Veteran Education Transfer Plan
Cover Sheet

Title of ETP	Elementary Statistics Workbook
Name of IISME Fellow	Alyson Clark
Fellow's year-round email	alyson_clark@westvalley.edu
Sponsor Company	Synopsys, Inc.
Name of Mentor	Katherine Neff
National Board Certificate Area	Adolescence and Young Adulthood Mathematics

I, the IISME Fellow named above, affirm that the ETP I am submitting is my own work, that I acknowledged sources where appropriate, and that I avoided including any proprietary information of the Sponsor Company. By my submission I am assigning to IISME my entire copyright in the ETP. I understand IISME is simultaneously granting me a license to use the ETP for pedagogical purposes.

Signature ______________________ Date ______________

Category	Math Curriculum – Community College Level
Objectives	Develop a statistics workbook that will enhance student comprehension of new concepts and provide immediate practice when learning key ideas. This tool will be used daily in class and will be a resource when students study for exams. In addition, since most problems in statistics involve large data sets and long word problems, a pre-printed workbook will allow the students to spend less time copying data or problems down and more time practicing and discussing.
Abstract	Create a workbook that will supplement lecture and the textbook on key concepts taught in an Elementary Statistics course. The workbook will help students develop a deeper understanding of the material by allowing them to discover ideas and to practice problems on their own during lecture, in addition to working on problems with their classmates in groups.
Describe how your ETP aligns with the National Board Standard stated in your proposal.	This ETP aligns with the Adolescence and Young Adulthood Mathematics National Board Standard VI: The Art of Teaching. The Elementary Statistics Workbook is intended to stimulate and facilitate student learning by using a format other than lecture to guide students' learning of mathematics.

Describe the connection between your ETP and the Summer Fellowship.	My summer fellowship project was to help develop the curriculum for a project management class to be taught at Synopsys by my mentor. My ETP is about developing curriculum to supplement and enhance the material taught in my Elementary Statistics class.
Growth-Measurement Devices	Implementation of the workbook in class will reduce the time I spend handing out daily worksheets and writing out problems on the board. It will increase the amount of time students have to work on problems individually and in groups, and allow me to spend more one-on-one time with them. The workbook will bring another dimension of learning to the classroom, and ideally result in improved exam scores.
Resources Needed	Workbook saved as MS Word document, paper to print it on, and binding materials (perhaps a comb bind).
Evaluation/Assessment Measures Used	I will use student surveys and faculty/peer reviews to assess the usefulness of the workbook. I will also compare students' exam scores from every semester prior to Fall 2004 (who did not have the workbook) to students' exam scores from this coming semester (who will have used the workbook).
Formatting specifications	PC – Microsoft Word
Submitted Copy	Soft and hard copy due to peer coach by the end of the summer fellowship. Also, a copy of the cover sheet signed by a school site administrator submitted to IISME Oct.3, 2004 to receive $300 grant.

I, the Mentor named above [please select one of the following],

- ☐ have read the attached ETP, and my comments, if any, appear below.
- ☐ have read the attached ETP, and, as outlined in the IISME-Company Fellowship Agreement, have reviewed it on behalf of the Sponsor Company, and have determined that the ETP does not contain any Sponsor-proprietary information. My additional comments, if any, appear below.

Comments:

____________________________ ____________________

Signature Date

Administrator's comments:

____________________________ ____________________

Signature Date

Elementary Statistics Workbook

An Education Transfer Plan
by
Alyson Clark

Mathematics Department
West Valley College
14000 Fruitvale Avenue
Saratoga, CA 95070
(408) 741-2590
alyson_clark@westvalley.edu

Sponsor Company
Synopsys, Inc.
700 East Middlefield Road
Mountain View, CA 94043

Mentor
Katherine Neff
Senior Program Manager
Information Technology

ETP Table of Contents

Introduction

Over the last four-and-a-half years, I have taught around two sections of Elementary Statistics per semester. During that time I have developed many activities and handouts that I now use in every class, every semester. Thus, I decided to compile those documents into a workbook for my Elementary Statistics class as my ETP for this summer. This is an extension of my ETP from my fellowship last summer, where I developed a workbook for my Beginning Algebra class.

As I began compiling the statistics workbook, I realized it would make sense to put my lecture notes into electronic form at the same time. This has increased the amount of time necessary to complete my ETP, but I believe it will be well worth it in the long run. Plus, I can actually use my laptop in the classroom which is why I applied for a Fund for Innovation grant last summer.

The high-level goals of my ETP include:

- Dynamic format for lecture
- Put lecture on web and in handouts
- Utilize technology
- Reduce time spent passing things out or writing problems on board
- Provide students tangible collection of problems for practice and review

I will measure the success of each of these goals through student surveys and peer surveys, in addition to comparing exam scores from previous semesters (which did not have the workbook or electronic lecture) to this upcoming semester (which will have the workbook and electronic lecture).

Note: The workbook submitted in this ETP only covers Chapters 1 – 7 of our textbook (and the first section of Chapter 8), although there are 10 chapters total that are covered in the statistics course. Due to time constraints, I was unable to complete all 10 chapters of lecture notes and workbook problems over the summer. My plan is to split the workbook into two or three pieces during the semester, so that by the time I will need those later chapters I will have them completed. This will also save printing money since we unfortunately lose students as the semester progresses, thus it will require fewer copies to be made if I break it up.

