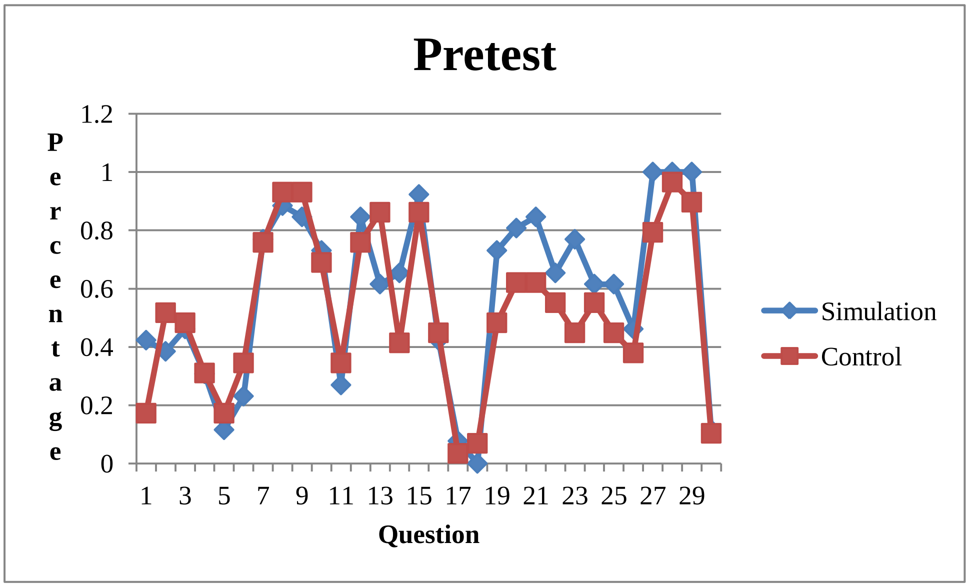



*Figure 4.8.* Pretest Proportion of Correct Responses By Question

The four most difficult questions for both groups (Questions 5, 17, 18 and 30) involved the shape of a binomial distribution given a real-life scenario, the most probable average number of defects of two cars given probabilities of the number of defects, the most probable outcome from a real-life geometric distribution, and a question illustrating the conjunction fallacy (Kahneman, Slovic, & Tversky, 1982). A summary of the correct response percentages along with the content strand of the question for each group are given in Table 4.5 below:

Table 4.5
*Summary of Pretest Correct Response Percentages by Question*

| Question | Control | Simulation | Content |
|---|---|---|---|
| 1 | 17% | 42% | compound events |
| 2 | 52% | 38% | sample space |
| 3 | 48% | 46% | shape |
| 4 | 31% | 31% | binomial distribution |
| 5 | 17% | 12% | binomial shape |
| 6 | 34% | 23% | sample space |
| 7 | 76% | 77% | shape |
| 8 | 93% | 88% | independence |
| 9 | 93% | 85% | independence |
| 10 | 69% | 73% | representativeness |
| 11 | 34% | 27% | sample space |
| 12 | 76% | 85% | shape |
| 13 | 86% | 62% | shape |
| 14 | 41% | 65% | representativeness |
| 15 | 86% | 92% | proportional reasoning |
| 16 | 45% | 42% | law of large numbers |
| 17 | 3% | 8% | compound events |
| 18 | 7% | 0% | geometric distribution |
| 19 | 48% | 73% | independence |
| 20 | 62% | 81% | independence |
| 21 | 62% | 85% | shape |
| 22 | 55% | 65% | compound events |
| 23 | 45% | 77% | randomness |
| 24 | 55% | 62% | conditional probability |
| 25 | 45% | 62% | geometric distribution |
| 26 | 38% | 46% | equiprobability |
| 27 | 79% | 100% | randomness |
| 28 | 97% | 100% | outcome approach |
| 29 | 90% | 100% | independence |
| 30 | 10% | 12% | conjunction fallacy |

A qualitative analysis using open coding and the constant comparative method revealed an abundance of common misconceptions throughout the pretest responses (Patton, 2002). The misconceptions were not readily apparent from all of the multiple-choice or true-false responses, but each question allowed a space for written feedback. These written responses from the students projected many of the misconceptions already noted in the literature and were coded accordingly. As the responses were mined, I compared the misconceptions to previously recorded ones seeking the emergence of any new misconceptions in the data.

Collectively, the students had a sound preliminary understanding of the shape of a distribution if given a picture of the graph as evidenced by the high percentages of correct responses for Questions 7, 12, 13 and 21. Shapes of distributions had previously been taught in the earlier weeks of the semester when covering exploratory data analysis. Very few students described the shape of the graph in terms of the association with probability, although one student responded with question 7 "I looked at how the frequency per week mainly decreased as the food costs per week increased creating a right-skewed distribution." The high percentage of correct responses to Question 21 along with explanations stating "the middle values are most likely to happen" confirmed the students' connection of shape and probability for the majority of the students in both groups. Students struggled to determine shape when this task involved independently calculating probabilities related to single or compound events as evidenced by the low percentages of correct responses in Questions 3 and 5. This low percentage was likely not due to the lack of understanding of shape but due to the lack of understanding how to calculate the probabilities themselves. Interestingly, several students took Question 3 and showed an overlay of Joe's right-skewed distribution with Tonya's symmetric distribution in an attempt to determine the shape of the distribution of the number of books they will buy together. Most likely, this stems from prior instruction on combination of functions which is actually an insightful strategy although ultimately they looked at the sum of the functions instead of the product and did not consider all possible combinations.

The pretest results indicated an excessive tendency to add probabilities and/or outcomes when determining compound event solutions regarding probabilities or sample space. This was resoundingly true for Questions 1, 2 and 22. One student even commented in the written feedback space "I know percentages but I don't know how to combine them." Difficulties with calculating probabilities of compound events has previously been noted in the literature (Polaki, 2005; Zimmermann, 2002). In addition, I saw a tendency to compute averages when this calculation was not appropriate. For example, with Question 1 when computing the probability that Joe and Tonya will purchase no books, some students took Joe's probability of buying 0 books and averaged it with Tonya's probability of buying 0 books adding 0.5 + 0.25 = 0.75 and then dividing this by either 2 or 6. A student that divided by 6 commented "because it showed 6 options." This tendency seems to stem from the tremendous preponderance to associate equal weights to all probabilistic calculations. I have seen a similar reliance on dividing by the total number of outcomes when teaching students to calculate expected value of a probability distribution. Cobb (1999) noted similar tendencies in his research and suggested this as a reflection of traditional, formalistic mathematics instruction. The students showed an attempt to answer the question quantitatively but with no logical arguments indicative of levels 1 and 2 in the Jones et al. probabilistic thinking framework (Jones et al., 1999).

As I anticipated, many students incorrectly relied on deterministic proportional calculations to answer some of the probability questions and even expressed their belief that some information was missing when they were not given part-whole information in the problem. On Question 4, numerous students showed the proportion 90/100 = x/4 which gave them an answer of x = 3.6 only to then completely ignore that the question asked for the LEAST PROBABLE value of x. Such reasoning is indicative of Level 2 in the probabilistic thinking framework (Jones et al., 1999). These students then put 3 or 4 for their answer based on this x value. This indication supports my theory of a reliance on deterministic calculations from prior math classes where students view the goal as coming up with the one correct answer to the question based on evaluating a function or solving an equation. Question 5 student comments also confirmed a dependency on being able to find the correct answer or specific value rather than consider all possible outcomes. The question seeks the shape of the distribution of the number of correct answers if you guess on a 20-question multiple-choice test with 5 choices per question. One student commented "I don't understand. How are you supposed to find the shape. Although you should draw a graph, on what basis. We don't have the correct answer. It's missing an entire variable." Similarly, another stated "Am I supposed to know how many are correct?" and another "You don't know how many are correct." These comments are an indication of the weakness and lack of understanding of a random variable with multiple outcomes and probabilities. Question 18 was a geometric distribution question stating that a basketball player makes 80% of her free throws and asking for the most probable number of throws it will take until she misses one. Students resorted to their familiar proportional rate calculations and set up the proportion 8/10 = 4/5 putting 4 as their answer even though the assumption of a fixed 5 trials was irrelevant to the question and to the type of distribution. Similarly, one student commented "I'm not sure how many she is throwing anyways." Question 23 was a true-false question stating that a 60% chance of rain for the next five days means that it should rain exactly 3 out of the 5 days. Students who were typically dependent on proportional rate calculations showed 60/100 = 3/5 and chose "true" for their answer.

Students also exhibited Konold's outcome approach on Question 23 with such responses as "Every day has 60% not the whole 5 days" and "Each day has a 60% chance, not the week" indicating their failure to see probability as a relative frequency application (Konold, 1989). This localized tendency and failure to see the more global aspect of outcomes in probability calculations can also be associated with Level 2 in the probabilistic thinking framework (Jones et al., 1999).

Collectively on the pretest, most of my students exhibited level 1 subjective and level 2 transitional probabilistic thinking (Jones et al., 1999) with such subjective reasons as "because I've played yatzee before" to justify the most likely roll of three dice in Question 11 and "I don't get caught" for the reasoning in Question 29 regarding getting caught for sneaking out of the house. An attempt to list outcomes in the calculation-oriented questions involving compound events and/or sample space typically showed an incomplete and unsystematic way of listing outcomes as noted in the framework for transitional thinking.

Amongst scattered evidence of the representativeness and independence misconceptions, the most protrusive misconception I noticed on the pretest was the equiprobability bias (Lecoutre, 1992). It was obvious through both calculations and comments that many students believed that probability meant "equal chance." On Question 6 regarding the most likely roll of a die with black and white faces, one student put "they are both possible" and put "equally likely" as her response. Another put "It's probability" and selected "equally likely." Question 5 regarding the shape of a binomial distribution with 20 multiple-choice questions and 5 choices per question summoned the following responses: "Randomly guessing makes there an equal chance for every question" and "An equal number of choices per question should result in no skewness" followed by "It seems you would maybe get equally as many right as wrong if guessed."

Comparison charts of percentage of questions correct within each group by level of student are given below:

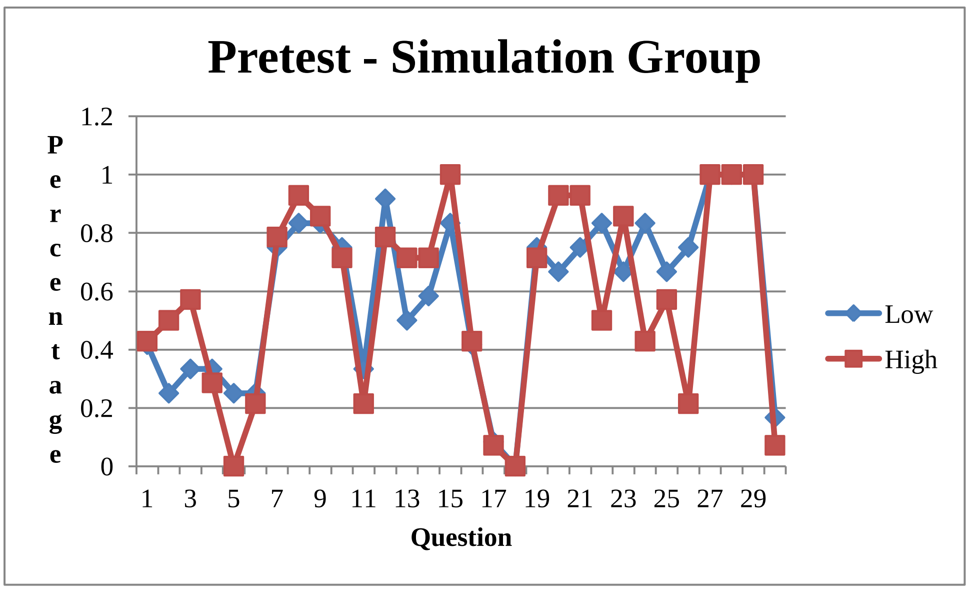



*Figure 4.9.* Pretest Proportion of Correct Responses By Question for Simulation Group

The largest percentage discrepancies between the low- and high-level students in the simulation group included questions regarding sample space, compound events, binomial distributions, shape, representativeness, independence and conditional probability.

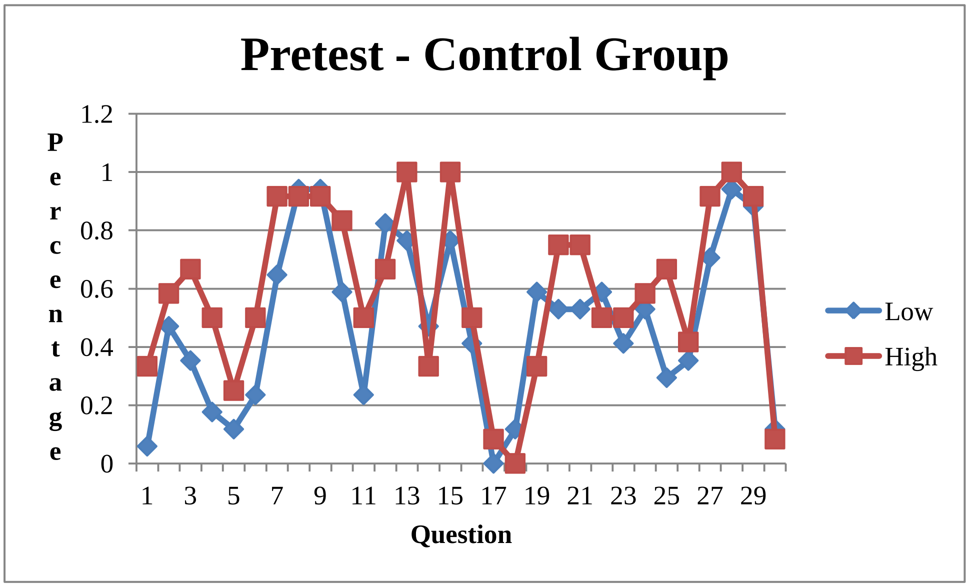



*Figure 4.10.* Pretest Proportion of Correct Responses By Question for Control Group

The largest percentage discrepancies between the low and high-level students in the control group included questions regarding compound events, binomial distributions, combinatorial reasoning, shape, independence, and equiprobability.

Day 1 - Sample Space

The goal of this first activity was to help students conceptualize the outcomes of an experiment and eventually to construct the sample space of all possible outcomes. Personally, I believe that this foundational concept in the study of probability distributions is not typically developed well or emphasized with students, yet understanding how to generate and analyze the possible outcomes helps bridge the understanding between possible values of a random variable, the graphical representation of the random variable and eventually probabilities and the theoretical distribution of the random variable. My hope was that the simulations would help the students discover outcomes they had not previously thought of, to notice that some outcomes occurred more frequently than others and to notice the combinatorial aspect of the outcomes. On Question 3 on the pretest, there were several students who listed all possible combinations of outcomes to arrive at their answer. These students got correct answers for this question compared to the majority of students who missed this question. The listing of outcomes is generally a simplistic, fundamental skill which should prove to be accessible to both low- and high-level students.

Both the control group and simulation group were given the same initial page of the activity sheet. Both were taught the fundamental counting principle and tree diagrams as methods for generating all possible outcomes and were asked to complete three exercises listing outcomes in the sample space. All classes were divided into small groups. The simulation group was then given instructions on how to generate random number sequences using their TI-83 calculator and was asked to simulate problems 1, 2, and 3 from the first page. The worksheet and lab for this activity can be found in Appendix E. Both groups were asked to compare their answers with their group members. Time constraints proved to be a deterrent for completing this activity in the simulation groups as the calculator simulations took longer than anticipated. As a result, some groups were only able to simulate problem 1 with the coins.

Both groups showed difficulty in their initial construction of the sample spaces. The problem involving the weather pattern proved more difficult for the students than the coin and spinner problems most likely due to familiarity with coins and spinners from previous math courses. The weather example will typically prove to be more difficult later as the chance of rain or no rain is not necessarily a 50-50 chance as many students will assume based on the equiprobability bias inherent in the pretest results (Lecoutre, 1992). Coin and spinner problems typically assume equal probabilities when given as examples to students.

The first problem presented the tossing of a coin 3 times. When asked for how many outcomes there were in the sample space, some students erroneously multiplied 2*3 = 6. As one student explained, "There are 2 things you can get everytime you flip a coin so 2 x 3 = 6." Subsequently, the list of outcomes from some students was then incorrectly written as

HT    HT    HT

or others incorrectly put {H, T, H, T, H, T} to support their incorrect total of 6 outcomes.

Similarly, with the weather problem, the students were told to assume on a given day, it either rains or it doesn't rain. After being asked to suppose they record the weather results over the next four days, they were asked to tell how many outcomes there were in the sample space and then to list those outcomes. Once again, many students incorrectly showed 2*4 = 8 outcomes, and their listings included such patterns as:

1R, 2R, 3R, 4R, 1N, 2N, 3N, 4N

or

| | |
|---|---|
| 1 | 1 Rain |
| | 2 No Rain |
| 2 | 1 Rain |
| | 2 No Rain |
| 3 | 1 Rain |
| | 2 No Rain |
| 4 | 1 Rain |
| | 2 No Rain |

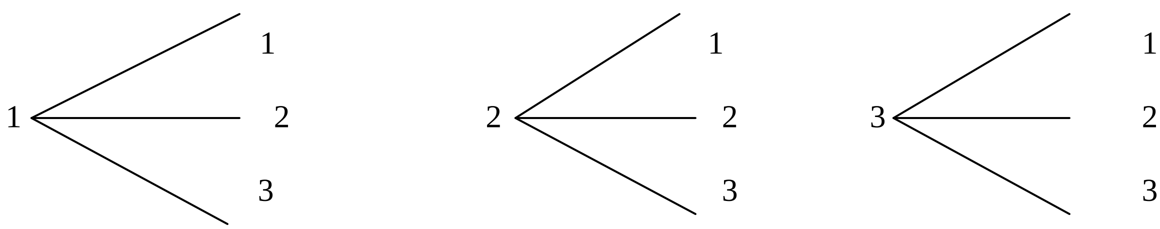
On the spinner problem, an activity from the textbook (Yates et al, 1999), students were asked to imagine spinning the spinner 3 times. When prompted for how many outcomes there were in the sample space, the students showed a similar additive pattern exhibited on the pretest and came up with 9 outcomes showing such patterns as:

1) 1, 2, 3          2) 1, 2, 3,          3) 1, 2, 3

or

S = {11, 12, 13, 21, 22, 23, 31, 32, 33}

or an incorrect tree diagram such as:




Many students displayed an incorrect use of the counting principle by multiplying two numbers together despite the fact that all three exercises involved three or four trials. I speculate that this could, in part, be due to the way the counting principle is presented in most textbooks. From Yates, Moore, & McCabe (1999, p. 320), the multiplication principle is summarized as:

If you can do one task in **a** number of ways and a second task in **b** number of ways, then both tasks can be done in **a x b** number of ways.

Though teachers would, of course, explain that the rule can be generalized, the presentation of the rule typically shows **a x b** or **m x n** promoting the product of two factors in the calculation.

Some students did not see their incongruencies between the number of outcomes and the list of outcomes in the sample space. For example, Anna said there were two outcomes for the weather question yet listed the sample space as {1R, 2R, 3R, 4R, 1N, 2N, 3N, 4N}. The study suggests this is, in part, due to students' lack of clarity on what is meant by an outcome and the inability to extend their understanding to situations beyond single and simple trials to compound events (Zimmermann, 2002). In her study with secondary students' understanding of probability simulations, Zimmermann also noted this struggle that students had in constructing models for calculating joint probabilities.

Once the simulation group continued the activity with their calculators, the students seemed comfortable with the dummy coding and use of binary numbers to represent heads and tails when simulating the coin problem which boded well for upcoming random number simulations using Minitab© in the computer lab. The students seemed to notice that the outcome 1, 2, 3 was different from 2, 1, 3 on the spinner problem circling both as separate outcomes and noting similar combinatorial aspects from the other questions as well. Adam interestingly noticed that "even though we did 20 trials all outcomes aren't necessarily going to show up" as he compared his simulation answer to the answer he obtained from using the counting principle. Like Adam, Donovan noticed he "only got 6 out of the outcomes the first time" on the coin problem that had 8 total outcomes. He tried it again and got all 8 outcomes among his 40 repetitions noting differences in the results. When asked about his learning anything new from the simulations, he commented "The calculator is better at getting outcomes." This introductory simulation exercise fostered follow-up discussion regarding variation in results and the number of repetitions necessary to see all possible outcomes which encouraged students to begin to conceptualize the law of large numbers, sampling variability and the discrepancy between empirical and theoretical distributions. Though some students claimed they did not learn anything new from the simulation and that "it was too long," others noted the benefits:

Molly had initially shown errors in her tree diagram on the spinner problem, and although she said "not really" for learning anything new with the simulations, she did note when comparing her pre- and post-simulation answers, "I have gotten 18 outcomes so far and in my previous I only had 9. I did my tree wrong. The calculator is really somewhat random, its giving me outcomes I hadn't thought about."

Similar to Donovan, Charles noticed that he did not find all his outcomes with 20 trials but did 40 more on his own and found them all indicative of students constructing their own knowledge and building their own understanding of the law of large numbers through active learning.

Cole commented that the simulations made it "easier to find answers and mistakes than manually," and Wade and Alan working in a group together realized their answers on the weather question did not agree with their first page because "we didn't do separate raining results for each day." This comment is indicative of the simulation students using new representations developed from the technology to connect to other conceptual representations.

Whereas no students in the treatment group were left confused after doing the simulation, working through their mistakes and discrepancies, and having their group discussions, Shanelle in the control group commented on the coin problem "I did 2*3 to get six and 2*2*2 to get 8 but I do not know which is right. A tree diagram will give me six also but I am still confused." Her tree diagram was wrong, but her list of outcomes was correct. Her peer-to-peer discussion failed to lead to a resolution, and without the aid of the representations developed from the technology, she was left unsure of the correct approach.

Day 2 - Probability Rules, Independence, Multiplication Rule

Students in the control group were presented with notes regarding probability models, the complement rule, the addition rule for disjoint events and the multiplication rule for independent events according to the Yates, Moore, & McCabe text (Yates et al, 1999). Students in the simulation group went to the computer lab and learned initially how to assign weights to events to develop an understanding beyond the equiprobability assumption (Lecoutre, 1992). The worksheets and lab for Day 2 can be found in Appendix F. Students began to see how Minitab© could assign a weight to simulate a basketball player being a 70% shooter and noticed that the random number generator produced binary numbers that favored the player making shots rather than missing them. Students repeated the sampling to get comfortable with the weighting procedure in Minitab© which also presented the opportunity to talk about variation in results and randomness. Students were asked to generate two columns of 100 shots each and compare their own results and then to compare with others sitting next to them. It was apparent that students were developing their own understanding of randomness. This developing conception of variation indicates the simulation activities were helping students progress from Level A to Level B noted in the GAISE framework (American

Statistical Association, 2007). When asked if everyone had the exact same answers, the responses were as such:

"It's random. And in real life you won't get exact same numbers of makes and misses. 70% and 30% are averages."

"Our answers are close. This is reasonable because both of our outcomes should have been close to 30 misses and 70 makes."

"It randomly generated the same numbers as before, which is highly unlikely. I wouldn't expect it to happen too often."

Students were then instructed to explore joint probabilities by simulating the player shooting twice. They were prompted to remember the generation of the sample space of outcomes from the previous day and then to use Minitab© to calculate the percentage of no shots made, one shot made and two shots made. Finally, they were then instructed to generate a dotplot of their results and make the connection to the multiplication rule for independent events. In this transfer from the generation of outcomes to the sum of the number of shots made to the eventual dotplot, it was partly my intention to help the students begin to develop the construct of a random variable and the distribution of a random variable.

Shane, a low-level student who had struggled with the tree diagrams and sample space the day before showed improvement in his tree diagram as well as the generation of outcomes using the 11, 10, 01 and 00 notation to help him then list make-make, make-miss, miss-make, and miss-miss outcomes. When describing the dotplots, students correctly associated the various frequencies with most likely and least likely number of baskets. Students compared their calculations of making both shots using the independence rule with their simulated outcomes. When asked if his simulated answer agreed with his answer using the multiplication rule, one student who got exactly the same simulated probability as the formula stated "Yes it is identical which is probably uncommon that it hits it exactly." Comments such as these provide evidence that the students may be able to understand sampling variability, point estimation and interval estimation as they are cognizant of expected variation in their results and appear to be making a transition from Level B to Level C in the GAISE framework (ASA, 2007). When asked if their dotplots were the same as their neighbors, one student responded "No because there is variation." Another responded "They are not the same but they are close. My neighbors are also working with the data for a 70% shooter." This indication of knowledge of the source of the data and the connections made between the sample data and the source suggests that students are beginning to make the bi-directional link between the data-driven distribution and the modeling distribution encouraged by Prodromou and Pratt in their research (Prodromou & Pratt, 2006).

## Day 3 - Independence, General Addition Rule, Joint Probability

In this section of the text, the authors presented the use of Venn diagrams as a visual representation of the relationships among events and event probabilities. The notions of complementary events, disjoint events, and independent events were also presented with discussion of the general addition rule for union of events and multiplication rule for joint probability of independent events: P(A and B) = P(A)P(B) if A and B are independent (Yates et al, 1999,). The topics of disjoint and independent events are typically confusing for students, and I personally struggled with the design of a simulation exercise that could properly embellish the students' understanding of these topics. Too many topics were presented in this one section of the text, and I believe if I had split up the simulation tasks across several days, I would have seen more success with this lab. In addition, it may be more logical to group the concept of independence with conditional probability to develop a better understanding of the concept. This activity lacked the flow of the previous ones and, in hindsight, did not encourage connections within the students' zones of proximal development.

Both the control group and simulation group were given notes with examples on formulas for P(at least one) using the complement rule, constructing Venn diagrams, and applying the general addition rule and the multiplication rule for independent events. The control group proceeded to work on exercises from the text whereas the simulation group went to the lab to use Minitab©. The worksheets and lab for Day 3 can be found in Appendix G. The simulation group was asked to simulate a company shipping boxes of computer chips that contained 4 chips assuming the company had a 6% defect rate. In hindsight, I should have picked a higher defect rate or another more appropriate simulation as many students got 0 defective items in all of their 100 boxes of 4 chips, and the magnitude of the conceptual understanding of the complement rule from the simulation was diminished. Students did understand from observation of their outcomes that "at least one" was inclusive of all outcomes except "none" thus bridging a conceptual understanding of the formula presented in the text.

The second part of the simulation was an attempt to use outcome analysis from the simulated results of rolling two dice to help the students distinguish among the terms mutually exclusive or disjoint, independent and dependent events, and to gain a deeper understanding of when and when not to apply the multiplication rule for compound events. Students struggled to mathematically calculate the probability of A and B assuming A = rolling a 6 and B = sum of 12 despite a discussion of all possible outcomes when rolling two dice. Two other AP Statistics teachers assisted me during this lab, and we all spent a great deal of time helping students with their calculations thus taking time away from the actual simulation and the central points of the simulation exercise. Students did notice from the outcomes that you could only get a sum of 12 if the first value in the outcome was a 6 and thus concluded that the sum of 12 did depend on what they got on the first roll. On the other hand, they struggled to make sense of the mathematical formula and why their experimental calculation of the probability of getting a 6 on the first die and a sum of 12 did not reconcile with their assumption of independence and their theoretical calculation of P(A)*P(B). A follow-up discussion with both groups the following day using tree diagrams to illustrate why we could not use 1/36 as the P(sum of 12) was conducted to aid the student's conceptual understanding of P(A and B) when A and B are dependent.

Day 4 - Conditional Probability

Aligned with the frequentist approach to probability (Konold, 1989) and using the simulation outcome analysis techniques developed thus far in the study, my intent with this lab was to help the students understand that with conditional probability, the scope of relevant outcomes was reduced due to the conditions involved in the problem. Both groups were presented with the conditional probability formula P(A|B) = P(A and B)/P(B) and examples to illustrate the use of the formula using tree diagrams. The control group then worked examples at their desks while the simulation group worked the same examples using Minitab© in the lab. The worksheet and lab for Day 4 can be found in

Appendix H. The problem was a continuation of the rolling dice problem from the previous day. Time constraints in the lab once again prevented this lab from running smoothly as the students continued to struggle with reconciling their simulation results with their theoretical calculations. They did seem to understand that P(A|B) was not the same as P(B|A) because "you were looking at two different groups" indicative of Level 3 reasoning in the probabilistic thinking framework (Jones et al., 1999).

Interviews following this lab did reveal important contrasts in the understanding of conditional probability from two high-level students, one from each of the two groups. Each was presented with the following question: Suppose the probability that a Hispanic marries another Hispanic is 97%. The probability that a non-Hispanic marries a Hispanic is 5%. 2% of a population is Hispanic. Find the probability that if a person marries a Hispanic that they are non-Hispanic? Shannon was a high-level student from the control group, and Alan was a high-level student from the simulation group. Both students calculated the probability correctly using the conditional probability formula and got an answer of 0.7164. When Shannon was asked to explain what her answer meant, she seemed untrusting of her answer and commented "I don't get it. Why isn't it .05?" She was ignoring the condition and thinking of the question of finding the probability that a Non-Hispanic marries a Hispanic. When probing her to explain her understanding of the question and solution, she continued, "What I'm getting is that 71.64% of people that marry Hispanics are not Hispanics, but this seems too high because not that many Non-Hispanic people marry Hispanics." When asked how he felt about his answer, Alan also commented that he felt his answer was too high. When probed to explain his understanding of the question and solution, he stated, "Well, I guess it could be that high because most of the people are not going to be Hispanic. I get it more than just the formulas. I can't do the formulas sometimes, but I get what it is saying." Although Alan had a harder time articulating his thinking compared to Shannon, his comments did seem to indicate he understood the concept of the smaller scope of people that were being considered for the answer to the question and what that scope would contain based on the given probabilities in the question.

## Day 5 - Discrete Random Variables

This task was designed to help students further develop their understanding of the construct of a random variable and to begin to associate the results with an empirical probability distribution. Both groups were given the first part of the worksheet to complete before then being given notes on probability distributions and the formula for expected value. The following day the control group worked on exercises from the text while the simulation group went to the lab. The worksheets and lab for Day 5 can be found in Appendix I. The text presented the construction of the probability distribution of a discrete random variable using a coin example with $X$ = the number of heads in 4 tosses of a coin and explained that one must find the total number of outcomes first. Then to find $P(X = 2)$, one needed to count the number of ways that $X = 2$ could occur. The weakness in using a coin example is the encouragement of an equiprobability bias as the two outcomes on a coin are equally likely. Unfortunately, given the pretest results, this was already a vulnerable area for these students.

The example on the worksheet assumed an 8% defect rate. The company was to ship four items to a customer. On the initial page of the worksheet, most students in both groups continued to show improvement in determining the number of possible outcomes using both tree diagrams and the counting principle. In addition, the majority of students in both groups were able to transfer from the list of outcomes to generating all possible values of the random variable though a few carelessly listed 1, 2, 3 and 4 omitting the possibility of all non-defective items for a value of X=0. The most salient discrepancy between the pre-simulation responses from both groups was in constructing the theoretical probability distribution of X. Most of the students in the control group incorrectly calculated the probabilities in their distribution assuming equally likely outcomes and using part-whole reasoning. For P(X = 2), one stated "there are 6 outcomes out of the 16 that could have X = 2" and then calculated the probability as 6/16 = 0.375. All but a few of the control group students did the same types of calculations to complete their probability distribution table. On the other hand, there was more variation in the simulation group's calculations, some using the same equiprobability bias as the control group, but many trying to apply the actual 8% probability given in the problem. For those students that did use the 8% defect rate, many erroneously calculated the probabilities due to lack of combinatorial reasoning. For example, the P(X = 2) calculations showed .92*.92*.08*.08 neglecting to account for the 6 different ways that the two defective items could occur. The author posits that the consistent integration with data from the simulation group triggered an increased awareness of using the probabilities in their calculations as opposed to the equiprobability use of part/whole. Most students in both groups incorrectly calculated the mean of the random variable by disregarding the probabilities and just adding up the five values of X and dividing by the 5 possible outcomes : 0 + 1 + 2 + 3 + 4 = 10/5 = 2.

The simulation generated meaningful discussion regarding the simulated probability distribution and the theoretical probability distribution. The exploration of data led to rich explanations regarding the true meaning of an 8% defect rate, a conceptual understanding of expected value of a probability distribution,  and a developing understanding of a probability distribution. Based on the simulated results, one student described the meaning of the 8% defect rate as "an average 8 out of each 100 items will be defective." At this point in the study after five days of exploring data, I was confident the simulation students were developing a strong sense of variation and randomness. Due to the low defect rate, some students only showed 0, 1, and 2 defectives in their simulated outcomes which created a teachable moment regarding simulated probability distributions and theoretical probability distributions. When asked to describe this discrepancy with their answer before the simulation, one student replied "The computer simulation does not include the possibility of 3 or 4 defective items because those results are so rare in real life." Comments such as these give evidence that students in the simulation group were beginning to see the connection between their distributions and the larger global theoretical distribution and were moving between Levels B and C in the GAISE framework (ASA, 2007) and between Levels 3 and 4 in the probabilistic thinking framework (Jones et al., 1999). When asked to calculate the mean of their number of defective items using the simulated results and compare to the answer they got on the initial worksheet, students realized their mistakes and many noticed their tendency to assume equal weights. One student commented, "I didn't take the different values into account." From reading his responses, I was able to discern he meant the different weights or probabilities when he said "values." Students calculated the mean of their simulated results by averaging all their values of the random variable X then comparing the answer to what they would have gotten if they had substituted the values in the expected value formula they were taught. When asked if the mean value made sense, one student replied, "Yes, because most of the values are 0, so the mean should be close to 0." Similarly, another stated, "Yes because the abundance of zeros will pull the mean down." Students were asked if they understood why their earlier predictions may have been incorrect. From the responses to this question, students realized such things as they forgot to include 0 as a possible value of X. This was readily apparent as most of the simulated results were 0 defective items. One student commented, "I learned that weight plays a huge role in determining which results are most likely and that not all probability distributions have equal weight." Another student said, "Yes, because the true probability of getting it effects the results." Others commented that they noticed the mean could be a decimal number even though the values of X were not decimals thus realizing that the mean is not necessarily the most probable value of a distribution. This was in contrast to a comment from a student in the control group. When discussing 2.6 for the expected value of the size of an American household, Allyson questioned "Doesn't it have to be a whole number?" When I asked her why she thought it had to be a whole number, she responded "Because it's discrete." Although the control group students could accurately calculate the expected value using the formula, the understanding of what the answer represented was not as clear without the simulation. When the treatment group was asked what the simulation helped them learn about probability distributions, one student wrote, "It gave me a better understanding of what the % and probability answers mean. Helps me <u>see</u> what's happening" and another said, "Yes, at first I didn't understand what I was finding for "value of X" but now I understand" confirming that the simulations were helping the students develop a conceptual understanding of the concept of a random variable and the formulas that are presented in a textbook. The descriptive explanations of what we mean by a "probability distribution" included such responses as "Possible outcomes for X and chance of it happening, how often we will get certain outcomes for X" and "A probability distribution is a graph that shows the distribution of probability, like it shows how much probability for an x value compared to other x values."

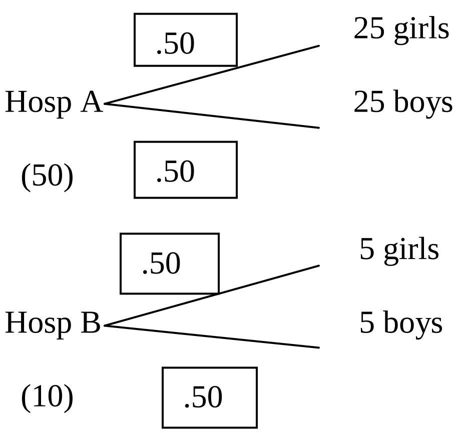
### Day 6 - Representativeness

To start this activity, both groups were given some sample questions from the pretest regarding the representativeness heuristic. The control group then explored their answers to the questions with a discussion of the assumptions and formulas that would apply to the calculation of the probabilities of the various sequences. The simulation group used their TI-83 calculators to simulate the coin sequences and then pooled their results with other group members to note the changes in the calculations and to hopefully see results that were approaching equal probabilities for each sequence. As labs using the calculator typically seemed to go, this one took longer than anticipated so students were not able to finish the entire lab activity. The worksheet and lab for Day 6 can be found in Appendix J.

When I began to analyze the results from this activity, one of the most salient discrepancies I noticed right away was the large percentage of students in the control group who answered that the sequences in all three questions were all equally likely. 79% of the students in the control group answered "all of the above sequences are equally likely" for each of the three questions compared to 33% of the treatment group. After looking more closely at the responses to the questions, I noticed that on Question 2 regarding three fair dice, the students in the simulation group were more likely to show a list of outcomes to arrive at their answer. I believe this inclination to be due to the representations developed from analyzing outcomes from conducting the simulations. In addition, I believe the previous simulations were helping to improve their awareness of the equiprobability bias and to make them more cognizant of situations that were not necessarily equally likely. The control group students who were limited primarily to using formulas in their understanding showed a propensity to fall back on the equiprobability bias. They didn't have the visual recollection of generated outcomes to remind them of this misconception. While many students had initially put that it was harder to get repeats in their explanations, after doing the simulation, students commented that it didn't matter what had happened previously as the calculator would still give an equal chance when generating the next list of numbers. Using the technology helped the students see there was no "memory" of what had happened previously thus aiding in their understanding and acceptance of the independence assumption. The pooling activity helped the students see that their results got closer to equal probabilities the more trials they observed and contributed to their developing understanding of the law of large numbers. Students acknowledged and responded that there was less variation in their pooled results and their pooled results were more accurate with what they expected would happen. One student who had time to conduct the simulation on the three dice problem had initially answered the question correctly but with incorrect reasoning. She had responded that the 5, 3 and 6 in any order would be more likely because it was less likely to get repeats. After conducting the simulation, she commented that she still agreed with her answer but in her explanation, she now wrote "Yes, because the 5, 3, 6 can be in any order, there are 6 ways but 555 has one way and 553 has 3."

## Day 7 - Law of Large Numbers

This activity presented a familiar "hospital" problem that prompted the students to choose whether a large or small hospital was more likely to record 80% or more female births (Fischbein & Schnarch, 1997). The worksheet and lab for Day 7 can be found in Appendix K. Surprisingly, after the previous day's lab on representativeness when the simulation group pooled results with other students to see the effects of the law of large numbers, 43% of these students said both hospitals were equally likely to have 80% or more female births for Hospital A with 50 births a day compared to Hospital B with 10 births a day. As noted in the literature, probability misconceptions are persistent and difficult to eradicate despite instruction (delMas, Garfield, & Chance, 1999), and the concept of variation was still elusive to many students shackled by their deterministic tendencies. In the development of the probabilistic thinking framework, it was noted that students could waver back and forth between levels especially with regard to common misconceptions (Jones et al., 1999). In the control group, 41% believed the two hospitals were equally likely to have 80% or more female births. Most of these students in both groups reasoned with such responses as "because boys and girls probability is 50/50 so it doesn't matter how many births there are a day." This response reflects the opposing development of the principle of equivalence of ratios and the understanding of the law of large numbers which Fischbein and Schnarch (1997) cited as justification of the sustainable strength of this misconception over time. The outcome approach was apparent with such responses as "because it is equally likely to have a boy or girl no matter how many births." Curiously, several students in the control group tried to "mathematically" calculate the answer to the question indicating an excessive reliance and dependency on formulas even when they may not be appropriate and still indicated probabilistic thinking at Level 2 according to the framework (Jones et al., 1999). This strategy reflects the emphasis on picking the right formula to get the correct final answer that is characteristic of traditional mathematics instruction (Cobb, 1999).  One student in the control group showed a type of tree diagram like the following:




and then followed with an attempt at applying a formula:

A = P(50% boy and 30% girl) + (more girls)

= .80 + .15 = .95

*Figure 4.11.* Tree diagram and formula solution approach for student in control group

She then wrote "I have no idea how to get that 80% cause they are all equal." I believe this excessive reliance on meaningless formulas in a particular situation is a danger when trying to teach probability to students without a conceptual understanding of the constructs that underlie the formulas.

The following day both groups were given notes on binomial random variables with associated formulas for the mean, variance, and the binomial formula for calculating probabilities and were introduced to foundational ideas regarding the formation of a sampling distribution of proportions by dividing the binomial formulas for mean and variance by the sample size. The simulation group then went to the lab to use Minitab© to illustrate the formulas and analyze comparison graphs for the two hospitals. Students clearly noticed the change in variation when converting from number of female births to proportion of female births. Although proportion plots for both hospitals were centered at 0.5, students noted the smaller variation of results in the plot of the hospital with more births. When asked what they learned from doing the simulation, the comments were rich and indicative of the development of distributional reasoning (Shaughnessy, 2006). One student commented, "I learned that variance is important in these studies. Before I didn't understand why we even needed to calculate variance." Another noted, "It helped me visualize the difference between short-run and long-run." Evidence of distributional reasoning stood out with such comments as "I thought that by just having a p=0.5 everything was equally likely. I forgot to consider sample size" and "I learned how to look at means and compare, how spread or predicted values match with the mean, and the difference between a lot of trials and fewer trials." As many students initially were giving explanations in terms of "number of" births, it was significant to read such comments as "while the variance may be higher in a larger group, probability can be more spread in the smaller one." Comments such as these illustrate a progression from additive reasoning to multiplicative reasoning as noted by Cobb (1999) and indicate the students' cognizance of reasoning proportionally in contrast to relying on statistical reasoning solely through counts and frequencies as admonished by Shaughnessy (2006).

## Day 8 - The Binomial Distribution

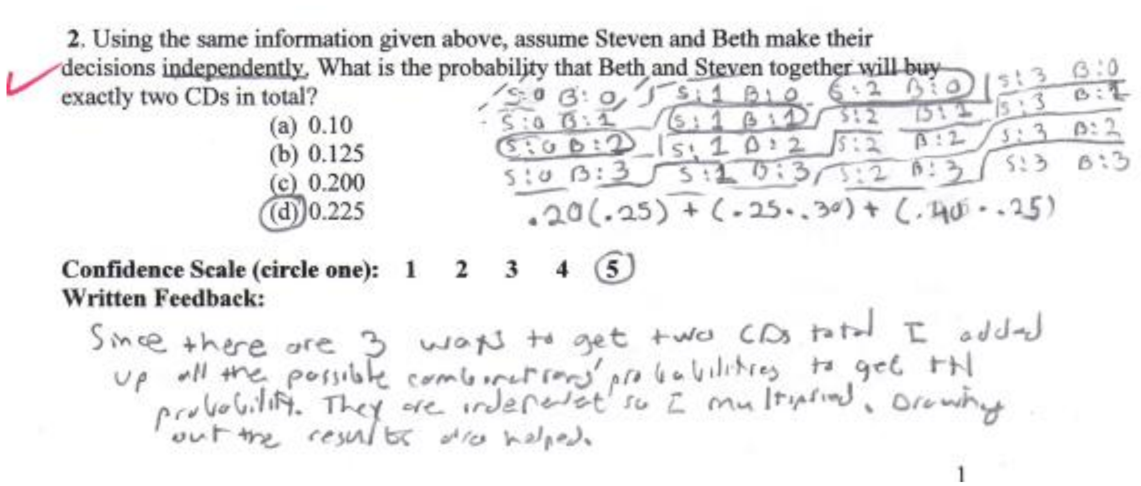
Following the sequence of topics presented in the text, the last two topics that the students explored for this study were the binomial and geometric distributions. Both groups looked more closely at the binomial formula for calculating probabilities. The simulation group was reminded of the lab on discrete probability distributions to determine if this fit the conditions of a binomial setting. The examination of outcomes was relived as students were reminded of the importance in considering how many ways a company could have two defective items out of 4. Subsequently, the simulation group took the opportunity to look at the relationship between shape, sample size, and the binomial distribution in the lab using Minitab©. The worksheet and lab for Day 8 can be found in Appendix L. These students used the software to generate random data from a binomial distribution and analyzed changes in the plots as the parameters changed from $p = .10$, $p = .50$ and $p = .90$ using $n = 10$ trials. They predicted shapes of the distributions and calculated means from the formula they had learned. Students were then asked to predict shapes for $p = .90$ as they changed the number of trials from $n = 5$, $n = 30$ to $n = 100$. Whereas the students noticed they had made correct predictions for the smaller number of trials, they were surprised to notice the changing shapes of the graphs as the number of trials increased from 5 to 100. This lab would hopefully have significant implications for the Central Limit Theorem for sampling distributions of proportions that would be taught in the following chapter.

Day 9 - The Geometric Distribution

For this final activity, both groups were presented with the conditions for a geometric distribution along with the formula for calculating the mean of a geometric distribution and the probability for a geometric random variable. The control group students looked at developing the distribution using the formulas whereas the simulation group looked at developing a simulated probability distribution first using a random number table and then comparing these calculations to their theoretical calculations. The worksheet and lab for Day 9 can be found in Appendix M. Prior to conducting the simulation, the students in both groups had been asked to predict what they believed would be the most probable number of trials until a success using a basketball player missing a basket and prizes in a cereal box examples. Most students had incorrect intuitions regarding the geometric distribution and predicted that it would take 3 or 4 trials until the player missed a shot. Many students correctly answered that the chance of getting the first prize in the tenth cereal box was not higher than the chance of getting the first prize in the first box, but some noticed their reasoning was incorrect after conducting the simulation. The simulations using the random number table went slowly, so unfortunately the students were not able to conduct enough simulations to make their simulated probabilities reconcile with the theoretical probabilities. They were, however, able to notice their initial predictions were incorrect as most successes were recorded in the first or second trials. Many students struggled to understand why this was so, but some were able to justify the right-skewed shape mathematically due to the formula. As one student commented about the shape of the distribution, "Right skewed theoretically b/c every probability is usually less than 1, so the #s should get smaller, but w/ simulation everything is random & different things can happen in the short run." Students were comfortable enough by now with the idea of the law of large numbers that they commented on the discrepancies between their empirical mean and theoretical mean as "there were only 20 trials in the simulation." At the end of the worksheet, students were asked to describe the differences in running simulations for a binomial with a basketball player taking three shots and the corresponding geometric distribution example. One student commented, "You're always going to be looking at 3 shots in a binomial. But in a geometric you do trials until a success. You would look for sets of 3 numbers instead of looking for numbers until you get a hit." Another student responded with a more descriptive answer based on the previous labs we had completed, "It would mean that you would have groups of three constantly. In a lab you would have three columns. All could be misses or a various combination with 2^3 options which is 8." Based on the pretest results with students' proclivity to look at part-whole calculations, I hoped that by conducting the actual simulation that this lab would help students notice the variation in the number of trials and understand the geometric distribution has a random variable that represents the total number of trials without it being a given, fixed value.

## The Posttest

A qualitative analysis of the posttest responses showed a heavy reliance on formulas, sometimes meaningless mathematical calculations, for many of the students in the control group. On the other hand, the responses from the simulation group showed a mélange of methods indicative of their exposure to a wider variety of representations to aid them in their conceptual understanding of probability and probability distributions. Though questions regarding compound events still presented a struggle for both groups, there was evidence that the simulations using the exploration of possible outcomes gave the students a starting point to build their solution. The following solution was given by a student in the simulation group on Question 2 of the posttest:


**2**. Using the same information given above, assume Steven and Beth make their decisions independently. What is the probability that Beth and Steven together will buy exactly two CDs in total?

(a) 0.10
(b) 0.125
(c) 0.200
(d) 0.225

.20(.25) + (.25 · .30) + (.40 · .25)

**Confidence Scale (circle one):** 1 2 3 4 5
**Written Feedback:**

Since there are 3 ways to get two CDs total I added up all the possible combinations' probabilities to get the probability. They are independent so I multiplied. Drawing out the results also helped.

1

*Figure 4.12.* Listing of outcomes solution approach for student in simulation group

An item-by-item analysis of the posttest peformance for both the simulation group and the control group showed very similar patterns again, but compared to the control group, the simulation group showed consistently higher percentages of correct responses than on the pretest item analysis.

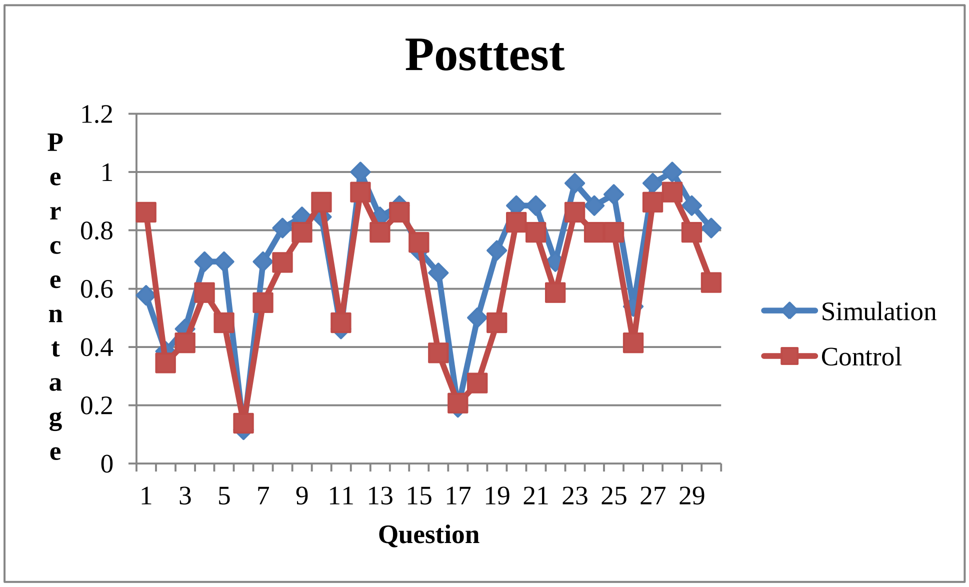



*Figure 4.13.* Posttest Proportion of Correct Responses by Question

A chart showing the percentage differences between the simulation group and control group (Simulation - Control) for both the pretest and posttest show the greatest gains for the simulation group on questions involving shape, law of large numbers, geometric distributions, and the conjunction fallacy. The simulation group had seen all of these topics from a process-oriented, data-centered perspective with the labs versus a formula-oriented, content approach used with the control group. The questions with negative differences, highlighted in yellow, reflect topics from labs that I deemed less successful in the day-by-day discussions including compound events and representativeness. In designing my labs, these topics seemed to disrupt the flow of developing the concept of sample space to random variable to distribution of a random variable, so they may have been beyond the students' zones of proximal development. Recall from the quantitative analysis, however, that GROUP did not show a significant main effect for posttest scores.

Table 4.6

*Summary of Percentage Correct Differences Between Groups (Simulation - Control) by Question*

| Question | Pretest | Posttest | Content |
|---|---|---|---|
| 1 | 25% | -29% | compound events |
| 2 | -13% | 4% | sample space |
| 3 | -2% | 5% | shape |
| 4 | -0.3% | 11% | binomial distribution |
| 5 | -6% | 21% | binomial shape |
| 6 | -11% | -2% | sample space |
| 7 | 1% | 14% | shape |
| 8 | -5% | 12% | independence |
| 9 | -8% | 5% | independence |
| 10 | 4% | -5% | representativeness |
| 11 | -8% | -2% | sample space |
| 12 | 9% | 7% | shape |
| 13 | -25% | 5% | shape |
| 14 | 24% | 2% | representativeness |
| 15 | 6% | -2% | proportional reasoning |
| 16 | -3% | 27% | law of large numbers |
| 17 | 4% | -1% | compound events |
| 18 | -7% | 22% | geometric distribution |
| 19 | 25% | 25% | independence |
| 20 | 19% | 6% | independence |
| 21 | 23% | 9% | shape |
| 22 | 10% | 11% | compound events |
| 23 | 32% | 10% | randomness |
| 24 | 6% | 9% | conditional probability |
| 25 | 17% | 13% | geometric distribution |
| 26 | 8% | 12% | equiprobability |
| 27 | 21% | 6% | randomness |
| 28 | 3% | 7% | outcome approach |
| 29 | 10% | 9% | independence |
| 30 | 1% | 19% | conjunction fallacy |

Recall from the quantitative analysis that the GROUP*LEVEL interaction was significant for posttest scores. An item-by-item analysis of the pretest and posttest performance for the low-level students and the high-level students from both groups accentuates the gains within each level.

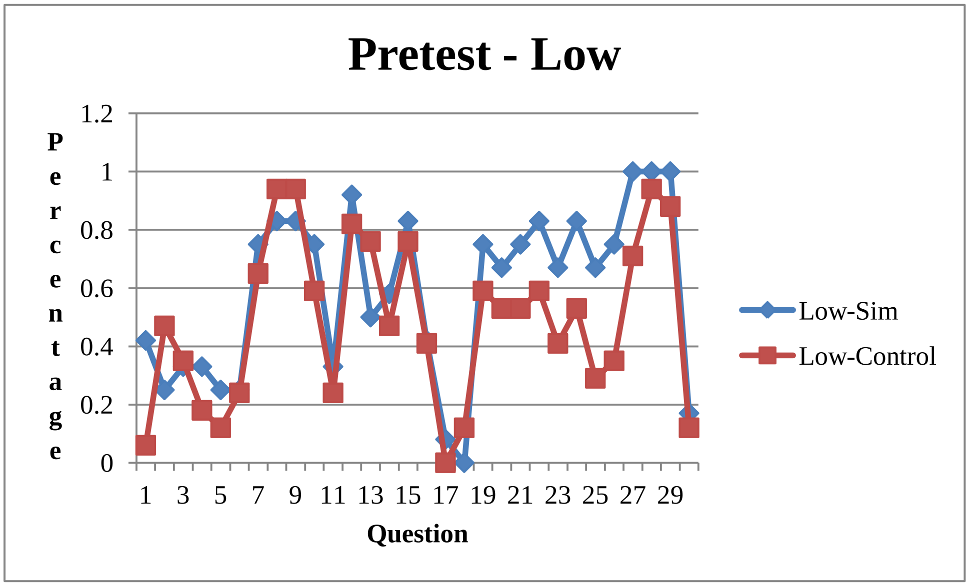



*Figure 4.14.* Pretest Proportion of Correct Responses by Question for Low-level Students

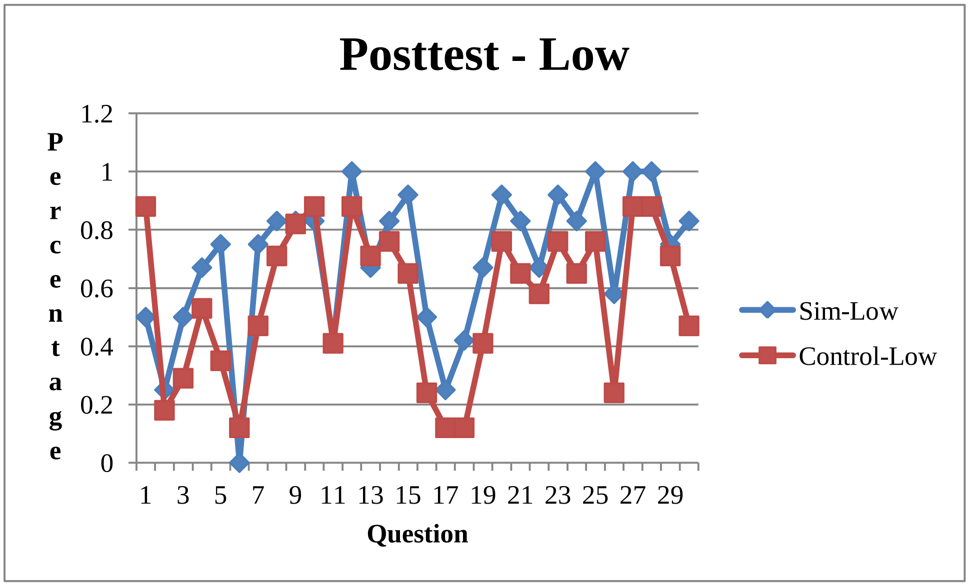



*Figure 4.15.* Posttest Proportion of Correct Responses by Question for Low-level Students

The low-level students in the simulation group surpassed the correct response percentages for the control group with substantial differences on nearly every item on the posttest assessment. This dominance is more transparent in the following table showing percentage correct differences on the pretest and posttest for low-level students. The yellow highlights show the questions with negative differences.

Table 4.7

*Summary of Percentage Correct Differences Between Groups (Simulation - Control) by Question for Low-level Students*

| Question | Pretest | Posttest | Content |
|---|---|---|---|
| 1 | 36% | -38% | compound events |
| 2 | -22% | 7% | sample space |
| 3 | -2% | 21% | shape |
| 4 | 15.0% | 14% | binomial distribution |
| 5 | 13% | 40% | binomial shape |
| 6 | 1% | -12% | sample space |
| 7 | 10% | 28% | shape |
| 8 | -11% | 12% | independence |
| 9 | -11% | 1% | independence |
| 10 | 16% | -5% | representativeness |
| 11 | 9% | 1% | sample space |
| 12 | 10% | 12% | shape |
| 13 | -26% | -4% | shape |
| 14 | 11% | 7% | representativeness |
| 15 | 7% | 27% | proportional reasoning |
| 16 | 1% | 26% | law of large numbers |
| 17 | 8% | 13% | compound events |
| 18 | -12% | 30% | geometric distribution |
| 19 | 16% | 26% | independence |
| 20 | 14% | 16% | independence |
| 21 | 22% | 18% | shape |
| 22 | 24% | 9% | compound events |
| 23 | 26% | 16% | randomness |
| 24 | 30% | 18% | conditional probability |
| 25 | 38% | 24% | geometric distribution |
| 26 | 40% | 34% | equiprobability |
| 27 | 29% | 12% | randomness |
| 28 | 6% | 12% | outcome approach |
| 29 | 12% | 4% | independence |
| 30 | 5% | 36% | conjunction fallacy |

The greatest gains for the low-level students in the simulation group from the pretest to the posttest were on questions involving shape of distributions, shape of binomial distributions, law of large numbers, geometric distributions, and the conjunction fallacy.

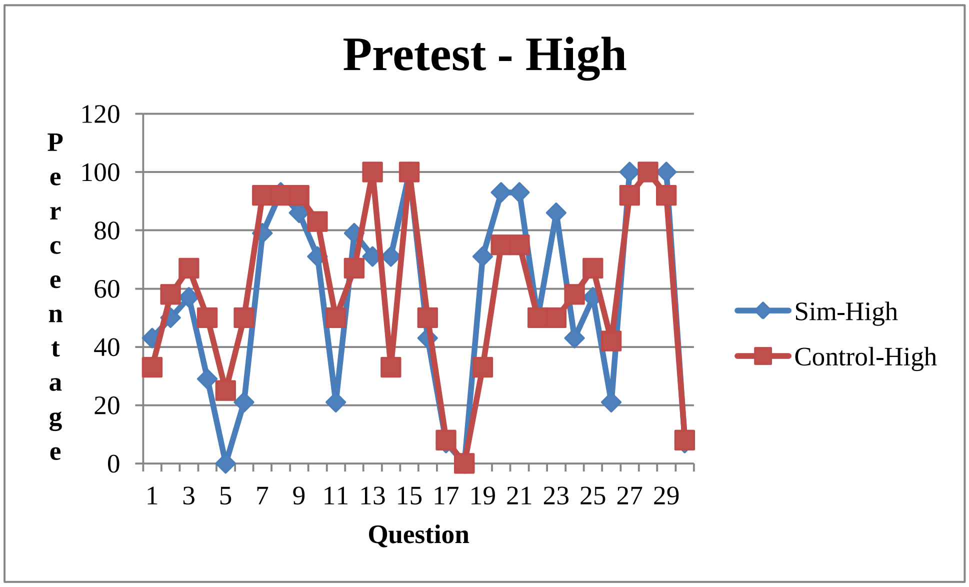



*Figure 4.16.* Pretest Proportion of Correct Responses by Question for High-level Students

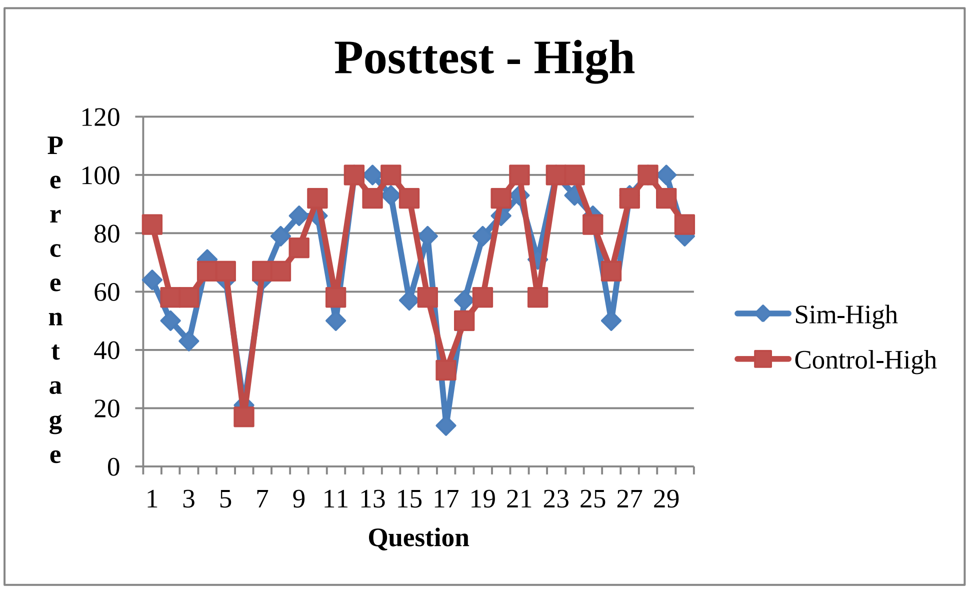



*Figure 4.17.* Posttest Proportion of Correct Responses by Question for High-level Students

Table 4.8
*Summary of Percentage Correct Differences Between Groups (Simulation - Control) by Question for High-level Students*

| Question | Pretest | Posttest | Content |
|---|---|---|---|
| 1 | 10% | -19% | compound events |
| 2 | -8% | -8% | sample space |
| 3 | -10% | -15% | shape |
| 4 | -21% | 4% | binomial distribution |
| 5 | -25% | -3% | binomial shape |
| 6 | -29% | 4% | sample space |
| 7 | -13% | -3% | shape |
| 8 | 1% | 12% | independence |
| 9 | -6% | 11% | independence |
| 10 | -12% | -6% | representativeness |
| 11 | -29% | -8% | sample space |
| 12 | 12% | 0% | shape |
| 13 | -29% | 8% | shape |
| 14 | 38% | -7% | representativeness |
| 15 | 0% | -35% | proportional reasoning |
| 16 | -7% | 21% | law of large numbers |
| 17 | -1% | -19% | compound events |
| 18 | 0% | 7% | geometric distribution |
| 19 | 38% | 21% | independence |
| 20 | 18% | -6% | independence |
| 21 | 18% | -7% | shape |
| 22 | 0% | 13% | compound events |
| 23 | 36% | 0% | randomness |
| 24 | -15% | -7% | conditional probability |
| 25 | -10% | 3% | geometric distribution |
| 26 | -21% | -17% | equiprobability |
| 27 | 8% | 1% | randomness |
| 28 | 0% | 0% | outcome approach |
| 29 | 8% | 8% | independence |
| 30 | -1% | -4% | conjunction fallacy |

Comparatively, we can see that the high-level students in the simulation group did not fare as well with significant gains as the low-level students did with many negative differences highlighted in yellow. This is not surprising as notes from interviews, discussion and written feedback indicate a preference for using mathematical formulas for the high-level students. A large discrepancy in favor of the control group on Question 15 prompted a closer inspection of the written feedback on that question. These responses revealed how the simulation on the law of large numbers actually hindered students' understanding of this particular question. Several students recalled the Hospital A versus Hospital B simulation and put that if one marble was chosen, Box B (with 60 red and 40 blue marbles) would have a greater chance of giving a blue marble than Box A (with 6 red and 4 blue marbles). One high-level student wrote: "Since there are a greater # of marbles in Box B, according to the law of large #'s the chance of picking a blue marble (even though its .40 for each box) should be closer since there are more marbles." Another wrote: "b/c although both boxes have the same theoretical probabilities, box B has a greater total amount of marbles which would produce imperical (sic) results that b/c they reflect "the long run" are more likely to reflect theoretical probability." Both of these students had mistakenly associated the number of marbles in the populations as the number of trials that had been observed and did not see the misapplication of the law of large numbers as only one marble was being picked from the box. This conflict illustrates the tension between the developing notions of equivalence of ratios and the law of large numbers as noted previously (Fischbein & Schnarch, 1997).

## The Retention Test

The retention test was given three weeks following the posttest once the students returned to school from their winter break.

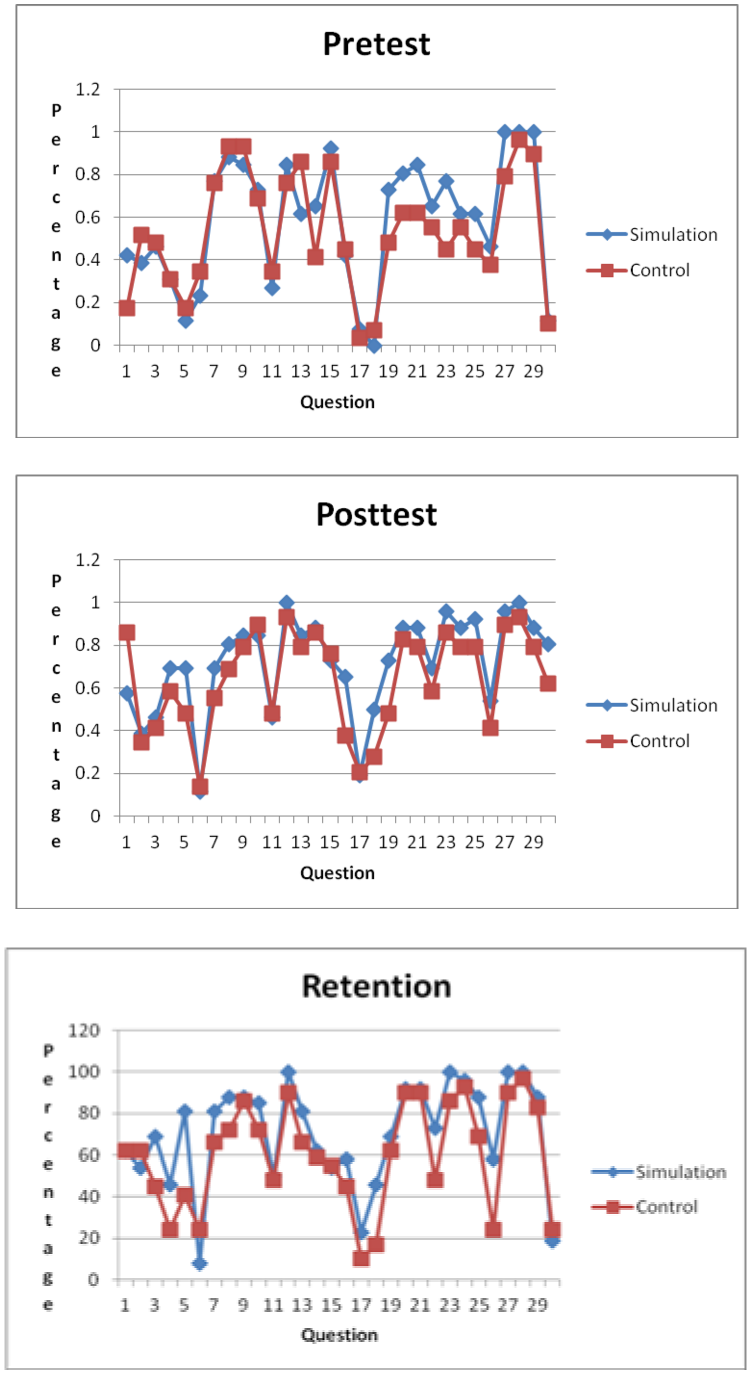



*Figure 4.18.* Summary Results of Proportion of Correct Responses by Question for All Students on Pretest, Posttest, and Retention Test

Comparing the retention item analysis results to the pretest and posttest results, we see a return to increased variation reminiscent of the pretest variation. This behavior reinforces the stubborn persistence of probabilistic misconceptions despite instruction noted in the literature (delMas et al, 1999).

Table 4.9
*Summary of Retention Test Correct Response Percentages by Question*

| Question | Simulation | Control | Content |
|---|---|---|---|
| 1 | 62% | 62% | compound events |
| 2 | 54% | 62% | sample space |
| 3 | 69% | 45% | shape |
| 4 | 46% | 24% | binomial distribution |
| 5 | 81% | 41% | binomial shape |
| 6 | 8% | 24% | sample space |
| 7 | 81% | 66% | shape |
| 8 | 88% | 72% | independence |
| 9 | 88% | 86% | independence |
| 10 | 85% | 72% | representativeness |
| 11 | 50% | 48% | sample space |
| 12 | 100% | 90% | shape |
| 13 | 81% | 66% | shape |
| 14 | 62% | 59% | representativeness |
| 15 | 54% | 55% | proportional reasoning |
| 16 | 58% | 45% | law of large numbers |
| 17 | 23% | 10% | compound events |
| 18 | 46% | 17% | geometric distribution |
| 19 | 69% | 62% | independence |
| 20 | 92% | 90% | independence |
| 21 | 92% | 90% | shape |
| 22 | 73% | 48% | compound events |
| 23 | 100% | 86% | randomness |
| 24 | 96% | 93% | conditional probability |
| 25 | 88% | 69% | geometric distribution |
| 26 | 58% | 24% | equiprobability |
| 27 | 100% | 90% | randomness |
| 28 | 100% | 97% | outcome approach |
| 29 | 88% | 83% | independence |
| 30 | 19% | 24% | conjunction fallacy |

The summarized assessment results by student for each of the two groups show some interesting results. The control group appears as a cacophony of scores seemingly as if the students just resorted to random guessing on each assessment. Several students had retention scores that dipped at or below pretest levels. Contrastingly, the simulation results show patterned, predictable, behavior with pretest results for most students lining the bottom layer and posttest and retention results closely skimming the top layer.

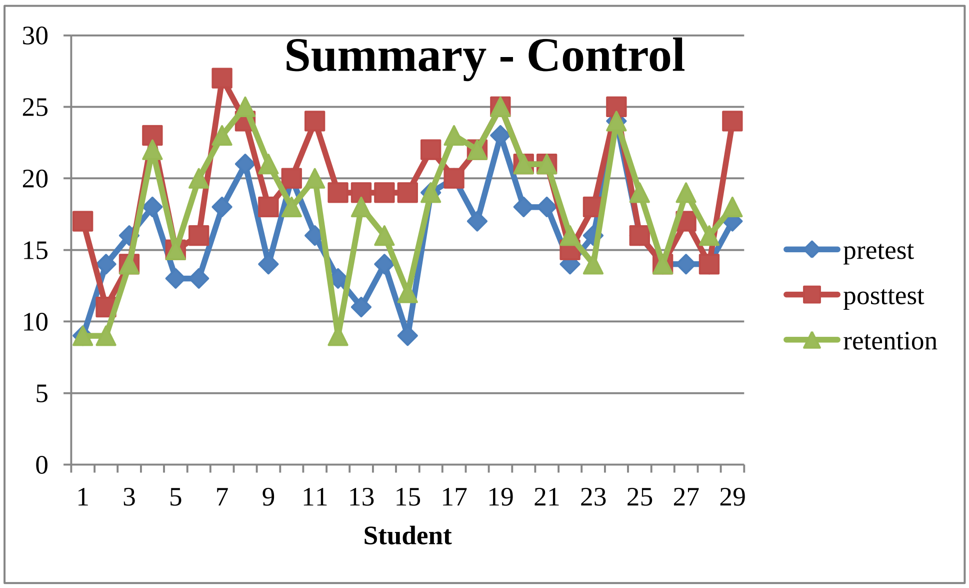



*Figure 4.19.* Summary Results of Number of Questions Correct by Student on Pretest, Posttest, and Retention Test for Control Group Students

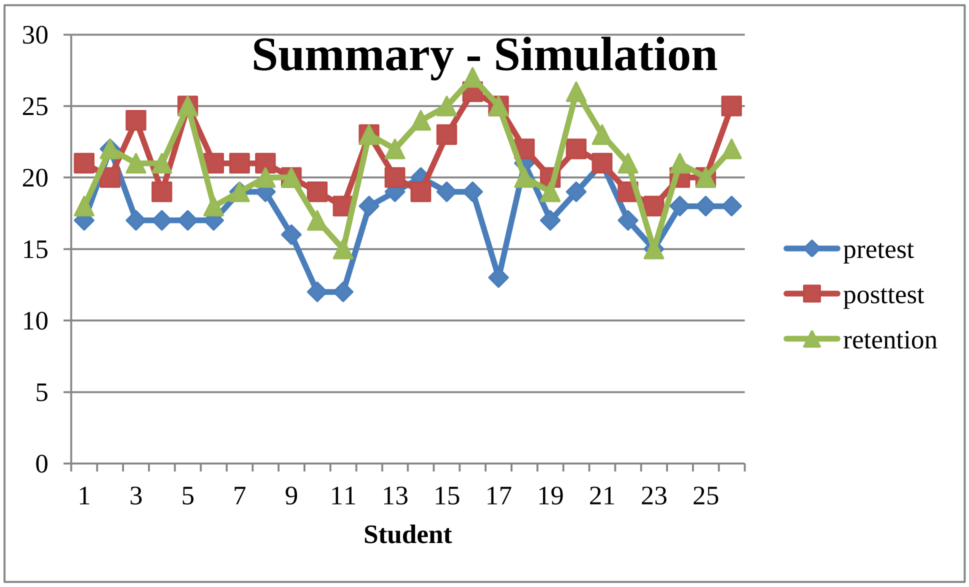



*Figure 4.20.* Summary Results of Number of Questions Correct by Student on Pretest, Posttest, and Retention Test for Simulation Group Students

Analogous to bringing a process under control, the researcher draws parallels with this graphical representation and posits this as evidence of conceptual change. The simulation group's performance reflects a stable, sustainable understanding of probability and probability distributions. Questions that indicated the greatest gains for the simulation students were ones involving shape, least and most probable events, geometric distributions, compound events, and equiprobability bias.

To conclude the results chapter of this study, I return to the voice of the student. Recall earlier queries that underlined student difficulties with probability: “not understanding what the question is asking, not being able to gauge whether the answer seems correct or not, and not understanding what I am really finding.” At the end of the posttest, students were asked to give their own definitions of terms and how they felt about probability. Students in the simulation group were subsequently asked if they trusted the simulations and how they felt about the simulations. Responses from the simulation group include the following:

- "Simulations are much easier than doing it by hand or actually collecting the data. Yes it helped because I could actually see what I was doing and if I was doing it right or wrong."
- "When we first started the simulations I didn't like them because they seemed so complicated. I didn't understand what I was really doing. I was just following directions. But, as I got used to it, I learned what the different columns meant. It helped me to visualize what was really happening. With just formulas you don't see what's happening. Simulation helps to see what's really happening."
- "I did enjoy the simulations. Formulas are great, but the simulations applied them to real-life situations. It did help by allowing me to manipulate the situation into a way I can understand it while still sticking to the formulas. I didn't like how time-consuming it turned out to be, but that's really the only negative I can think of."

Responses from the control group rendered an entirely different tone:

- "I thought before I took this post test that I understood probability more than I actually do. It seems like I don't know how to use or remember which formula goes with which problem or why. Because I don't understand it, I don't really like it right now."

- "I think it's hard and confusing.  Everything runs together. I hate it, because everything looks like the same thing to me and I can't tell the difference."
- "I'm confused about it.  I don't fully understand it, and I like to understand fully how something works instead of blindly following formulas."

# CHAPTER 5

## SUMMARY AND DISCUSSION

In this final chapter I will summarize the findings from my study relative to the initial questions posed at the outset of the research, situate my results within the existing body of research, and discuss the limitations and implications of the study.

### Summary of the Results

With this study I sought to provide empirical evidence and to embellish the developing literature on the use of simulation as an instructional tool in the teaching of statistics, specifically in the area of probability distributions. Simulations are lauded in the existing research base and curriculum frameworks, though claims are justified primarily through smaller case studies in isolated settings and qualitative analysis. As both teacher and researcher, my Advanced Placement Statistics classes were used in the study. I randomly assigned two classes to the treatment group using simulations in their instruction and two classes to the control group using traditional, textbook-guided instruction. I also subjectively designated the students as either low- or high-level performance based on previous course histories and performance in the class.

Both groups of students began the study with an abundance of common probabilistic misconceptions noted in the literature (Kahneman, Slovic, & Tversky, 1982; Konold, 1989; Lecoutre, 1992). Both groups initially showed a limited understanding of probability dominated by a goal-oriented, deterministic view of using formulas to get a correct answer. The two groups were immersed in a two-week series of teaching units that centered on the development of an understanding of probability distributions with the primary difference in the two classroom environments being the technological tool of simulation.

The study addressed the following research questions:

1. Does the use of simulations as an instructional tool aid in improving secondary students' understanding of probability distributions?
2. How does the use of simulations as an instructional tool help and/or hinder secondary students' understanding of probability distributions?

To answer the first question I look initially at the quantitative results from the study. A 2x2 factorial MANCOVA showed group differences (simulation vs. control) were not significant on the posttest scores but were significant on the retention test scores. Level differences (low vs. high), however, were significant on the posttest scores but not on the retention scores, and the interaction of group and level was significant for the posttest scores but not for the retention test scores.

Kaput's passion with his work was in the "democratisation of access to mathematical reasoning" (Kaput & Roschelle, 1997, p. 1; Kaput & Schorr, 2002). Kaput et al. believed that technological tools were the pathway to learning complex mathematical ideas that historically were accessible only to those few who could understand the formal abstractions of these topics. He pushed the representational and modeling potential of the technological tools as a way for the multitude of other students to access the knowledge of these concepts that involved change and variation. Though few studies on using simulations in probabilistic and statistical instruction have addressed performance level, the results from my study are emancipatory. Probabilistic concepts that could easily evade lower-level students can be made accessible to these students and have an equalizing effect on the traditional disequilibrium within the classroom. The quantitative results illuminate the performance of the low-level students in the simulation group. The low-level students in the simulation group made significant gains over the low-level students in the control group and nearly matched the mean performance of the high-level students in both the control group and the simulation group on the posttest. The low-level students in the simulation group made significant gains over the low-level students in the control group, surpassed the mean performance of the high-level students in the control group and nearly matched the mean performance of the high-level students in the simulation group on the retention test. Qualitative results indicate that higher-level students tend to show a resistance to using simulations in their solutions to the problems and a preference for using mathematical formulae. This was noted from interviews, observation logs, as well as written responses dating back to the initial pilot study. This resistance possibly explains the insignificant or lesser gains from the high-level students in the simulation group as compared to the low-level students. Despite this penchant for mathematical formulae, the significantly higher retention test scores for the simulation group indicate that possibly the high-level students relied on their conceptual knowledge from the simulations when the recollection of appropriate formulas failed them. Kaput and Schorr (2002) noted similar retention discoveries in their work with students from low-performing schools. The students made significant gains after using technology to learn complex algebraic concepts. Years later, follow-up stories showed these students still relied on visual representations developed during prior instruction with the technology to trigger solution processes.

The significant results on the retention test suggest that students who use simulations in their learning are likely to experience conceptual change. Posner, Strike, Hewson, and Gertzog (1982) proposed three, and later four, conditions necessary for conceptual change to take place. First, one must experience dissatisfaction with their current conception. Subsequently, the new conception must make sense and fit well within their current schema and offer productive ways of processing information. Instructional decisions and social context of the classroom may also play a role in successful change (Tyson, Venville, Harrison, & Treagust, 1997; Pintrich, Marx, & Boyle, 1993). The comparative design of my study, the inquiry-based nature of the lab sessions I constructed, and  the inclusion of a retention test provided a setting where conceptual change could be monitored. The collective quantitative results reflect a stable, sustainable understanding of probability and probability distributions for the students in the simulation group compared to unstable, highly variable results in the control group with many students reverting back to pretest performance. Qualitative analysis revealed that many students in the simulation group used visual representations developed during the simulation exercises and recollection imagery from the simulation experience to aid them in answering the questions on the posttest. Though qualitative feedback was not available for the retention test, it is plausible to conclude that the students continued to access these visual representations on this test as well.

To answer my second research question, I rely on a social constructivist perspective and retain a focus on the social nature of the classroom, the interactions between the teacher and the student, the peer interactions, and, most important to this study, the interactions with the simulation tool itself. It is significant to describe what the simulation activities brought to the social experience and how they influenced student reasoning throughout the study. The probabilistic thinking framework developed by Jones, Langrall, Thornton, and Mogill (1999) provided a framework to trace the progression and development in students' reasoning of probabilistic constructs and misconceptions relevant to the notion of probability distributions. These constructs included sample space, probability of an event, probability comparisons, conditional probability, and independence. Though the GAISE curriculum framework (American Statistical Association, 2007) was not initially considered in the design of my study due to the concurrent nature of the two events, this instructional framework proved relevant in the analysis of my findings as the three levels in this framework were more closely matched to the evolution of distributional thinking which was a focus of my study.

From the very beginning of the study with the development of the concept of sample space, students showed Level 1 and 2 reasoning in the probabilistic thinking framework (Jones et al., 1999) with incomplete listings of outcomes and an unsteady grasp of the concept of sample space. Through the simulations, the students in the treatment group gained confidence in their understanding of sample space as the technology prompted them to create outcome representations to guide them in their listings of multistage outcomes. These students began to develop a one-to-one mapping correspondence between the various representations of outcome listings, tree diagrams, and the fundamental counting principle which allowed them to move to Levels 3 and 4 in the framework. The control group students typically moved to Levels 3 and 4 through the instructional units with regard to sample space although the frequency of errors was greater and confidence was low when confronted with conflicting results from their tree diagrams and the counting principle.

All students initially showed many probabilistic misconceptions, some even using subjective judgments in their pretest responses for some of the probability questions thus starting at Levels 1 and 2 in the framework. The simulation group quickly adapted to a frequentist orientation in their reasoning with probability as they made associations between the weights of the probabilities and the randomly generated outcomes from the technology. Both groups showed part-whole reasoning in their probability calculations, although the control group students showed a greater vulnerability to the equiprobability bias than the simulation students. Again, I believe the visualization resulting from the outcomes representation reminded the students that not every outcome would be given equal weight. The students in the simulation group illustrated a test-retest function of the simulations to validate their conjectures. Although I initially encouraged this interaction with the simulation based on the theory of conceptual change (Posner, Strike, Hewson & Gertzog, 1982), the students began to spontaneously incorporate this validation process even when they were not prompted. The validation procedure proved successful in promoting student recognition of probabilistic misconceptions, errors in proportional reasoning, and errors in combinatorial reasoning. This was in stark contrast to the control group students who experienced cognitive conflict within their class discussions but often could not reach a resolution. Ultimately the connections made between the various representations from the simulations and the resulting evolution from outcomes to values of a random variable to the resulting probability distribution graph allowed the simulation students to move to Level 4 with confidence whereas the control group students struggled more with judgments involving "most likely" and "least likely" without the visual representations to guide them. The simulation activities for independent events, conditional probability, and representativeness proved too time-consuming for significant learning patterns to emerge. These concepts have been noted in the literature as areas where difficulties persist despite instruction (Batanero & Sanchez, 2005; Tarr & Lannin, 2005), and further research needs to be conducted in these particular areas.

Qualitative evidence suggests that students in the simulation group reasoned differently about probability distributions than students in the control group. Responses from students in the simulation group suggest a weaponry of representations utilized by these students with conceptual constructions including outcome analysis, drawings, tables and formulas. These students articulated a sense of continuity and depth in their responses developed from their consistent, daily exposure to the generation of outcomes to the recognition of patterns and to the drawing of conclusions. The group comparisons of the simulation output allowed the students to construct their own knowledge of randomness, variation, central tendency, distribution, and the law of large numbers, all considered the bigger ideas to gain from a statistics class according to Scheaffer, Watkins, and Landwehr (1998). Like myself, these researchers see probability distributions as a critical factor in developing notions of statistical inference. By starting with outcome analysis and development of sample space for a specific example, the students had started at Level A in the GAISE framework. The easy movement among representations with the technology including spreadsheet sums allowed the students to then connect the outcomes to the concept of a random variable with its associated probability and then to a graphical representation. Here, along with peer-to-peer comparisons, the simulation students were able to transfer to Level B with a developing understanding of the concept of an empirical probability distribution. The group comparisons and ability of the technology to generate a large number of trials continued the evolution with a developing understanding of the law of large numbers, and most seemed poised and ready to move to Level C as they began to understand generalizations to theoretical distributions. Limited by formulas and without the visual representations generated by the simulations, the control group struggled to make sense out of questions that required connected thinking about distributions. This is evidenced by the results as the most significant gains for the simulation group were on questions related to binomial distributions, geometric distributions and the law of large numbers.

It is significant to look at the results from the control group instruction. As noted earlier, I have taught for many years, and the earlier years I followed a traditional, teacher-centered approach. This is a natural tendency for me in my teaching as I believe it is for many mathematics instructors. There was still significant learning that took place in the control group - the high-level students performed just as well, and in some cases better, than the high-level students in the simulation group. Discussions were still oriented towards ways to develop conceptual representations for the probabilistic concepts being taught. The control group students found success in their representational systems including tree diagrams, tables, and formulas. Some of the assessment questions could be solved easily with the appropriate use of formulas, and as evidenced by the qualitative analysis, many of the control group students used this strategy. Calculations for probabilities involving simple events, fundamental counting principle, and expected value are fairly easy calculations mathematically, although these students could not articulate the meaning of their answers as well as the simulation group in terms of how the solution fit into the bigger concept of a probability distribution. Many of the low-level students in the control group used formulas that indicated a tendency to revert to deterministic calculations involving setting up a proportion, adding, or taking averages which were inappropriate formulas to use for the probabilistic problems they were given.

## Limitations of the Study

As the sole teacher-researcher involved in this study, this type of study naturally presents the possibility of researcher bias. With a preconceived notion that simulations were helpful, there was a risk that I may inadvertently favor the simulation group in my instruction. To the best of my ability, I tried to equalize the instruction to both groups as much as possible except for the simulations themselves so as not to confound the treatment effects. Both groups received the same notes, and as the instructor for all of the classes, this eliminated the possibility of confounding effects from the variability of different instructors. Although I was involved in discussions with both groups, I purposefully strived to allow the students to develop their own ideas. The choice as teacher-researcher presented some advantages as the students were in their natural domain with no disruption from outside researchers conducting the study which may detract from the natural state of the classroom environment. As their regular teacher, there was a preexisting intimacy and comfort already set in place. In addition, there were two AP Statistics teachers at my school who served as observer-witnesses for some of the lab sessions. These teachers recorded field notes of their observations as well so that comparisons could be made and triangulation could be achieved.

I believe the assessments used in this study could have been improved to reduce ambiguity of questions and to reduce the possibility of getting a question correct but with incorrect reasoning. The incorporation of written feedback and multiple choices that probed reasoning served well to reduce this risk, and the interviews allowed me to probe further into the nature of the students' reasoning to offset the objective assessment drawbacks. Researchers in statistics education have noted the assessment challenge in monitoring appropriate statistical reasoning and the need to develop better instruments (Garfield, 2002). All of my questions used in the pretest, posttest, and retention tests were either matched items or the same items to assure equality of assessments thus increasing the chance of confounding from learning effects from prior assessments. On the other hand, this seemed a better choice than using different assessments and running the risk of not testing the same objectives.

The interviews were unstructured, informal interviews that prompted students to think out loud so I could monitor the students' reasoning. Due to time constraints in the students' and my schedule, I was unable to conduct interviews after each lab. More frequent, structured interviews would have possibly produced a deeper look into how the simulations affected the evolution of the students' probabilistic reasoning.

For the most part, I believe the labs that I designed were beneficial and fell within the students' zones of proximal development to allow for learning to take place. A few of the labs such as the independence lab, the conditional probability lab, and the representativeness lab were either poorly designed or turned out to be too time-consuming for any true learning benefits to be detected within the allotted timeframe. I believe that simulations in these areas can be helpful, but based on the limitations from the instructional design of this study, these benefits could not be substantiated by the results.

## Implications for Teachers

Common Core State Standards (CCSSI, 2011) continue to push the influx of data analysis and probability within the secondary curriculum. Many teachers currently teaching these courses are unfamiliar with the teaching and learning of probabilistic and statistical thinking. These teachers are teaching courses such as algebra, geometry and precalculus and are familiar with a deterministic approach to teaching math. As a result, they may choose a formulaic approach to teaching the data analysis segments within their courses rather than an approach that emphasizes simulations and exploratory methods. This study generates ideas for teachers on how they can design simulations that foster the development of new representational systems to encourage a deeper understanding of probability and probability distributions. The study enlightens awareness to the different, diverse performance levels within our classroom but with a viable option as to how to address these differentiated levels in instructional design. My earlier pilot study and previous teaching experience with simulations indicates that students are typically uncomfortable with learning through exploration and simulation at first. The results of my study give confirming evidence that students will most likely come to adjust, and even like, the experience and increase their confidence once they gain familiarity with the approach.

## Implications for Students

The implication of this study for students is one of liberation. Qualitative data analysis revealed a renewed interest, enjoyment, and confidence associated with the study of probability due to the use of the simulations. Although participants in the study were in an AP Statistics course, the less advanced students showed the most benefit from the simulations. This is especially significant as more and more non-advanced students are being exposed to data analysis, probability, and statistics in the secondary curriculum. Evidence from the study supports that shallow instruction in data analysis and probability with no sense of the concept of variation or distribution produces only an isolated, temporary understanding for many students. It is with hope and optimism that I say with new standards focused on modeling activities and the implementation of technology that students will begin to gain an understanding of complex concepts which they may not have been capable of learning in the traditional school mathematics curriculum and to develop a positive self-concept of their mathematical abilities.

## Implications for Further Research

Although instinctively I knew that using data exploration and simulations would benefit my students, I was dubious that my results would be statistically significant. Prior to the analysis of the data, I worried that the results would be disappointingly unconvincing as other noted attempts using simulations in the literature. The results of this study have been convincing. Researchers should be encouraged to conduct whole-class experiments to infuse the literature with undisputable evidence of the benefits of incorporating simulations in the teaching of statistics. The study highlighted successful simulations as well as ones that did not seem successful. Time constraints hinder the development of successful simulations that can be utilized in a realistic classroom setting. Further studies in realistic settings are encouraged to devise workable and successful simulations within the typical constraints. The results of my study indicate that simulations can both help and hinder students' understanding of probability distributions. Simulations cannot be an isolated task but rather need to be a sequence of data-oriented lessons, pedagogically connected, so as to allow for the conceptual changes to take place. There is a need for research on suitable curriculum materials. Based on the results of the study, I suggest that simulations should be taught in conjunction with the formulas that apply but need to be well-developed enough to allow the students to reconcile their results with the outcomes from the formulas so a deeper understanding can occur. I believe the computer simulations allow for a more connected understanding of the construct of a probability distribution from generation of outcomes to development of a random variable to the graphical representation, but caution must be taken to assure the students understand what the graphs represent and the source of the representation. I suggest that in designing simulation activities relative to probability distributions researchers start with the notion of outcomes and sample space as the technology allows the students to develop representations that instill visualizations for the students when making necessary connections between the concepts.

This study has provided both quantitative and qualitative evidence of significant differences in secondary students' reasoning when using simulations to learn about probability and probability distributions. Results indicate that lower-level students may benefit more from the use of simulations than higher-level students perhaps due to higher-level students' preference for formal mathematical rules and procedures. Future researchers who are looking at the effect of simulations on student learning should note this differential when designing their studies to avoid masking significant differences with aggregate data. Both groups continued to show difficulty with compound events, conditional probability, and independence. Further explorations with simulations in these areas are necessary with acknowledgement that these simulations will require a greater time commitment than others. Difficulties with compound events, conditional probability, and independence have been noted in the literature (Batanero & Sanchez, 2005; Zimmermann, 2002; Jones et al., 1999). Frameworks such as GAISE (ASA, 2007) and the probabilistic thinking framework of Jones and colleagues (Jones et al., 1999; Tarr and Jones, 1997) are useful for detecting and monitoring conceptual development as students progress from one level to another in their reasoning. Efforts to improve assessments for statistical reasoning should continue to ensure that appropriate learning has taken place. Finally, similar studies should be conducted that continue the progression noted in this study and monitor the reasoning from the simulations and this prerequisite knowledge on distributions with regard to sampling distributions and statistical inference (Chance et al., 2004).

The results of my study were exciting to me. I hope that these results will make a positive contribution to the limited, yet growing and developing, body of research in statistics education in the secondary setting. I believe simulations and a modeling approach to teaching mathematics show great promise in improving the learning and motivation of our students as it offers them a process-oriented approach to learning that makes their learning more successful and applicable to their lives.