Assessment

Student Survey

Please respond to each question with as much detail as possible. This survey is completely anonymous and has no bearing on your class grade whatsoever. Thank you for your time!

1. What did you like about the lecture notes being presented electronically?

2. What did you not like about the lecture notes being presented electronically?

3. How do you feel the workbook helped you during class, or did it?

4. How do you feel the workbook helped you outside of class, or did it?

5. What would you change about the way lecture was presented?

6. What would you change about the format or utilization of the workbook?

Peer Survey

Please respond to each question with as much detail as possible. Thank you for your time!

1. Have you ever taught a class through Power Point instead of traditional board lecture? If so, what made it successful? What made it challenging?

2. Have you ever developed a workbook for supplemental use in class? If so, what format do you feel worked? How did you use it?

3. Have you ever placed your lecture notes or materials for students online? If so, what do you feel is the best way to do this? Were the students comfortable and willing to print things on their own before class? What would you have done differently?

Name ______________________________

Elementary Statistics Workbook

Math 10
West Valley College
Fall 2004

Table of Contents

Chapter 1 – Getting Started

Section 1.1 – What is Statistics?

Determine if the following variables are quantitative or qualitative.

(a) Hair color ______________________________

(b) Number of students with red hair ______________________________

(c) Height ______________________________

(d) Whether or not a student has red hair ______________________________

(e) Zip code of home town ______________________________

(f) Gender ______________________________

(g) Length of signature ______________________________

(h) Political identification ______________________________

(i) Phone number ______________________________

Section 1.2 – Random Samples

Suppose you want to know the average number of hours a student who attends West Valley College works per week.

(a) What is the population of interest?

(b) What is the variable of interest?

(c) Is it possible to obtain a random sample from this population? If so, describe in detail how you would obtain a random sample.

Section 1.3 – Introduction to Experimental Design

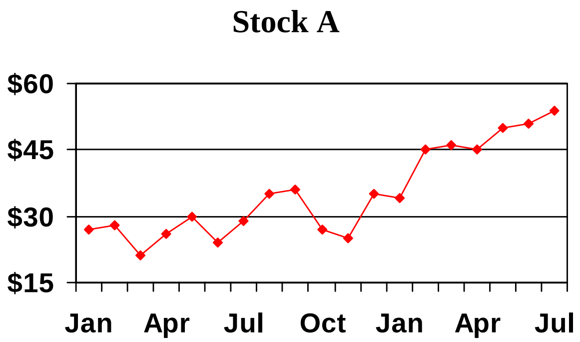
1. As a medical researcher, you are trying to determine different causes of breast cancer. What would be the best method of collecting data to help you conduct this research – an observational study or an experiment? Why?

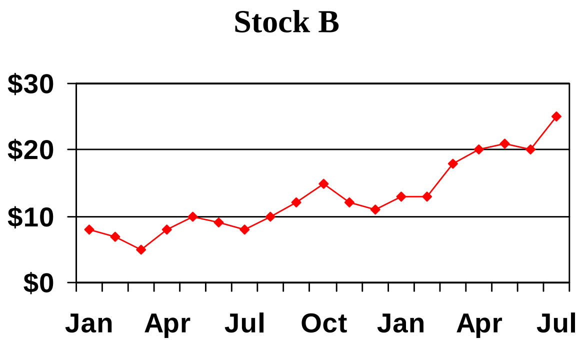
2. Suppose you are Billy Blanks, creator of Tae Bo. If you would like to show that using Tae Bo as a form of exercise will help you loose a significant amount of weight, what would be the best method of collecting data to support this statement – an observational study or an experiment? Why?

Chapter 2 – Organizing Data

Section 2.1 – Bar Graphs, Circle Graphs, and Time Plots

Consider the following time plots.





Suppose you invested $5000 on January 1. With which stock would you have earned more money?

Section 2.2 – Frequency Distributions and Histograms

1. Write down your height in inches, rounding to the nearest inch. Report your height when asked.

2. Make a frequency table, listing frequencies and relative frequencies, for the set of students' heights. Use five classes.

3. Draw a histogram for the set of students' heights. Label each scale carefully.

4. Describe the distribution of the data set.

Section 2.3 – Stem-and-Leaf Displays

Consider the following table.

Game	Date	Result	Spread
XXXVIII	Feb. 1, 2004	New England 32, Carolina 29	3
XXXVII	Jan. 26, 2003	Tampa Bay 48, Oakland 21	27
XXXVI	Feb. 3, 2002	New England 20, St. Louis 17	3
XXXV	Jan. 28, 2001	Baltimore 34, N.Y. Giants 7	27
XXXIV	Jan. 30, 2000	St. Louis 23, Tennessee 16	7

XXXIII	Jan. 31, 1999	Denver 34, Atlanta 19	15
XXXII	Jan. 25, 1998	Denver 31, Green Bay 24	7
XXXI	Jan. 26, 1997	Green Bay 35, New England 21	14
XXX	Jan. 28, 1996	Dallas 27, Pittsburgh 17	10
XXIX	Jan. 29, 1995	San Francisco 49, San Diego 26	23
XXVIII	Jan. 30, 1994	Dallas 30, Buffalo 13	17
XXVII	Jan. 31, 1993	Dallas 52, Buffalo 17	35
XXVI	Jan. 26, 1992	Washington 37, Buffalo 24	13
XXV	Jan. 27, 1991	N.Y. Giants 20, Buffalo 19	1
XXIV	Jan. 28, 1990	San Francisco 55, Denver 10	45
XXIII	Jan. 22, 1989	San Francisco 20, Cincinnati 16	4
XXII	Jan. 31, 1988	Washington 42, Denver 10	32
XXI	Jan. 25, 1987	N.Y. Giants 39, Denver 20	19
XX	Jan. 26, 1986	Chicago 46, New England 10	36
XIX	Jan. 20, 1985	San Francisco 38, Miami 16	22
XVIII	Jan. 22, 1984	L.A. Raiders 38, Washington 9	29
XVII	Jan. 30, 1983	Washington 27, Miami 17	10
XVI	Jan. 24, 1982	San Francisco 26, Cincinnati 21	5
XV	Jan. 25, 1981	Oakland 27, Philadelphia 10	17
XIV	Jan. 20, 1980	Pittsburgh 31, L.A. Rams 19	12
XIII	Jan. 21, 1979	Pittsburgh 35, Dallas 31	4
XII	Jan. 15, 1978	Dallas 27, Denver 10	17
XI	Jan. 9, 1977	Oakland 32, Minnesota 14	18
X	Jan. 18, 1976	Pittsburgh 21, Dallas 17	4
IX	Jan. 12, 1975	Pittsburgh 16, Minnesota 6	10
VIII	Jan. 13, 1974	Miami 24, Minnesota 7	17
VII	Jan. 14, 1973	Miami 14, Washington 7	7
VI	Jan. 16, 1972	Dallas 24, Miami 3	21
V	Jan. 17, 1971	Baltimore 16, Dallas 13	3
IV	Jan. 11, 1970	Kansas City 23, Minnesota 7	16
III	Jan. 12, 1969	N.Y. Jets 16, Baltimore 7	9
II	Jan. 14, 1968	Green Bay 33, Oakland 14	19
I	Jan. 15, 1967	Green Bay 35, Kansas City 10	25

1. Make a stem-and-leaf display for the set of spreads of Super Bowl box-scores.

2. Describe the distribution of the data set.

3. Of all the distribution shapes you know, which one would advertisers most want this data set to follow? Why?

Chapter 3 – Averages and Variation

Section 3.1 – Measures of Central Tendency: Mode, Median, and Mean

1. Complete each of the frequency tables below. Then determine the mode.

Eye Color	**Frequency**
Brown	
Blue	
Green	
Hazel	

Age	**Frequency**
18 and under	
19	
20	
21	
22	
23 and over	

2. For each data set below, find the mode, median, and mean (assume all are samples).

   (a) 1 1 2 3 4

   (b) 1 2 3 4 5 6

   (c) 1 1 2 2 3 4 50

<u>Section 3.2 – Measures of Variation</u>

1. The heights of a random sample of 8 students in this classroom were recorded. The data set is given below. Calculate the range, variance, and standard deviation for this data set.

   66 71 69 61 75 66 72 64

2. In 1994, the average MLB salary was $1,183,417 with standard deviation $1,390,922. In 2000, the average was $1,988,034 with standard deviation $2,511,397. Calculate the CV for each year, and interpret the results.

3. (a) The average man is 69 inches tall with a standard deviation of 2.8 inches. Find an interval of heights in which you would expect at least 93.8% of the male population to fall.

   (b) The average age of a person causing a car accident is $\mu = 20$ years with $\sigma = 2$ years. What percent of car accidents are caused by drivers aged 14 to 26?

4. For the last two weeks I kept track of the number of minutes it took me to get home from

work. The results from this sample are shown below.

8	10	9	13	13
19	8	8	12	10

(a) Find $\bar{x}$, the median, and the mode for this data set.

(b) Find the range, s^2, and s for this data set.

(c) Use Chebyshev's theorem to find an interval centered about the mean for the times in which you would expect at least 75% of the travel times to fall. Write a complete sentence that interprets this interval.

(d) Are there any extreme data values? Can you provide a reasonable explanation for their presence? Would you describe the distribution of this data set as symmetric, uniform, left/right skewed, or bimodal?

Section 3.4 – Percentiles and Box-and-Whisker Plots

1. (a) If you score at the 89th percentile on the SAT, what percent of the student population did worse than you? Better than you?

   (b) If you score at the 95th percentile on the ACT, which test (SAT or ACT) did you perform better on?

2. Use the set of times for me to get home from work (see section 3.2 #4) and find the five-number summary and the IQR.

3. Create a box-and-whisker plot for the data set in #2.

4. Seventeen clerical staff at a hospital were surveyed to see how long they had been at their current position. Their responses (in months) were:

25	22	7	24	22	3	29	15	72
31	42	6	26	31	18	17	20	

   (a) Find the five-number summary.

   (b) Find the IQR.

   (c) Draw the box-and-whisker plot.

Chapter 4 – Elementary Probability Theory

Section 4.1 – What is Probability?

1. Suppose we polled 200 students and asked them the color of the car they drive, with the following results: 64 said red, 42 said blue, 36 said black, 30 said white, and 28 said some other color.