# REFERENCES

American Statistical Association. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A PreK-12 curriculum framework.* Retrieved March 16, 2012, from http://www.amstat.org/education/gaise/GAISEPreK-12_Full.pdf

Bakker, A. & Gravemeijer, K.P.E. (2004). Learning to reason about distribution. In J.B. Garfield & D. Ben-Zvi (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking.* (pp. 147-167). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Batanero, C. & Sanchez, E. (2005). What is the nature of high school students' conceptions and misconceptions about probability? In G.A. Jones (Ed.), *Exploring Probability in School: Challenges for teaching and learning.* (pp. 241-266). Springer.

Becker, B.J. (1996). A look at the literature (and other resources) on teaching statistics. *Journal of Educational and Behavioral Statistics, 21,* 71–90.

Biggs, J.B., & Collis, K.F. (1991). Multimodal learning and the quality of intelligent behavior. In H.A.H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 57-76). Hillsdale, NJ: Erlbaum.

Borovcnik, M. & Peard, R. (1996). Probability. In A. Bishop et al. (Eds.), *International handbook of mathematics education* (Part 1, pp. 239-287). Dordrecht, The Netherlands: Kluwer.

Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking.* (pp. 295-323). Kluwer Academic Publishers.

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning 1*(1), 5-43.

Cobb, P. (2000). Conducting teaching experiments in collaboration with teachers. In A.E. Kelly & R.A. Lesh (Eds.), *Handbook of research design in mathematics and science education,* (pp. 307-334). Mahwah, NJ: Erlbaum.

Common Core State Standards Initiative. (2011). Common core state standards. Retrieved March 21, 2012, from https://www.corestandards.org

Davis, J. (2001). Conceptual Change. In M. Orey (Ed.), *Emerging perspectives on learning, teaching, and technology*. Retrieved January 29, 2012, from http://projects.coe.uga.edu/epltt/

delMas, R., Garfield, J., and Chance, B. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education* [Online], 7(3). Retrieved February 4, 2012, from www.amstat.org/publications/jse/secure/v7n3/delmas.cfm

delMas, R., Garfield, J., Ooms, A. & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal, 6*(2), 28-58.

Dewey, J. (1938/1997). *Experience and Education.* New York: Simon & Schuster.

Doerr, H., & Tinto, P. (2000). Paradigms for teacher-centered, classroom-based research. In A.E. Kelly & R.A. Lesh (Eds.), *Handbook of research design in mathematics and science education,* (pp. 403-427). Mahwah, NJ: Erlbaum.

Feldman, A., & Minstrell, J. (2000). Action research as a research methodology for the study of the teaching and learning of science. In A.E. Kelly & R.A. Lesh (Eds.) *Handbook of research design in mathematics and science education.* Mahwah, NJ: Lawrence Erlbaum Associates.

Fischbein, E. (1975). The intuitive sources of probabilistic thinking in children. Dordrecht, The Netherlands: Reidel.

Fischbein, E. (1999). Psychology and mathematics education. *Mathematical Thinking and Learning, 1*(1), 47-58.

Fischbein, E., & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? *Educational Studies in Mathematics, 15,* 1-24.

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education, 28*(1), 96–105.

Garfield, J. (1995). How students learn statistics. *International Statistical Review, 63*(1), 25-34.

Garfield, J. (1998). The statistical reasoning assessment: Development and validation of a research tool. *Proceedings of the Fifth International Conference on Teaching Statistics,* 781-786. Singapore: International Statistical Institute.

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education, 10*(3). Retrieved February 4, 2012, from www.amstat.org/publications/jse/v10n3/garfield.html

Garfield, J. & Gal, I. (1999). Teaching and assessing statistical reasoning. *NCTM 1999 Yearbook: Developing Mathematical Reasoning in Grades K-12,* 207-219.

Georgia Department of Education. (2011, August). Common Core Georgia Performance Standards. Retrieved March 21, 2012, from https://www.georgiastandards.org/Common-Core/Pages/Math.aspx

Green, D.R. (1990). A longitudinal study of pupils' probability concepts. *Proceedings of the Third International Conference on Teaching Statistics.*

Greene, J.C., Caracelli, V.J., & Graham, W.F. (1989). Toward a conceptual framework for mixed-method evaluation design. *Educational Evaluation and Policy Analysis, 11*(3), 255-274.

Horvath, J.K., & Lehrer, R. (1998). A model-based perspective on the development of children's understanding of chance and uncertainty. In S.P. LaJoie (Ed.), *Reflections on statistics: agendas for learning, teaching, and assessment in K-12* (pp. 121-148). Mahwah, NJ: Lawrence Erlbaum Associates.

International Statistical Literacy Project. Retrieved April 24, 2012, from http://www.stat.auckland.ac.nz/~iase/islp/hist

Jones, G.A., & Thornton, C.A. (2005). An overview of research into the teaching and learning of probability. In G.A. Jones (Ed.), *Exploring Probability in School: Challenges for teaching and learning.* (pp. 65-92). Springer.

Jones, G.A., Langrall, C.W., Thornton, C.A., & Mogill, A.T. (1999). Students' probabilistic thinking in instruction. *Journal for Research in Mathematics Education, 30*(5), 487–519.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Kaput, J. J., & Roschelle, J. (1997). Deepening the impact of technology beyond assistance with traditional formalisms in order to democratize access to ideas underlying calculus. In E. Pehkonen (Ed.), *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education* (pp. 105-112). Lahti, Finland: University of Helsinki.

Kaput, J. J., & Roschelle, J. (1998). The mathematics of change and variation from a millennial perspective: New content, new context. In C. Hoyles, C. Morgan, & G. Woodhouse (Eds.), *Rethinking the mathematics curriculum* (pp. 155-170). London: Springer-Verlag.

Kaput, J. & Schorr, R. (2002). Changing representational infrastructures changes most everything: The case of simcalc, algebra, and calculus. In G. Blume & K. Heid (Eds.), *Research on Technology in the Learning and Teaching of Mathematics: Synthesis and Perspectives.* Mahwah, NJ: Erlbaum.

Klein, E. E., & Herskovitz, P. J. (2005). Philosophical foundations of computer simulation validation. *Simulation & Gaming, 36*(3), 303-329.

Koehler, J. (1996). The base-rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19*, 1-53.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6*(1), 59–98.

Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics. *Journal of Statistics Education*, *3*(1).

Lecoutre, M.P. (1992). Cognitive models and problem spaces in “purely random” situations. *Educational Studies in Mathematics, 23*, 557-568.

Lincoln, Y.S., & Guba, E.G. (1985). *Naturalistic inquiry.* Beverly Hills, CA: Sage.

Liu, H. J. (1998). *A cross-cultural study of sex differences in statistical reasoning for college students in Taiwan and the United States.* Doctoral dissertation, University of Minnesota, Minneapolis.

Metz, M. L. (2010). Using GAISE and NCTM standards as frameworks for teaching probability and statistics to pre-service elementary and middle school mathematics teachers. *Journal of Statistics Education, 18*(3).

Mills, J.D. (2002). Using computer simulation methods to teach statistics: A review of the literature. *Journal of Statistics Education, 10*(1).

Moore, D. & Cobb, G. (2000). Statistics and mathematics: Tension and cooperation. *American Mathematical Monthly, 107*(7), 615-630.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA.

Nemirovsky, R., Kaput, J., & Roschelle, J. (1998). Enlarging mathematical activity from modeling phenomena to generating phenomena. Paper presented at the *Proceedings of the 22nd International Conference of the Psychology of Mathematics Education*. Stellenbosch, South Africa.

Patton, M.Q. (2002). *Qualitative Research & Evaluation Methods.* Thousand Oaks, CA: Sage Publications, Inc.

Piaget, J. & Inhelder, B. (1975). *The Origin of the Idea of Chance in Children.* New York: W.W. Norton & Company, *Inc*.

Pintrich, P.R., Marx, R.W., & Boyle, R.A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 6*, 167-199.

Polaki, M.V. (2002). Using instruction to identify key features of Basotho elementary students' growth in probabilistic thinking. *Mathematical Thinking and Learning, 4,* 285-314.

Polaki, M.V. (2005). Dealing with compound events. In G.A. Jones (Ed.), *Exploring Probability in School: Challenges for teaching and learning.* (pp. 191-214). Springer.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Towards a theory of conceptual change. *Science Education. 66*(2), 211-227.

Pratt, D., & Noss, R. (2002). The microevolution of mathematical knowledge: The case of randomness. *The Journal of the Learning Sciences, 11*(4), 453-488.

Prodromou, T. (2007a). Making connections between the two perspectives on distribution. *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education.* Larnaca, Cyprus.

Prodromou, T. (2007b). The coordination of distributional thinking. Paper presented at the *56th session of the International Statistical Institute.* Lisbon, Portugal.

Prodromou, T., & Pratt, D. (2006). The role of causality in the co-ordination of the two perspectives on distribution within a virtual simulation. *Statistics Education Research Journal, 5*(2), 69-88.

Rubel, L. (2007). Middle school and high school students' probabilistic reasoning on coin tasks. *Journal for Research in Mathematics Education, 38*(5), 531-556.

Scheaffer, R.L., Watkins, A.E., & Landwehr, J.M. (1998). What every high school graduate should know about statistics. In S.P. LaJoie (Ed.), *Reflections on statistics: agendas for learning, teaching, and assessment in K-12* (pp. 3-31). Mahwah, NJ: Lawrence Erlbaum Associates.

Schoenfeld, A.H. (2002). A highly interactive discourse structure. In J. Brophy (Ed.), *Social constructivist teaching: Its affordances and constraints* (Vol. 9, pp. 131-170). Amsterdam: JAI Press.

Shaughnessy, J.M. (2006). Research on students' understanding of some big ideas in statistics. In G. Burrill (Ed.), *NCTM 2006 Yearbook: Thinking and reasoning with data and chance* (pp. 77-98). Reston, VA: NCTM.

Shaughnessy, J.M., & Bergman, B. (1993). Thinking about uncertainty: Probability and statistics. In P.S. Wilson (Ed.), *Research ideas for the classroom: High school mathematics* (pp. 177–197). New York: Macmillan.

Sternstein, M. (2000). *Barron's How To Prepare for the AP Statistics Advanced Placement Examination.* New York: Barron's Educational Series, Inc.

Tang, E. & Ginsburg, H. (1999). Young children’s mathematical reasoning: A psychological view. In L. Stiff & F. Curio (Eds.), *Developing Mathematical Reasoning in Grades K-12.* (pp. 45-61). NCTM: Reston, Virginia.

Tarr, J.E. & Jones, G.A. (1997). A framework for assessing middle school students' thinking in conditional probability and independence. *Mathematics Education Research Journal, 9*(1), 39-59.

Tarr, J.E., & Lannin, J.K. (2005). How can teachers build notions of conditional probability and independence? In G.A. Jones (Ed.), *Exploring Probability in Schools: Challenges for teaching and learning.* (pp. 215-238). Springer.

Truran, J. (2001). Postscript: Researching stochastic understanding – the place of a developing research field in PME. *Educational Studies in Mathematics, 45,* 9-13.

Tyson, L.M., Venville, G.J., Harrison, A.G., & Treagust, D.F. (1997). A multidimensional framework for interpreting conceptual change events in the classroom. *Science Education, 81*, 387-404.

Vere-Jones, D. (1995). The coming of age of statistical education. *International Statistical Review, 63*(1), 3–23.

Vygotsky, L. (1978). Interaction between learning and development. From: *Mind and Society.* Reprinted in M. Gauvain & M. Cole (Eds.) *Readings in the Development of Children* (2nd. ed.). W.H. Freeman and Company: New York.

Wilensky, U. (1995). Paradox, programming, and learning probability: A case study in a connected mathematics framework. *Journal of Mathematical Behavior, 14,* 253–280.

Yates, D., Moore, D., & McCabe, G. (1999). *The Practice of Statistics.* New York: W.H. Freeman and Company.

Zimmerman, G.M. (2002). *Students' reasoning about probability simulation during instruction.* Unpublished doctoral dissertation, Illinois State University.

APPENDIX A - PILOT STUDY QUESTIONNAIRE AND LAB SHEET

**AP Statistics** **Name ________________________**

**Concepts of Randomness**

For each of the following questions, circle the answer that you consider most appropriate. After each question, summarize your thoughts on each question and state why you chose the particular answer you circled. If you guessed, then state that. If you were able to narrow down your choices, then give details regarding how you were able to narrow them down. If you knew for sure which answer was correct, then state how you knew. Finally, for each question, circle a confidence scale number from 1 to 5 (with 1 = not confident at all to 5 = very confident) regarding the confidence you have in your circled answer.

1. Five faces of a fair die are painted black, and one face is painted white. The die is rolled six times. Which of the following results is more likely?
    (a) Black side up on five of the rolls; white side up on the other roll
    (b) Black side up on all six rolls
    (c) Choices a and b are equally likely

| Summary for Question 1: |
|---|

Confidence rating for Question 1: 1 2 3 4 5
(not confident at all) (very confident)

2. Half of all newborn babies are girls and half are boys. Hospital A records an average of 50 births a day. Hospital B records an average of 10 births a day. On a particular day, which hospital is more likely to record 80% or more female births?
   (a) Hospital A (with 50 births a day)
   (b) Hospital B (with 10 births a day)
   (c) The two hospitals are equally likely to record such an event

| Summary for Question 2: |
|---|
| |

Confidence rating for Question 2: 1 2 3 4 5
(not confident at all) (very confident)

3. When two dice are thrown simultaneously, it is possible that one of the following results may occur. **Result 1**: A 5 and a 6 are obtained. **Result 2**: A 5 is obtained twice. Select the response that you agree with the most.
   (a) The chances of obtaining each of these results is equal.
   (b) There is a higher chance of obtaining **Result 1**
   (c) There is a higher chance of obtaining **Result 2**
   (d) It is impossible to give an answer

| Summary for Question 3: |
|---|
| |

Confidence rating for Question 3: 1 2 3 4 5
(not confident at all) (very confident)

4. When tossing a fair coin, there are two possible outcomes: heads or tails. Ronni flipped a coin three times, and in all cases heads came up. Ronni intends to flip the coin again. What is the chance of getting heads the fourth time?
    (a) smaller than the chance of getting tails
    (b) equal to the chance of getting tails
    (c) greater than the chance of getting tails

| Summary for Question 4: |
|---|
| |

Confidence rating for Question 4: 1 2 3 4 5
(not confident at all) (very confident)

**Simulation Activity – Question 1**

**Name**: ________________

Group Members:

________________________

________________________

________________________

**Exercise #1**: Five faces of a fair die are painted black, and one face is painted white. The die is rolled six times. Which of the following results is more likely?

(a) Black side up on five of the rolls; white side up on the other roll
(b) Black side up on all six rolls
(c) Choices a and b are equally likely

**Materials:**
One fair die with five faces painted black and one face painted white

**Have each student in your group repeat this experiment 20 times. Each set of 6 tosses is considered one repetition. Record the results of each roll on the attached spreadsheet.**

(a) Based on your individual simulation results, which do you think is more likely, a or b, or do you think they are equally likely? Why do you think that?

(b) Does this answer agree with your answer prior to doing the simulation activity?

(c) Is there anything that the simulation helped you realize that you didn't think of when you answered the question prior to the simulation? Would that lead to a change in your answer to the question? Why would it lead to a change in your answer? (verify what you think the answer to the question is now)

(d) Now compare your results to the results of your other group members? Did you all get similar results? If so, why do you think that happened? If not, why do you think that happened? (record each member's results on attached spreadsheet)

(e) Pool your results with all of the other group members, and answer the question based on your combined results. Did the combined results cause you to change your answer from part c? If so, why? If not, why not? (verify what you think the answer to the question is now)

(f) If you pool your group results with the whole class, what do you think will happen? Why do you think this will happen?

(g) After looking at the other group results, how do the different group results compare to each other? Are the results similar to the comparison of your individual results with your other group members? (See step e). Record results on attached spreadsheet.

(h) After comparing group results, do you want to change your answer to the question from what you put in part e? (verify what you think the answer to the question is now)

(i) On a scale from 1 to 5 with 1 = not confident at all to 5 = very confident, how confident are you in your answer to part h? Why are you or are you not confident with this answer?

(j) Do you think the answer would depend on the person rolling the die? Explain.

(k) Summarize what you have learned from this simulation activity. Be as specific as possible.

| | 1st toss | 2nd toss | 3rd toss | 4th toss | 5th toss | 6th toss |
|---|---|---|---|---|---|---|
| example | **B** | **W** | **B** | **B** | **W** | **B** |
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | | | | | |
| 4 | | | | | | |
| 5 | | | | | | |
| 6 | | | | | | |
| 7 | | | | | | |
| 8 | | | | | | |
| 9 | | | | | | |
| 10 | | | | | | |
| 11 | | | | | | |
| 12 | | | | | | |
| 13 | | | | | | |
| 14 | | | | | | |
| 15 | | | | | | |
| 16 | | | | | | |
| 17 | | | | | | |
| 18 | | | | | | |
| 19 | | | | | | |
| 20 | | | | | | |

| | % of choice A | % of choice B |
|---|---|---|
| | | |
| **Individual Result** | | |
| | | |
| **Group Member 1** | | |
| **Group Member 2** | | |
| **Group Member 3** | | |
| | | |
| **Group A Results** | | |
| **Group B Results** | | |
| **Group C Results** | | |
| **Group D Results** | | |
| **Group E Results** | | |

## APPENDIX B - INSTRUMENT A PRETEST

INSTRUMENT A

For each question, record your answer on the answer sheet provided. Then for each question, give a rating on the appropriateness of the question as a measure of secondary student understanding of probability distributions (1 = LOWEST to 5 = HIGHEST). So, for example, if you feel the question is **not** appropriate at all, you would give it a low rating. If possible, then give written feedback on the appropriateness and difficulty of the items. You may write your rating and feedback on the actual test rather than on the answer sheet.

**1.** Joe and Tonya plan to visit a bookstore. Based on their previous visits to this bookstore, the probability distributions of the number of books they will buy are given below:

| # of books Joe will buy | 0 | 1 | 2 |
|---|---|---|---|
| Probability | 0.50 | 0.25 | 0.25 |

| # of books Tonya will buy | 0 | 1 | 2 |
|---|---|---|---|
| Probability | 0.25 | 0.50 | 0.25 |

Assume that Joe and Tonya make their decisions independently, what is the probability that they will purchase no books on this visit to the bookstore?

(a) 0.0625
(b) 0.1250
(c) 0.1875
(d) 0.2500
(e) 0.7500

**Confidence Scale (circle one): 1 2 3 4 5**

**Written Feedback:**

**2.** Using the same scenario in Question #1, how many outcomes in the sample space are there?

(a) 3
(b) 5
(c) 6
(d) 9

**Confidence Scale (circle one): 1 2 3 4 5**

**Written Feedback:**

**3.** Using the same scenario in Question #1 and letting the random variable represent the total number of books Joe and Tonya buy together, what shape does the probability distribution have?

(a) uniform
(b) right-skewed
(c) left-skewed
(d) mound-shaped

**Confidence Scale (circle one): 1 2 3 4 5**

**Written Feedback:**

**4.** Circuit boards are assembled by selecting 4 computer chips at random from a large batch of chips. In this batch of chips, 90 percent of the chips are acceptable. Let X denote the number of acceptable chips out of a sample of 4 chips from this batch. What is the least probable value of X?

(a) 0
(b) 1
(c) 2
(d) 3
(e) 4

**Confidence Scale (circle one): 1 2 3 4 5**

**Written Feedback:**

**5.** Assuming you randomly guess on a 20 question multiple-choice test with 5 choices per question, what shape does the distribution of the number of questions correct have?

(a) left-skewed
(b) right-skewed
(c) uniform
(d) symmetric, mound-shaped

**Confidence Scale (circle one): 1 2 3 4 5**

**Written Feedback:**

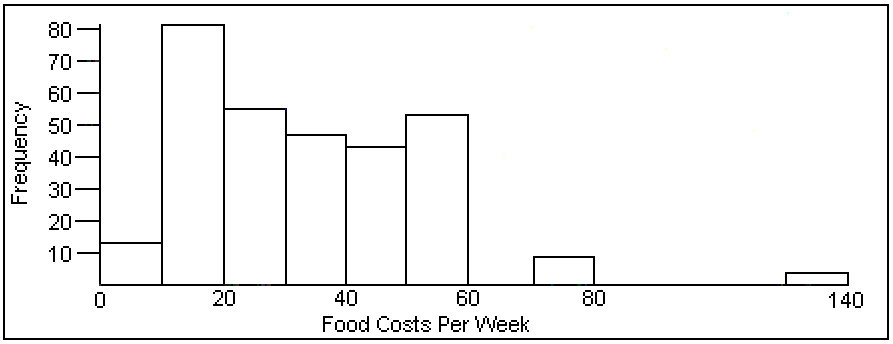
**6.** Five faces of a fair die are painted black, and one face is painted white. The die is rolled six times. Which of the following results is more likely?

(a) Black side up on five of the rolls; white side up on the other roll
(b) Black side up on all six rolls
(c) a and b are equally likely

**Confidence Scale (circle one): 1 2 3 4 5**

**Written Feedback:**

**7.** This is a distribution of how much money was spent per week for a random sample of college students. The following statistics were calculated: mean = $31.52; median = $30.00; interquartile range = $34.00; standard deviation = $21.60; range = $132.50.




(question continued on next page)

Given the statement: **The distribution of food costs basically looks bell-shaped, with one outlier.** Do you:

(a) Agree, it looks pretty symmetric if you ignore the outlier.
(b) Agree, most distributions are bell-shaped.
(c) Disagree, it looks more skewed to the left.
(d) Disagree, it looks more skewed to the right.

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**8.** A fair coin is tossed, and it lands heads up. The coin is to be tossed a second time. What is the probability that the second toss will also be a head?

(a) 1/4
(b) 1/2
(c) 1/3
(d) Slightly less than 1/2
(e) Slightly more than 1/2

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**9.** Which of the following best describes the reason for your answer to the preceding question?

(a) The second toss is less likely to be heads because the first toss was heads.
(b) There are four possible outcomes when you toss a coin twice. Getting two heads is only one of them.
(c) The chance of getting heads or tails on any one toss is always 1/2.
(d) There are three possible outcomes when you toss a coin twice. Getting two heads is only one of them.

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**10.** If a fair coin is tossed five times, which of the following ordered sequence of heads (H) and tails (T), if any, is MOST LIKELY to occur?

(a) H T H T T
(b) T H H H H
(c) H T H T H
(d) Sequences (a) and (c) are equally likely.
(e) All of the above sequences are equally likely.

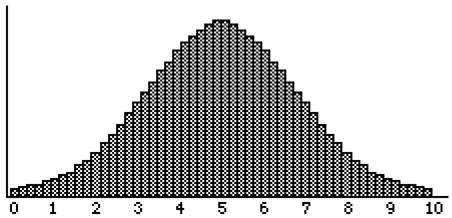
**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**11.** When three fair dice are simultaneously thrown, which of the following results is MOST LIKELY to be obtained?

(a) Result 1: A 5, a 3 and a 6 in any order
(b) Result 2: Three 5's
(c) Result 3: Two 5's and a 3
(d) All three results are equally likely.

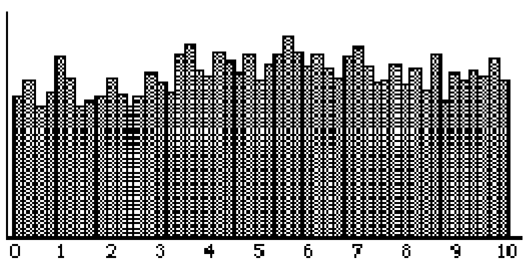
**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**12.** Two distributions of test scores (questions 12 and 13) are presented below. For each distribution, select the one descriptor that best represents the shape of the distribution.




(a) Normal
(b) Skewed
(c) Bimodal
(d) Uniform

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**




**13.** .

(a) Normal
(b) Skewed
(c) Bimodal
(d) Uniform

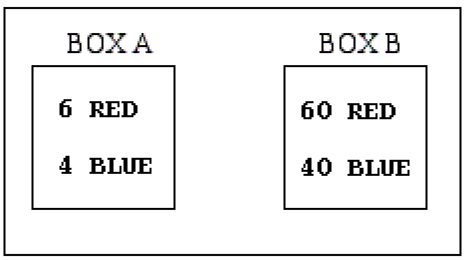
**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**14.** If a fair die is rolled five times, which of the following ordered sequence of results, if any, is MOST LIKELY to occur?

(a) 3 5 1 6 2
(b) 4 2 6 1 5
(c) 5 2 2 2 2
(d) Sequences (a) and (b) are equally likely.
(e) All of the above sequences are equally likely.

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**15.** Box A and Box B are filled with red and blue marbles as follows. Each box is shaken. You want to get a blue marble, but you are only allowed to pick out one marble without looking.




Which box should you choose?

(a) Box A (with 6 red and 4 blue)
(b) Box B (with 60 red and 40 blue)
(c) It does not matter

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**16.** Half of all newborn children are girls and half are boys. Hospital A records an average of 50 births a day. Hospital B records an average of 10 births a day. On a particular day, which hospital is more likely to record 80% or more female births?

(a) Hospital A (with 50 births a day)
(b) Hospital B (with 10 births a day)
(c) The two hospitals are equally likely to record such an event

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**17**. Suppose a car manufacturer makes cars with a 0.60 probability of having no defects, a 0.30 probability of having one defect, and a 0.10 probability of having 2 defects. If two cars were shipped to a dealer, what is the most probable average number of defects of the two cars?

(a) 0
(b) 0.5
(c) 1
(d) 1.5
(e) 2

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**18.** A basketball player makes 80% of his free throws. If this player is asked to shoot free throws until he misses one, which of the following is the MOST PROBABLE number of throws it will take?

(a) 0
(b) 1
(c) 2
(d) 3
(e) 4

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**19.** True or False: A weatherman predicts 70% chance of rain for the next five days. This means the most likely outcome is rain for the next five days.

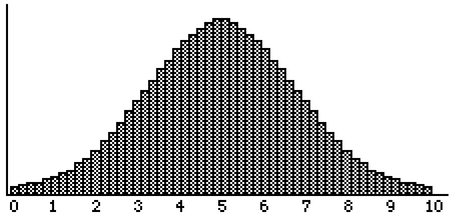
(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**20**. True or False: A basketball player has a 60% field goal percentage. If he has made four shots in a row in a game, the likelihood of him making his next shot is higher than at the start of the game.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**21.** 


True or False: Given the distribution above, it is equally likely to get any one unit interval between 0 and 10 (ex. 0-1, 1-2, 2-3, etc. are all equally likely).

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**22.** Joe and Tonya plan to visit a bookstore. Based on their previous visits to this bookstore, the probability distributions of the number of books they will buy are given below:

| # of books Joe will buy | 0 | 1 | 2 |
|---|---|---|---|
| Probability | 0.50 | 0.25 | 0.25 |

| # of books Tonya will buy | 0 | 1 | 2 |
|---|---|---|---|
| Probability | 0.25 | 0.50 | 0.25 |

True or False: Letting the random variable represent the total number of books Joe and Tonya buy together, it is equally likely to get 0 books or 2 books total.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**23.** True or False: A weatherman predicts 60% chance of rain for the next five days. This means that it should rain exactly 3 out of the 5 days.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**24.** True or False: Consider a standard 52-card deck and suppose you pull one card from the deck. The probability that it is a jack, given that it is a heart is the same as the probability that it is a heart, given that it is a jack.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**25.** True or False: The probability of getting a prize in a cereal box is 0.10. The chance of getting your first prize in the tenth box you buy is higher than the chance of getting your first prize in the first box you buy.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**26**. True or False: Every morning when you arrive at school, you are either late or not late. Thus, the probability of arriving late or not late is 50%-50%.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**27**. The weather report on the news said there was an 80% chance of rain tomorrow. This means it will rain tomorrow.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**28**. The weather report said there was an 80% chance of rain tomorrow. The next day it did not rain. The weather report was incorrect.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**29**. You snuck out of your house the last 4 weekends in a row and did not get caught. This means you are bound to get caught this weekend.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**30**. A woman is bright, single, 31 years old, outspoken, and concerned with issues of social justice. Which of the following is more likely?

(a) She is a bank teller and a feminist
(b) She is a bank teller
(c) They are equally likely

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

APPENDIX C - INSTRUMENT B POSTTEST

**Name: ______________________________ Date: _______________**

**INSTRUMENT B**

For each question, record your answer on the answer sheet provided. Then for each question, give a rating on your CONFIDENCE LEVEL for each answer (1 = LEAST CONFIDENT to 5 = MOST CONFIDENT). Then give a written explanation detailing your thought process as to how you arrived at your answer. You may write your rating and feedback on the actual test rather than on the answer sheet.

Steven and Beth plan to visit a record store. Their frequent trips to the store result in the following probability distribution of the number of compact discs they buy.

| Number of CDs Steven will buy | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Probability | 0.20 | 0.25 | 0.40 | 0.15 |

| Number of CDs Beth will buy | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Probability | 0.25 | 0.30 | 0.25 | 0.20 |

**1.** Consider the sample space for the number of CDs Steven will buy followed by the number of CDs Beth will buy. How many outcomes are there in the sample space?

(a) 4
(b) 8
(c) 12
(d) 16

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**2**. Using the same information given above, assume Steven and Beth make their decisions independently. What is the probability that Beth and Steven together will buy exactly two CDs in total?

(a) 0.10
(b) 0.125
(c) 0.200
(d) 0.225

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**3**. Using the same scenario above and letting the random variable represent the total number of CDs Steven and Beth buy together, what shape does the probability distribution have?

(a) uniform
(b) right-skewed
(c) left-skewed
(d) mound-shaped

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**4.** Robots are assembled by selecting 3 computer chips at random from a large batch of chips. In this batch of chips, 90 percent of the chips are acceptable. Let X denote the number of acceptable chips out of a sample of 3 chips from this batch. What is the least probable value of X?

(a) 0
(b) 1
(c) 2
(d) 3

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**5.** Assuming a class of 20 randomly guesses on a 10 question multiple-choice test with 5 choices per question, what shape does the distribution of the number of questions correct most likely have?

(a) left-skewed
(b) right-skewed
(c) uniform
(d) symmetric, mound-shaped

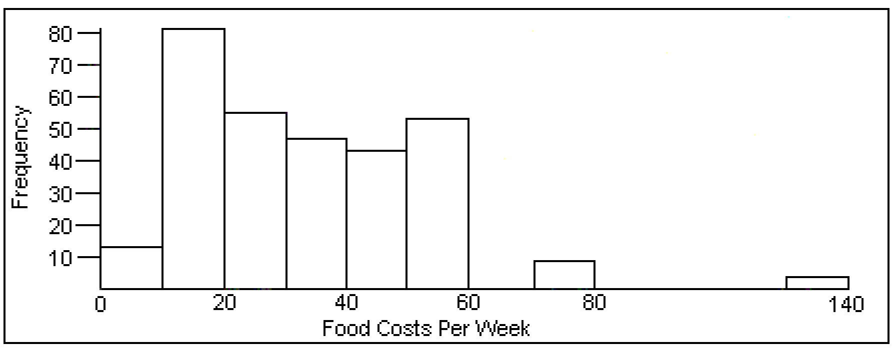
**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**6.** Five faces of a fair die are painted white, and one face is painted black. The die is rolled six times. Which of the following results is more likely?

(a) White side up on five of the rolls; black side up on the other roll
(b) White side up on all six rolls
(c) a and b are equally likely

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**7.** This is a distribution of how much money was spent per week for a random sample of college students. The following statistics were calculated: mean = \$31.52; median = \$30.00; interquartile range = \$34.00; standard deviation = \$21.60; range = \$132.50.




Given the statement: **The distribution of food costs basically looks bell-shaped, with one outlier.** Do you:

(a) Agree, it looks pretty symmetric if you ignore the outlier.
(b) Agree, most distributions are bell-shaped.
(c) Disagree, it looks more skewed to the left.
(d) Disagree, it looks more skewed to the right.

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**8.** A fair coin is tossed, and it lands tails up. The coin is to be tossed a second time. What is the probability that the second toss will also be a tail?

(a) 1/4
(b) 1/2
(c) 1/3
(d) Slightly less than 1/2
(e) Slightly more than 1/2

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**9.** Which of the following best describes the reason for your answer to the preceding question?

(a) The second toss is less likely to be tails because the first toss was tails.
(b) There are four possible outcomes when you toss a coin twice. Getting two tails is only one of them.
(c) The chance of getting heads or tails on any one toss is always 1/2.
(d) There are three possible outcomes when you toss a coin twice. Getting two tails is only one of them.

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**10.** If a fair coin is tossed five times, which of the following ordered sequence of heads (H) and tails (T), if any, is MOST LIKELY to occur?

(a) T H T H T
(b) T H T H H
(c) H T T T T
(d) Sequences (a) and (c) are equally likely.
(e) All of the above sequences are equally likely.

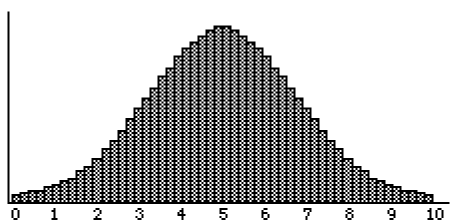
**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**11.** When three fair dice are simultaneously thrown, which of the following results is MOST LIKELY to be obtained?

(a) Result 1: Three 2's
(b) Result 2: A 3, a 6 and a 4 in any order
(c) Result 3: Two 1's and a 5
(d) All three results are equally likely.

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

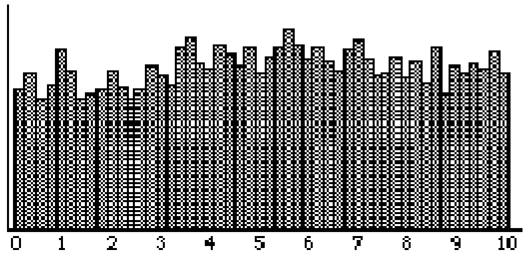
**12.** Two distributions of test scores (questions 12 and 13) are presented below. For each distribution, select the one descriptor that best represents the shape of the distribution.




(a) Normal
(b) Skewed
(c) Bimodal
(d) Uniform

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**13.** .




(a) Normal
(b) Skewed
(c) Bimodal
(d) Uniform

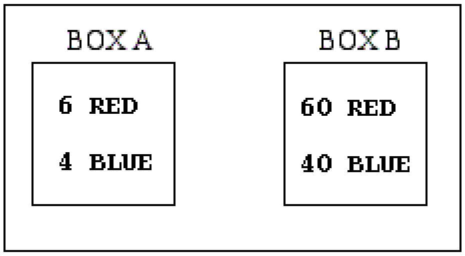
**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**14.** If a fair die is rolled five times, which of the following ordered sequence of results, if any, is MOST LIKELY to occur?

(a) 3 6 6 6 6
(b) 5 3 1 6 4
(c) 2 1 3 4 6
(d) Sequences (a) and (b) are equally likely.
(e) All of the above sequences are equally likely.

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**15.** Box A and Box B are filled with red and blue marbles as follows. Each box is shaken. You want to get a blue marble, but you are only allowed to pick out one marble without looking.




Which box should you choose?

(a) Box A (with 6 red and 4 blue)
(b) Box B (with 60 red and 40 blue)
(c) It does not matter

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**16.** In Game A, you toss a coin 60 times. In Game B, you toss a coin 8 times. You win the game if you toss at least 70% or more heads. If you play the game once, which game are you more likely to win?

(a) Game A
(b) Game B
(c) It is equally likely to win either game

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**17**. Suppose a car manufacturer makes cars with a 0.70 probability of having no defects, a 0.20 probability of having one defect, and a 0.10 probability of having 2 defects. If two cars were shipped to a dealer, what is the most probable average number of defects of the two cars?

(a) 0
(b) 0.5
(c) 1
(d) 1.5
(e) 2

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**18.** Assume it is equally likely to have a boy or girl. A couple decides to have children until they have a boy. Which of the following is the MOST PROBABLE number of children they will have?

(a) 1
(b) 2
(c) 3
(d) 4

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**19.** True or False: A weatherman predicts 70% chance of rain each day for the next four days. This means the most likely outcome is rain every day for the next four days. Assume independence.

(a) True
(b) False

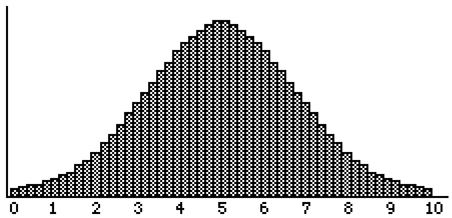
**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**20**. True or False: A basketball player has a 75% field goal percentage. If she has made three shots in a row in a game, the likelihood of her making her next shot is higher than at the start of the game.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**21.**




True or False: Given the distribution above, it is equally likely to get any one unit interval between 0 and 10 (ex. 0-1, 1-2, 2-3, etc. are all equally likely).

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**22**. Steven and Beth plan to visit a record store. Their frequent trips to the store result in the following probability distribution of the number of compact discs they buy.

| Number of CDs Steven will buy | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Probability | 0.20 | 0.25 | 0.40 | 0.15 |

| Number of CDs Beth will buy | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Probability | 0.25 | 0.30 | 0.25 | 0.20 |

True or False: Letting the random variable represent the total number of CDs that Steven and Beth buy together and assuming independence, it is less likely for them to buy 2 CDs total than 4 CDs total.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**23**. True or False: A weatherman predicts 80% chance of rain for the next five days. This means that it will rain exactly 4 out of the 5 days.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**24**. True or False: Consider a standard 52-card deck and suppose you pull one card from the deck. The probability that it is a king, given that it is a face card is the same as the probability that it is a face card, given that it is a king.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**25**. True or False: The probability of getting a prize in a cereal box is 0.20 or 1/5. The chance of getting your first prize in the fifth box you buy is higher than the chance of getting your first prize in the first box you buy.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**26.** True or False: Every morning when you arrive at school, you are either late or not late. Thus, the probability of arriving late or not late is 50%-50%.

(a) True
(b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**27.** The weather report on the news said there was a 90% chance of rain tomorrow. This means it will rain tomorrow.
  (a) True
  (b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**28.** The weather report said there was a 90% chance of rain tomorrow. The next day it did not rain. The weather report was incorrect.
  (a) True
  (b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**29.** You sped to work the last four days and did not get caught. This means you are more likely to get caught today.
  (a) True
  (b) False

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**30.** You roll a die twice and get a sum of 7. Which of the following is more likely?
  (a) You roll a 4 on the first roll and a 3 on the second one.
  (b) You roll a 4 on the first roll.
  (c) They are equally likely

**Confidence Scale (circle one): 1 2 3 4 5**
**Written Feedback:**

**Explain, in your own words, your understanding of the following. Give detailed explanations.**

Law of large numbers –

Randomness –

Expected value –

Probability distribution –

How do you feel about the concept of probability? Do you understand it? Do you like it? Explain.

If you are in the simulation group, do you trust the simulations from the random number generators (calculator, computer, random number table, rolling dice, etc.)? Do you think they can be used to model real life? EXPLAIN.

If you are in the simulation group, how do you feel about simulations? Do you feel they helped you in your learning of probability and randomness, and, if so, how? What were some positives and negatives about doing the simulations?

APPENDIX D - INSTRUMENT C RETENTION TEST

**Name: ______________________________ Date: _______________**

**INSTRUMENT C**

For each question, record your answer on the answer sheet provided.

Jacob and Emily plan to visit a candy store. Their frequent trips to the store result in the following probability distribution of the number of pieces of candy they buy.

| Number of Candies Jacob will buy | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Probability | 0.1 | 0.1 | 0.2 | 0.3 | 0.3 |

| Number of Candies Emily will buy | 0 | 1 | 2 |
|---|---|---|---|
| Probability | 0.2 | 0.3 | 0.5 |

**1**. Consider the sample space for the number of candies Jacob will buy followed by the number of candies Emily will buy. How many outcomes are there in the sample space?

(a) 5
(b) 8
(c) 12
(d) 15

**2**. Using the same information given above, assume Jacob and Emily make their decisions independently. What is the probability that Jacob and Emily together will buy exactly two candies in total?

(a) 0.03
(b) 0.04
(c) 0.05
(d) 0.12

**3**. Using the same scenario above and letting the random variable represent the total number of candies Jacob and Emily buy together, what shape does the probability distribution have?

(a) uniform
(b) right-skewed
(c) left-skewed
(d) mound-shaped

**4.** Robots are assembled by selecting 3 computer chips at random from a large batch of chips. In this batch of chips, 90 percent of the chips are acceptable. Let X denote the number of acceptable chips out of a sample of 3 chips from this batch. What is the least probable value of X?

(a) 0
(b) 1
(c) 2
(d) 3

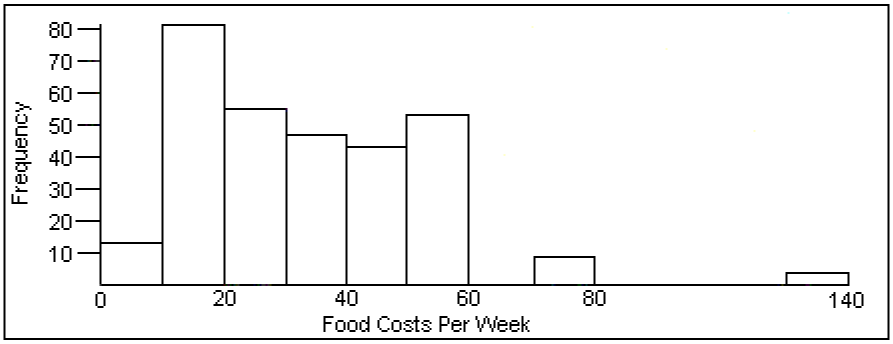
**5.** Assuming a class of 20 randomly guesses on a 10 question multiple-choice test with 5 choices per question, what shape does the distribution of the number of questions correct most likely have?

(a) left-skewed
(b) right-skewed
(c) uniform
(d) symmetric, mound-shaped

**6.** Five faces of a fair die are painted white, and one face is painted black. The die is rolled six times. Which of the following results is more likely?

(a) White side up on five of the rolls; black side up on the other roll
(b) White side up on all six rolls
(c) a and b are equally likely

**7.** This is a distribution of how much money was spent per week for a random sample of college students. The following statistics were calculated: mean = $31.52; median = $30.00; interquartile range = $34.00; standard deviation = $21.60; range = $132.50.




Given the statement: **The distribution of food costs basically looks bell-shaped, with one outlier.** Do you:

(a) Agree, it looks pretty symmetric if you ignore the outlier.
(b) Agree, most distributions are bell-shaped.
(c) Disagree, it looks more skewed to the left.
(d) Disagree, it looks more skewed to the right.

**8.** A die is rolled, and an odd number comes up. The die is to be rolled a second time. What is the probability that the second toss will also be an odd number?

(a) 1/4
(b) 1/3
(c) 1/2
(d) Slightly less than 1/2
(e) Slightly more than 1/2

**9.** Which of the following best describes the reason for your answer to the preceding question?

(a) The chance of getting an odd or even number on any one roll is always ½.
(b) The second roll is less likely to be odd because the first roll was odd.
(c) There are 36 possible outcomes when you roll a die twice. Getting two odds is only one of them.
(d) There are 12 possible outcomes when you roll a die twice. Getting two odds is only one of them.

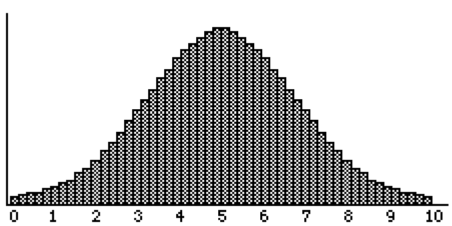
**10.** If a fair coin is tossed five times, which of the following ordered sequence of heads (H) and tails (T), if any, is MOST LIKELY to occur?

(a) T H T H T
(b) T H T H H
(c) H T T T T
(d) Sequences (a) and (c) are equally likely.
(e) All of the above sequences are equally likely.

**11.** When three fair dice are simultaneously thrown, which of the following results is MOST LIKELY to be obtained?

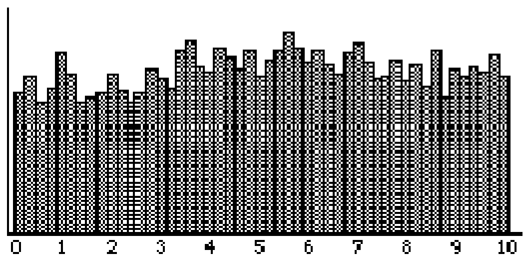
(a) Result 1: Three 2's
(b) Result 2: A 3, a 6 and a 4 in any order
(c) Result 3: Two 1's and a 5
(d) All three results are equally likely.

**12.** Two distributions of test scores (questions 12 and 13) are presented below. For each distribution, select the one descriptor that best represents the shape of the distribution.




(a) Normal
(b) Skewed
(c) Bimodal
(d) Uniform

**13.** .

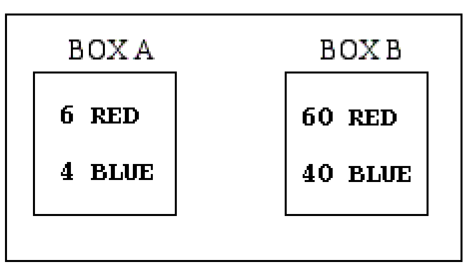



(a) Normal
(b) Skewed
(c) Bimodal
(d) Uniform

**14.** If a fair die is rolled five times, which of the following ordered sequence of results, if any, is MOST LIKELY to occur?

(a) 3 6 6 6 6
(b) 5 3 1 6 4
(c) 2 1 3 4 6
(d) Sequences (a) and (b) are equally likely.
(e) All of the above sequences are equally likely.

**15.** Box A and Box B are filled with red and blue marbles as follows. Each box is shaken. You want to get a blue marble, but you are only allowed to pick out one marble without looking.




Which box should you choose?

(a) Box A (with 6 red and 4 blue)
(b) Box B (with 60 red and 40 blue)
(c) It does not matter

**16.** In Game A, you toss a coin 60 times. In Game B, you toss a coin 8 times. You win the game if you toss at least 70% or more heads. If you play the game once, which game are you more likely to win?

(a) Game A
(b) Game B
(c) It is equally likely to win either game

**17**. Suppose a car manufacturer makes cars with a 0.70 probability of having no defects, a 0.20 probability of having one defect, and a 0.10 probability of having 2 defects. If two cars were shipped to a dealer, what is the most probable average number of defects of the two cars?

(a) 0
(b) 0.5
(c) 1
(d) 1.5
(e) 2

**18.** Assume it is equally likely to have a boy or girl. A couple decides to have children until they have a boy. Which of the following is the MOST PROBABLE number of children they will have?