   (a) Find the probability a student drives a red car. Is this experimental or theoretical?

   (b) Find the probability a student drives a black car.

   (c) What is the probability a student does not drive a white car?

2. One card is to be drawn from a standard deck.

   (a) Find *P*(draw a king). Is this experimental or theoretical?

   (b) Find *P*(draw a diamond).

   (c) Find *P*(do not draw an ace).

3. Three slips of paper, with the numbers 1, 2, and 3 on them, are placed in a box. One slip is randomly selected, the number recorded, then returned to the box. A second slip is then selected and the number recorded.

   (a) Write out the sample space for this probability experiment.

   (b) What is the probability the sum of the two numbers drawn is 4?

   (c) Find the probability the sum of the two numbers is 8.

4. A coin and a die are tossed together once.

   (a) Write out the sample space.

   (b) Find *P*(see heads on the coin).

   (c) Find *P*(see more than 4 spots on the die).

   (d) Find *P*(see heads on the coin <u>and</u> more than 4 spots on the die).

5. If a quality control test at a printer manufacturer shows there are 5 defective printers and 15 good printers, what is the probability a randomly selected printer is defective? Is not defective?

6. How do people want to be treated when they have the flu? A poll of 1000 people gave the following information:

Want to be…	*Number of People*
Left alone	770
Waited on hand and foot	160
Treated differently	70

(a) Use the data to assign probabilities to each of these events. Do the probabilities add up to 1? Should they?

(b) Find the probability that someone does not want to be left alone.

(c) Find the probability that someone does not want to be waited on hand and foot.

Section 4.2 – Compound Events

1.

Experiment	**Probability**	**Independent?**	**Mutually Exclusive?**
Roll two die	*P*(roll a 2 and roll a 3)	Yes or No	Yes or No
Roll one die	*P*(roll a 2 or roll a 3)	Yes or No	Yes or No
Draw two cards	*P*(draw a King and a Queen)	Yes or No	Yes or No
Draw one card	*P*(draw a King or a Queen)	Yes or No	Yes or No
Draw two cards	*P*(draw a Kind and a King)	Yes or No	Yes or No

2. Suppose a friend will soon give birth to **one** child. Define events *A* and *B* as follows:

Event *A*: the child is a girl
Event *B*: the child is a boy

(a) Are these two events mutually exclusive?

(b) Are these two events independent?

3. Suppose one card is drawn from a deck.

(a) What is the probability of selecting a king or queen?

(b) What is the probability of selecting a jack, queen, or king?

(c) What is the probability of selecting a jack or a club?

4. Suppose a single die is rolled twice.

(a) What is the probability of rolling a 2 and then another 2?

(b) What is the probability of rolling an even number and then a number greater than 2?

5. Suppose an ice chest contains 18 cans of soda: 6 cans of Coke, 8 cans of Sprite, and 4 cans of Root Beer. You randomly draw out 2 cans.

   (a) What is the probability of selecting a Root Beer and a Coke?

   (b) What is the probability of selecting 2 cans of Sprite?

6. The admissions and records department at a university keeps a large amount of data about student demographics. The table below shows the number of males and females in each class level at a small university. Suppose one student is randomly chosen from all students at this university.

		GENDER		
		Male	Female	
	Freshman	106	110	216
CLASS	Sophomores	97	108	205
LEVEL	Juniors	88	92	180
	Seniors	75	106	181
	Graduates	25	30	55
		391	446	837

   (a) What is the probability that the student is a freshman?

   (b) What is the probability that the student is a female?

   (c) What is the probability that the student is a male and a freshman?

   (d) What is the probability that the student is a female or a freshman?

   (e) What is the probability that the student is a freshman, given that the student is a male?

6. In the book *Chances: Risk and Odds in Every Day Life*, James Burke says that 56% of the general population wears eyeglasses, while only 3.6% wears contacts. He also says that of those who do wear glasses, 55.4% are women and 44.6% are men. Of those who wear contacts, 63.1% are women and 36.9% are men. Assume that no one wears both glasses and contacts. For the next person you encounter at random, what is the probability that this person is

   Define the following probabilities using the given information:

   (a) P(person wears eyeglasses) =

   (b) P(person wears contacts) =

   (c) P(person is a woman, given that the person is wearing glasses) =

(d) *P*(person is a man, given that the person is wearing glasses) =

(d) *P*(person is a woman, given that the person is wearing contacts) =

(e) *P*(person is a man, given that the person is wearing contacts) =

For the next person you encounter at random, what is the probability that this person is…

(a) wearing glasses and is a woman?

(b) wearing glasses and is a man?

(c) wearing contacts and is a woman?

(d) wearing contacts and is a man?

(e) none of the above (i.e. is not wearing glasses or contacts)?

Introduction to the TI-83

HOW TO ENTER DATA

1. [STAT]
2. [ENTER] to get in EDIT mode
3. Choose a list to enter your data set (use the arrow keys to move around)
4. Type in your numbers with [ENTER] after each

HOW TO DELETE A LIST OF DATA

1. [STAT]
2. [ENTER] to get in EDIT mode
3. Move cursor to the list you wish to delete
4. Use up arrow key to highlight the list name
5. [CLEAR]
6. [ENTER]

HOW TO GET THE SUMMARY STATISTICS

1. Enter the data set in a list
2. [STAT]
3. Right arrow key once to CALC menu
4. [ENTER] to get One-Variable Statistics (this will take you back to the home screen)
5. Give the appropriate list to use (e.g. [2nd] [1] to get L1)
6. [ENTER]
7. Scroll up and down with the arrow keys to see the mean, standard deviation (sample or

population), sample or population size, median, and quartiles

HOW TO CREATE A HISTOGRAM

1. Enter the data set in a list
2. 2nd Y= to get STAT PLOTS
3. 1 to get Plot 1
4. ENTER to turn On
5. Use arrow keys to move cursor to the picture of a histogram
6. ENTER to select the histogram
7. Use arrow keys to move cursor to Xlist:
8. Give the appropriate list to use (e.g. 2nd 1 to get L1)
9. GRAPH
10. ZOOM 9 to zoom in on the statistical plot

Chapter 5 – The Binomial Probability Distribution

Section 5.1 – Introduction to Random Variables and Probability Distributions

1. Determine if the following random variables are discrete or continuous.

   (a) Inches of rainfall in the month of April ____________________

   (b) Number of units a student takes in a semester ____________________

   (c) Height of players on Sacramento Kings ____________________

   (d) Weight of players on San Francisco 49ers ____________________

   (e) Number of siblings a student has ____________________

   (f) Amount of change in a student's wallet ____________________

   (g) Time it takes instructor to get to school ____________________

2. Are we influenced to buy a product because we saw it advertised on TV? Let the random variable x be the number of times a potential buyer saw a commercial for a product on TV. The probability distribution of x is given below. Find its mean and standard deviation.

Number of times saw commercial, x	1	2	3	4	5
Probability, $P(x)$	0.27	0.31	0.18	0.09	0.15

3. What does an insurance company expect to pay on average for a claim submitted by a driver aged 16 to 21? Let x be the amount of the claim. The probability distribution of x is as follows:

Amount of claim, *x*	$0	2000	4000	6000	8000	10000
Probability, *P*(*x*)	0.70	0.15	0.08	0.05	0.01	0.01

4. A carpenter can build up to 5 desks per day of work depending on the availability of materials, time spent with customers, etc. Let the random variable *x* be the number of desks built on one working day. The probability distribution of *x* is as follows:

Number of desks built, *x*	0	1	2	3	4	5
Probability, *P*(*x*)	0.06	0.10	0.20	0.19	0.20	

(a) Find the probability that the carpenter builds 5 desks in one day.

(b) Graph the probability distribution histogram.

(c) What is the probability the carpenter builds 3 or fewer desks in one day?

(d) What is the probability the carpenter builds more than 3 desks in one day?

(e) Find the mean of *x*, the number of desks the carpenter expects to build per day. If the carpenter makes a profit of $200 per desk, what is the expected profit per day?

(f) Find the standard deviation of *x*.

The Normal Distribution and the TI-83

STANDARD NORMAL PROBABILITY CALCULATIONS (*z*)

1. [2nd] [VARS] to get Distributions menu
2. [2] to get normalcdf(
3. Enter lower bound [,] upper bound [)]
   (a) For negative infinity as the upper/lower bound use [(–)] [1] [2nd] [EE] [9] [9]
   (b) For positive infinity as the upper/lower bound use [1] [2nd] [EE] [9] [9]

NORMAL PROBABILITY CALCULATIONS (*x*)

1. [2nd] [VARS] to get Distributions menu
2. [2] to get normalcdf(
3. Enter lower bound [,] upper bound [,] value of μ [,] value of σ [)]
   (a) For negative infinity as the upper/lower bound use [(–)] [1] [2nd] [EE] [9] [9]
   (b) For positive infinity as the upper/lower bound use [1] [2nd] [EE] [9] [9]

DETERMINING *Z*-SCORES OR *X*-VALUES FROM A GIVEN PROBABILITY

1. Determine total area under the curve to the left of the desired *z*-score or *x*-value
2. [2nd] [VARS] to get Distributions menu
3. [3] to get invNorm(
4. If trying to find *z*, enter area to left then [)]
5. If trying to find *x*, enter area to left [,] value of μ [,] value of σ [)]