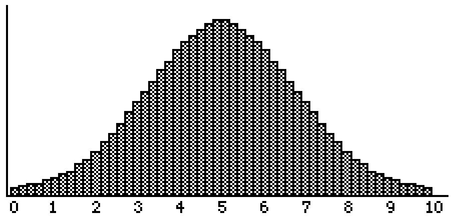
(a) 1
(b) 2
(c) 3
(d) 4

**19.** True or False: A weatherman predicts 70% chance of rain each day for the next four days. This means the most likely outcome is rain every day for the next four days. Assume independence.

(a) True
(b) False

**20**. True or False: A basketball player has a 75% field goal percentage. If she has made three shots in a row in a game, the likelihood of her making her next shot is higher than at the start of the game.

(a) True
(b) False

**21.** 


True or False: Given the distribution above, it is equally likely to get any one unit interval between 0 and 10 (ex. 0-1, 1-2, 2-3, etc. are all equally likely).

(a) True
(b) False

**22**. Jacob and Emily plan to visit a record store. Their frequent trips to the store result in the following probability distribution of the number of compact discs they buy.

| Number of CDs Jacob will buy | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Probability | 0.20 | 0.25 | 0.40 | 0.15 |

| Number of CDs Emily will buy | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Probability | 0.25 | 0.30 | 0.25 | 0.20 |

True or False: Letting the random variable represent the total number of CDs that Jacob and Emily buy together and assuming independence, it is less likely for them to buy 2 CDs total than 4 CDs total.

(a) True
(b) False

**23**. True or False: A weatherman predicts 80% chance of rain for the next five days. This means that it will rain exactly 4 out of the 5 days.

(a) True
(b) False

**24**. True or False: Consider a standard 52-card deck and suppose you pull one card from the deck. The probability that it is a king, given that it is a face card is the same as the probability that it is a face card, given that it is a king.

(a) True
(b) False

**25**. True or False: The probability of getting a prize in a cereal box is 0.20 or 1/5. The chance of getting your first prize in the fifth box you buy is higher than the chance of getting your first prize in the first box you buy.

(a) True
(b) False

**26**. True or False: Every morning when you arrive at school, you are either late or not late. Thus, the probability of arriving late or not late is 50%-50%.

(a) True
(b) False

**27**. The weather report on the news said there was a 90% chance of rain tomorrow. This means it will rain tomorrow.

(a) True
(b) False

**28**. The weather report said there was a 90% chance of rain tomorrow. The next day it did not rain. The weather report was incorrect.

(a) True
(b) False

**29**. You sped to work the last four days and did not get caught. This means you are more likely to get caught today.

(a) True
(b) False

**30**. You roll a die twice and get a sum of 7. Which of the following is more likely?

(a) You roll a 4 on the first roll and a 3 on the second one.
(b) You roll a 4 on the first roll.
(c) They are equally likely

APPENDIX E - SAMPLE SPACE WORKSHEET

**Sample Space** **Name: _________________Date: ______**

**Answer the following WITHOUT doing a simulation:**

1. Suppose you toss a coin 3 times.
   How many outcomes are there in the sample space?
   List the outcomes in the sample space:

   Describe how you got your answer to the previous questions.

2. Go to Activity 6 on p. 310 in your book (spinning the spinner 3 times).
   How many outcomes are there in the sample space?
   List the outcomes in the sample space:

   Describe how you got your answer to the previous questions.

3. On a given day, it either rains or it doesn't rain. Suppose you record the weather results over the next four days.
   How many outcomes are there in the sample space?
   List the outcomes in the sample space:

   Describe how you got your answer to the previous questions.

Now go back and use your calculator to SIMULATE the previous 3 exercises. (Do 20 repetitions with your calculator and compare to your group members to reach your answers). Describe how you simulated the situation. Do your conclusions from the simulation agree with your previous answers? Did you learn anything new from doing the simulation?

**Problem 1:**
Explain how you simulated:

Do your answers agree with previous answers? Explain.

Did you learn anything new from doing the simulation? Explain.

**Problem 2:**
Explain how you simulated:

Do your answers agree with previous answers? Explain.

Did you learn anything new from doing the simulation? Explain.

**Problem 3:**
Explain how you simulated:

Do your answers agree with previous answers? Explain.

Did you learn anything new from doing the simulation? Explain.

# APPENDIX F - PROBABILITY, INDEPENDENCE, MULTIPLICATION RULE LAB AND WORKSHEET

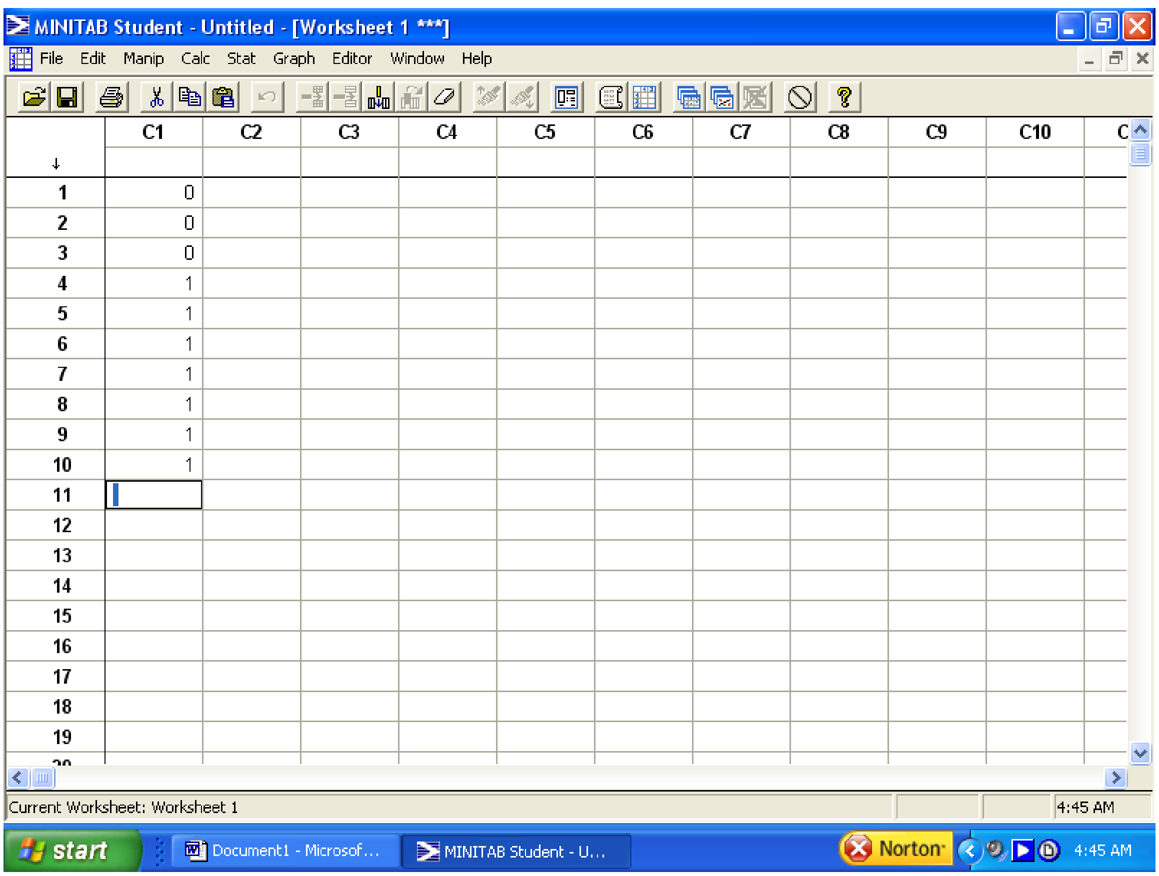
**Probability; Independence; Multiplication Rule**

So far, in using the TI83 and RandInt, we have given equal probability to all the integers generated. For example, doing RandInt(1, 6, 20) gave equal probability to the numbers 1 through 6 on being generated.

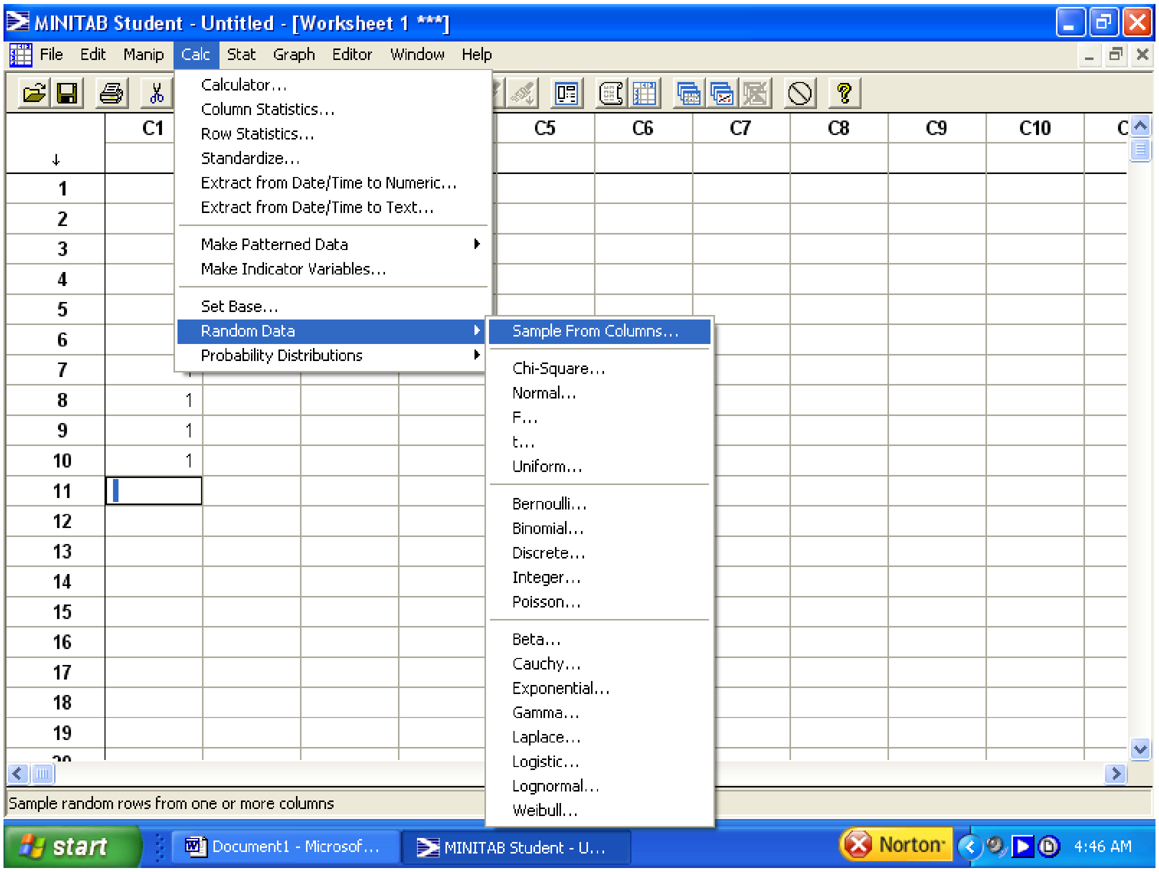
Sometimes values aren't equally likely and thus we assign specific probabilities (weights) to those values. To illustrate the concept, let's consider a specific example:

**Ex. 1** Suppose the probability that a basketball player makes a free throw is 70%. Let's illustrate what the random behavior for 100 possible free throws could look like. Let 0 = miss and 1 = make.

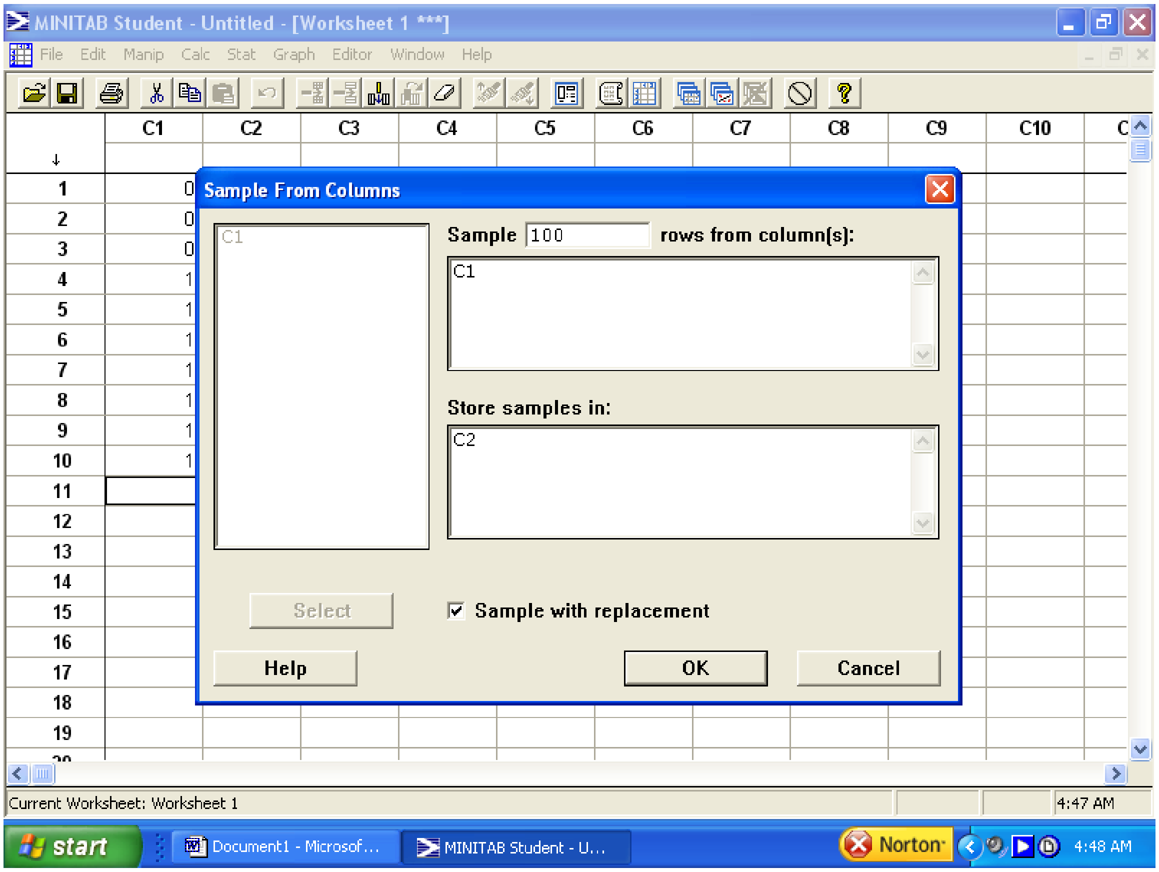
Go to PROGRAMS, then MINITAB and open this program. Down C1 (column 1) type in 3 zeros and 7 ones (this weights the miss as 30% and the make as 70%).



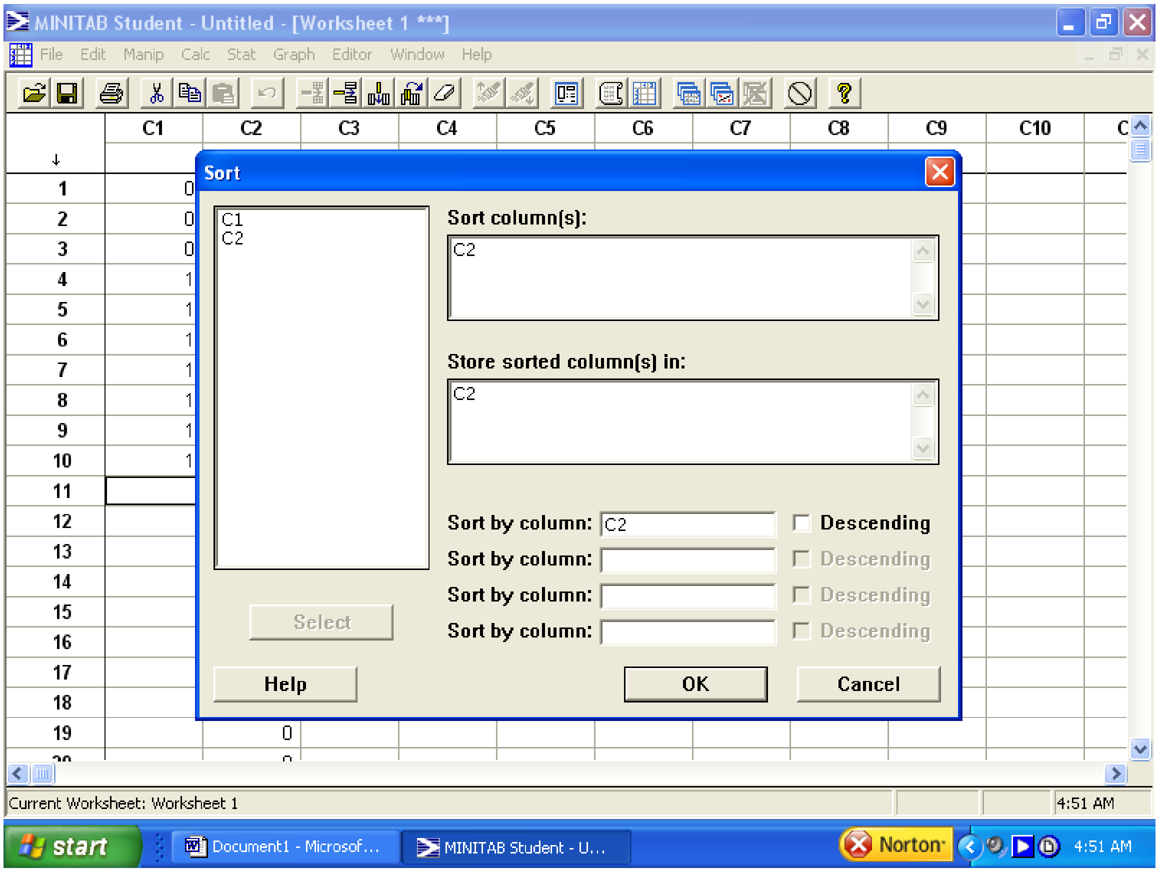

Now click CALC at the top, then RANDOM DATA, then SAMPLE FROM COLUMNS as shown:




When the following window comes up, type in the appropriate columns (you will sample from C1 and store your 100 results in C2). Make sure you check SAMPLE WITH REPLACEMENT. Then press OK.




Now Sort C2 by clicking MANIP at the top, then type in the following in the window, then press OK.




**PART I.** How many cells contain zeros? ________ How many cells contain ones? ______

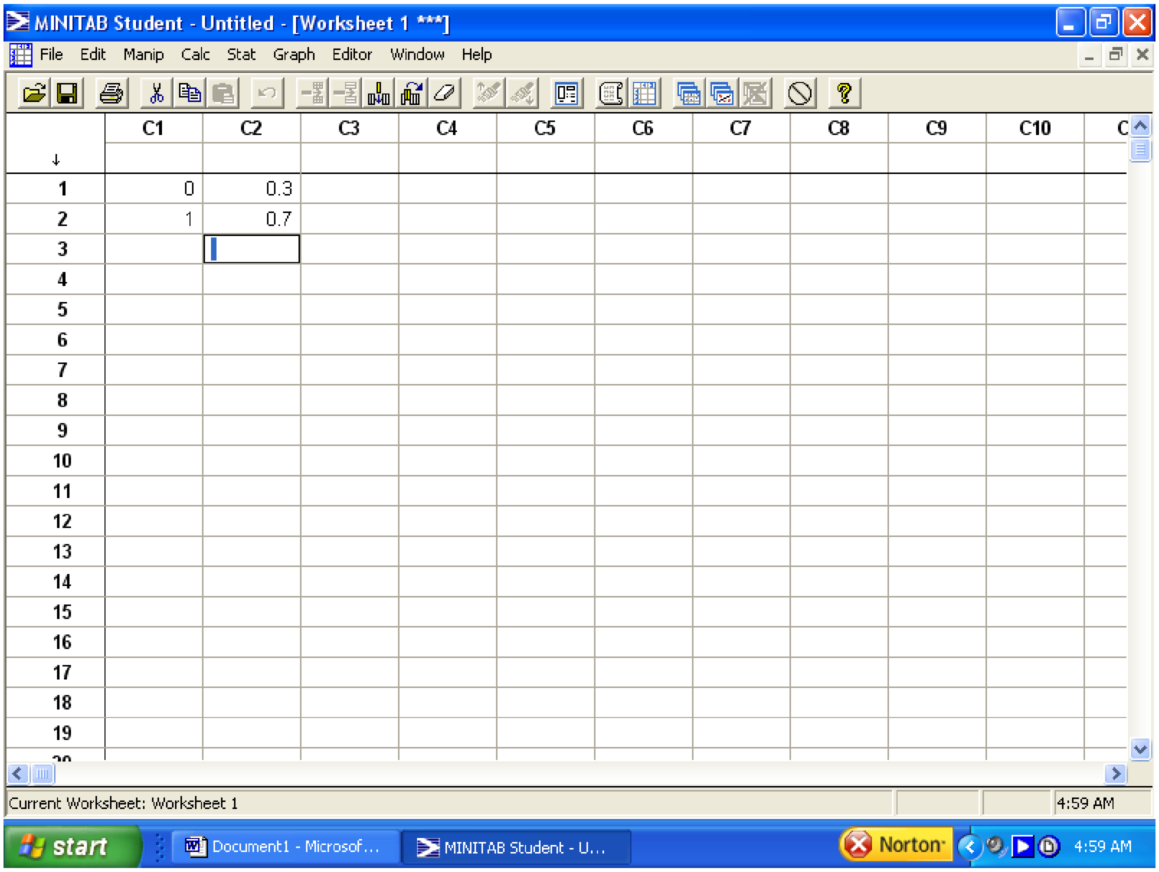
Now convert to percentages (since you simulated 100 shots this should be easy ie. if 22 cells contained zeros then this means 22% of the shots were misses and 78% were makes).
Percent of Misses ___________

Percent of Makes ___________

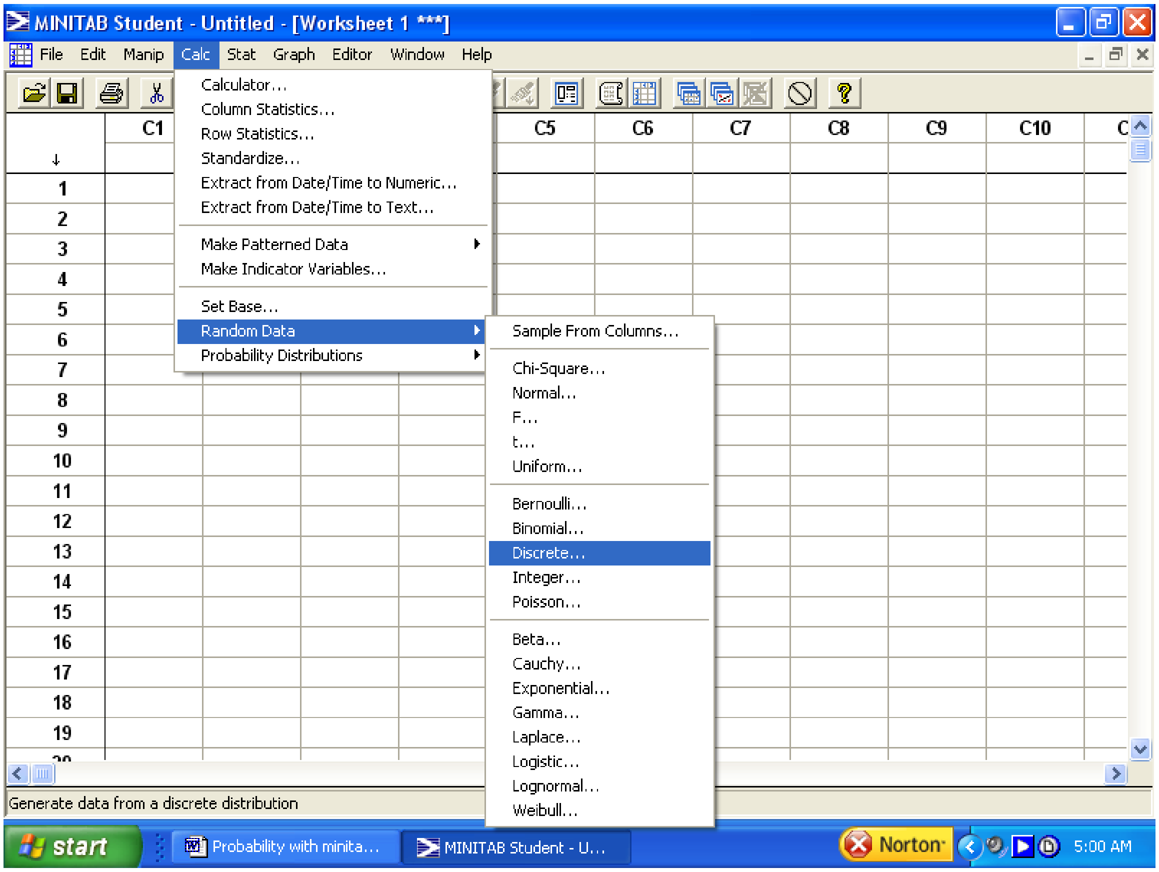
Check with the people sitting next to you. Do they have the exact same answers? _______ Why or why not?

Are you and your neighbors' answers close? _______ Comment.

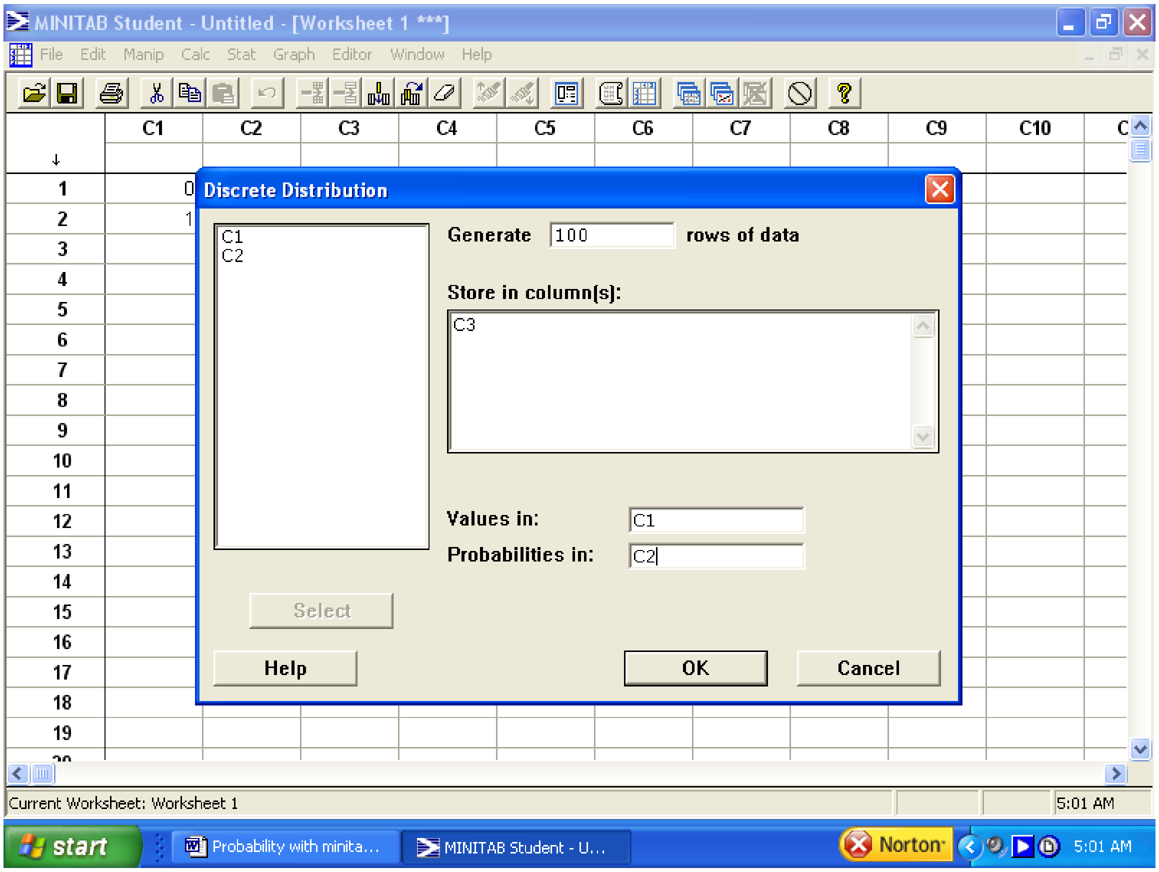
Now we do not always want to have to type in a lot of numbers to assign our weights, so there is a shortcut we can take. To do the same problem as above, we can type in the values of our variable in C1 (in this case 0 and 1). In C2, we can assign the weights (SEE BELOW).



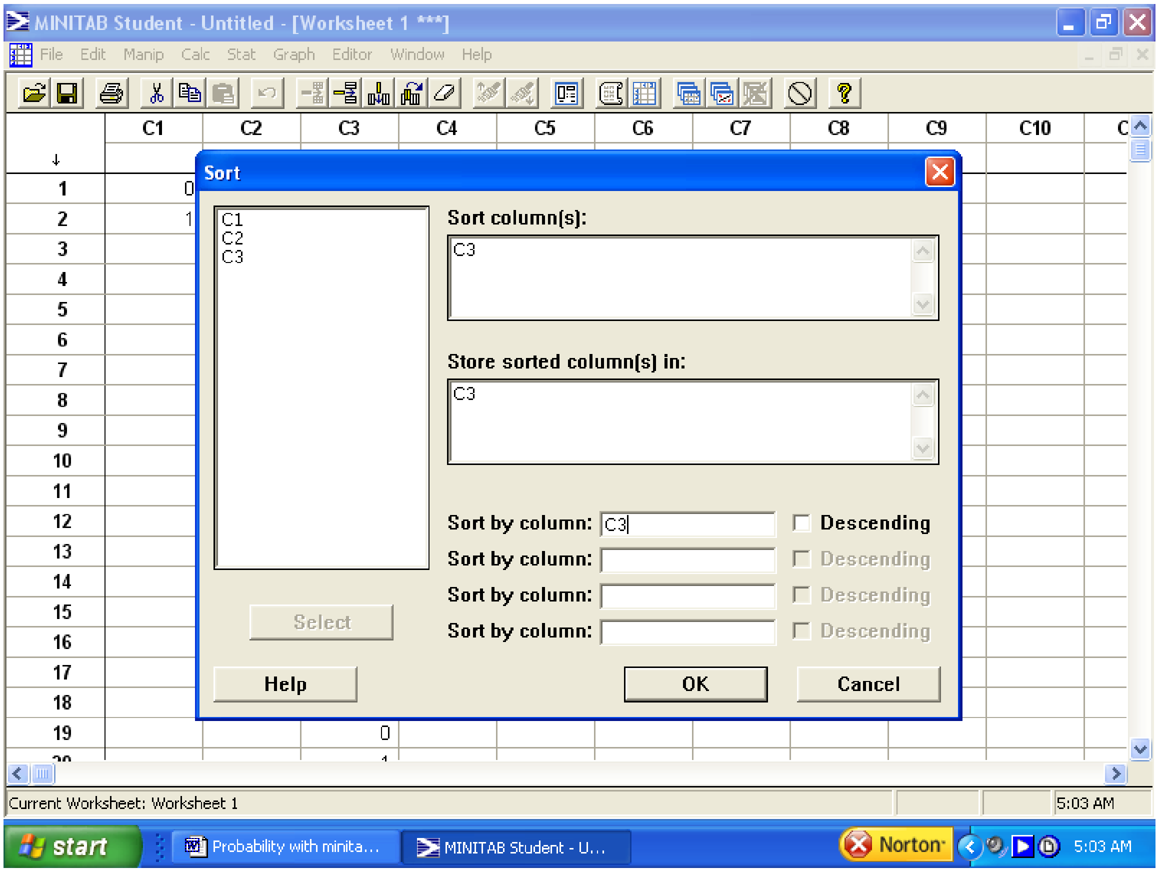

Next, click CALC, RANDOM DATA, then DISCRETE.




When the following window appears, store values in C3 as below. Then press OK.




Sort C3 now (press MANIP, SORT, and enter C3 in all the spaces in the window). Then press OK.




**PART II**. What percentage of the shots are misses? _____ What percentage are makes? _____ Are these answers the same as your previous answers? _________ Why or why not?

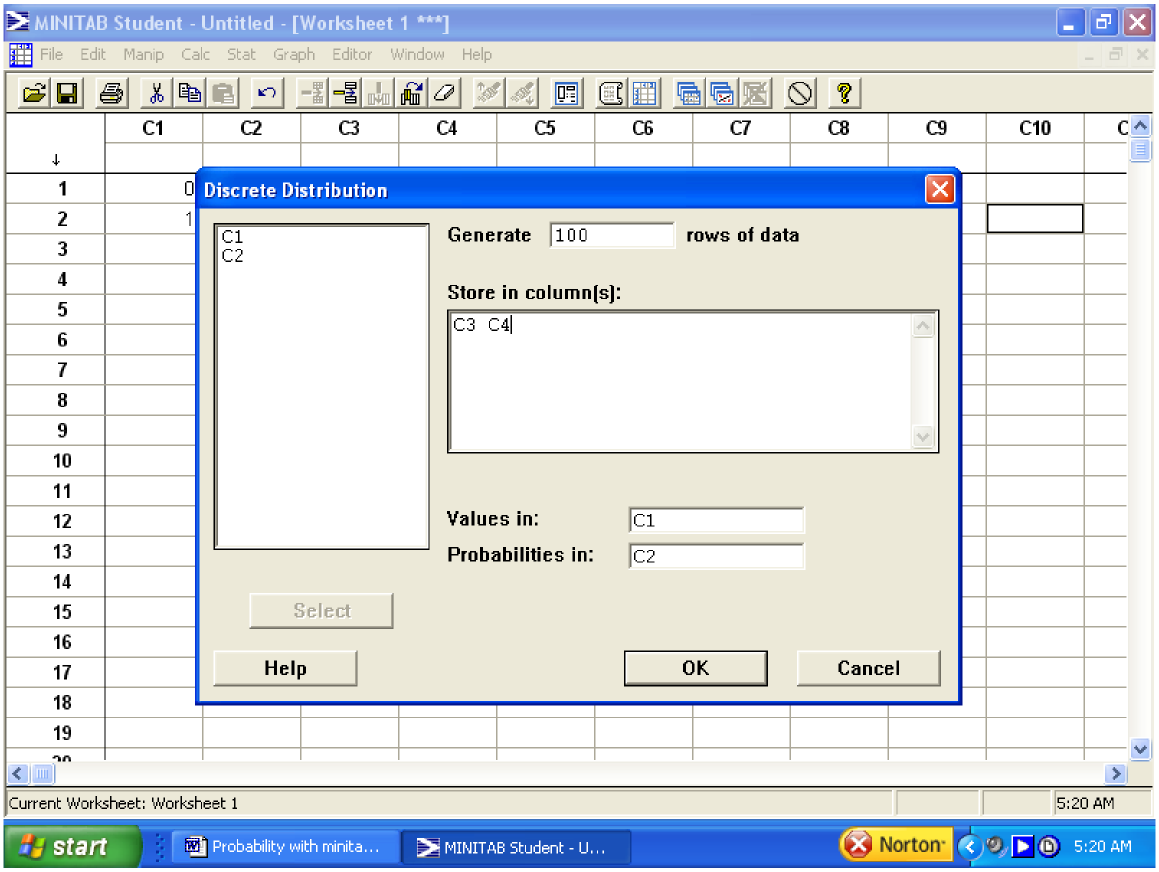
Now let's look at compound events, specifically joined by the word AND (which means the events may not be disjoint if they can possibly both occur).

Definition of **INDEPENDENT**: Two events A and B are **independent** if knowing that one occurs does not change the probability that the other occurs. If A and B are independent, then P(A and B) = P(A) * P(B) (so we multiply probabilities with "AND"). This is the multiplication rule for independent events.

**PART III**. Let's go back to our basketball problem. Suppose the player shoots twice. Based on what you have learned over the past few days, how many different outcomes are there? __________ List the possible outcomes in the sample space:

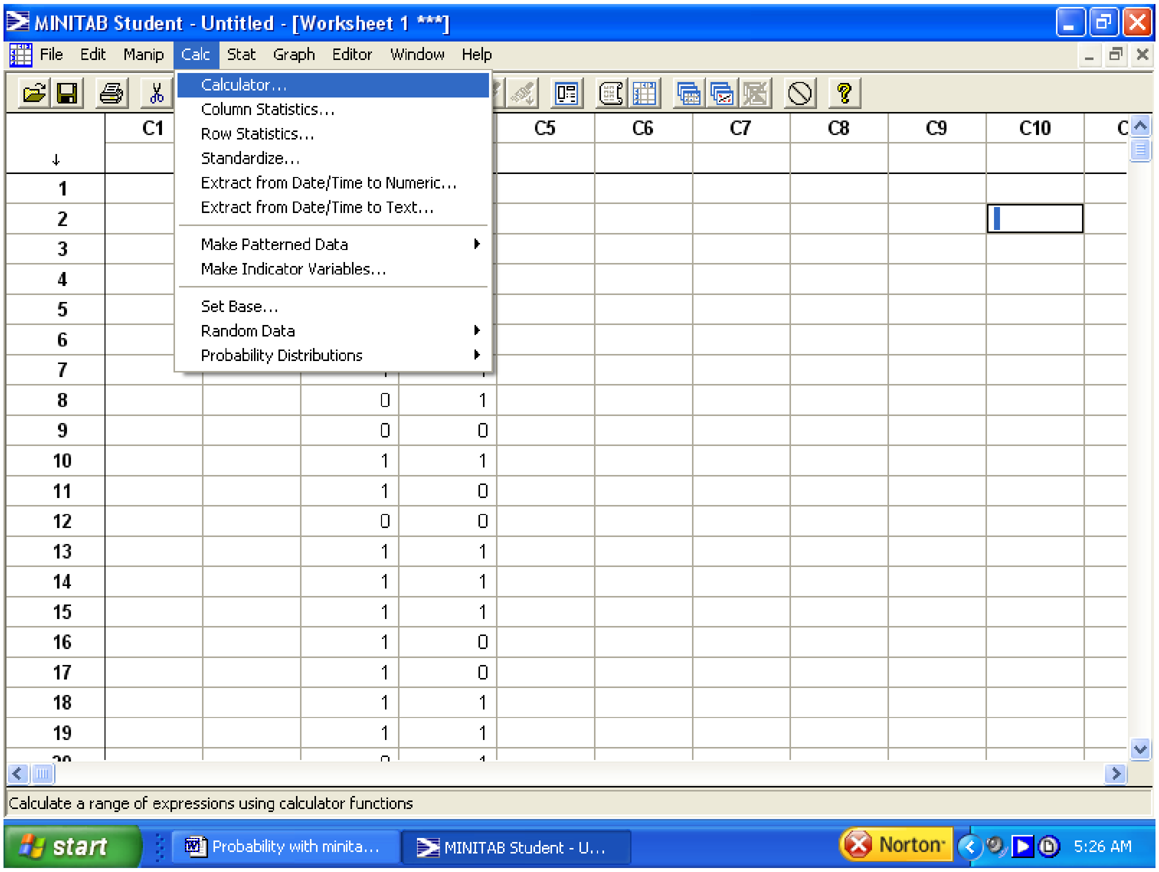
___________________________________

Use the same assigned values and probabilities in C1 and C2, but delete C3 values (just click at the top of C3 until the whole column is highlighted, then click DELETE on the keyboard). Now go and click CALC at the top of the screen, then RANDOM DATA, then DISCRETE again. Store 100 values in BOTH C3 and C4 this time (representing TWO basketball shots) as shown below:

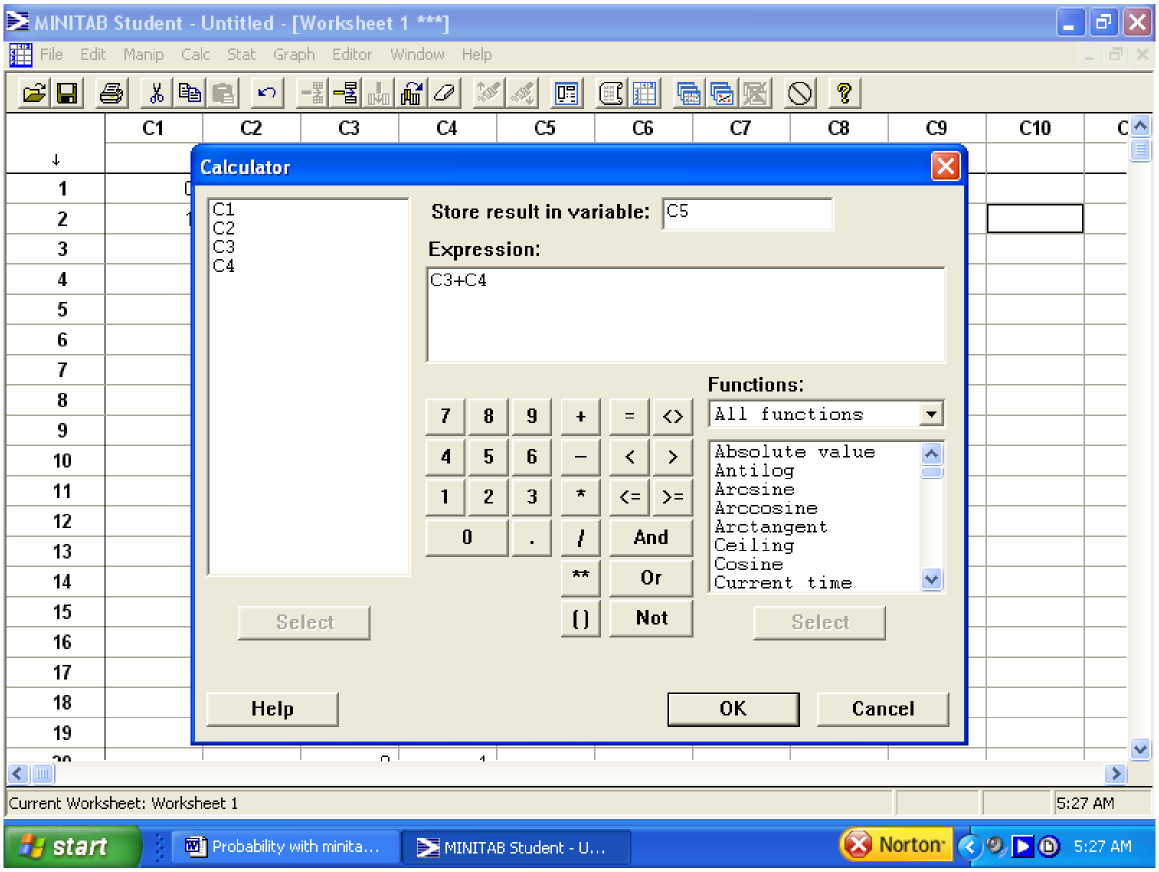



**PART IV.** You should now have 100 values down both C3 and C4. Suppose we want to calculate the probability that a player makes both shots so P(make AND make) which is the multiplication rule assuming the two shots are independent. Using the multiplication rule, what should our answer be theoretically (recall P(make) = .7)? SHOW WORK ___________________________

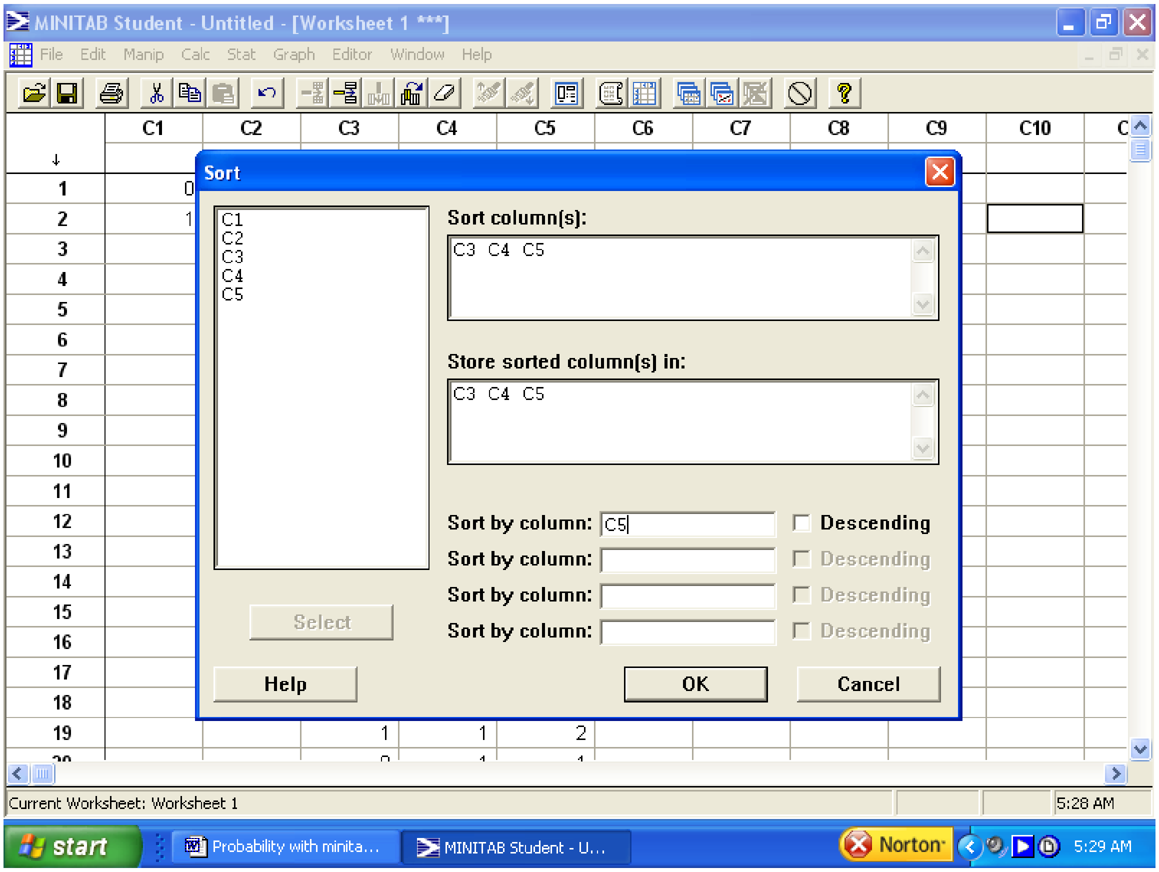
To check and see if our simulation results confirm this, we will add up the C3 and C4 pairs of numbers and see which ones give us a sum of 2 (which means MAKE = 1 + MAKE = 1 and 1+1 = 2). We will add the numbers and store our sums in C5. To do this, click CALC at the top, then CALCULATOR as shown below:




When the next window pops up, type in the following (you will add C3+C4 and store results in C5):

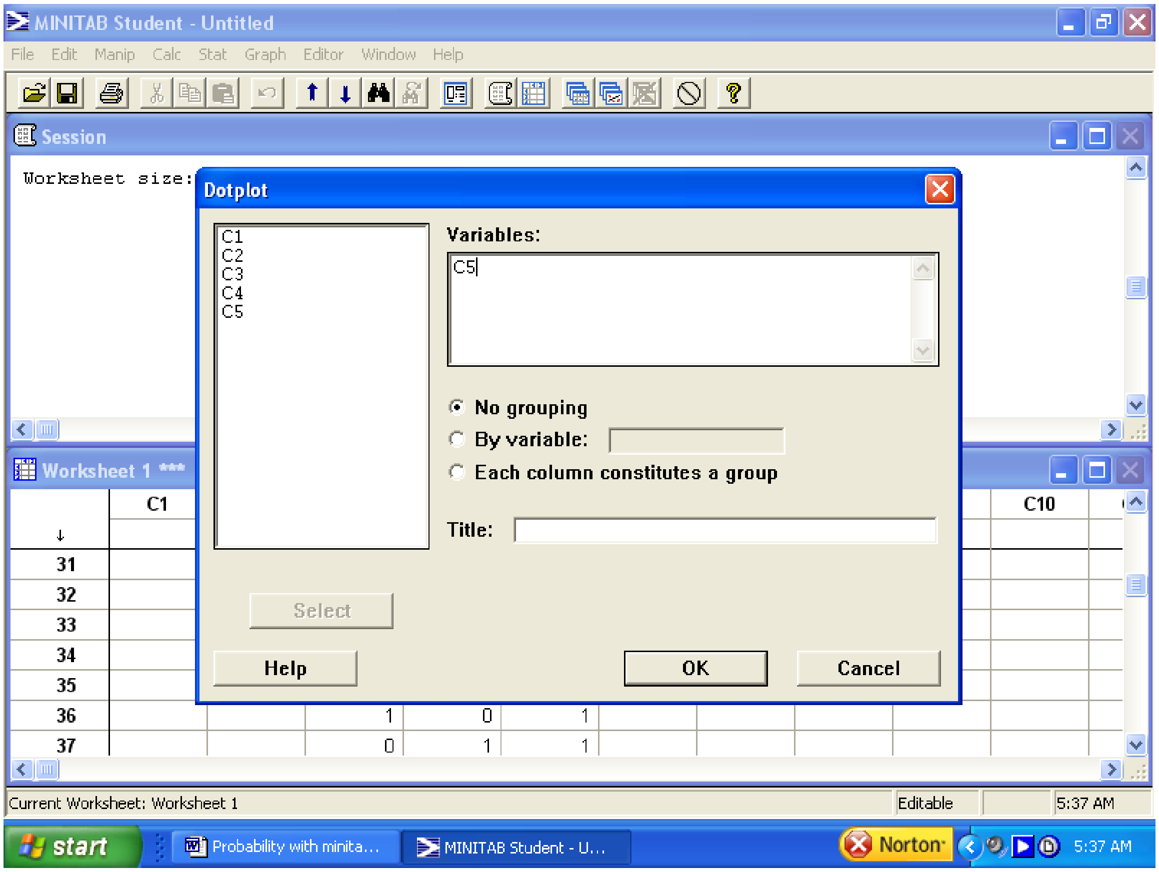



Now sort C5 by clicking MANIP then SORT then typing in the following:




**PART V**. Now count how many cells contain a 2 in C5 (which implies BOTH baskets were made) and convert to a percent (since it was based out of 100 this should be easy). What percent were both makes? __________ Does this agree with your answer based on the multiplication rule? Explain.

Graph your results by clicking GRAPH at the top, then DOTPLOT and filling in the window with C5 as below:




**PART VI**. Then press OK. By looking at your dotplot, if a person is a 70% free throw shooter and takes 2 shots, how many is he/she MOST likely to make? ___________ how many is he/she LEAST likely to make? _____________ How can you tell by looking at the dotplot?

Do your neighbors necessarily show the same dotplot? Why or why not?

**Probability/Independence Multiplication Rule** **Name: __________________ Date: ___________**

**PART I**. How many cells contain zeros? ________ How many cells contain ones? ______

Now convert to percentages (since you simulated 100 shots this should be easy ie. if 22 cells contained zeros then this means 22% of the shots were misses and 78% were makes).
Percent of Misses ___________

Percent of Makes ___________

Check with the people sitting next to you. Do they have the exact same answers? _______ Why or why not?

Are you and your neighbors' answers close? _______ Comment.

**PART II**. What percentage of the shots are misses? _____ What percentage are makes? _____ Are these answers the same as your previous answers? ________ Why or why not?

**PART III**. Let's go back to our basketball problem. Suppose the player shoots twice. Based on what you have learned over the past few days, how many different outcomes are there? _________ List the possible outcomes in the sample space: _______________________________

**PART IV.** You should now have 100 values down both C3 and C4. Suppose we want to calculate the probability that a player makes both shots so P(make AND make) which is the multiplication rule assuming the two shots are independent. Using the multiplication rule, what should our answer be theoretically (recall P(make) = .7)? SHOW WORK _________________________

**PART V**. Now count how many cells contain a 2 in C5 (which implies BOTH baskets were made) and convert to a percent (since it was based out of 100 this should be easy). What percent were both makes? __________ Does this agree with your answer based on the multiplication rule? Explain.

**PART VI**. Then press OK. By looking at your dotplot, if a person is a 70% free throw shooter and takes 2 shots, how many is he/she MOST likely to make? ___________ how many is he/she LEAST likely to make? _____________ How can you tell by looking at the dotplot?

Do your neighbors necessarily show the same dotplot? Why or why not?

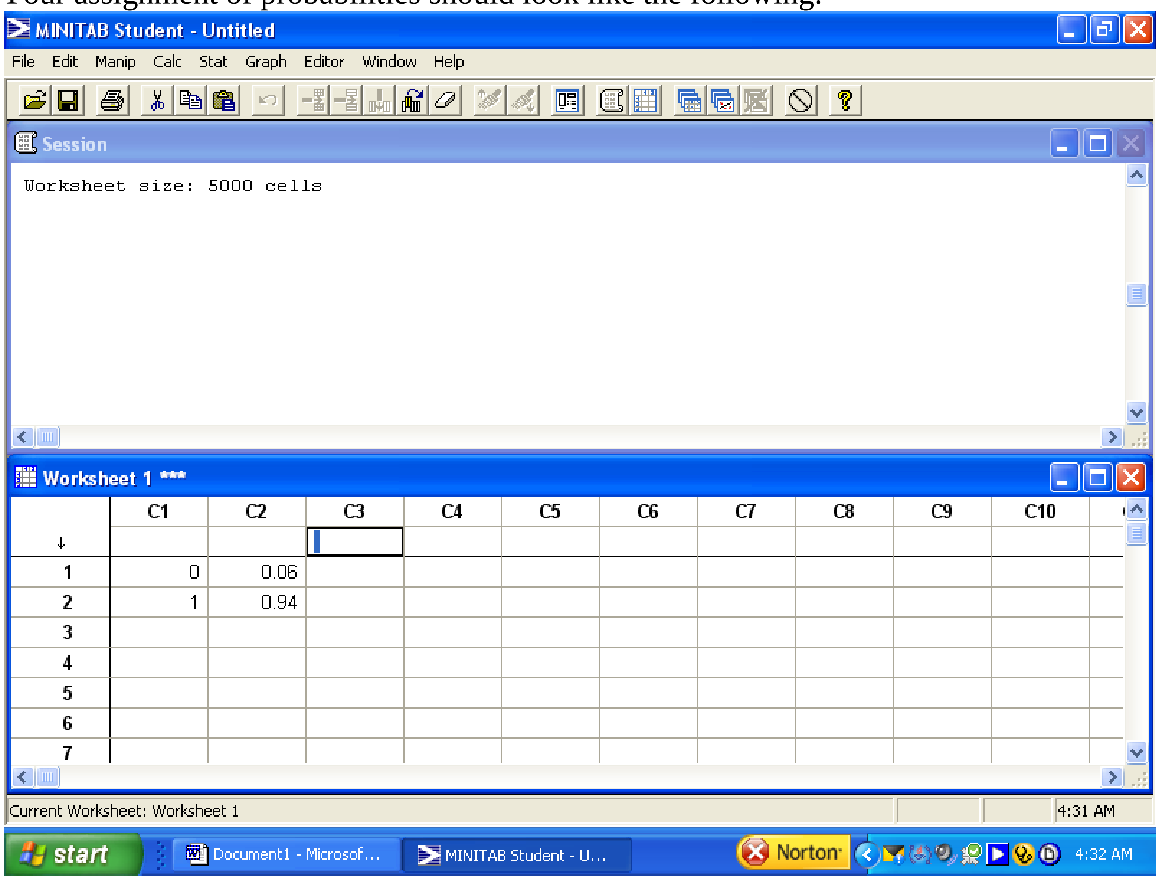
APPENDIX G - INDEPENDENCE, GENERAL ADDITION RULE, JOINT PROBABILITY LAB AND WORKSHEET

## 6.2/6.3 Independence/General Addition Rule/Joint Probability

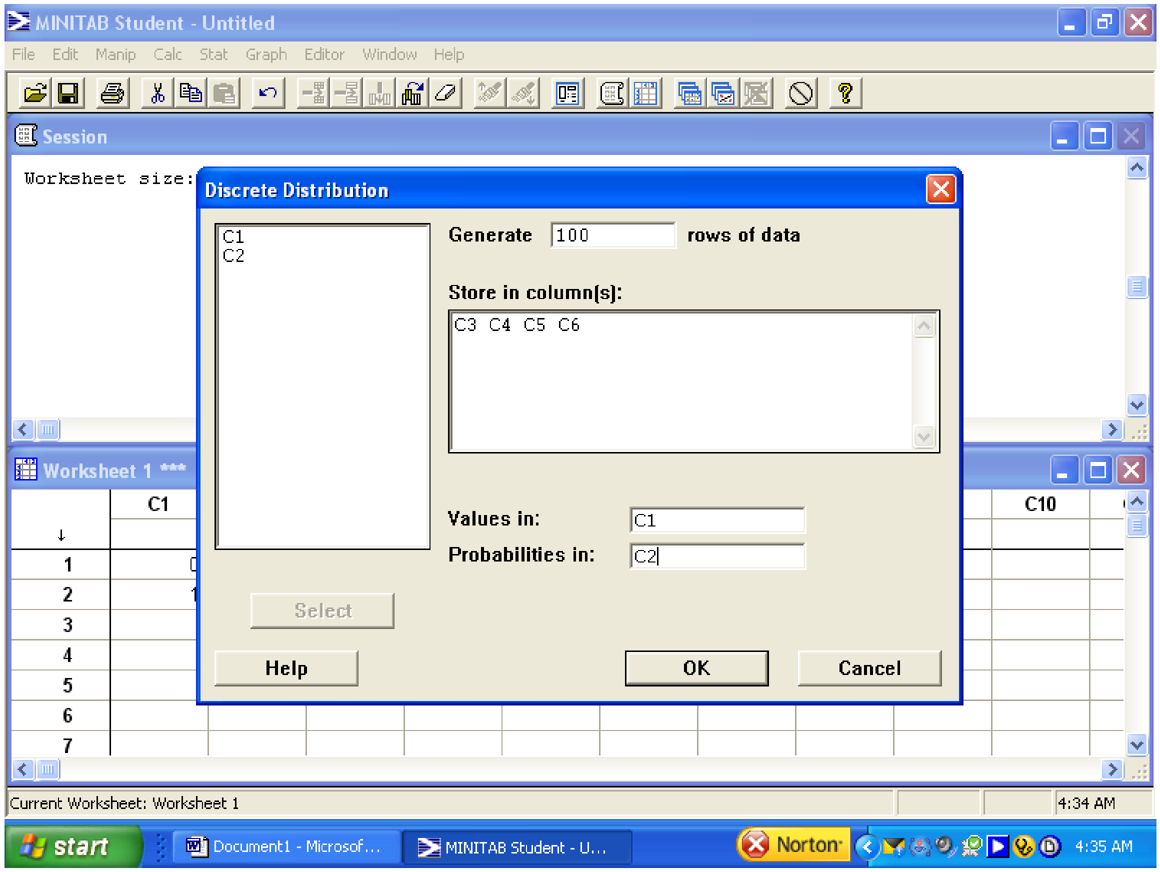
**Part I**: A company ships boxes of 4 computer chips to its customers. Suppose from historical data, the company assumes a 6% defect rate. What is the probability that in a box of 4 that at least one of the chips is working properly? ___________ Predict your answer on your sheet (without performing any simulation).

Now using a simulation, open Minitab program. Assign probabilities letting 0 = defective and 1 = working properly. Since P(defective) = .06 is given, what is the P(working properly)? _________

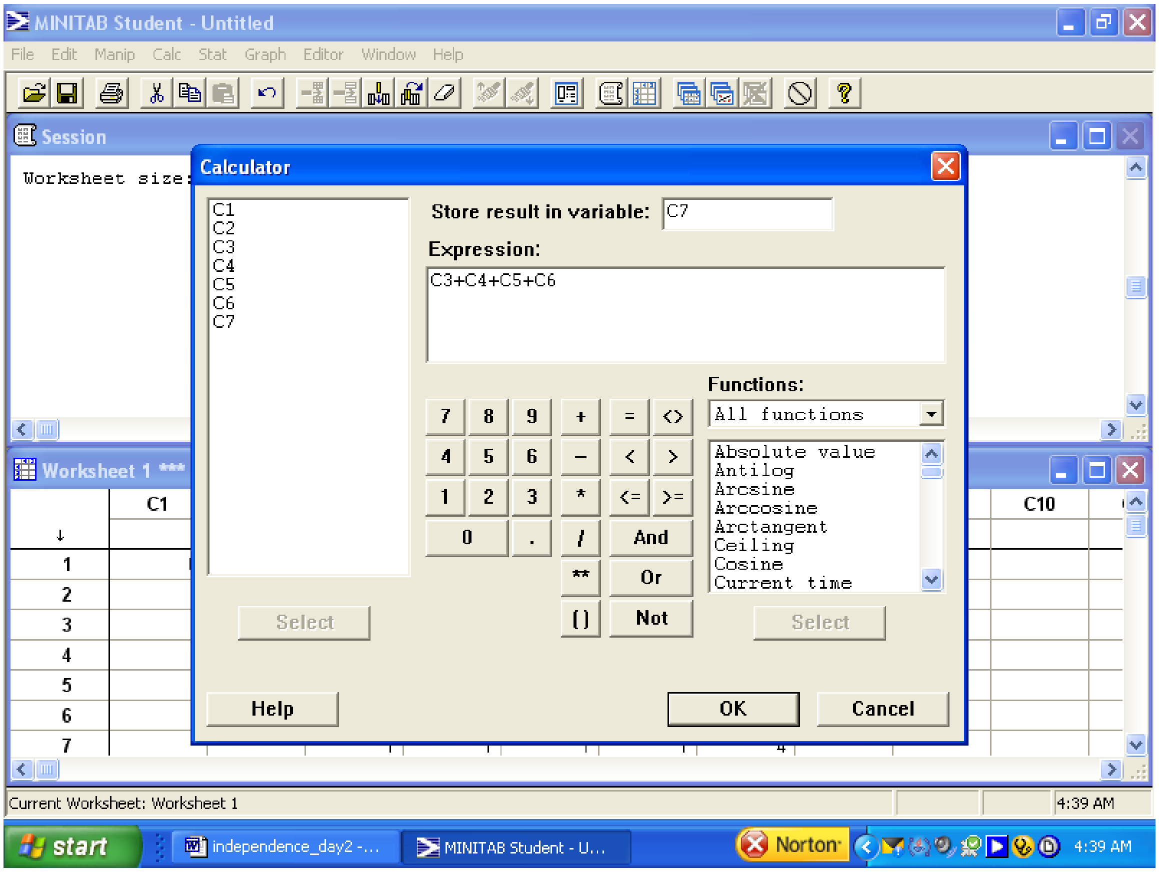
Your assignment of probabilities should look like the following:



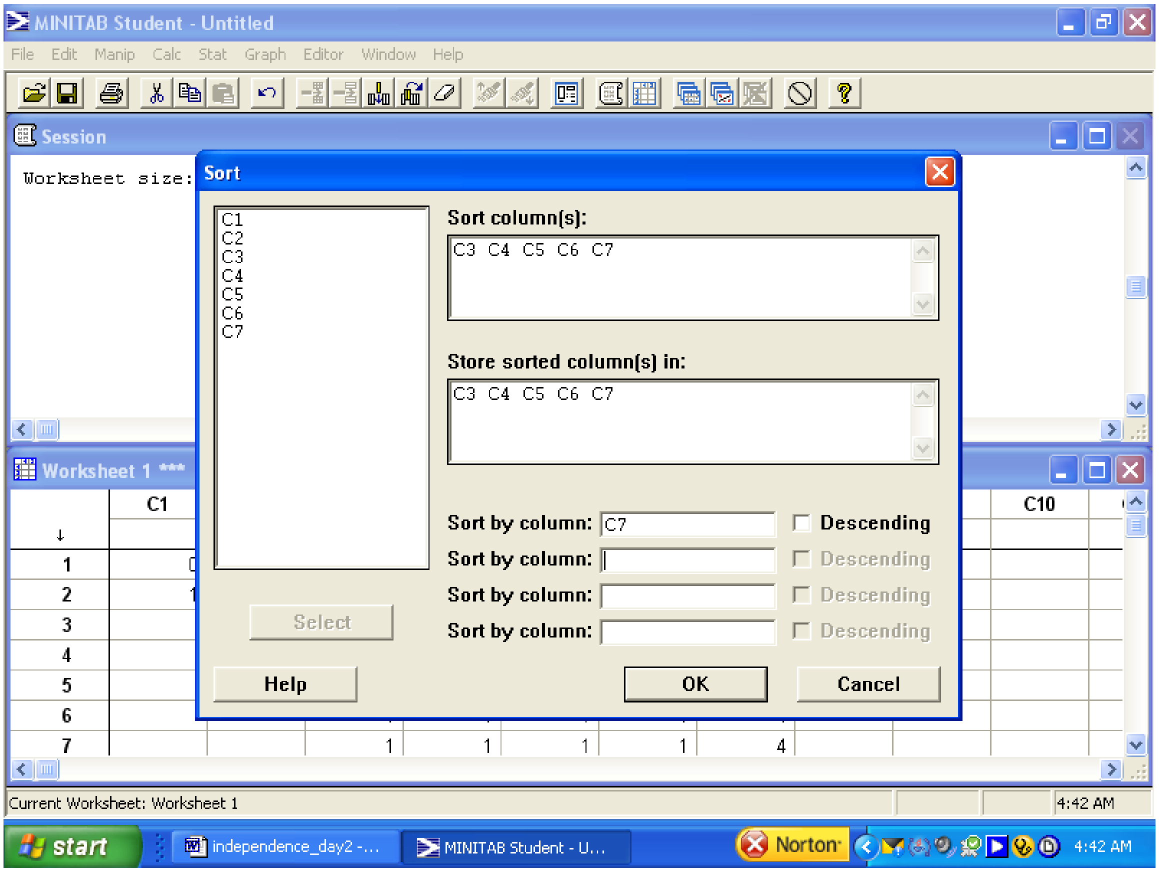

Now simulate 100 chips down C3, C4, C5, and C6 (we are using 4 columns since there are 4 chips). Remember we do this by clicking CALC at the top, then RANDOM DATA, then DISCRETE. Then when the following window appears, type:




Then press OK and you should have simulated data in all 4 columns. Next, we want to add up the four columns to see how many chips are working properly in each outcome (since working properly = 1 then our sum should show how many are working properly in each batch). To add C3 through C6, click CALC, then CALCULATOR, then type the following window:




Now sort C3 through C7 by clicking MANIP, SORT, then typing the following window:




**PART I cont'd** Now since we based this out of 100 to make it easy to calculate percents, calculate what percent of the outcomes had at least one working properly __________. (This will probably be a lot of them if not all of them). How many outcomes had none of them working? __________ Since the sum of these two percents adds up to 100%, then mathematically we should be able to find P(at least one) by using the formula **P(at least one) = 1 – P(none).**

Example: Suppose the probability of passing a test is 0.8. The result of each test is independent of another test. If you take 3 tests, find the probability of passing at least one test. Realize that P(none) = fail AND fail AND fail and that AND means to multiply probabilities. What would your answer be and show how you got it?

______________________________________________________________________

**PART II.** Realize that when we toss 2 dice, there are 36 total outcomes. The following chart shows all possible sums when we roll 2 dice:

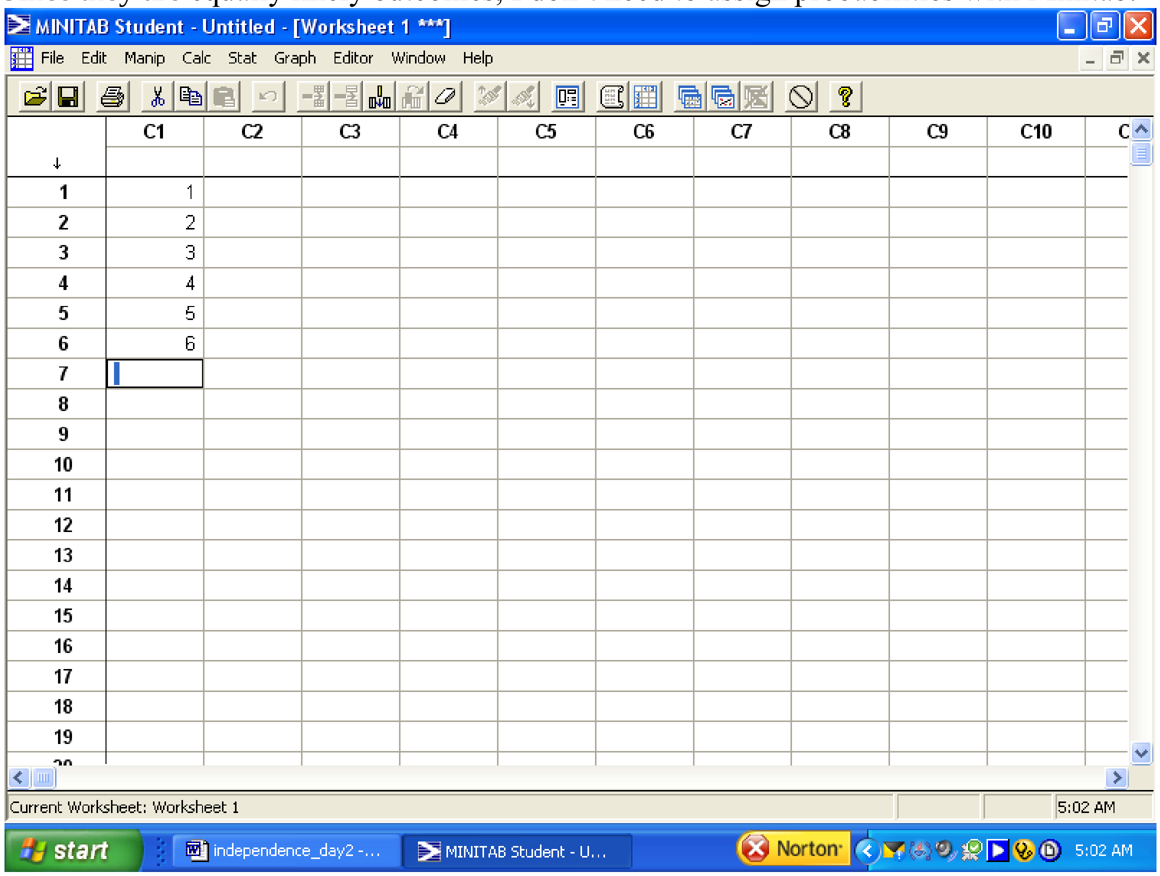
| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 6 | 7 | 8 | 9 | 10 | 11 | 12 |

Now try to answer the following question: Suppose you roll two dice. If event A = rolling a 6 on the first die and event B = sum of 12, try to calculate the probability of rolling a 6 on the first die AND getting a sum of 12 by using formulas:

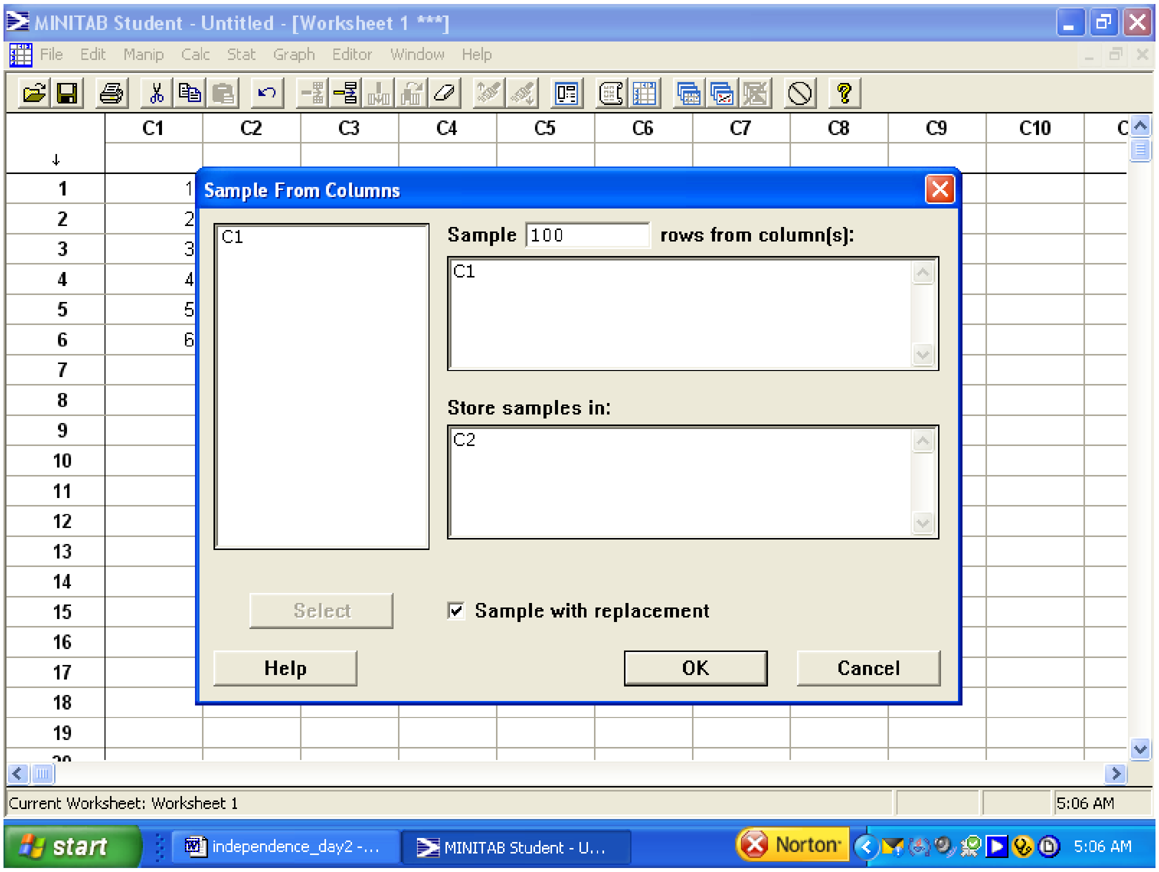
_______________________

Now, let's test your calculation by using a simulation (we should get fairly close to the correct answer if we simulate 100).

Clear out the columns from the previous data and type the numbers 1 through 6 down C1. Since they are equally likely outcomes, I don't need to assign probabilities with Minitab.

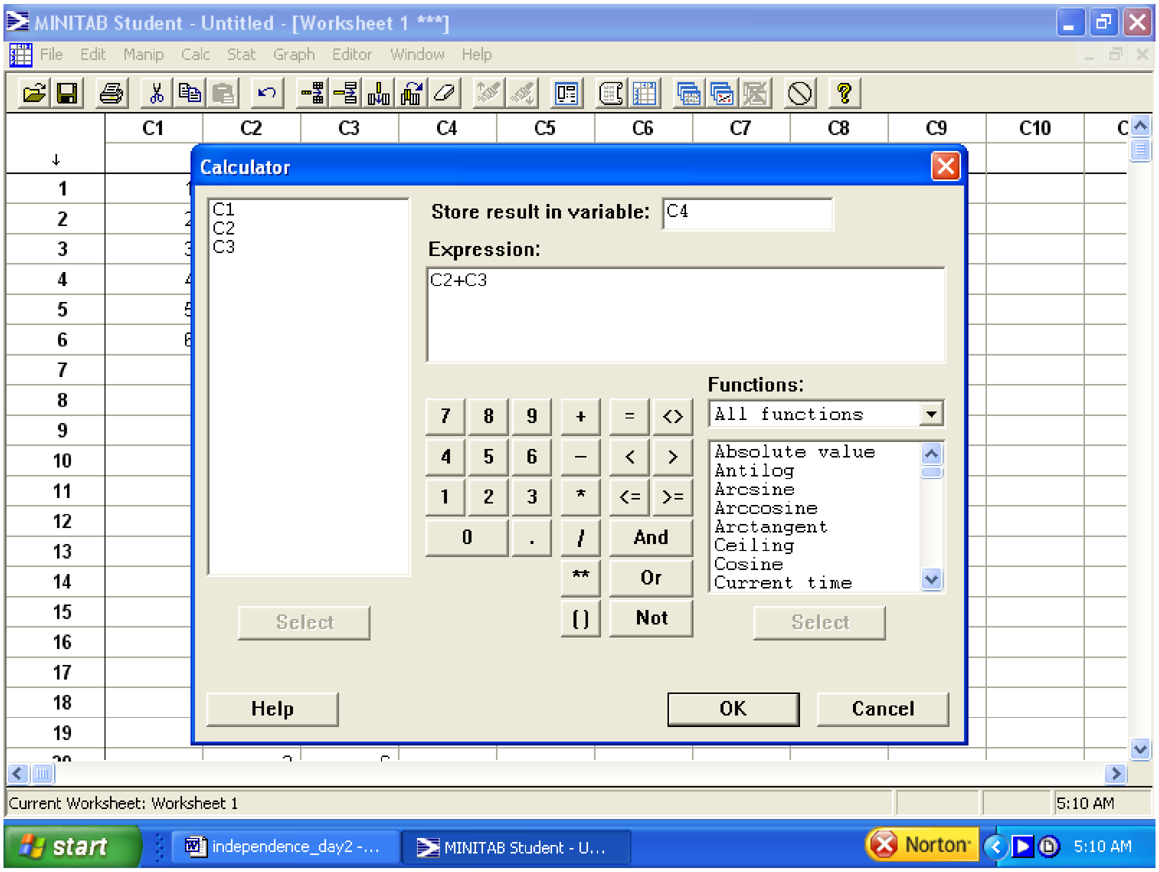



Now simulate rolling two dice (100 times) by clicking CALC, RANDOM DATA, SAMPLE FROM COLUMNS. The following window should appear so type in as follows. Make sure you check SAMPLE WITH REPLACEMENT below:

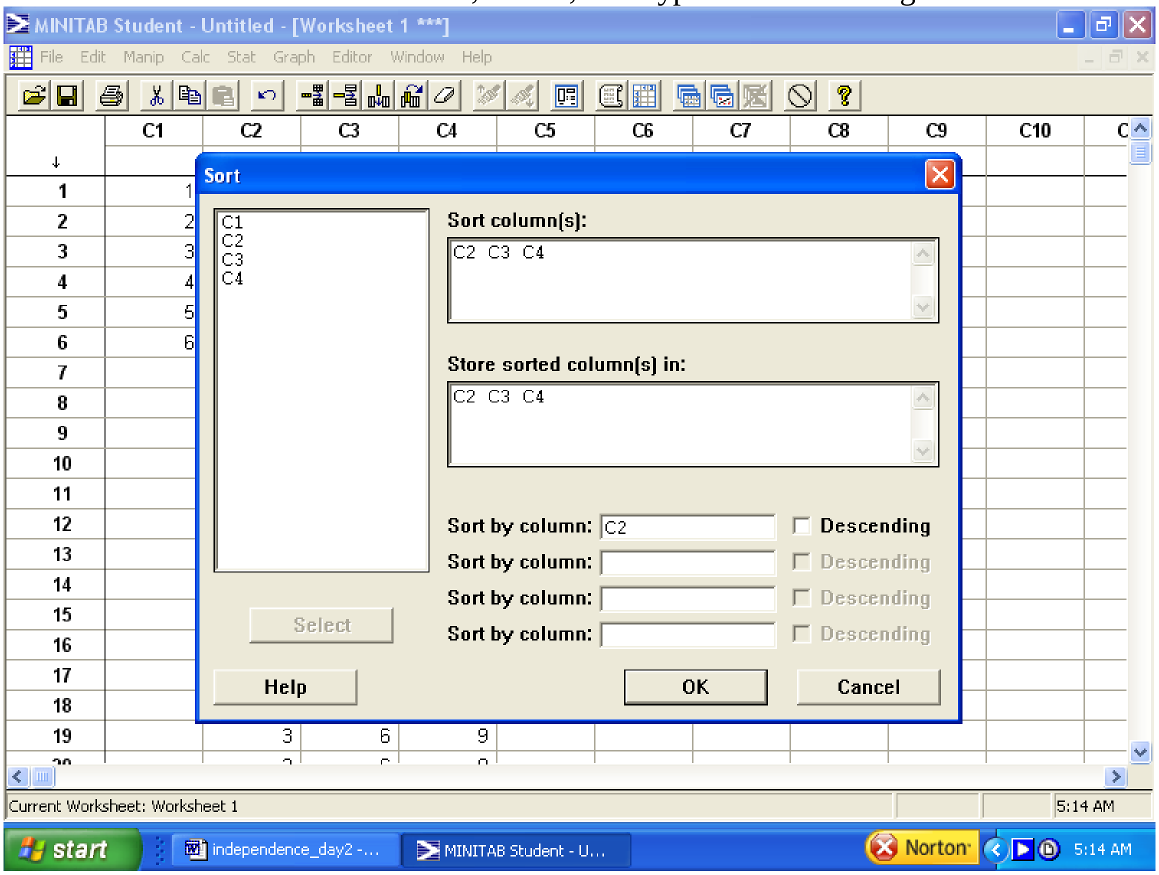



Now redo those same instructions: CALC, RANDOM DATA, SAMPLE FROM COLUMNS, but Store Samples in C3 rather than C2 (still check SAMPLE WITH REPLACEMENT). Then press OK. You should have simulations of tosses down both C2 and C3.

Now add C2 and C3 (to get the sum of your dice) by clicking CALC, CALCULATOR and typing in the following:




Now recall we are finding the probability of rolling a 6 on the first toss AND getting a sum of 12 (which means both must occur), so let’s sort the C2 and C3 according to C2 which is our first toss. Click MANIP, SORT, then type in the following:




Now go find the percent of outcomes that have a 6 in C2 (for tossing a 6 on the first die) and have a sum of 12 in C4.__________ If you count how many outcomes show this, then since it was based out of 100 this will be your percent. Is this answer close to your predicted answer? __________

Are events A = getting a 6 on the first toss and B = getting a sum of 12 INDEPENDENT ie. the occurrence of the first does not affect or change the occurrence of the second? __________ To help you answer this, start at the top and scroll down noticing the types of sums you got when you rolled a one first, then when you rolled a two first, etc. Does what you roll on the first die seem to influence what sum you will get? _____________ If so, then the events are NOT independent and thus you CANNOT use the multiplication rule P(A and B) = P(A)*P(B).

**PART III.** Next, using our same results from PART II, let's look at the following: Event A = rolling a 3 on the first toss and Event B = getting a sum of 7 with the two dice. Find the following probabilities from your simulated results (realize different students will get different answers):

P(rolling a 3 on the first die) = ___________

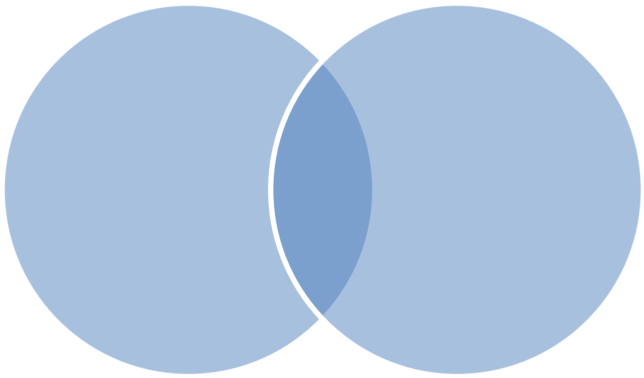
P(getting a sum of 7 with both dice) = _____________

P(rolling a 3 on the first die AND getting a sum of 7) = ______________

P(rolling a 3 on the first die but NOT getting a sum of 7) = ______________

P(getting a sum of 7 but NOT rolling a 3 on the first die) = ___________________

Using the information you calculated, see if you can complete the following Venn diagram:



Now predict the answer to P(getting a 3 on first die OR getting a sum of 7) = _________
Check your simulated results by counting how many have either a 3 on the first die or a sum of 7 (careful that you do not double count any outcomes). Are these events mutually exclusive? __________ Why or why not? ___________________________________
Your answer when you count using your simulated results should agree with the answer you get if you add up the 3 numbers in your Venn diagram.

**Name: ______________________________ Date: ________________**

**6.2/6.3 Independence/General Addition Rule/Joint Probability**

**Part I**: A company ships boxes of 4 computer chips to its customers. Suppose from historical data, the company assumes a 6% defect rate. What is the probability that in a box of 4 that at least one of the chips is working properly? ___________ Predict your answer on your sheet (without performing any simulation).
Now using a simulation, open Minitab program. Assign probabilities letting 0 = defective and 1 = working properly. Since P(defective) = .06 is given, what is the P(working properly)? _________

**PART I cont'd** Now since we based this out of 100 to make it easy to calculate percents, calculate what percent of the outcomes had at least one working properly _________. (This will probably be a lot of them if not all of them). How many outcomes had none of them working? _________ Since the sum of these two percents adds up to 100%, then mathematically we should be able to find P(at least one) by using the formula **P(at least one) = 1 – P(none).**

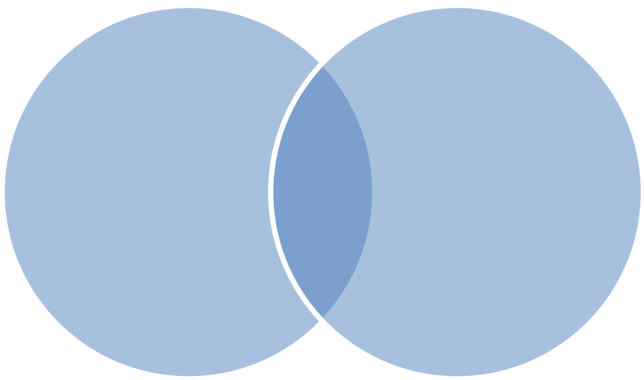
Example: Suppose the probability of passing a test is 0.8. The result of each test is independent of another test. If you take 3 tests, find the probability of passing at least one test. Realize that P(none) = fail AND fail AND fail and that AND means to multiply probabilities. What would your answer be and show how you got it?
____________________________________________________________________

Now try to answer the following question: Suppose you roll two dice. If event A = rolling a 6 on the first die and event B = sum of 12, try to calculate the probability of rolling a 6 on the first die AND getting a sum of 12 by using formulas:
_______________________

Now go find the percent of outcomes that have a 6 in C2 (for tossing a 6 on the first die) and have a sum of 12 in C4.__________ If you count how many outcomes show this, then since it was based out of 100 this will be your percent. Is this answer close to your predicted answer? __________
Are events A = getting a 6 on the first toss and B = getting a sum of 12 INDEPENDENT ie. the occurrence of the first does not affect or change the occurrence of the second? __________ To help you answer this, start at the top and scroll down noticing the types of sums you got when you rolled a one first, then when you rolled a two first, etc. Does what you roll on the first die seem to influence what sum you will get? _____________ If so, then the events are NOT independent and thus you CANNOT use the multiplication rule P(A and B) = P(A)*P(B).

**PART III.** Next, using our same results from PART II, let's look at the following: Event A = rolling a 3 on the first toss and Event B = getting a sum of 7 with the two dice. Find the following probabilities from your simulated results (realize different students

will get different answers):
P(rolling a 3 on the first die) = ___________
P(getting a sum of 7 with both dice) = _____________
P(rolling a 3 on the first die AND getting a sum of 7) = ______________
P(rolling a 3 on the first die but NOT getting a sum of 7) = ______________
P(getting a sum of 7 but NOT rolling a 3 on the first die) = ___________________
Using the information you calculated, see if you can complete the following Venn diagram:



Now predict the answer to P(getting a 3 on first die OR getting a sum of 7) = _________
Check your simulated results by counting how many have either a 3 on the first die or a sum of 7 (careful that you do not double count any outcomes). Are these events mutually exclusive? __________ Why or why not? _____________________________________
Your answer when you count using your simulated results should agree with the answer you get if you add up the 3 numbers in your Venn diagram.

APPENDIX H - CONDITIONAL PROBABILITY LAB AND WORKSHEET

**Conditional Probability**

Recall from the lab last week, when rolling two dice, let A = rolling a 6 on first die and B = sum of two dice, it appeared that P(A and B) $\neq$ P(A)*P(B). Why do you think this rule did not apply? ____________________________________________________________

This is an example of **conditional probability**, or the probability of one event under the condition that we know another event. The symbol for "the probability of B occurring **given that** A has occurred" is $P(B|A)$ so the "condition" appears after the slash mark. The formula for calculating $P(B|A)$ is the following:

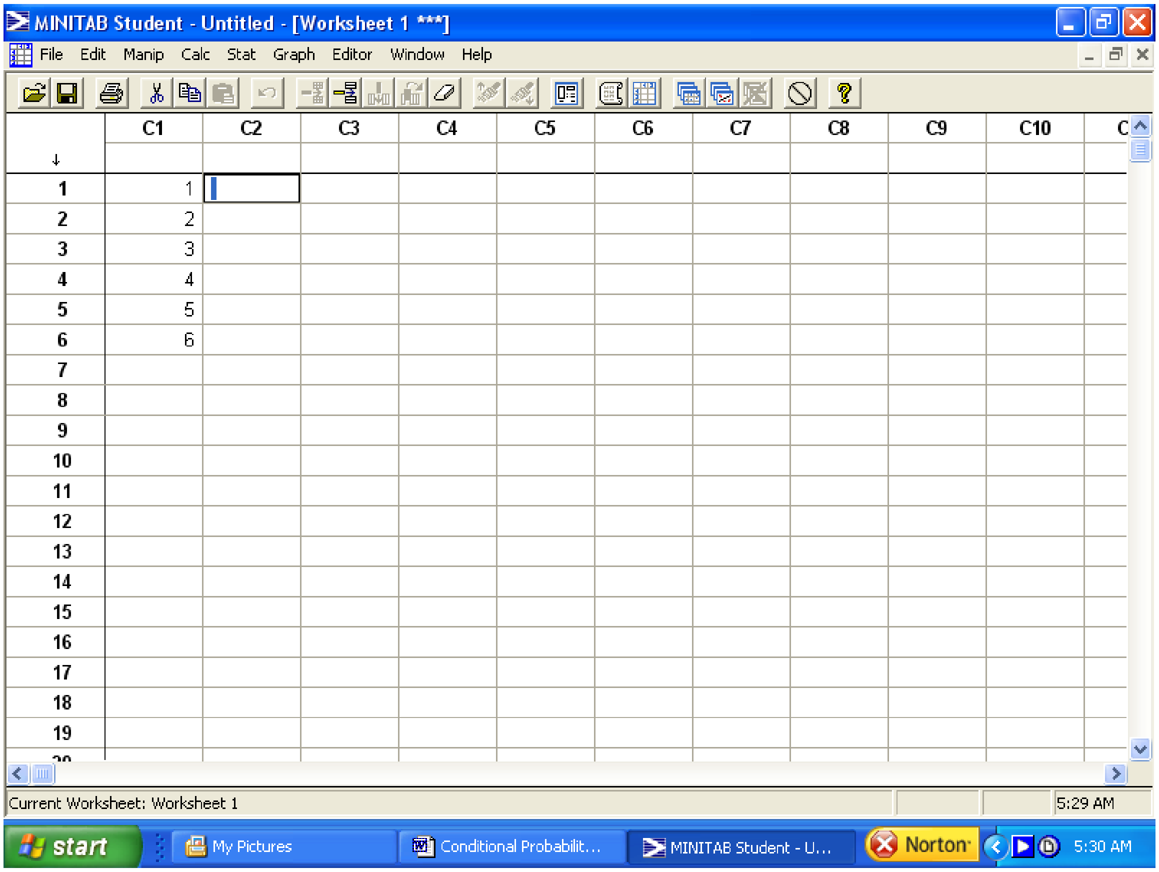
$$P(B|A) = \frac{P(A and B)}{P(A)}$$

Notice that if we do some algebra and rearrange the above equation, solving for P(A and B), we get:

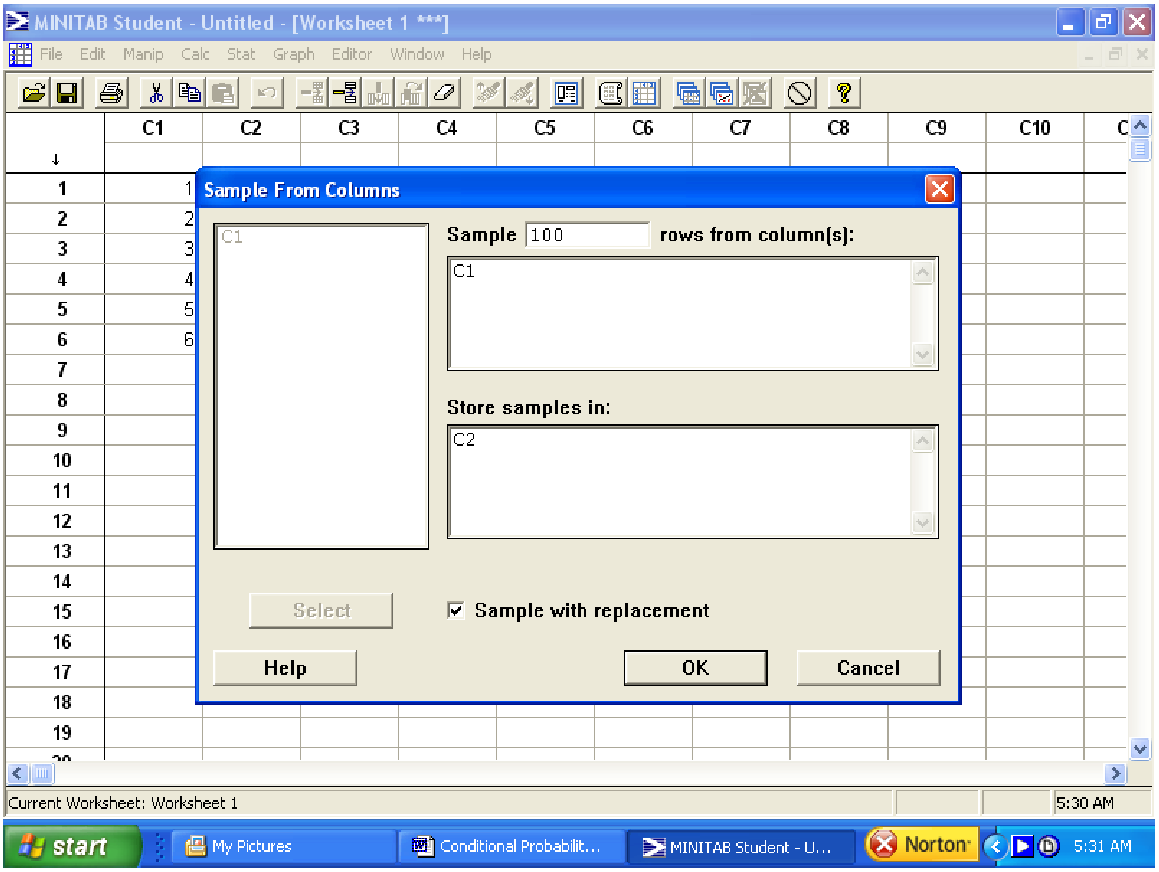
P(A and B) = P(A)*P(B|A). This is the formula for "and" when dealing with events that are not necessarily independent. Recall, if independent events, then P(A and B) = P(A) * P(B). This must allow us to conclude that if events are independent, then P(B|A) = P(B).