Chapter 6 – Normal Distributions

Section 6.1 – Graphs of Normal Probability Distributions

1. Sketch normal curves with:

   (a) $\mu = 4$ and $\sigma = 2$
   (b) $\mu = 4$ and $\sigma = 4$

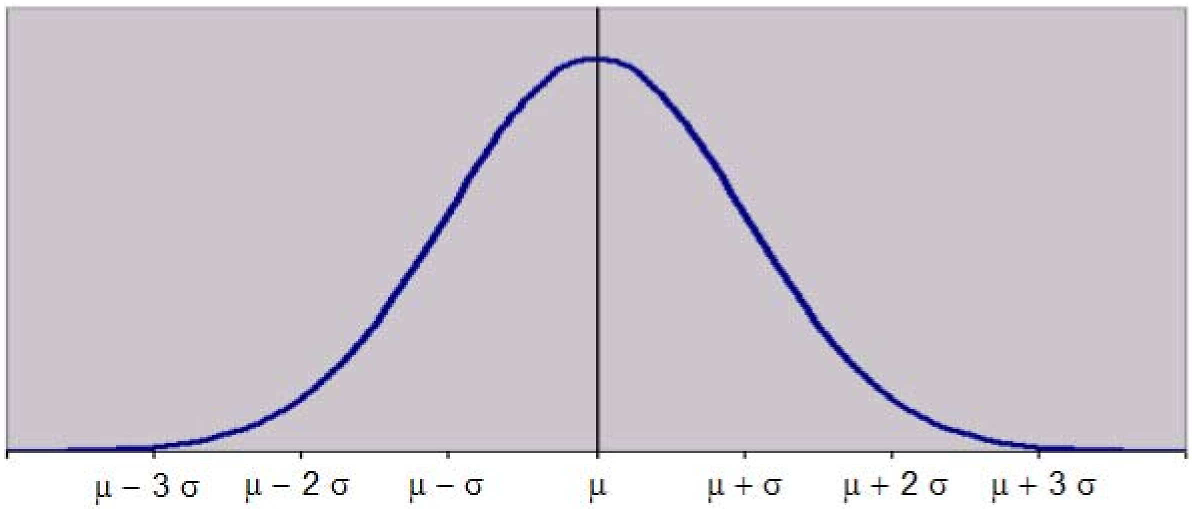
2. Sketch normal curves with:

   (a) $\mu = 4$ and $\sigma = 4$
   (b) $\mu = 12$ and $\sigma = 4$

3. Sketch normal curves with:

   (a) $\mu = 4$ and $\sigma = 6$
   (b) $\mu = 12$ and $\sigma = 2$

4. Shade the areas under the curve that correspond to the Empirical Rule.



5. What is the area under the normal curve…

(a) to the left of μ?

(b) between $\mu - \sigma$ and $\mu + \sigma$?

(c) between $\mu - 3\sigma$ and $\mu + 2\sigma$?

6. A vending machine automatically pours soda into 8 ounce cups. The amount of soda dispensed into a cup is normally distributed with mean $\mu = 7.6$ ounces and standard deviation $\sigma = 0.4$ ounces.

   (a) Draw a normal curve for this random variable. Label the mean, and 1, 2, and 3 standard deviations above and below the mean.

   (b) What is the probability an 8 ounce cup overflows?

   (c) What is the probability an 8 ounce cup does not overflow?

   (d) If the machine is loaded with 850 cups, how many do you expect will overflow?

<u>Section 6.2 – Standard Units and Areas Under the Standard Normal Distribution</u>

1. A college admissions office needs to compare scores of students who take the SAT with those who take the ACT. Among the college's applicants who take the SAT, scores are normally distributed and have a mean of 896 and a standard deviation of 174. Among the college's applicants who take the ACT, scores are also normally distributed but with a mean of 20.6 and a standard deviation of 5.2.

   (a) If Steve scored 1080 on the SAT, how many points above the SAT mean did he score?

   (b) If Michelle scored 28 on the ACT, how many points above the ACT mean did she score?

   (c) Is it sensible to conclude that since your answer to (a) is greater than your answer to (b), Steve outperformed Michelle on the admissions test? Explain.

   (d) Determine how many standard deviations above the mean Steve scored by dividing your answer to (a) by the standard deviation of the SAT scores.

   (e) Determine how many standard deviations above the mean Michelle scored by dividing your answer to (b) by the standard deviation of the ACT scores.

> This activity illustrates the use of the standard deviation to make comparisons of individual values from different distributions. One calculates a ***z*-score**, or **standardized score**, by subtracting the mean from the value of interest and then dividing by the standard deviation.

These z-scores indicate how many standard deviations above (or below) the mean a particular value falls. One should use z-scores only when working with Normal distributions.

(f) Which applicant has the higher z-score for his or her admissions test score?

(g) Explain in your own words which applicant performed better on his or her admissions test.

(h) Calculate the z-score for Chris, who scored 740 on the SAT, and for Beth, who scored 19 on the ACT.

(i) Which of Chris and Beth has the higher score? Explain.

(j) Under what conditions does a z-score turn out to be negative? Postive? Zero?

(k) Consider Morgan, whose score on the SAT was 3 standard deviations above the mean. What was his raw score?

(l) Consider Becca, whose score on the ACT was 2 standard deviations below the mean. What was her raw score?

2. Find the following probabilities using the empirical rule.

(a) $P(z > 0)$

(b) $P(z < -2)$

(c) $P(1 < z < 2)$

3. Let x be a normal random variable with $\mu = 27.2$ and $\sigma = 4.3$.

(a) Convert the interval $x < 30$ into the z-scale.

(b) Convert the interval $-1.99 < z < 1.44$ into the x-scale.

<u>Section 6.3 – Areas Under Any Normal Curve</u>

1. Suppose x is a normal random variable with mean 15 and standard deviation 4.

(a) Find the probability that x is between 10 and 26.

(b) Find the probability that x is greater than 14.

2. Suppose x is a normal random variable with $\mu = 5$ and $\sigma = 1.2$.

   (a) Find $P(x < 4)$.

   (b) Find $P(7 < x < 9)$.

3. Find z ($z > 0$) so that 34.1% of the area under the standard normal curve lies between 0 and z.

4. Find z so that 5.2% of the area under the standard normal curve lies to the left of z.

5. Find z so that 95% of the area under the standard normal curve lies to the right of z.

6. Birth weights of babies in the U.S. can be modeled by a normal distribution with mean 3250 grams and standard deviation 550 grams. Babies weighing less than 2500 grams are considered to be of low birth weight.

   (a) Sketch the distribution for this population. Label the scale and the value of the mean. Shade in the region whose area corresponds to the probability that a baby will have a low birth weight.

   (b) What is the probability a baby will have low birth weight?

   (c) Find the probability a baby weighs more than 10 pounds (4536 grams) at birth.

   (d) Determine the probability that a randomly selected baby weighs between 3000 and 4000 grams at birth.

   (e) Data from the *National Vital Statistics Report* indicate that there were 3,880,894 births in the U.S. in 1997. A total of 291,154 babies were of low birth weight, while 2,552,852 babies weighed between 3000 and 4000 grams. Use this data to calculate the probability of low birth weight and the probability of a baby weighing between 3000 and 4000 grams in 1997. Compare these values to those obtained in parts (b) and (d). Are they different? Why? Which probabilities are theoretical and which are experimental?

   (f) How little would a baby have to weigh to be among the lightest 2.5% of all newborns?

   (g) How much would a baby have to weigh to be among the heaviest 10% of all newborns?

7. Mary scores 680 on the mathematics part of the SAT. The distribution of SAT math scores in a reference population is normal with a mean of 500 and a standard deviation of 100. Bryan takes the ACT mathematics test and scores 27. ACT math scores in a reference population are normally distributed with a mean of 18 and a standard deviation of 6.

   (a) Find the standardized scores for both students. Assuming that both tests measure the same kind of ability, who had the higher score on the math part of the test?

(b) Find the probability a student scores above 637 on the SAT math test.

(c) Find the probability a student scores below 10 on the ACT math test.

(d) Find the probability a student scores between 556 and 656 on the SAT math test.

(e) Find the probability a student scores between 11 and 25 on the ACT math test.

(f) If a university wants to admit only those students who score in the top 10% of all students taking the test, what would be the minimum score needed on the SAT for admission? On the ACT?

Chapter 7 – Introduction to Sampling Distributions

<u>Section 7.1 – Sampling Distributions</u>

Discovering the Sampling Distribution of $\bar{x}$

- Let x be a uniformly distributed random variable whose sample space is the set of integers from 1 to 20. The probability distribution of x is given below, with mean $\mu = 10.5$ and standard deviation $\sigma = 5.7663$.

x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
P(x)	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$

- To determine the sampling distribution of $\bar{x}$, we need to look at the different values $\bar{x}$ can take on (think of $\bar{x}$ as our new random variable). First, we'll each take a random sample of the same size from the population given above and calculate our sample mean. Then, we'll look at the distribution of these sample means together.

- **Step 1**: Each student takes a random sample of size $n = 50$ from the population above.
  [MATH] [←] PRB [5]:randInt(1,20,50)

  Step 2: Store this list in L1.
  [STO→] [2nd] [L1] [ENTER]

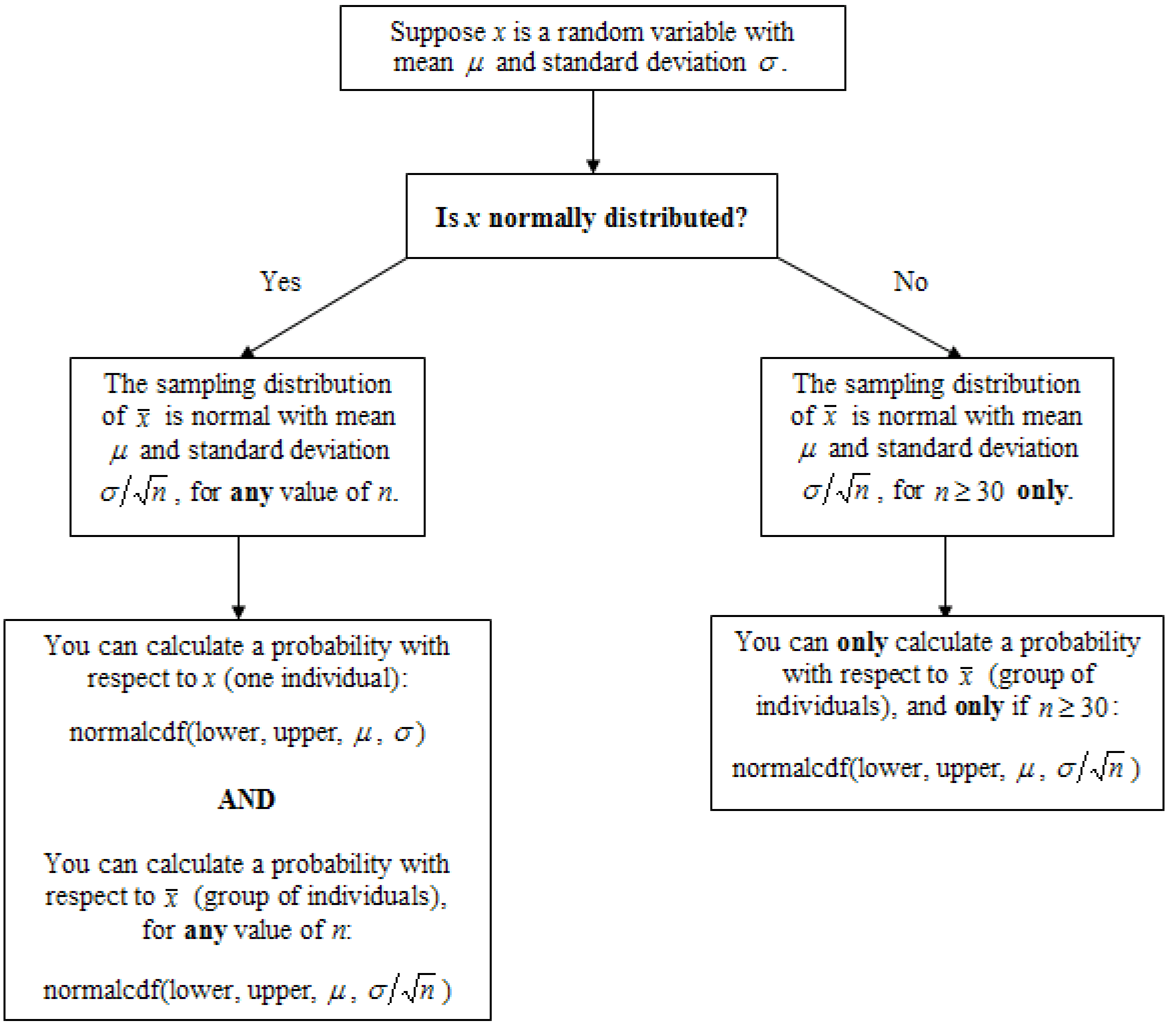
  Step 3: Calculate $\bar{x}$ for your sample, and report it when asked. (For future use, please record both $\bar{x}$ and *s*.)
  [STAT] [→] CALC [1]:1-Var Stats [2nd] [L1] [ENTER]