**PART I.** Let's illustrate with our previous example. Assume event A = rolling a 6 on the first die and event B = getting a sum of 12 on both die. Using the discussed formula above, predict the answer to P(getting a sum of 12 given that the first die is a 6). Show how you arrive at your answer: ______________________________________________

Now let's simulate the results and calculate our empirical result. Open MINITAB. Down C1, type the numbers 1 through 6 as shown below:

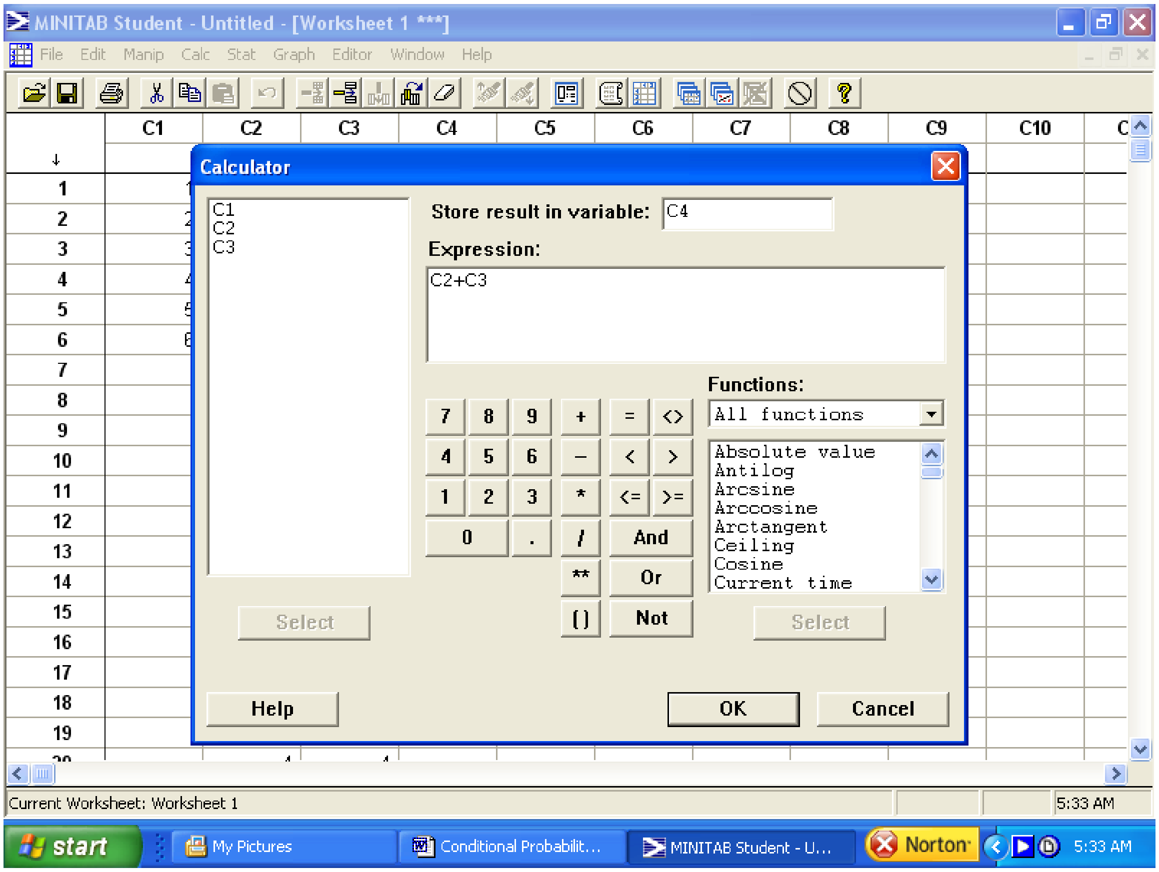



Next, click CALC, RANDOM DATA, SAMPLE FROM COLUMNS, and type the following:

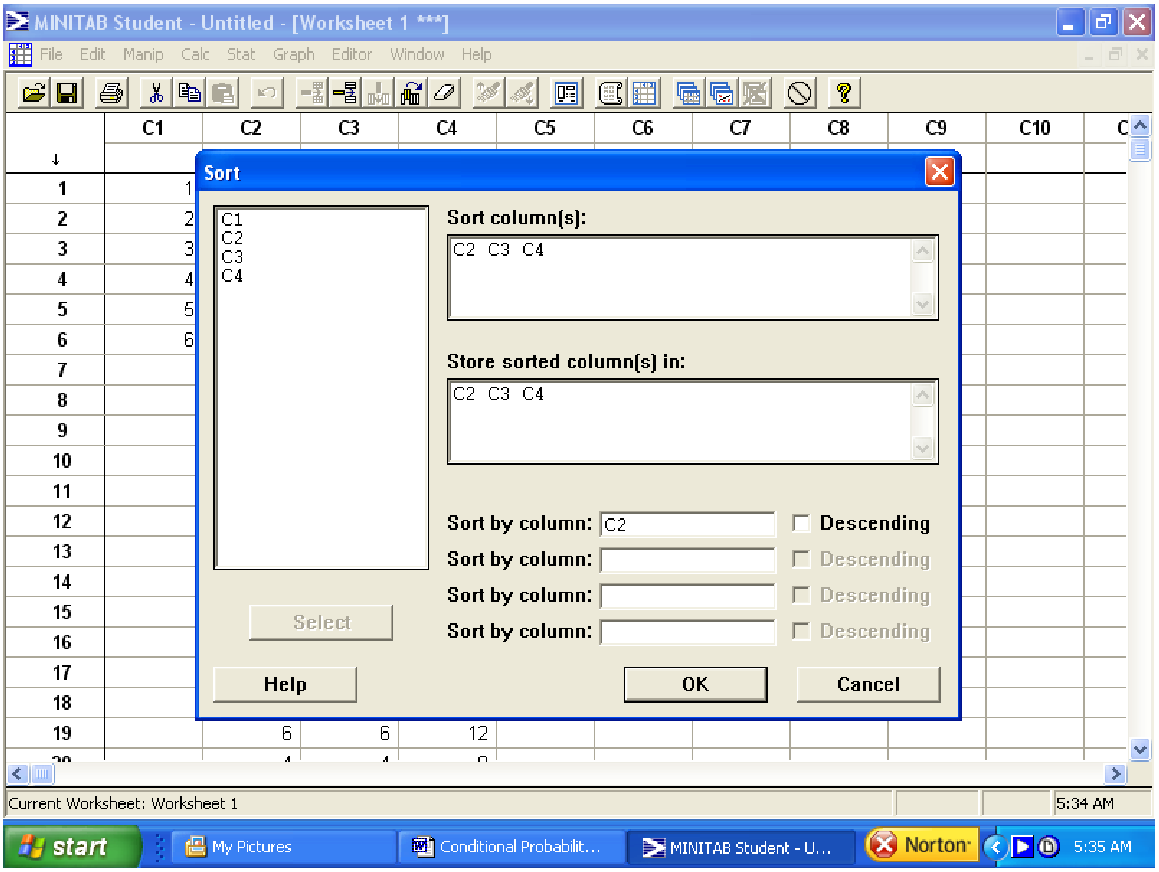



Repeat the previous step but instead of the bottom window storing values in C2, make the C2 say C3 the next time. (DON'T FORGET TO CHECK SAMPLE WITH REPLACEMENT).

You should now have two columns, C2 and C3 which have simulated tossing a die. Now calculate the sums of the two dice by clicking CALC, CALCULATOR and the following:


MINITAB Student - Untitled - [Worksheet 1 ***]
File Edit Manip Calc Stat Graph Editor Window Help
Calculator
C1
C2
C3
Store result in variable: C4
Expression:
C2+C3
Functions:
All functions
Absolute value
Antilog
Arcsine
Arccosine
Arctangent
Ceiling
Cosine
Current time
Select
Select
Help
OK
Cancel
Current Worksheet: Worksheet 1
5:33 AM

Now SORT according to C2 by clicking MANIP, SORT, then the following:




To estimate P(sum of 12| first die is a 6), count how many outcomes had a first roll of 6 (this would mean how many of your C2 entries are 6's). This will become your denominator. The numerator will be how many of your outcomes had both a 6 on the first toss (so C2 is 6) and also had a sum of 12 (so C4 is 12). Using the formula and your counted results, calculate:

$$P(\text{sum of } 12 \mid \text{first toss is a } 6) = \frac{P(\text{first toss is a 6 AND sum is 12})}{\text{P(first toss is a 6)}} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

Do these results agree (come close) with your initial prediction? ____________________

Do you think that $P(A \mid B) = P(B \mid A)$? ____________Why or why not?

______________________________________________________________________________

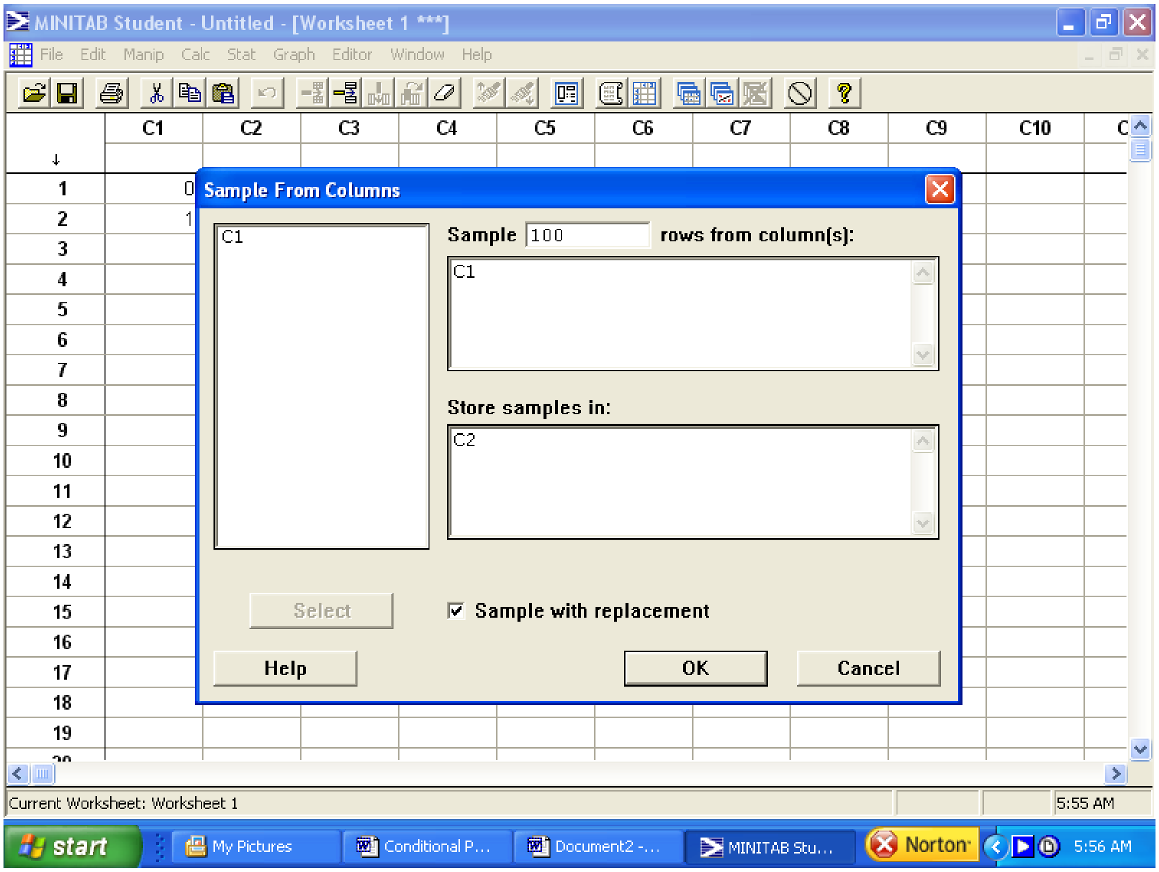
Use your simulated results to show how you would calculate P(B|A). SHOW WORK on how you arrived at your answer: _________________________________

**PART II.** Try the following example. Assume A = getting a heads on the first toss and B = getting a heads on the second toss. Calculate P(A and B) and show work:

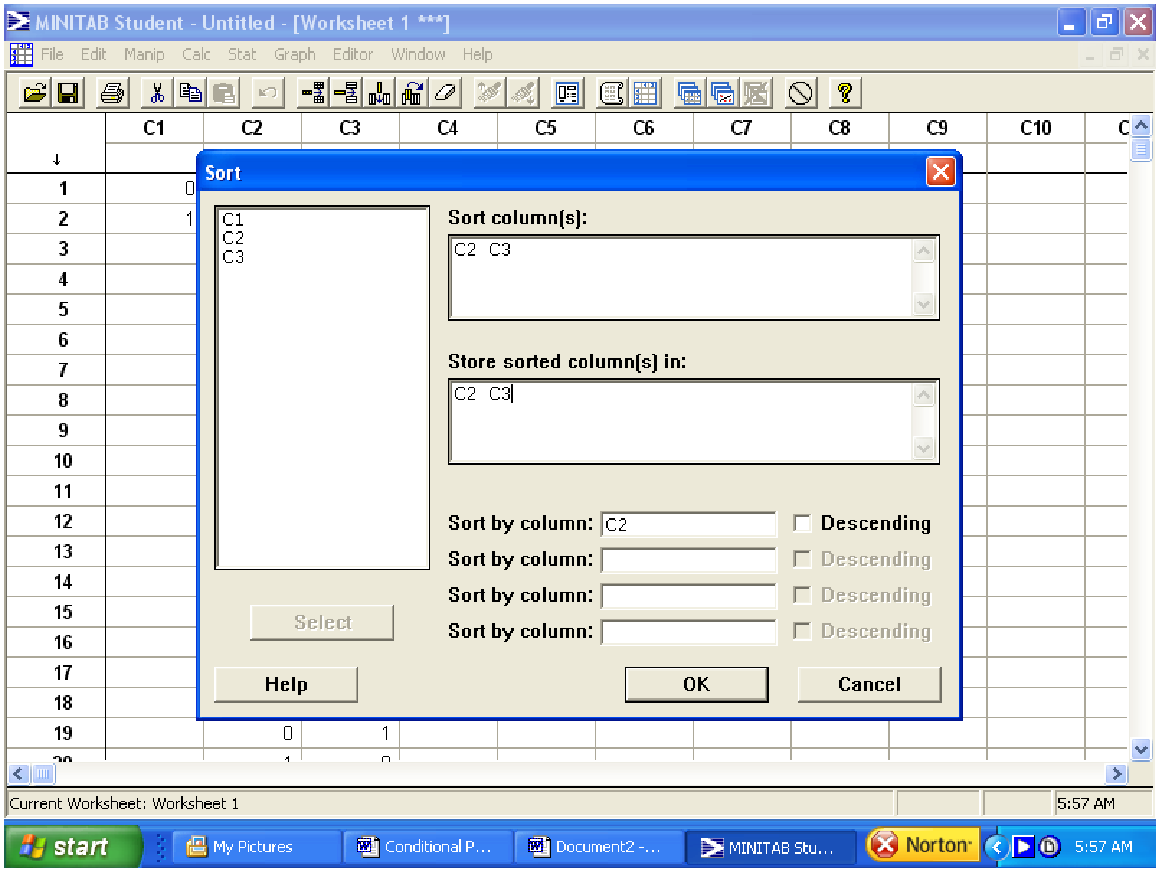
____________________________________________________________

Now calculate P(B|A) and show work: ________________________________________

To simulate, you would type the numbers 0 and 1 down C1. Then click CALC, RANDOM DATA, SAMPLE FROM COLUMNS and type:




Repeat but STORE SAMPLES in C3 rather than C2. Sort according to C2 by clicking MANIP, SORT, then the following:




To find P(A and B), count how many outcomes showed a 0 (Heads) in C2 and 0 in C3. Divide this by 100. __________

To find P(B|A), count how many outcomes showed A and B (a 0 in C2 and a 0 in C3) and divide this by how many outcomes had a 0 in C2. ________________

To find P(B) count how many outcomes have a 0 in C3 (you may want to sort according to C3) and divide this by 100. ____________________

Are your answers to P(B|A) and P(B) the same? ___________ If so, what does this mean? ____________________________________

**Conditional Probability Worksheet** **Name: ______________________ Date: ________**

**<u>Conditional Probability</u>**

Recall from the lab last week, when rolling two dice, let A = rolling a 6 on first die and B = sum of two dice, it appeared that P(A and B) $\neq$ P(A)*P(B). Why do you think this rule did not apply? __________________________________________________

**PART I.** Let's illustrate with our previous example. Assume event A = rolling a 6 on the first die and event B = getting a sum of 12 on both die. Using the discussed formula in the handout, predict the answer to P(getting a sum of 12 given that the first die is a 6). Show how you arrive at your answer:

___________________________________________

To estimate P(sum of 12| first die is a 6), count how many outcomes had a first roll of 6 (this would mean how many of your C2 entries are 6's). This will become your denominator. The numerator will be how many of your outcomes had both a 6 on the first toss (so C2 is 6) and also had a sum of 12 (so C4 is 12). Using the formula and your counted results, calculate:

$$P(sum\ of\ 12\ |\ first\ toss\ is\ a\ 6) = \frac{P(\text{first toss is a 6 AND sum is 12})}{\text{P(first toss is a 6)}} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

Do these results agree (come close) with your initial prediction? __________________

Do you think that $P(A|B) = P(B|A)$? ___________Why or why not?

___________________________________________________________________

Use your simulated results to show how you would calculate P(B|A). SHOW WORK on how you arrived at your answer: ______________________________

**PART II.** Try the following example. Assume A = getting a heads on the first toss and B = getting a heads on the second toss. Calculate P(A and B) and show work:

____________________________________________________

Now calculate P(B|A) and show work: ________________________________________

To find P(A and B), count how many outcomes showed a 0 (Heads) in C2 and 0 in C3. Divide this by 100. __________

To find P(B|A), count how many outcomes showed A and B (a 0 in C2 and a 0 in C3) and divide this by how many outcomes had a 0 in C2. ________________

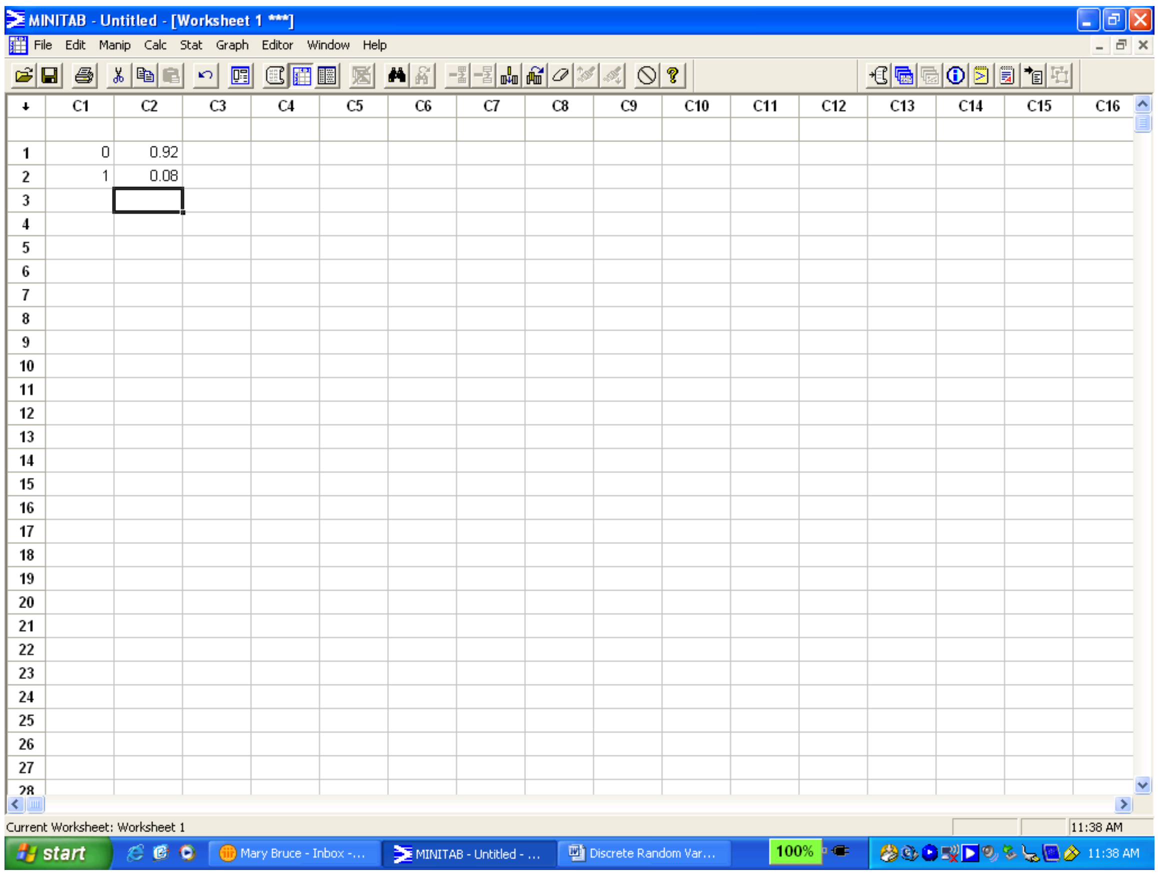
To find P(B) count how many outcomes have a 0 in C3 (you may want to sort according to C3) and divide this by 100. ____________________

Are your answers to P(B|A) and P(B) the same? ___________ If so, what does this mean? ______________________________________

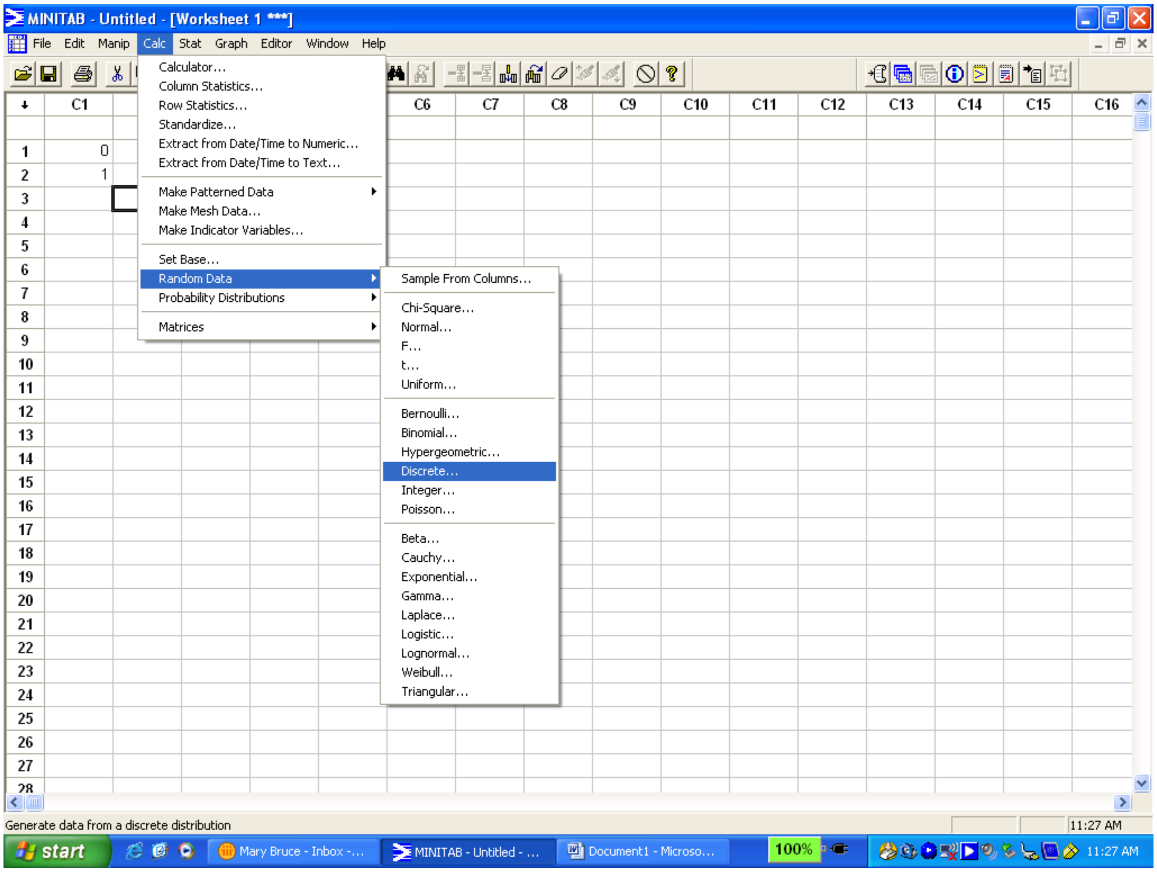
## Discrete Random Variables – Probability Distributions

Now check your predicted answers by doing a simulation.
Open MINITAB.

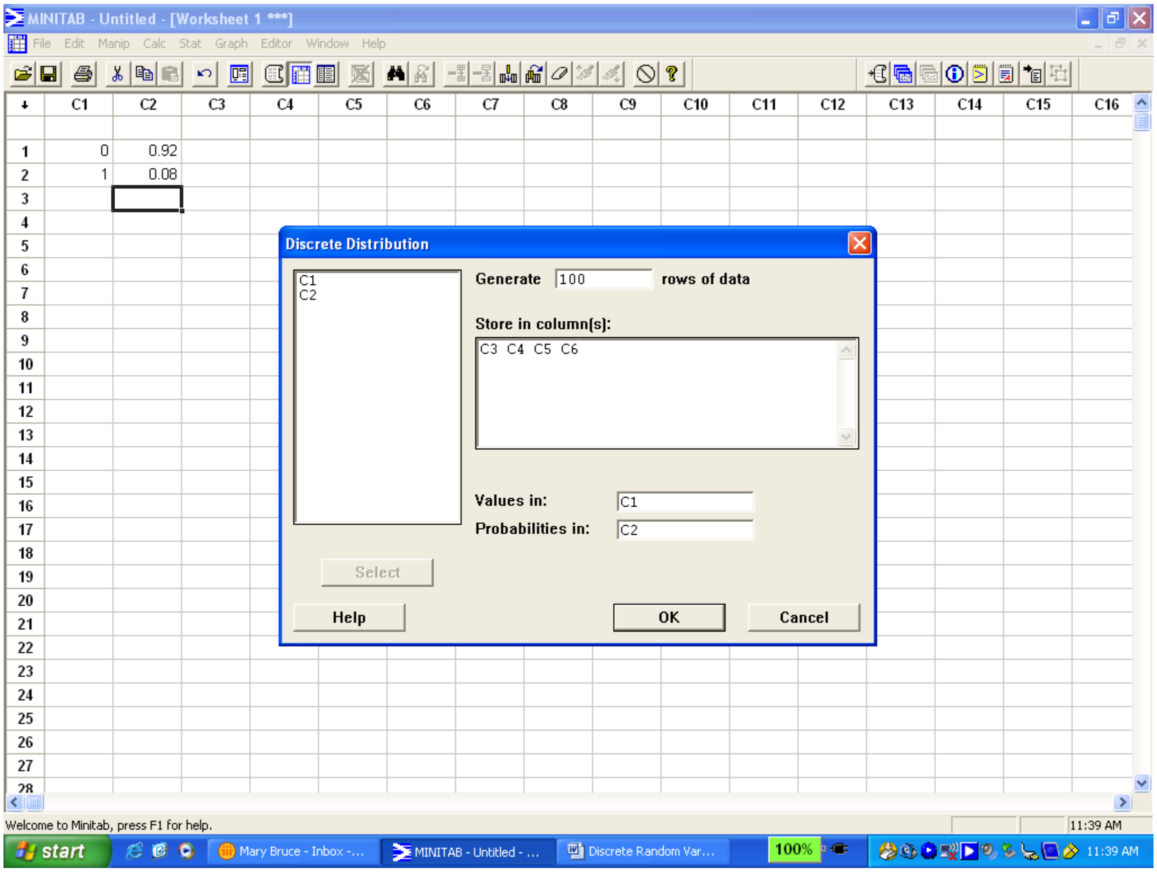
Down C1, type 0 and 1 (0 will represent a non-defective item and 1 will represent a defective item). Down C2 type the probability weights as shown below:



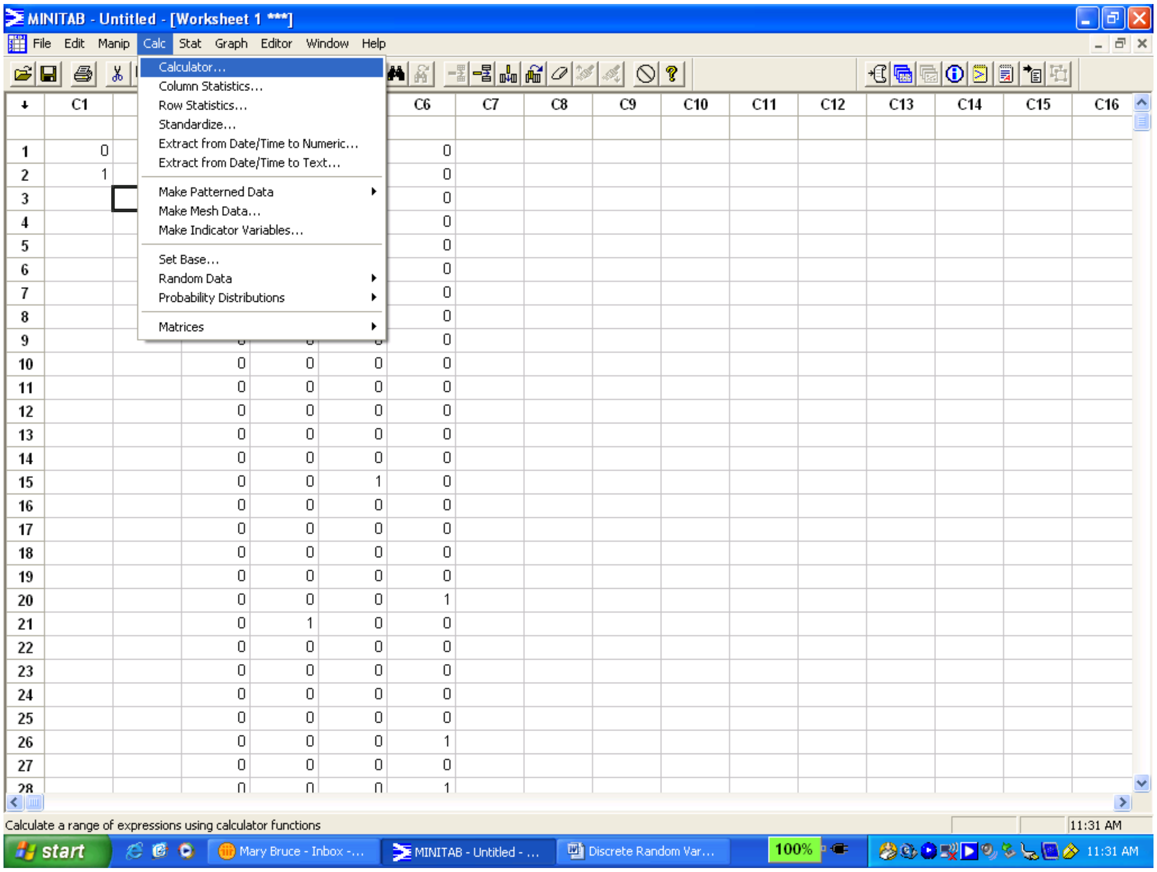

Now simulate a shipment of 4 items by using C3, C4, C5, and C6 and repeating this 100 times. Click CALC, RANDOM DATA, DISCRETE as shown below:


MINITAB - Untitled - [Worksheet 1 ***]
File Edit Manip Calc Stat Graph Editor Window Help
Calculator...
Column Statistics...
Row Statistics...
Standardize...
Extract from Date/Time to Numeric...
Extract from Date/Time to Text...
Make Patterned Data
Make Mesh Data...
Make Indicator Variables...
Set Base...
Random Data
Probability Distributions
Matrices
Sample From Columns...
Chi-Square...
Normal...
F...
t...
Uniform...
Bernoulli...
Binomial...
Hypergeometric...
Discrete...
Integer...
Poisson...
Beta...
Cauchy...
Exponential...
Gamma...
Laplace...
Logistic...
Lognormal...
Weibull...
Triangular...
Generate data from a discrete distribution
11:27 AM

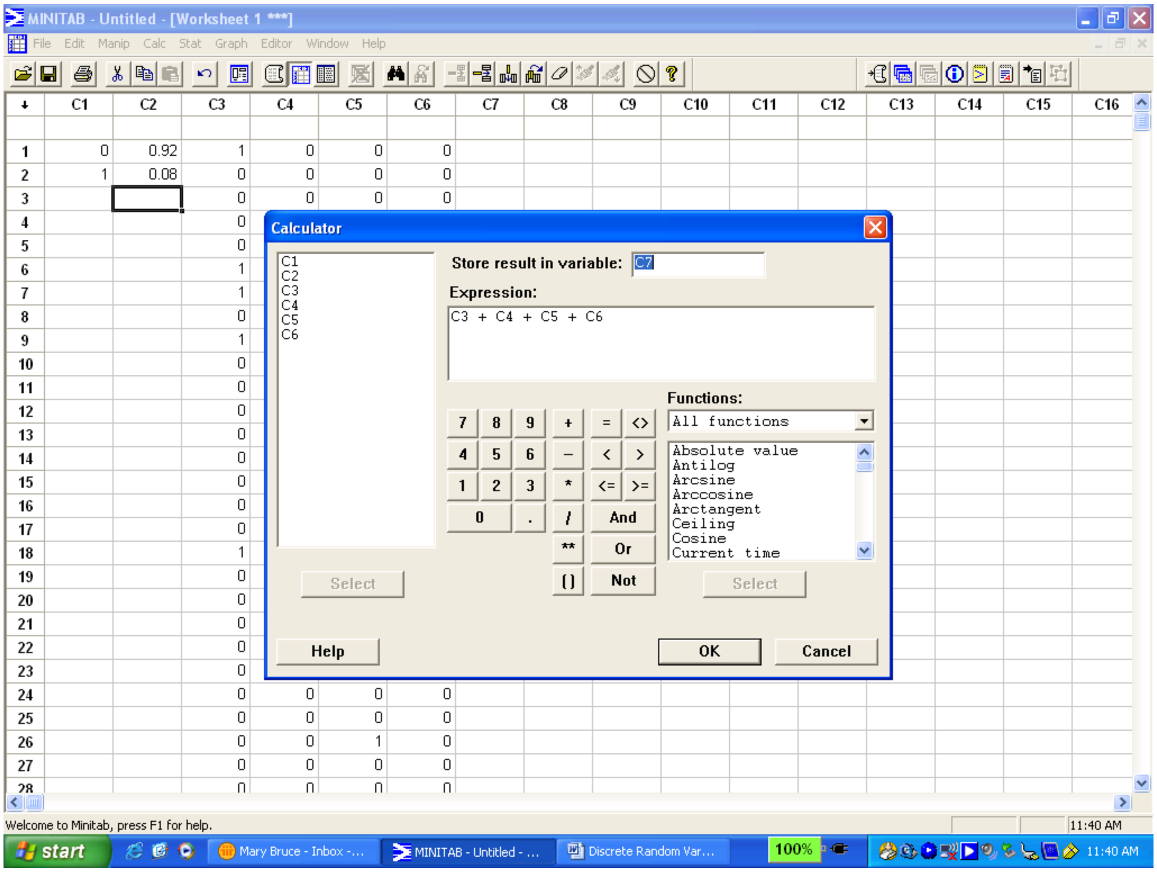
Then when the window appears, type the following:



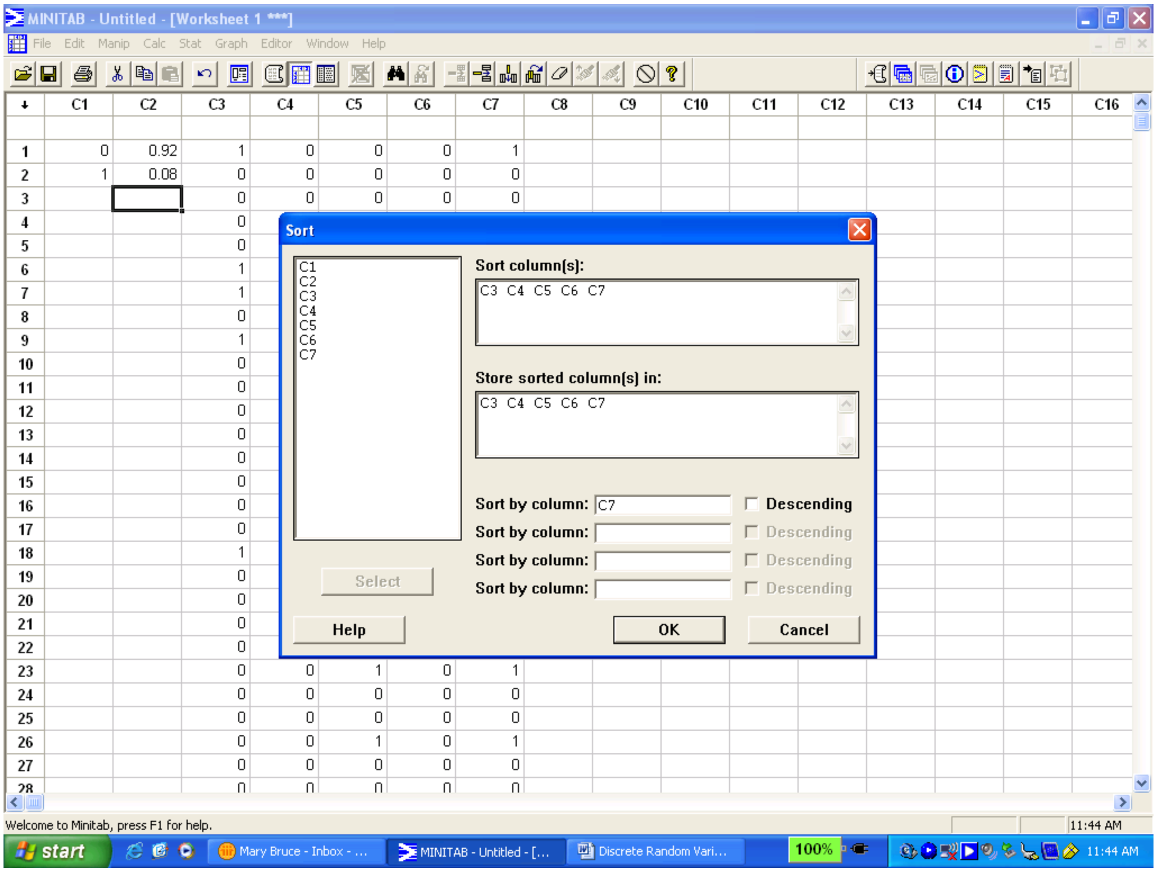

Press OK and you should have 4 columns of 0s and 1s representing the 4 items in the shipment. Since a 1 represents a defective item, then by calculating sums of the 4 columns, we can see how many defective items are in a shipment. To calculate sums, click CALC, CALCULATOR as shown below:




When window appears, type the following:

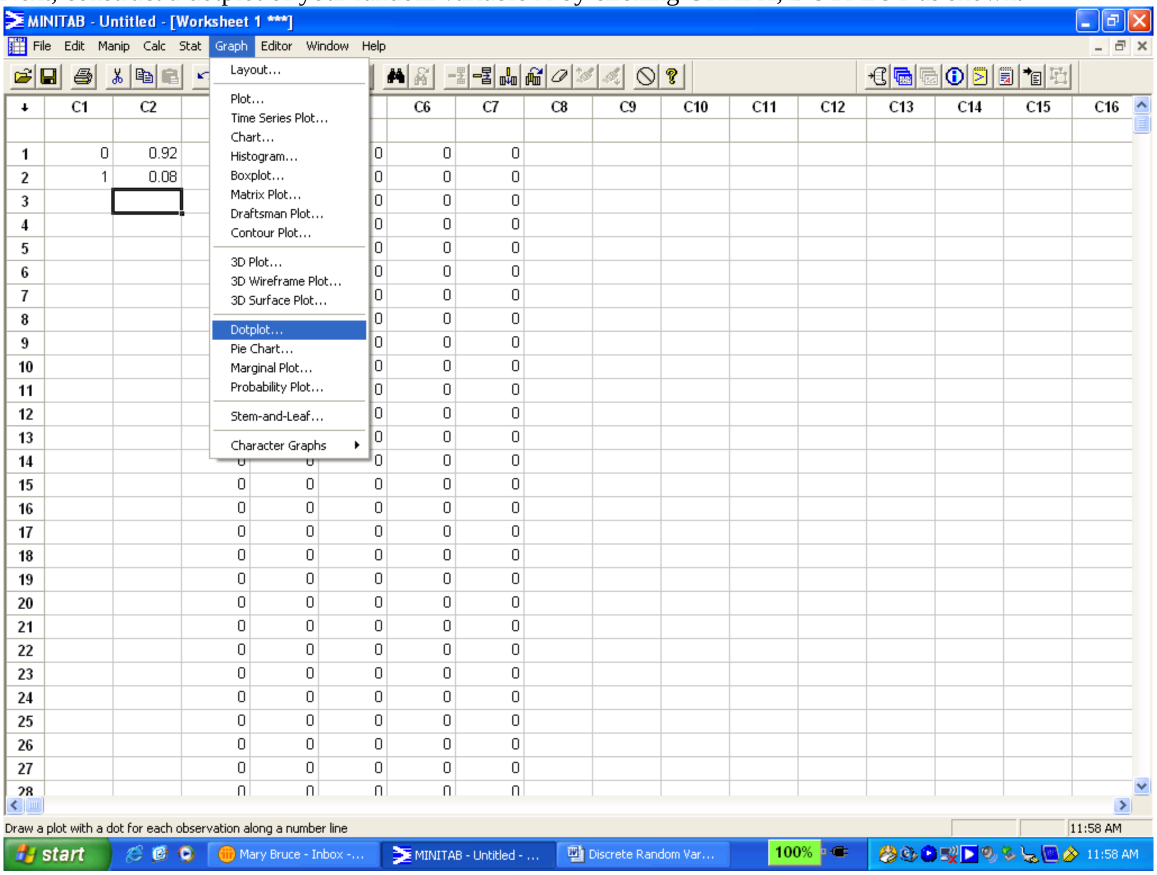



Then Click OK and sums should appear in C7.
Now Sort the sums by clicking MANIP, SORT, and when window appears type the following:

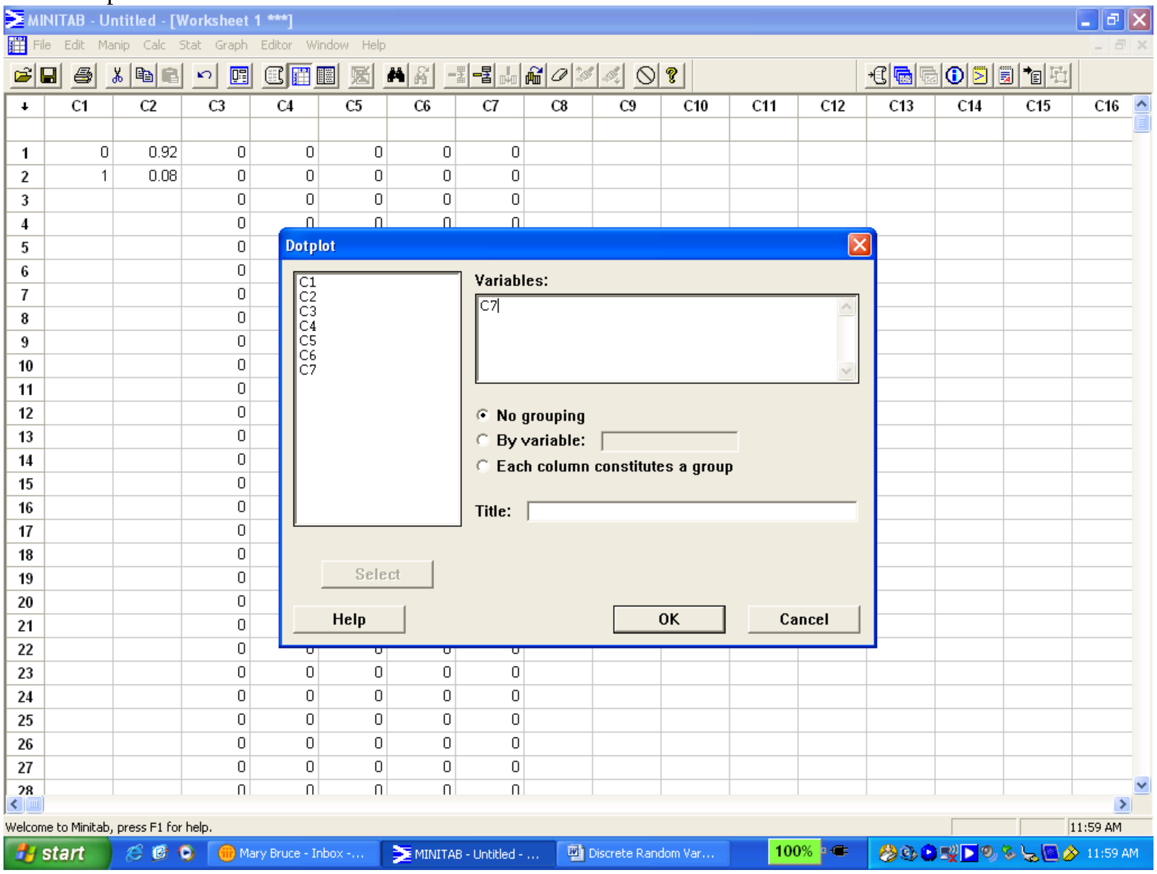



Press OK and calculate the relative frequency or probability (part/whole) of getting X = 2 defective items in the shipment. Compare this to your predicted answer to question #4 before doing the simulation. Now calculate relative frequencies for values of X = 0, X = 1, X = 2, X = 3, X = 4 and give an empirical probability distribution of X using a table (compare this to your predicted probability distribution answer to question #5 before doing the simulation and record result on sheet).

Next, construct a dotplot of your random variable X by clicking GRAPH, DOTPLOT as shown:

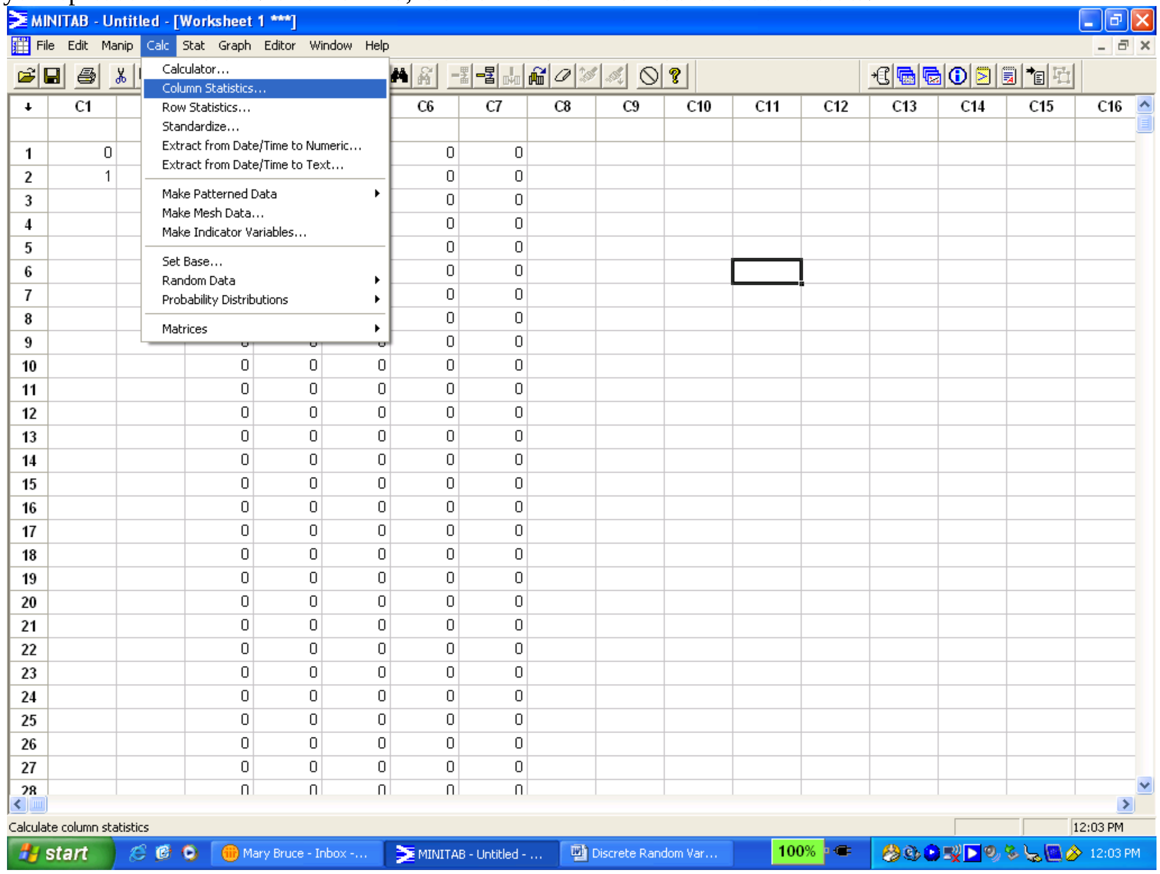


Then complete the window as shown:

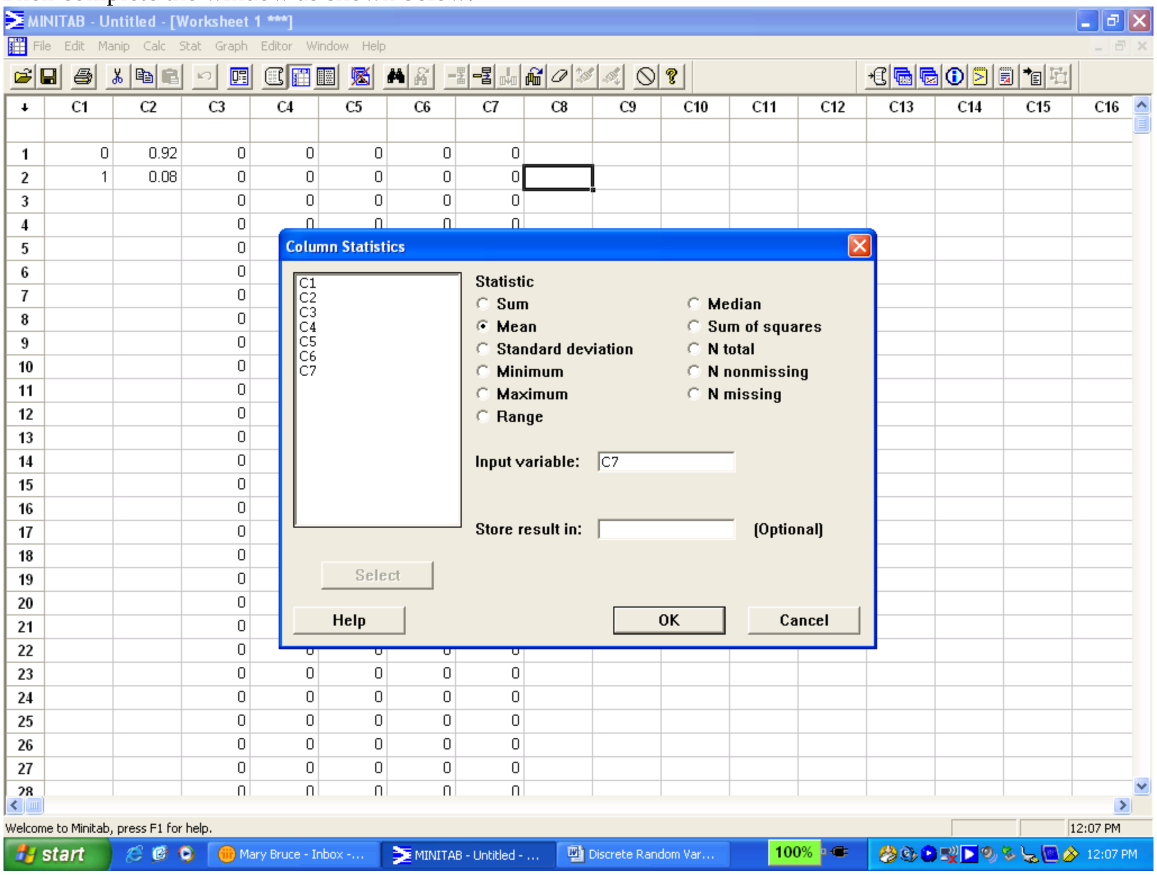



Press OK and use your dotplot to compare predicted answers to questions 6 and 8 before doing the simulation (record your answer on sheet).

To calculate the mean value of X based on your simulated results, minimize the dotplot screen to return to your spreadsheet screen. Click CALC, COLUMN STATISTICS as shown below:



Then complete the window as shown below:




Press OK and compare the mean to your predicted mean in question #7 before the simulation (record result on sheet). Does it make sense that the mean would be this value based on the shape of your probability distribution? EXPLAIN (on sheet).

Based on your simulated probability distribution in question #13, if you were to multiply the X value times the associated probability then add it to the next x value times the probability and do this for all values of X (ie the formula $\sum X \bullet P(X)$) (for example, 0*.68 + 1*.30 + 2*.02 + 3*0 + 4*0 = 0.34) do you get the same answer as the mean you calculated in question #15? (record answer on worksheet). With probability distributions, does each X value then, in a sense, get "weighted" by its probability to give a "weighted mean?" (record answer on worksheet)

**Discrete Random Variables** **Name: ___________________________**
**Probability Distributions** **Date: _________________**

Answer #1 – 8 **BEFORE** doing your simulation.

1. A company has an 8% defect rate. What does this mean?

2. If a company ships 4 items to a customer, how many possible outcomes are there? __________
   List the outcomes:

   How did you arrive at your answer?

3. Letting the random variable X = the number of defective items in the shipment of 4, how many possible values of X are there? _________ List the values: ______________________________________

4. What is the probability that X = 2 assuming independence? _______________ How did you get this answer?

5. Give the theoretical probability distribution of X using a table:

| | |
|---|---|
| Value of X | |
| Probability of X | |

How did you get your answers for each value?

6. What shape would the probability distribution of X have? ______________________How do you know?

7.What is the mean value of X? __________________ How do you know?

8. What is the MOST probable value of X? ________________ How do you know?

**AFTER DOING SIMULATION:** Answer similar questions as you did before but now use your simulated results to answer. Some of your answers may be the same as they were before.

9. A company has an 8% defect rate. Based on your simulated results down each individual column, what does this mean?

10. If a simulated result gives 0010 that means an outcome of NNDN. If a company ships 4 items to a customer, how many different possible outcomes are there? __________ Is this answer the same as you thought before doing the simulation? EXPLAIN.

11. Letting the random variable X = the number of defective items in the shipment of 4, how many possible values of X are there? _________ List the values: ______________________________________
Are these the same answers you got before doing the simulation? EXPLAIN.

12. Based on your empirical results, what is the probability that X = 2? _______________ Is this result close to your answer you got before doing the simulation? EXPLAIN.

13. Based on your simulation, give the empirical probability distribution of X using a table:

| | |
|---|---|
| Value of X | |
| Probability of X | |

Is this close to the same answers you got before doing the simulation? _________

14. What shape does your probability distribution of X have? ______________________Does this agree with your prediction before doing the simulation? EXPLAIN.

15.What is the mean value of X? __________________ Based on the shape of the dotplot, does it make sense that the mean could be this value? EXPLAIN.

16. Based on your simulated probability distribution in question #13, if you were to multiply the X value times the associated probability then add it to the next x value times the probability and do this for all values of X (ie the formula $\sum X \bullet P(X)$) (for example, 0*.68 + 1*.30 + 2*.02 + 3*0 + 4*0 = 0.34) do you get the same answer as the mean you calculated in question #15? ________. With probability distributions, does each X value then, in a sense, get “weighted” by its probability to give a “weighted mean?” _____________

17. What is the MOST probable value of X? ________________ Based on your answer to the previous question, do you think the most probable value of a discrete random variable is always the same as the mean value? EXPLAIN.

18. If you reached any contradictions to your original predictions, do you understand why your predictions may have been incorrect? EXPLAIN.

19. What did the simulation help you learn about probability distributions that you didn't realize before doing the simulation?

20. Give your own explanation of what we mean by a "probability distribution." What does a probability distribution tell us?

APPENDIX J - REPRESENTATIVENESS WORKSHEET

**Representativeness** **Name: ____________________ Date: __________**

**Answer the following BEFORE doing simulation.**

If a fair coin is tossed five times, which of the following ordered sequence of heads (H) and tails (T), if any, is MOST LIKELY to occur?

a. H T H T T
b. T H H H H
c. H T H T H
d. Sequences (a) and (c) are equally likely.
e. All of the above sequences are equally likely.

Why do you feel that you chose the answer that you did? EXPLAIN.

When three fair dice are simultaneously thrown, which of the following results is MOST LIKELY to be obtained?

a. Result 1: A 5, a 3 and a 6 in any order
b. Result 2: Three 5's
c. Result 3: Two 5's and a 3
d. All three results are equally likely.

Why do you feel that you chose the answer that you did? EXPLAIN.

If a fair die is rolled five times, which of the following ordered sequence of results, if any, is MOST LIKELY to occur?

a. 3 5 1 6 2
b. 4 2 6 1 5
c. 5 2 2 2 2
d. Sequences (a) and (b) are equally likely.
e. All of the above sequences are equally likely.

Why do you feel that you chose the answer that you did? EXPLAIN.

**SIMULATION**

Have each one in your group simulate 100 repetitions of the first question and record tally marks of each occurrence. You can get one repetition by clicking MATH PRB RANDINT(0, 1, 5) enter. Then keep pressing enter until you have done 100 repetitions. Let 0 = H and 1 = T so that an outcome of 01011 would be HTHTT. You may get a lot of outcomes that don't fit one of the 3 patterns so just don't put a tally mark in that case.

<u>HTHTT</u> <u>THHHH</u> <u>HTHTH</u>

Once you have recorded tally marks, based on your INDIVIDUAL results, calculate empirical probabilities for each outcome by dividing the number of tally marks by 100. Do you still agree with your answer you put BEFORE the simulation? EXPLAIN.

HTHTT probability =

THHHH probability =

HTHTH probability =

Now compare your individual results with others in the group. Did you get similar results?

Calculate probabilities of each outcome by pooling your tally totals with other group members and now dividing by 400 (if 4 in a group) or 500 (if 5 in a group) etc.

HTHTT probability =

THHHH probability =

HTHTH probability =

Do you still agree with your answer you put BEFORE the simulation? EXPLAIN.

Do you think your pooled results or individual results are closer to the theoretical probability? EXPLAIN. See if you can calculate theoretical probability for each outcome.

Have each one in your group simulate 100 repetitions of the second question and record tally marks of each occurrence. You can get one repetition by clicking MATH PRB RANDINT(1, 6, 3) enter. Then keep pressing enter until you have done 100 repetitions. You may get a lot of outcomes that don't fit one of the 3 patterns so just don't put a tally mark in that case.

| 536 (any order) | 555 (any order) | 553 (any order) |
|---|---|---|

Once you have recorded tally marks, based on your INDIVIDUAL results, calculate empirical probabilities for each outcome by dividing the number of tally marks by 100. Do you still agree with your answer you put BEFORE the simulation? EXPLAIN.
536 probability =

555 probability =

553 probability =

Now compare your individual results with others in the group. Did you get similar results?

Calculate probabilities of each outcome by pooling your tally totals with other group members and now dividing by 400 (if 4 in a group) or 500 (if 5 in a group) etc.
536 probability =

555 probability =

553 probability =

Do you still agree with your answer you put BEFORE the simulation? EXPLAIN.

Do you think your pooled results or individual results are closer to the theoretical probability? EXPLAIN. See if you can calculate theoretical probability for each outcome.

Have each one in your group simulate 100 repetitions of the third question and record tally marks of each occurrence. You can get one repetition by clicking MATH PRB RANDINT(1, 6, 5) enter. Then keep pressing enter until you have done 100 repetitions. You may get a lot of outcomes that don't fit one of the 3 patterns so just don't put a tally mark in that case.

| 35162 (in order) | 42615 (in order) | 52222 (in order) |
|---|---|---|

Once you have recorded tally marks, based on your INDIVIDUAL results, calculate empirical probabilities for each outcome by dividing the number of tally marks by 100. Do you still agree with your answer you put BEFORE the simulation? EXPLAIN.

35162 probability =

42615 probability =

52222 probability =

Now compare your individual results with others in the group. Did you get similar results?

Calculate probabilities of each outcome by pooling your tally totals with other group members and now dividing by 400 (if 4 in a group) or 500 (if 5 in a group) etc.

35162 probability =

42615 probability =

52222 probability =

Do you still agree with your answer you put BEFORE the simulation? EXPLAIN.

Do you think your pooled results or individual results are closer to the theoretical probability? EXPLAIN. See if you can calculate theoretical probability for each outcome.

What do your results from this simulation exercise say about short-run behavior versus long-run behavior? EXPLAIN.

APPENDIX K - LAW OF LARGE NUMBERS LAB AND WORKSHEET

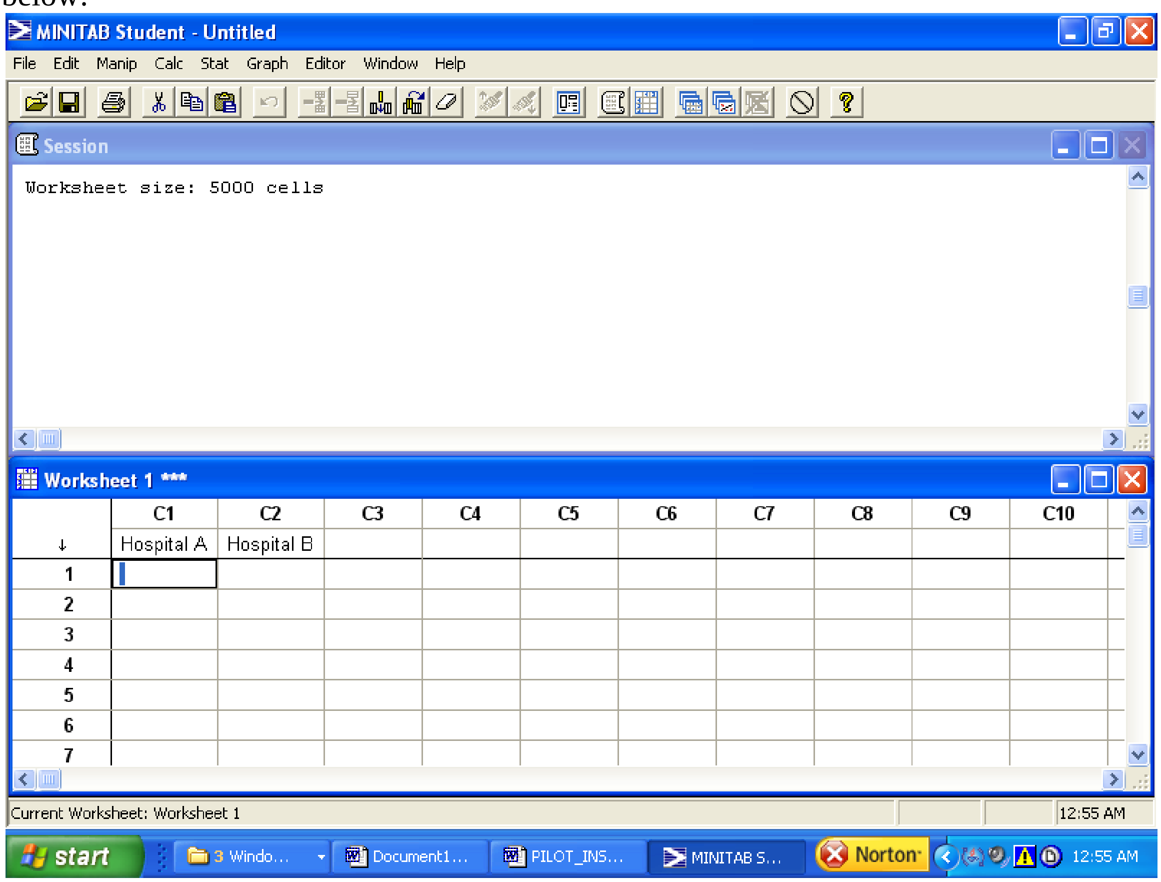
**Simulation**
**Law of Large Numbers**

Recall the problem from yesterday:
Half of all newborn children are girls and half are boys. Hospital A records an average of 50 births a day. Hospital B records an average of 10 births a day. On a particular day, which hospital is more likely to record 80% or more female births?

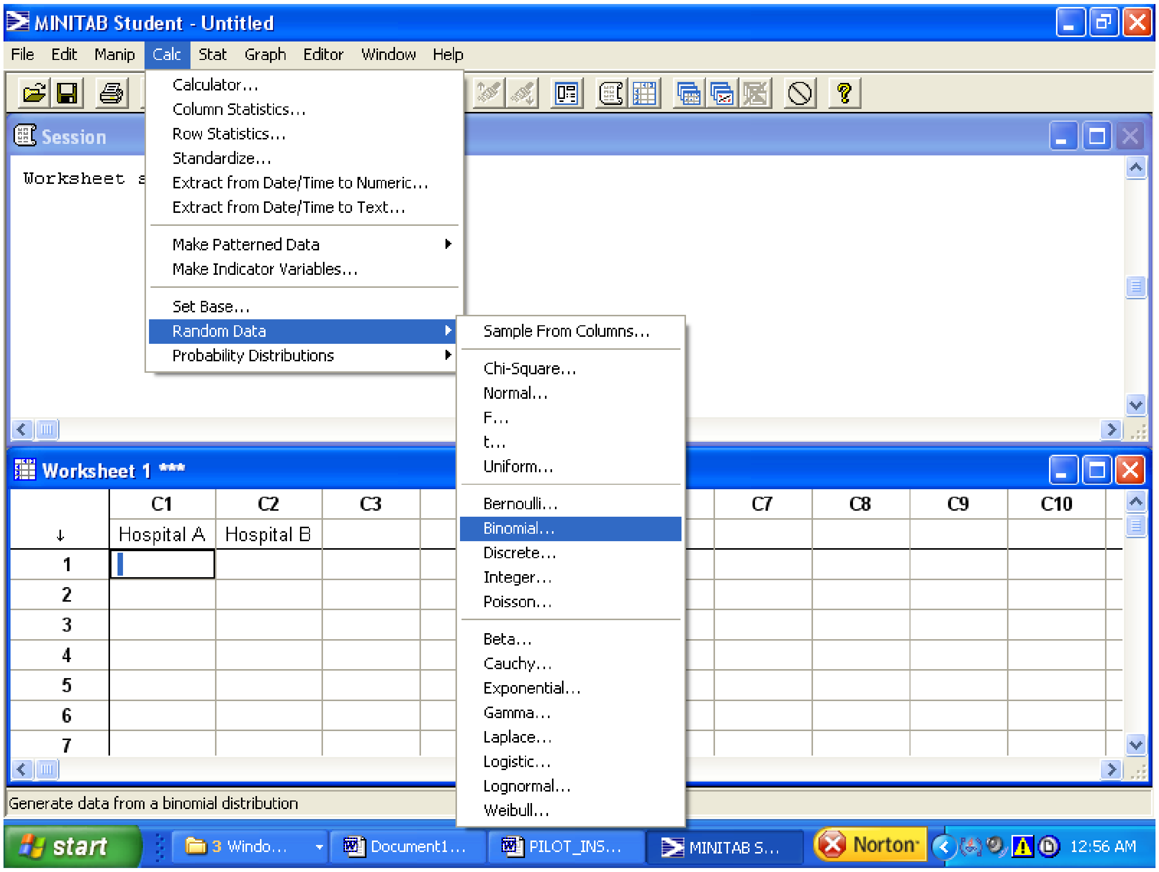
(A) Hospital A (with 50 births a day)
(B) Hospital B (with 10 births a day)
(C) The two hospitals are equally likely to record such an event

This is a binomial random variable. We can get the computer to simulate births from both hospitals by randomly generating how many girls out of 10 would be born on a given day and how many births out of 50 would be born on a given day. We will do 100 repetitions imitating what could happen on 100 different days given both situations.

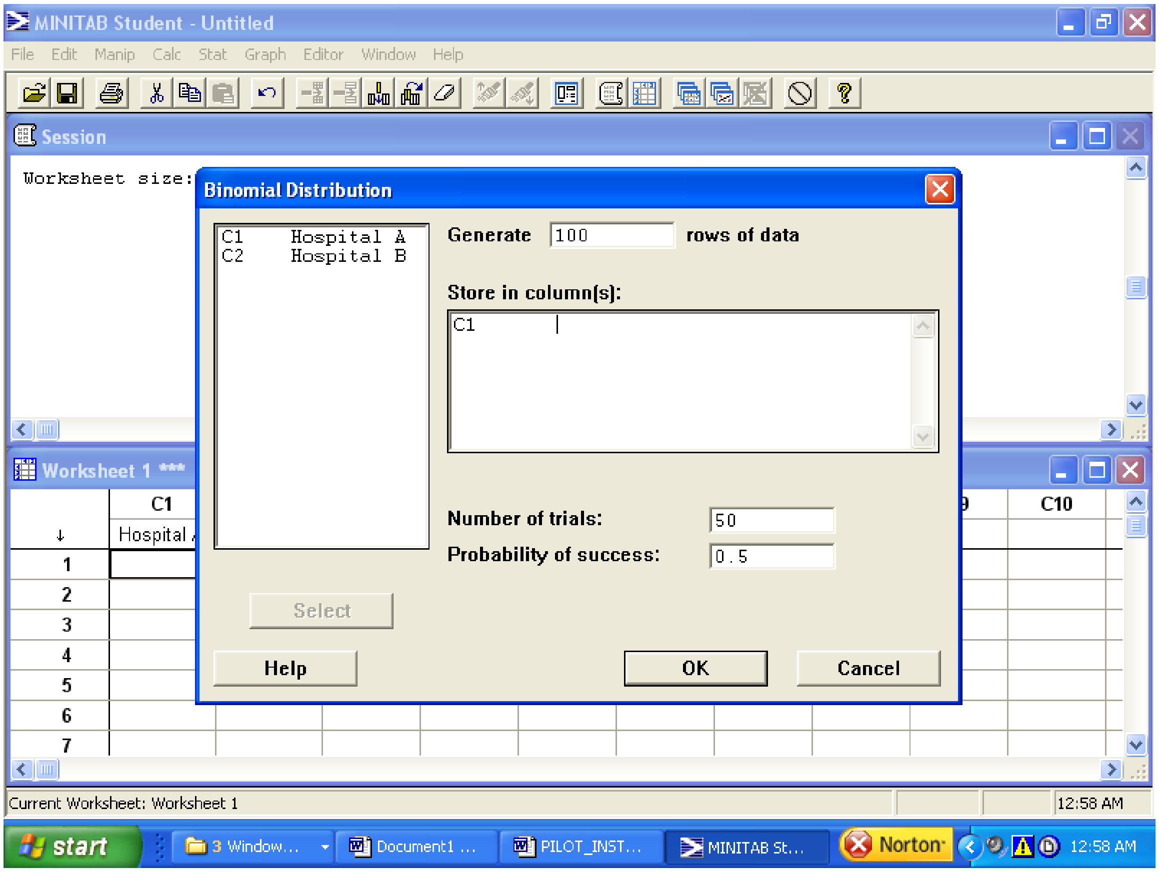
Open MINITAB. Type Hospital A and Hospital B in the column headings as shown below:



To simulate Hospital A (with 50 births a day), click CALC, RANDOM DATA, BINOMIAL as shown below:

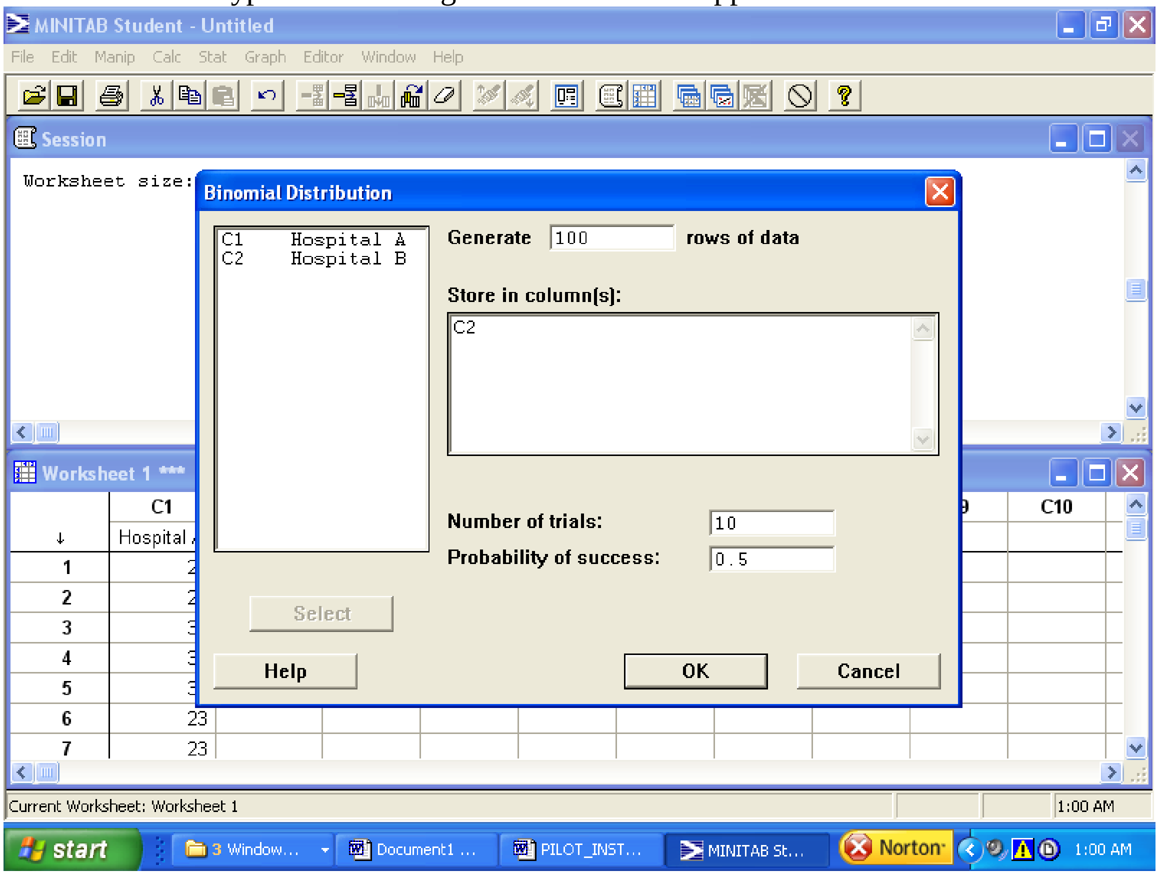


When the window appears, type the following for 50 births and a 0.5 probability of success:

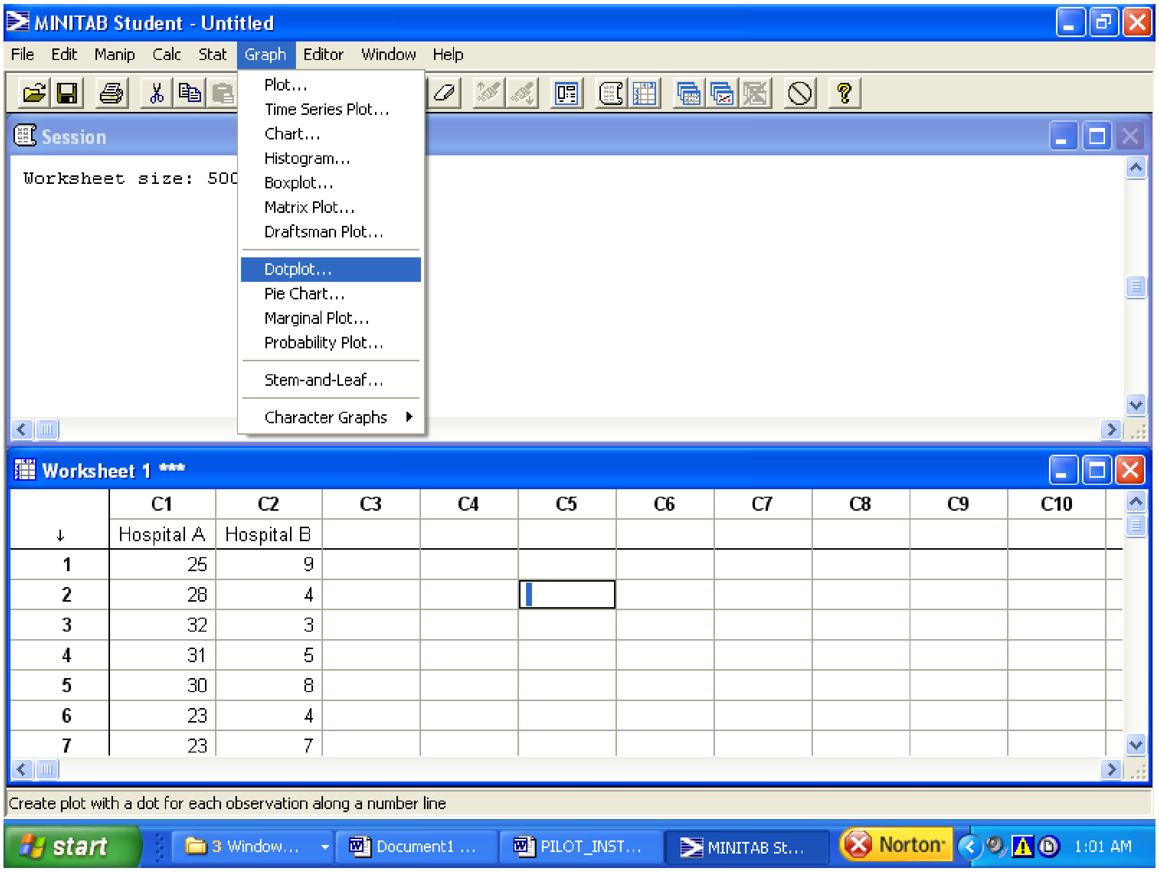



Press OK. You should end up with numbers in C1 that represent how many girls are born that day out of the 50 births.

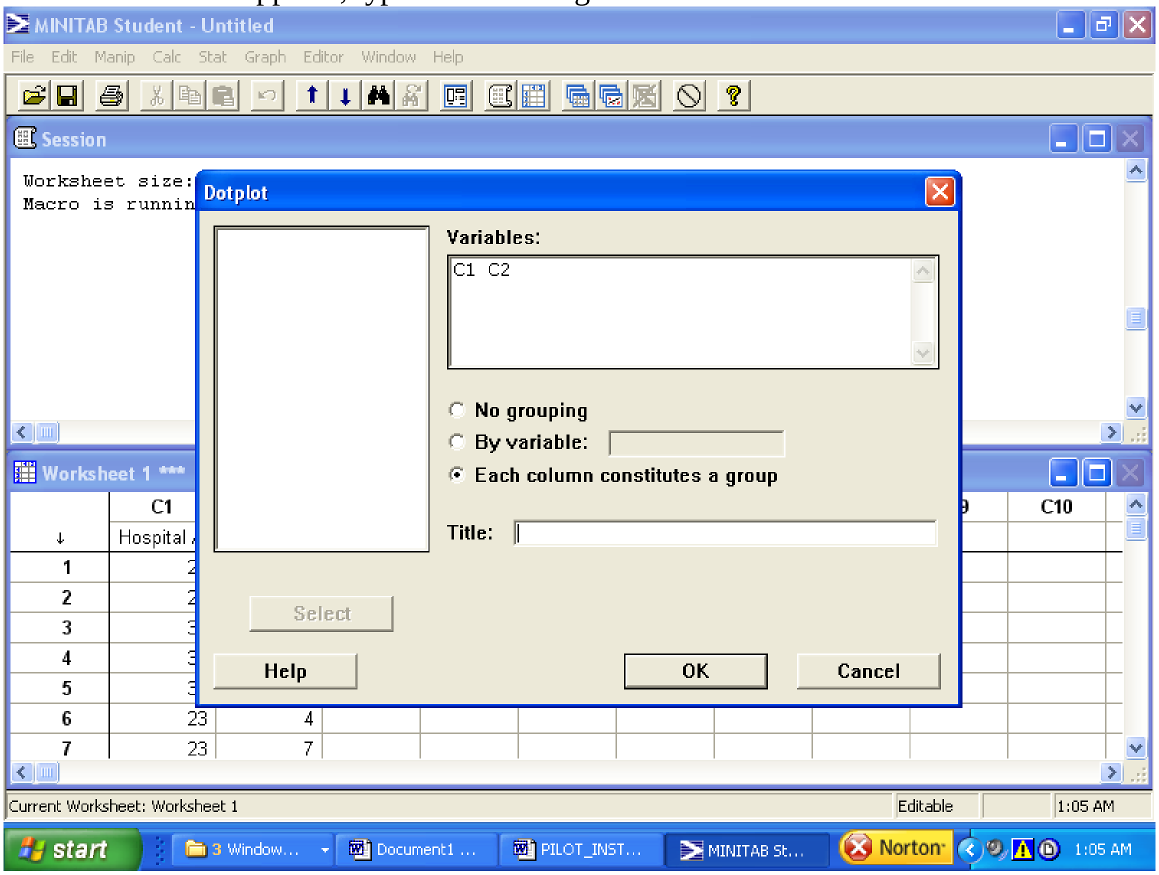
Now to simulate Hospital B (with 10 births a day), click CALC, RANDOM DATA, BINOMIAL and type the following when the window appears:



Then press OK. To analyze dotplots of the data, click GRAPH, DOTPLOT as shown below:



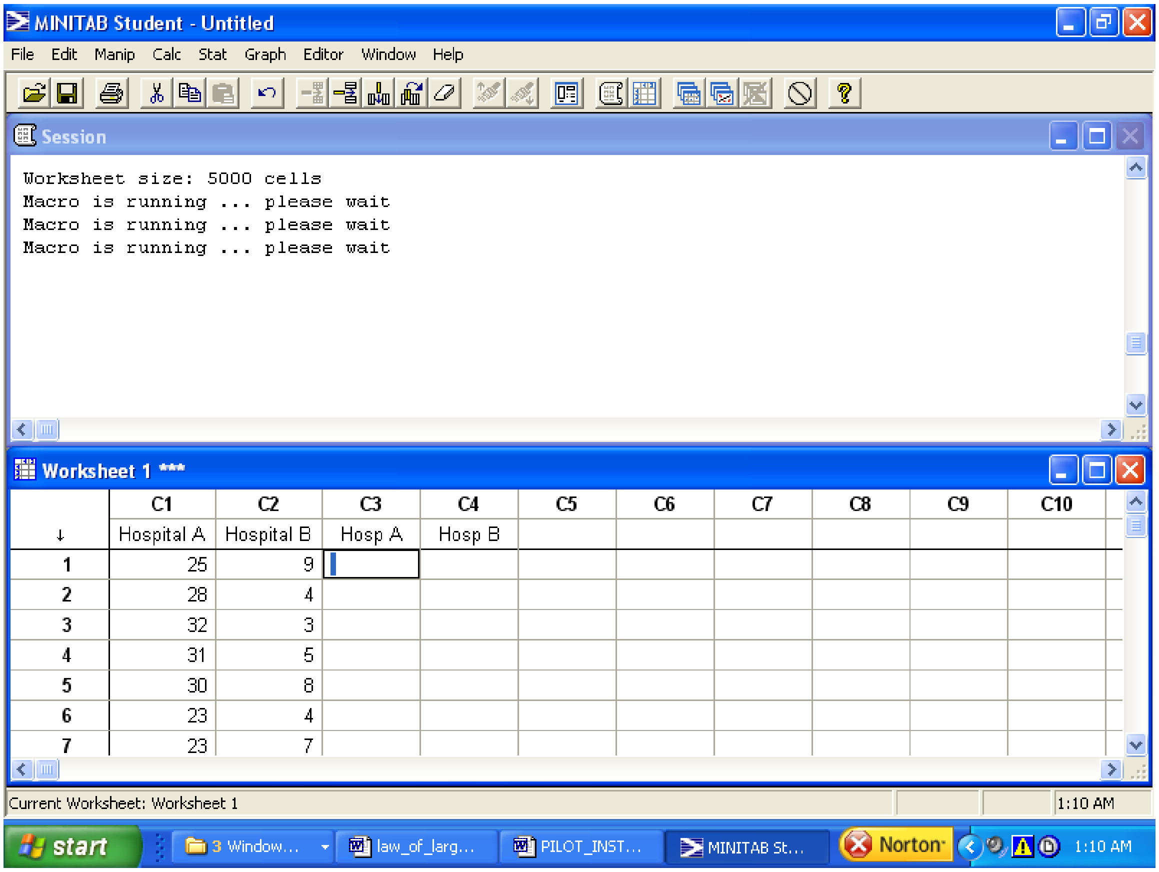
When the window appears, type the following:



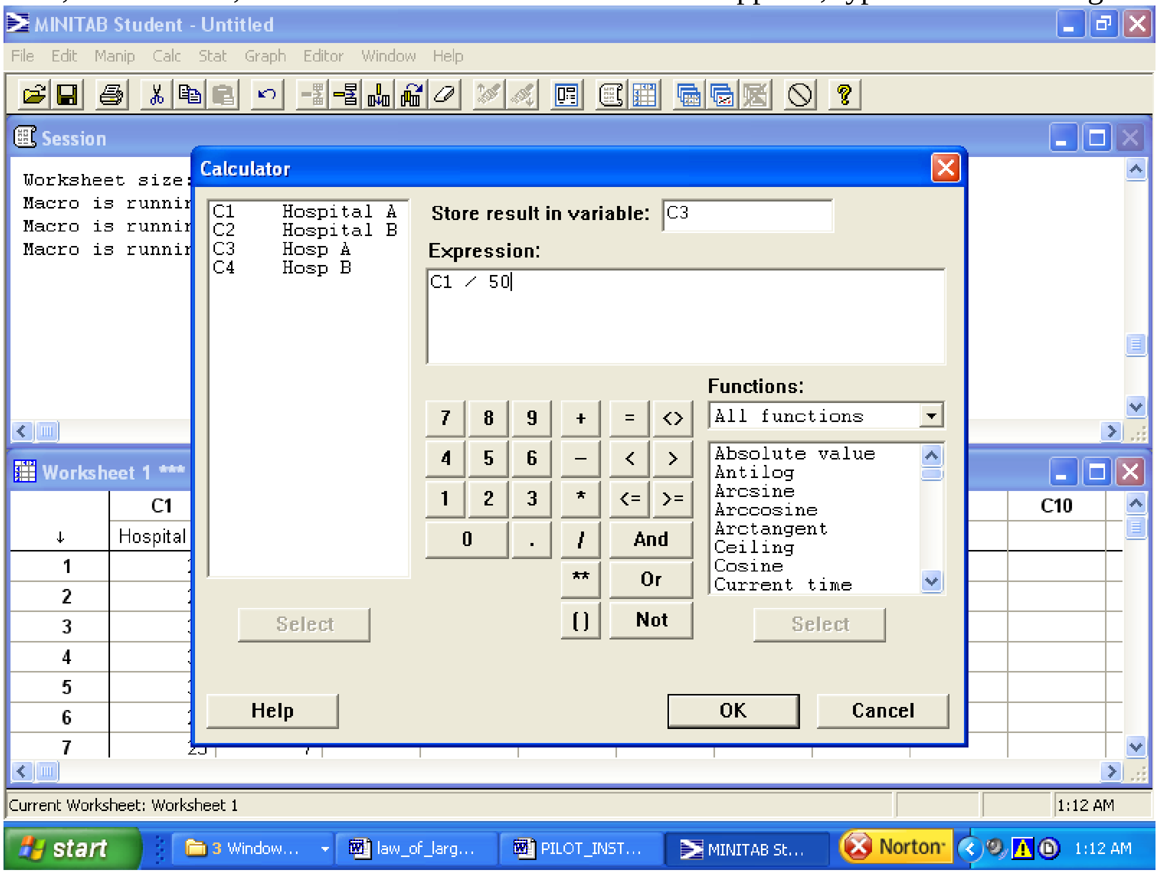

Then press OK.
Now go to your worksheet and answer questions.

Next, we will convert the values of our random variable into relative frequencies (probabilities) by dividing them by 50 (for 50 total babies) and 10 (for 10 total babies). Remember the theoretical probability for having a boy or girl is 0.50.

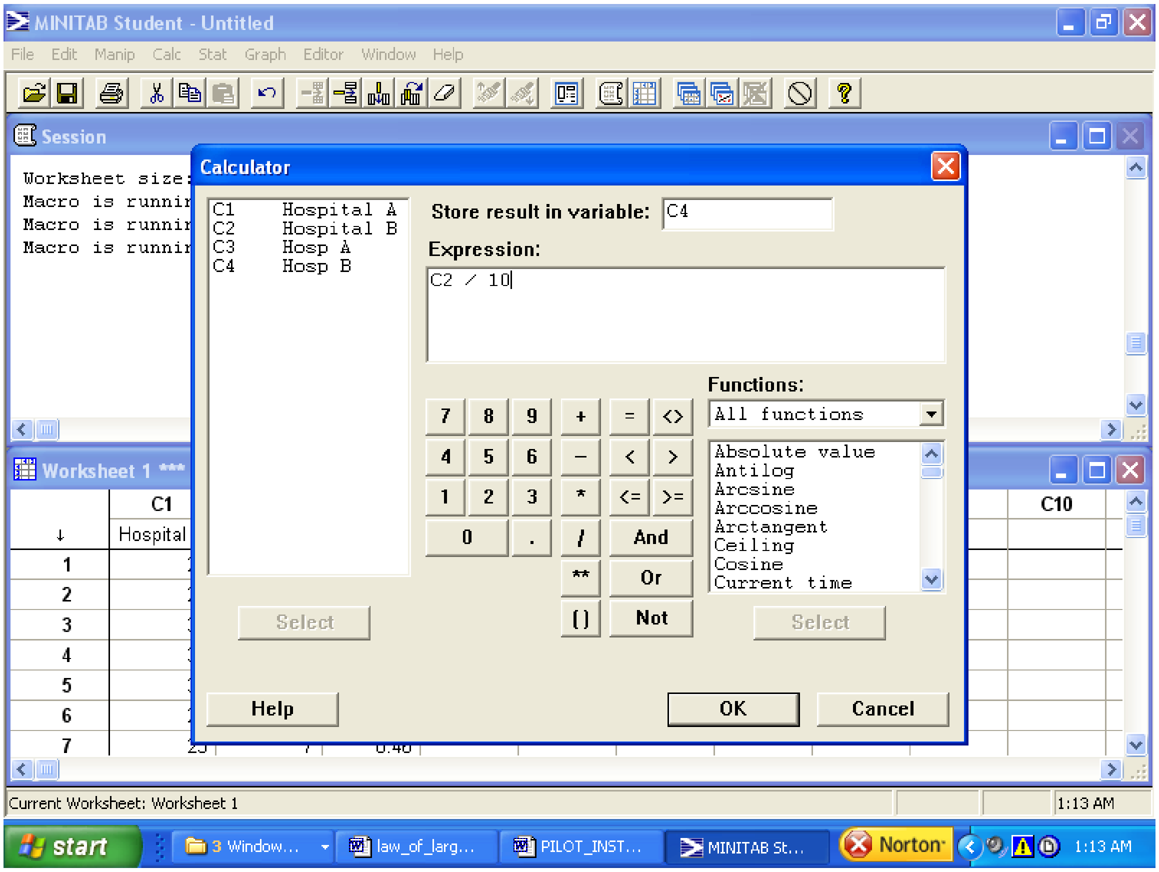
Type in column headings in C3 and C4 as shown:


MINITAB Student - Untitled
File Edit Manip Calc Stat Graph Editor Window Help
Session
Worksheet size: 5000 cells
Macro is running ... please wait
Macro is running ... please wait
Macro is running ... please wait
Worksheet 1 ***
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10
Hospital A Hospital B Hosp A Hosp B
1 25 9
2 28 4
3 32 3
4 31 5
5 30 8
6 23 4
7 23 7
Current Worksheet: Worksheet 1
1:10 AM
start
3 Window...
law_of_larg...
PILOT_INST...
MINITAB St...
Norton
1:10 AM

Next, click CALC, CALCULATOR. When the window appears, type in the following:

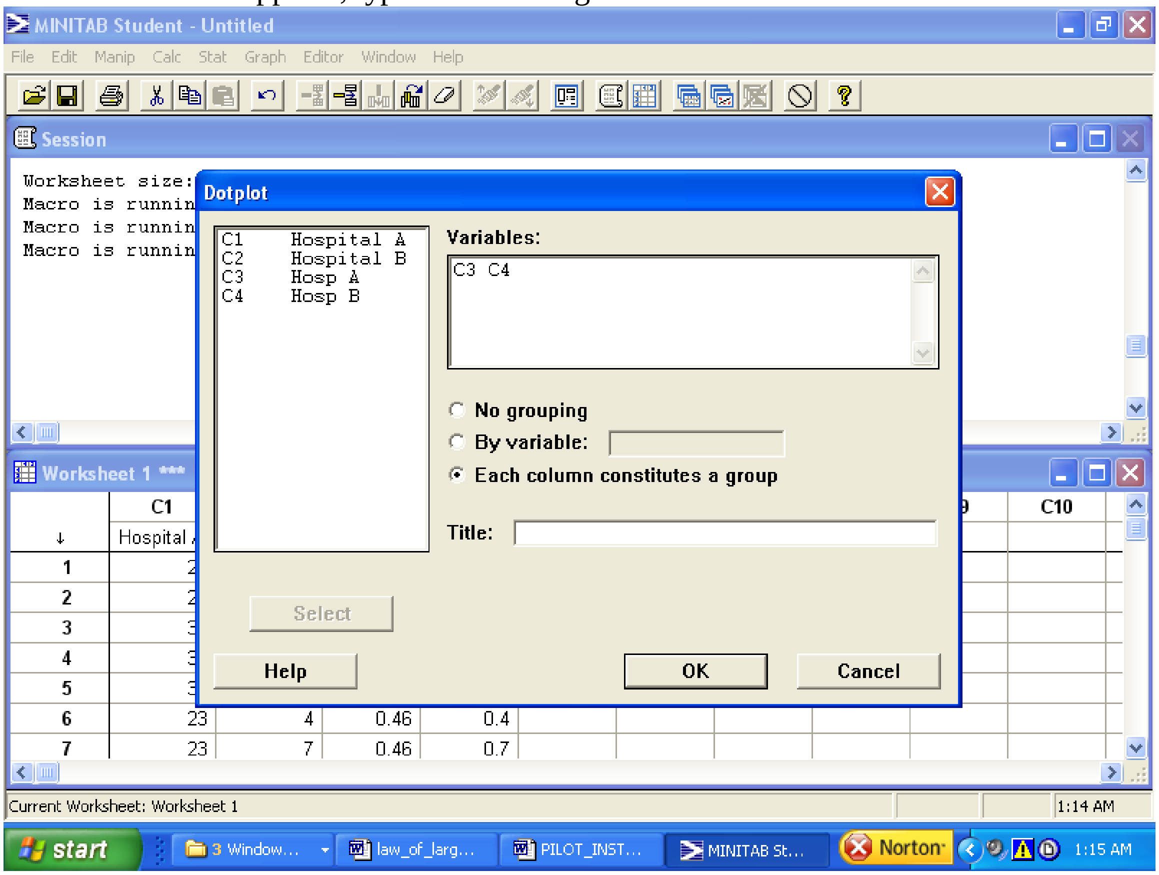



Then press OK.
Next click CALC, CALCULATOR. Then when the window appears, type the following:




Then press OK.
Now do dotplots of the relative frequencies by clicking GRAPH, DOTPLOT.

When the window appears, type the following:




Then press OK.
Now go to worksheet and answer the questions.

**Simulation** **Name: _________________Date: _____________**
**Law of Large Numbers**

Compare dotplots of C1 and C2, and letting X = number of girls, which Hospital shows a higher variance in terms of the random variable? ________________Explain why you put this answer:
____________________________________________________________________________________
__________________________________________________________ The formula for variance of a binomial random variable is $Variance = \sqrt{np(1-p)}$ where
n = total number of trials and p = probability of having a girl in this case. Since both hospitals have p = 0.5, does it make sense that the hospital with the larger number of births would show a larger spread? EXPLAIN.
____________________________________________________________________________________
__________________________________________________________

The mean of a random variable is $\mu = np$.
Calculate the mean of Hospital A with 50 births. _______________ Now calculate the mean of Hospital B with 10 births ______________.

Does it make sense that the means would be these two numbers? EXPLAIN.

Were both hospitals centered at their means? _______________

What shape did both distributions show? ________________________

Does it make sense why the distributions would show this shape? EXPLAIN.

Now comparing the dotplots for C3 and C4 after converting to relative frequencies (probabilities), which hospital shows results closer to the theoretical probability?
_______________________________________________________________

Can you explain why?

If you had to predict the probability of having a girl from taking a sample from one of the hospitals, which hospital would you pick? ____________________
Why?

If you answered the question about the hospitals now, would your answer be different than what you said yesterday? ______________

Explain what you have learned from doing the simulation:

APPENDIX L - BINOMIAL PROBABILITY DISTRIBUTIONS
LAB AND WORKSHEET

**Binomial Probability Distributions**

Recall conditions of a binomial distribution:

(a) each trial must have one of two possible outcomes called "success" and "failure"
(b) each trial is independent of the others
(c) there is a fixed number of trials
(d) the probability of success is the same on each trial

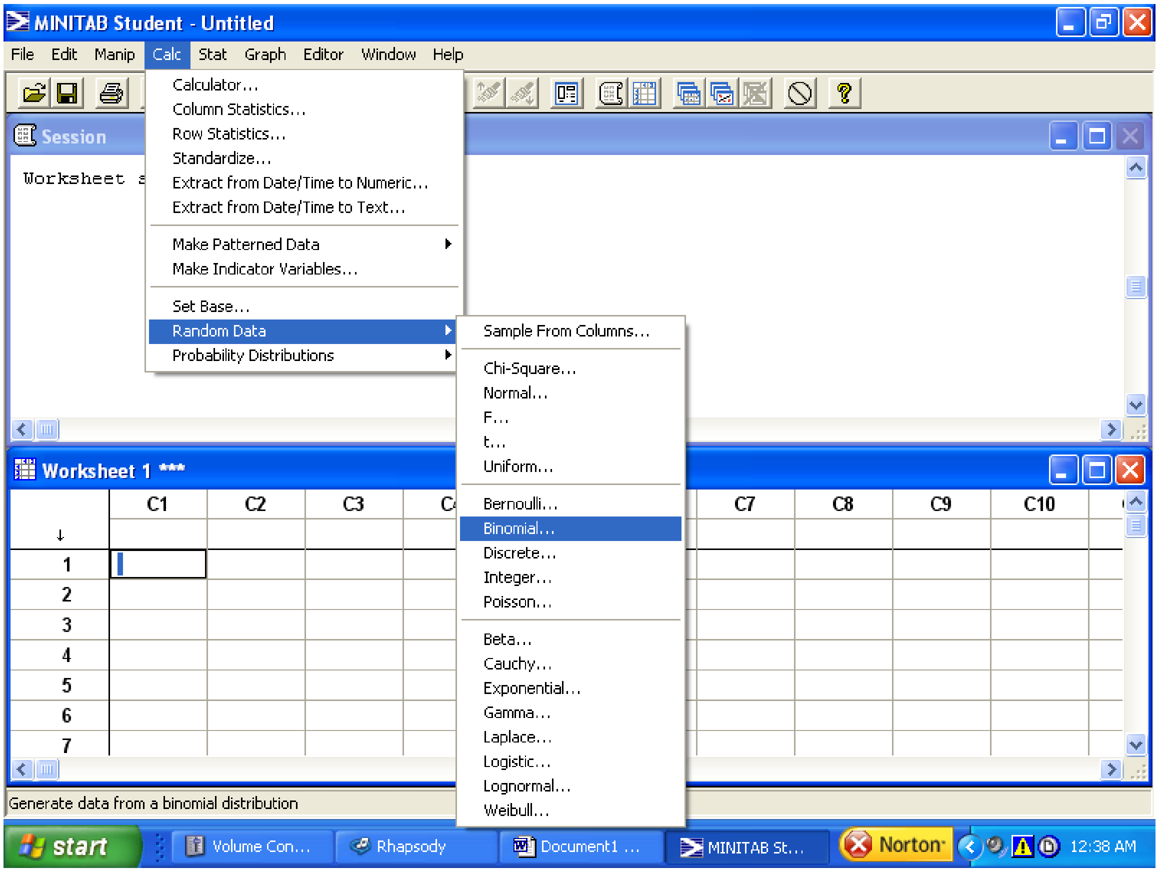
Recall from Friday's lab that the mean of a binomial is $\mu = np$ and the standard deviation is $\sigma = \sqrt{np(1-p)}$

Given the following scenarios,

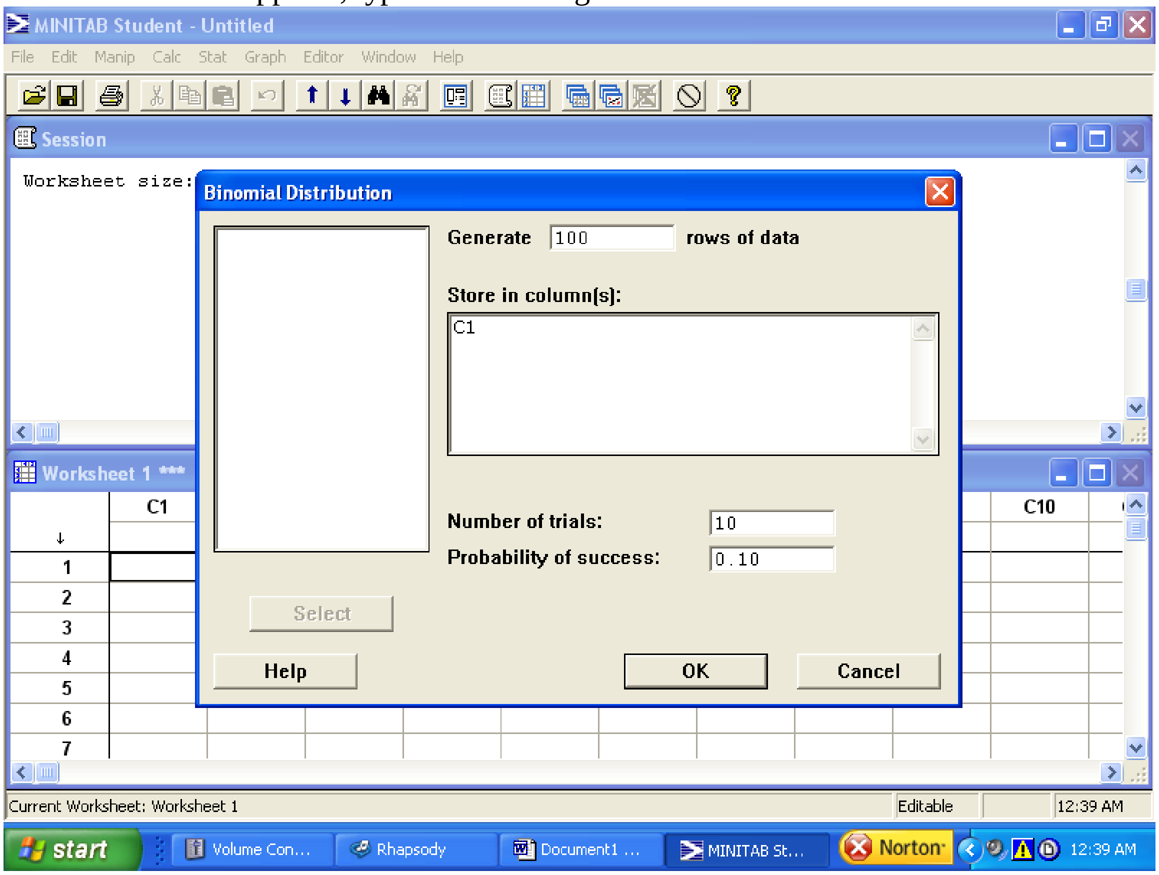
(i) a basketball player makes 10% of her free throws
(ii) a basketball player makes 50% of her free throws
(iii) a basketball player makes 90% of her free throws

**Go to your worksheet and answer the questions about these scenarios BEFORE doing the simulation.**

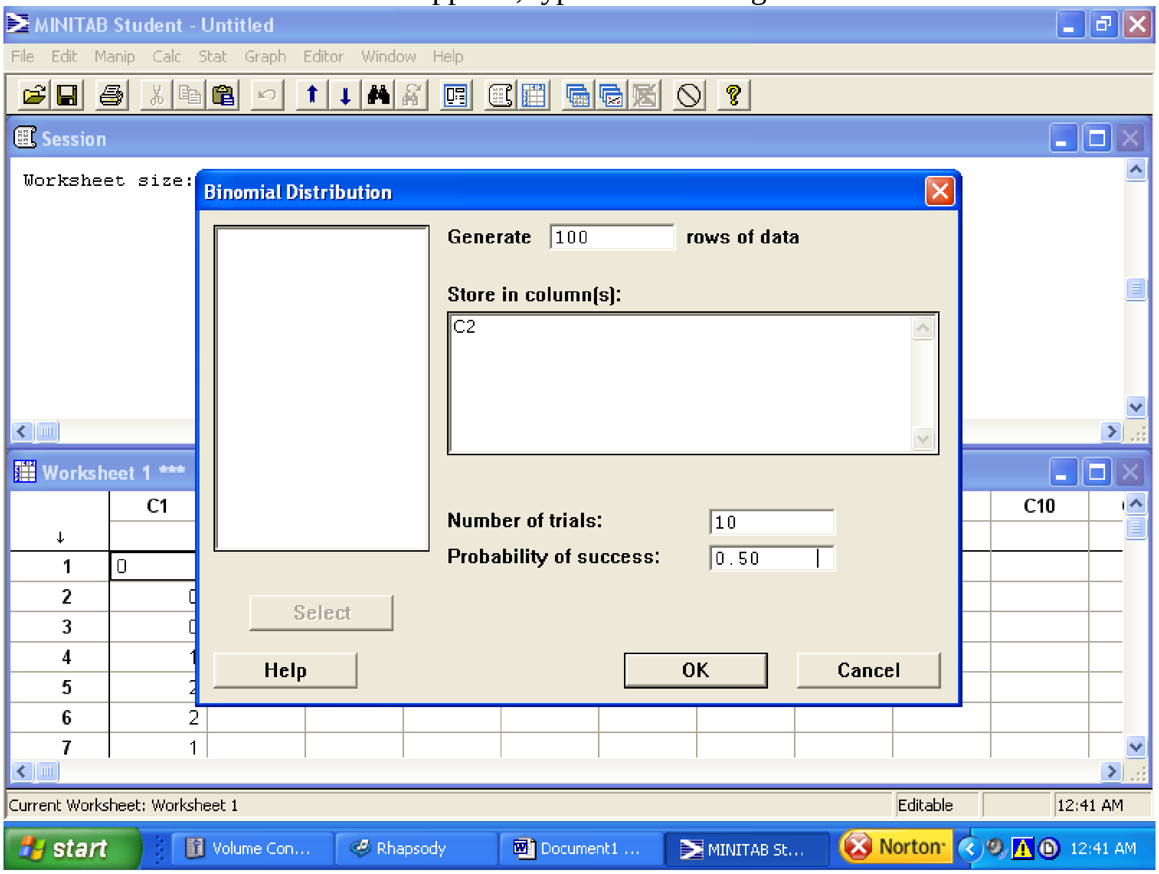
Now simulate the above example. Open MINITAB.
Let C1 represent scenario 1 with the 10% shooter by clicking CALC, RANDOM DATA, BINOMIAL as shown:


MINITAB Student - Untitled
File Edit Manip Calc Stat Graph Editor Window Help
Calculator...
Column Statistics...
Row Statistics...
Standardize...
Extract from Date/Time to Numeric...
Extract from Date/Time to Text...
Make Patterned Data
Make Indicator Variables...
Set Base...
Random Data
Probability Distributions
Sample From Columns...
Chi-Square...
Normal...
F...
t...
Uniform...
Bernoulli...
Binomial...
Discrete...
Integer...
Poisson...
Beta...
Cauchy...
Exponential...
Gamma...
Laplace...
Logistic...
Lognormal...
Weibull...
Session
Worksheet 1 ***
C1 C2 C3 C7 C8 C9 C10
Generate data from a binomial distribution
start
Volume Con...
Rhapsody
Document1 ...
MINITAB St...
Norton
12:38 AM

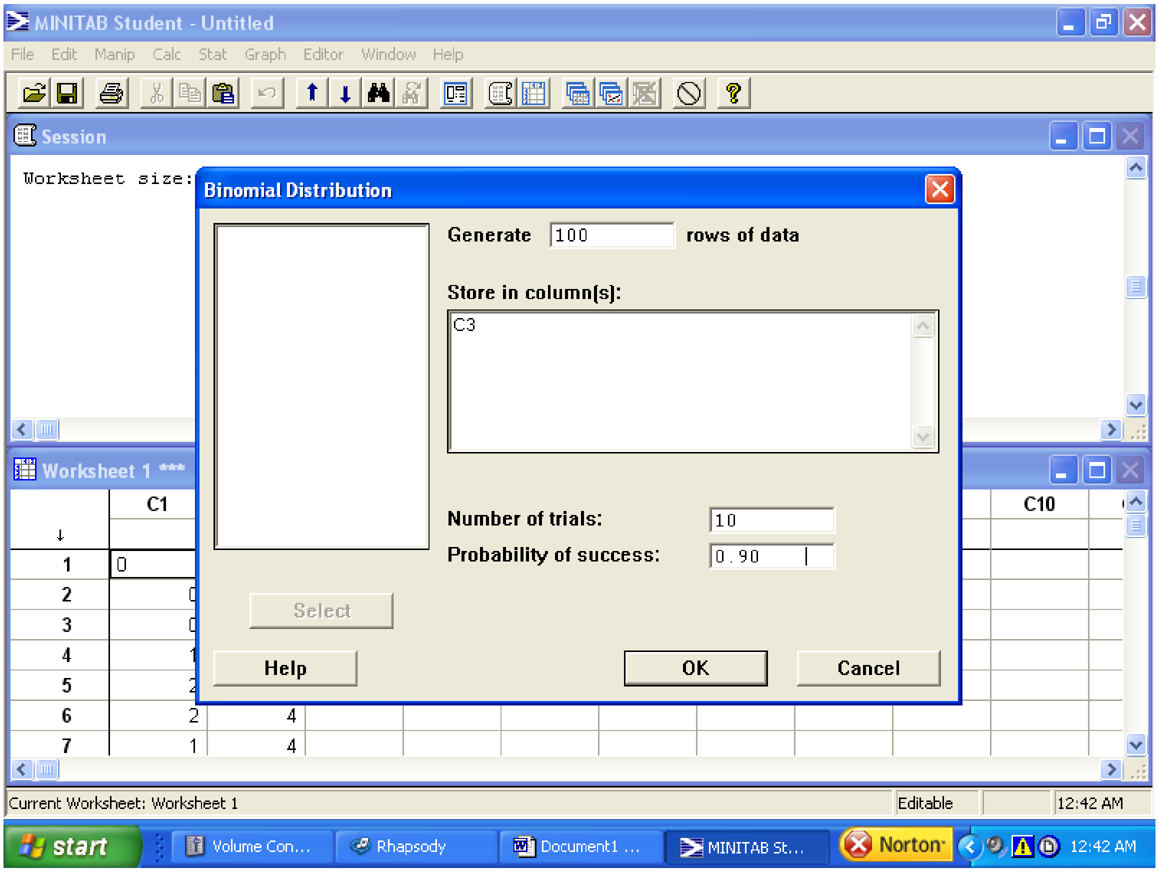
When the window appears, type the following:



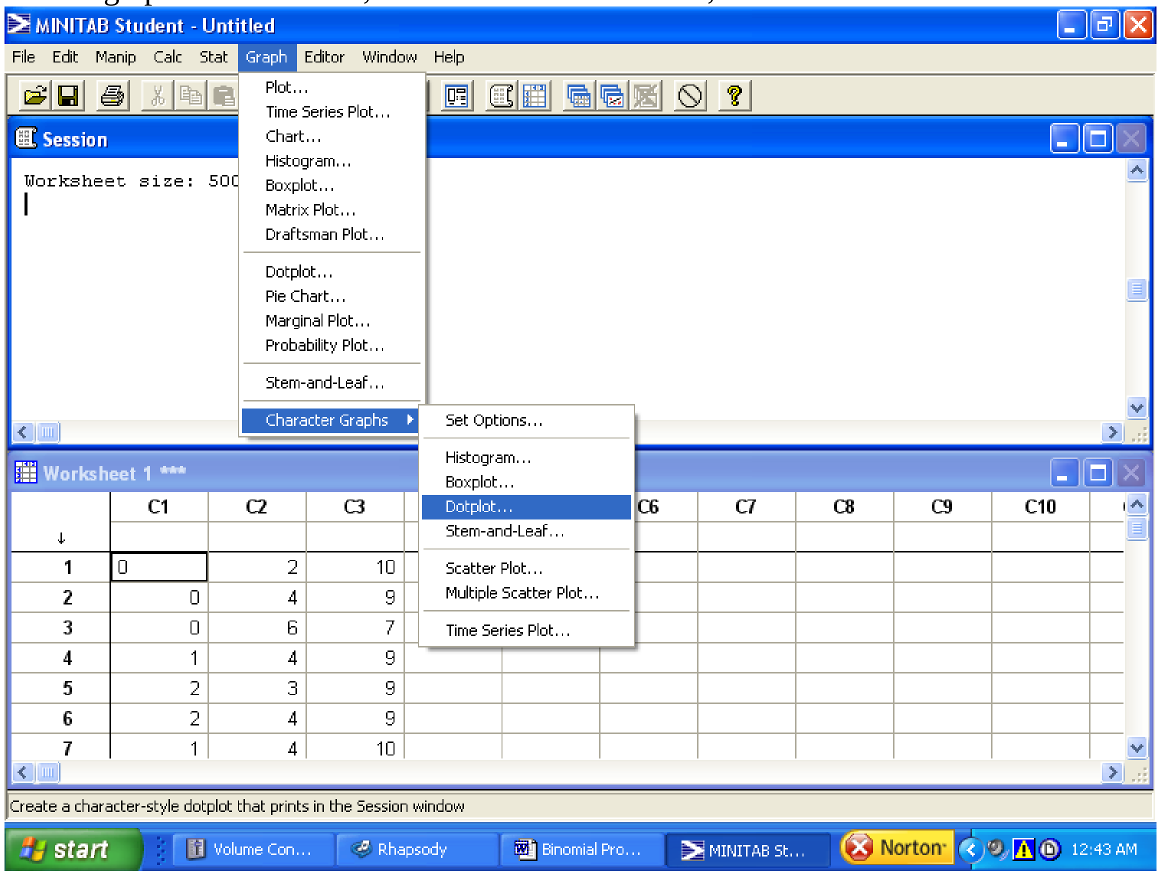
Let C2 represent scenario 2 with the 50% shooter by clicking CALC, RANDOM DATA, BINOMIAL. When the window appears, type the following:



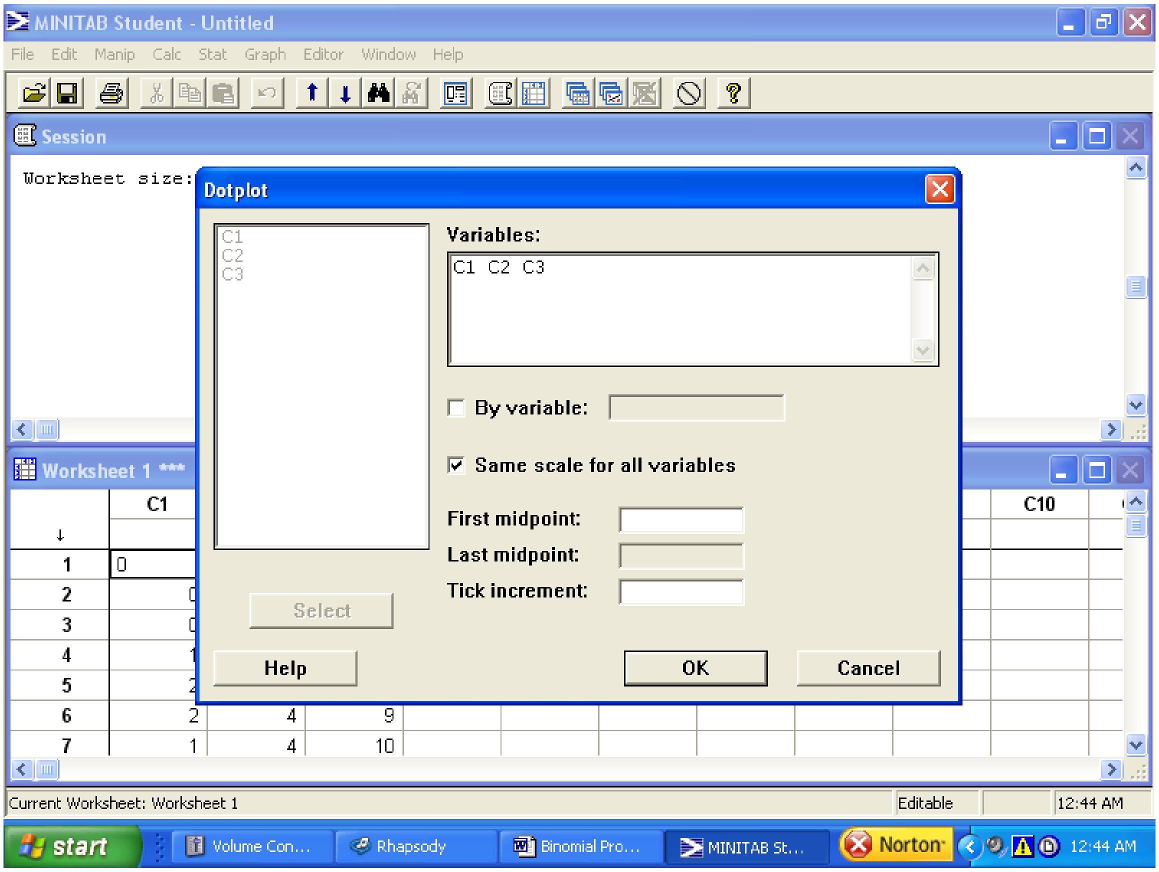

Let C3 represent scenario 3 with the 90% shooter by clicking CALC, RANDOM DATA, BINOMIAL. When the window appears, type the following:


MINITAB Student - Untitled
File Edit Manip Calc Stat Graph Editor Window Help
Session
Worksheet size:
Binomial Distribution
Generate 100 rows of data
Store in column(s):
C3
Number of trials: 10
Probability of success: 0.90
Select
Help
OK
Cancel
Worksheet 1 ***
C1
C10
Current Worksheet: Worksheet 1
Editable
12:42 AM
start
Volume Con...
Rhapsody
Document1 ...
MINITAB St...
Norton

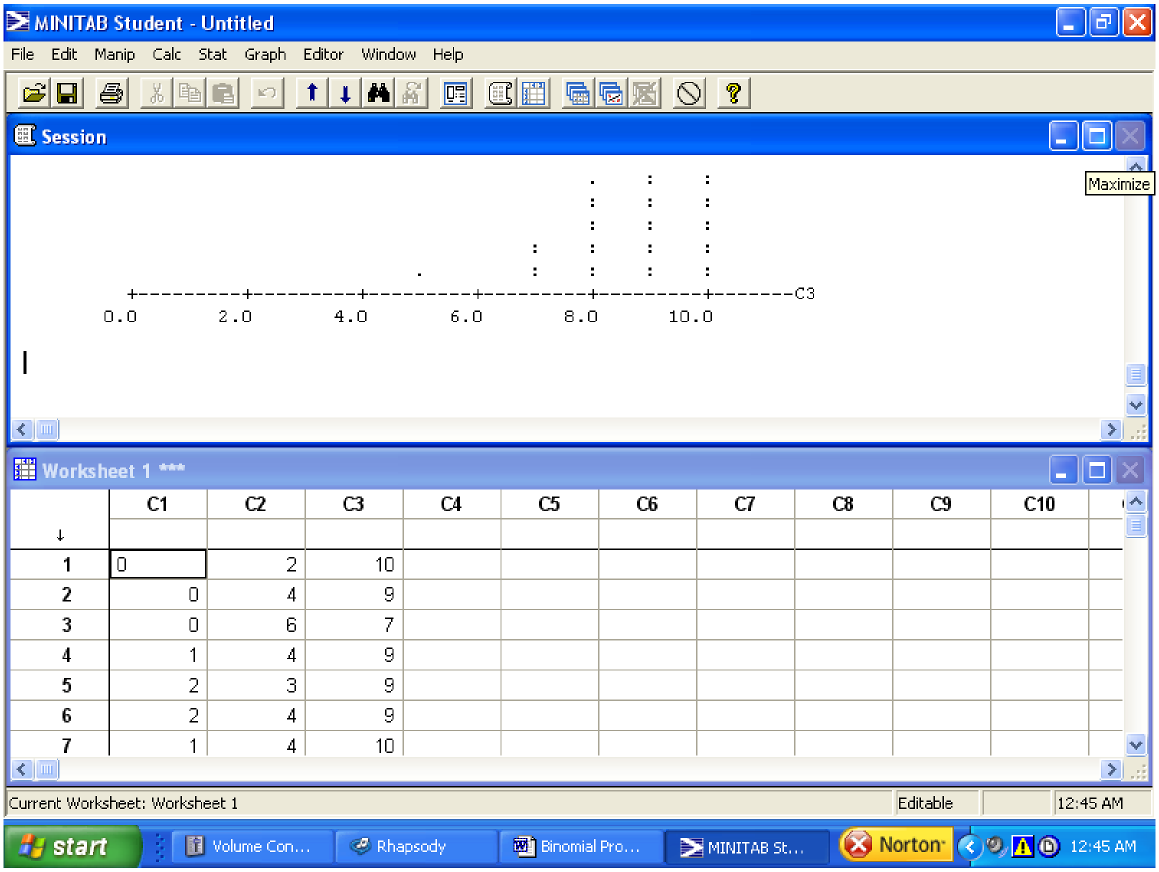
Now to graph click GRAPH, CHARACTER GRAPHS, DOTPLOT as shown:




When the window appears, type the following:




Maximize the sheet with the plots by clicking the maximize button as shown below and scroll up to see all the graphs.



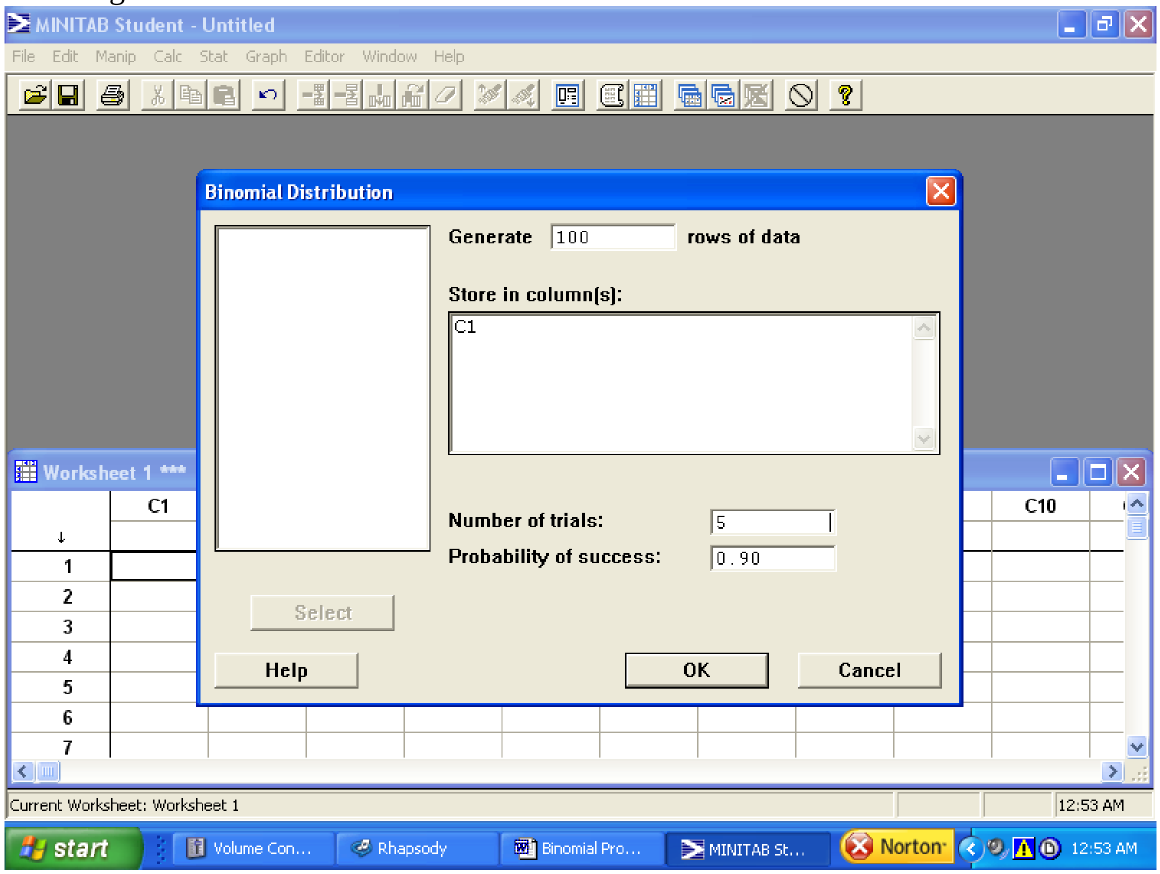
**Now go to your worksheet and answer the questions based on your graphs.**

**Now go to SECTION II on the worksheet and answer the questions BEFORE doing the next simulation.**

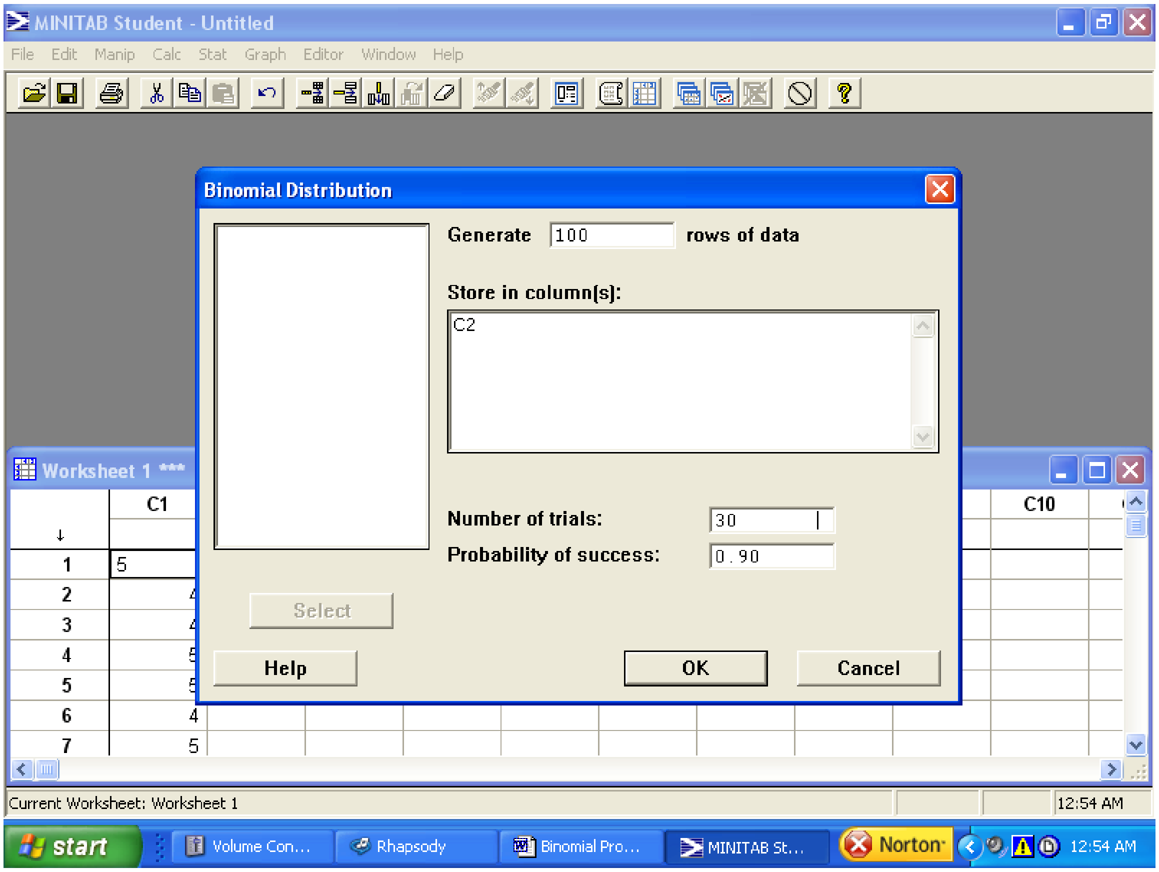
To simulate and see how the number of trials affects the shape of a binomial distribution, we will take the 90% shooter and compare n = 5 trials, n = 30 trials and n = 100 trials.

Minimize or close out the dotplot sheet. Delete the previous columns used on the original spreadsheet. You can do this by clicking and dragging across C1, C2, and C3 then pressing DELETE on the keyboard.

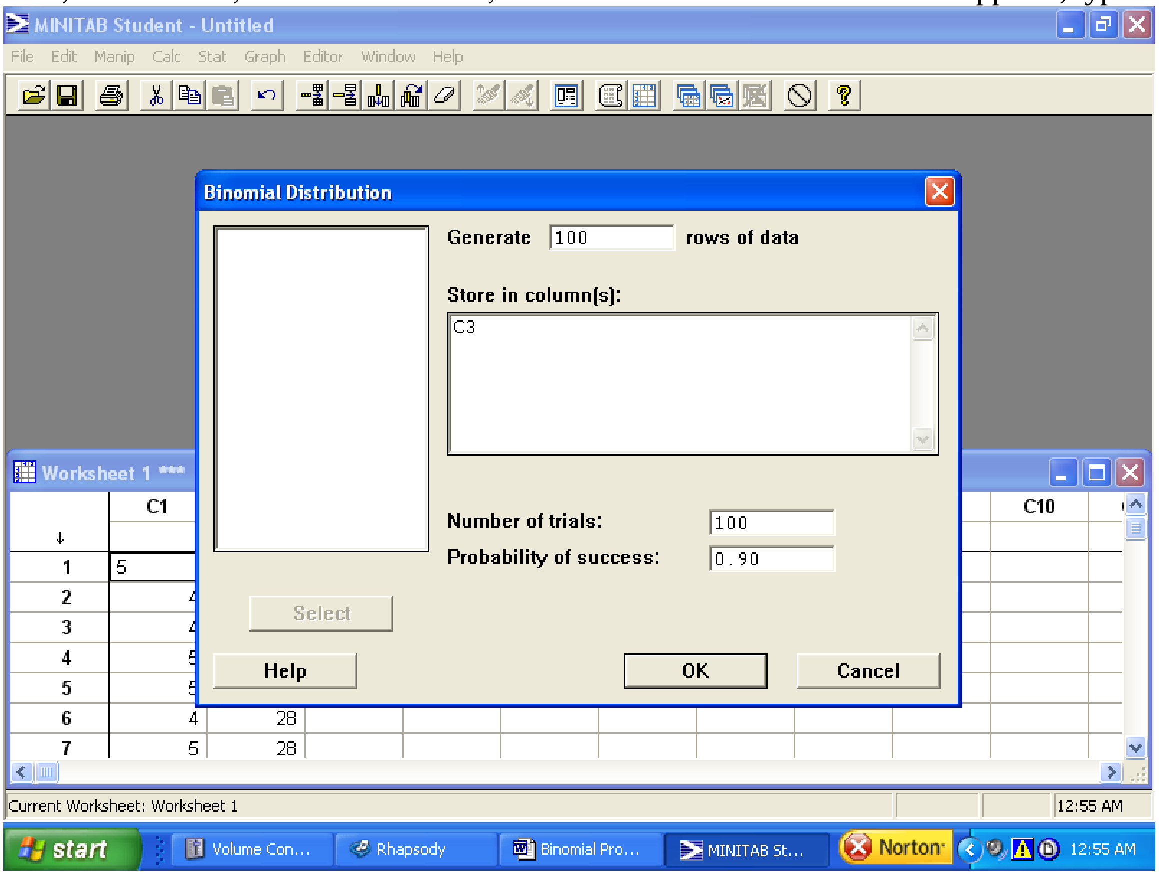
Now click CALC, RANDOM DATA, BINOMIAL. When the window appears, type the following:




Next, click CALC, RANDOM DATA, BINOMIAL and when the window appears, type:


MINITAB Student - Untitled
File Edit Manip Calc Stat Graph Editor Window Help
Binomial Distribution
Generate 100 rows of data
Store in column(s):
C2
Number of trials: 30
Probability of success: 0.90
Select
Help
OK
Cancel
Worksheet 1 ***
C1
C10
Current Worksheet: Worksheet 1
12:54 AM
start
Volume Con...
Rhapsody
Binomial Pro...
MINITAB St...
Norton
12:54 AM

Next, click CALC, RANDOM DATA, BINOMIAL and when the window appears, type:




To graph, click GRAPH, CHARACTER GRAPHS, DOTPLOT and do the same with the graphs as you did before.

**Now go to your worksheet and answer the questions based on your graphs.**

**Name: __________________________ Date: _______________**

**Binomial Probability Distributions**

Given the following scenarios,

(i) a basketball player makes 10% of her free throws
(ii) a basketball player makes 50% of her free throws
(iii) a basketball player makes 90% of her free throws

If the basketball player took 10 free throw shots (so n = 10 trials), what shape do you think the graphs of the 3 scenarios would look like? EXPLAIN.

What do you think the means of each of the 3 scenarios would be:

(i) ___________ (ii) _____________ (iii) ______________

EXPLAIN:

**AFTER** doing the simulation, did you notice anything different than what you predicted? Are the shapes what you expected? Are the means most likely where you expected them to be? EXPLAIN.

**SECTION II.**

**BEFORE SIMULATION**: How do you think the shape of a distribution is affected by the number of trials? Predict the following shapes of the distributions assuming the shooter is a 90% shooter.

(a) shape if the player takes 5 shots (n = 5)? EXPLAIN.

(b) shape if the player takes 30 shots (n = 30)? EXPLAIN.

(c) shape if the player takes 100 shots (n = 100)? EXPLAIN.

**AFTER DOING SIMULATION:** Did the shapes end up as you predicted above? EXPLAIN.

As the number of trials, n, increases, does the shape of the distribution become more "normal" (just base this on your observation even though we are not actually checking for normality)? EXPLAIN.

APPENDIX M - GEOMETRIC PROBABILITY DISTRIBUTIONS WORKSHEET

**Name: ___________________________ Date: ______________________**

**Geometric Probability Distributions**

BEFORE doing your simulation, answer the following questions and EXPLAIN why you put the answer you did:

A basketball player makes 80% of his free throws. If this player is asked to shoot free throws until he misses one, which of the following is the MOST PROBABLE number of throws it will take?

(a) 0 (b) 1 (c) 2 (d) 3 (e) 4

EXPLAIN:

True or False: The probability of getting a prize in a cereal box is 0.10. The chance of getting your first prize in the tenth box is higher than the chance of getting your first prize in the first box you buy.

(A) True
(B) False

EXPLAIN:

Recall conditions for the geometric setting are the following:
(i) each observation falls into one of two categories, "success" and "failure"

(ii) the probability of "success," p, is the same for each observation

(iii) the observations are all independent

(iv) the variable of interest is the number of trials required to obtain the first success

We will simulate the first example: A basketball player makes 80% of his free throws. If this player is asked to shoot free throws until he misses one, which of the following is the MOST PROBABLE number of throws it will take?

How will you assign the numbers in the random number table to fit the assumption?

How does using the random number table serve as a model for this geometric setting? In other words, EXPLAIN how the assumptions of the geometric setting are being met.

Now simulate the situation by starting at a random starting place in your random number table. Simulate 20 repetitions and then compare to other members of your group. Show your tally marks below:

Do your simulation results as well as your group members confirm your predicted answer BEFORE doing the simulation or do the results want to make you change your answer? EXPLAIN.

Use your simulation results to construct a probability distribution for the random variable X = # of shots until the first miss

List the outcomes associated with your probability distribution above ie. M, HM, HHM, etc. How could you find the theoretical probabilities for these outcomes?

What shape would the distribution have? Will all geometric distributions have this similar shape? EXPLAIN.

Use your calculator to find the mean number of trials it takes until the first miss (BASED on your SIMULATION probability distribution above). Explain how you calculated your mean.

The theoretical mean of a geometric distribution is $\mu = \frac{1}{p}$ where p is probability of "success" in this case missing a shot. Is your empirical mean close to this theoretical mean? EXPLAIN.

Let's suppose this was a binomial setting with the same shooting probabilities. If a player takes 3 shots and you simulate this situation to find the empirical probabilities for X = # of shots the player misses. Would you still use the same assignment of random digits? EXPLAIN.

How would running the simulation be DIFFERENT than the geometric setting? EXPLAIN.

What did you learn from doing this simulation?