  Step 4: Create a frequency and relative frequency table, with 5 classes, for the set of means.

  Step 5: Draw the probability histogram for x and then for $\bar{x}$. Compare the distribution shapes for each variable, and the center and spread. The second histogram shows the approximate sampling distribution of $\bar{x}$.

- Results:
    1. When we take repeated random samples of size *n* from a population and calculate the sample mean each time, the sampling distribution of the random variable $\bar{x}$ will be approximately normal, or bell-shaped.
    2. As the sample size increases for each sample, the distribution will become more and more perfectly bell-shaped and symmetric.
    3. The mean of the sampling distribution for $\bar{x}$, denoted by $\mu_{\bar{x}}$, is at the center of the sampling distribution and should be close in value to μ, the mean of the original random variable *x*.
    4. The standard deviation of the sampling distribution for $\bar{x}$, denoted by $\sigma_{\bar{x}}$, will be less than σ, the standard deviation of the original random variable *x*. It should be approximately equal to $\sigma/\sqrt{n}$.
    5. Why is this all so important?
        - We were able to take a non-normal random variable (*x*) and turned it into a normal random variable ($\bar{x}$) by simply taking a sample from the original population and calculating the mean. Following this process for *any* random variable *x*, we can always convert it into an approximately normal random variable $\bar{x}$ (so long as our sample size is large enough – discussed further in 7.2).
        - Thus, even if the original random variable *x* that we started with is not normally distributed (and so we cannot calculate probabilities using the *normalcdf* feature on the calculator), this process now allows us to calculate a probability for $\bar{x}$ using the *normalcdf* feature since we know that $\bar{x}$ is approximately normal.
        - In addition, since the analysis methods of statistics require the variable of interest be normal, we can now conduct the analysis on $\bar{x}$ and be confident our random variable is normal and our conclusions are accurate. Without this knowledge we would not be able to estimate population parameters (Ch. 8), formulate decisions about population parameters (Ch. 9), or make predictions (Ch. 10).

Section 7.2 – The Central Limit Theorem



1. For women aged 18-24, systolic blood pressures (in mm Hg) are normally distributed with a mean of 114.8 and a standard deviation of 13.1 (data from the National Health Survey).

   (a) Find the probability that a randomly selected woman has a systolic blood pressure greater than 120.

   (b) If 12 women in that age bracket are randomly selected, describe the distribution of the sample mean $\bar{x}$, the mean systolic blood pressure for the sample of 12 women, and give the mean and standard deviation of this distribution.

   (c) For the sample of 12 women, find the probability that the mean systolic blood pressure is greater than 120.

   (d) For a sample of 20 women, find the probability that the mean systolic blood pressure is between 112 and 118.

2. Assume that cans of Coke are filled so that the actual amount of soda in the filled cans has a mean of 12 ounces and a standard deviation of 0.11 ounces.

   (a) Suppose we randomly select one can of Coke. Can we calculate the probability that it contains 12.19 ounces of soda or more? Why or why not?

   (b) Suppose we randomly select a sample of 36 cans. Can we describe the sampling distribution of $\bar{x}$? If so, describe the shape of the distribution and give its mean and standard deviation.

   (c) Find the probability that the sample of 36 cans will have a mean amount of 12.19 ounces of soda or more. Why do you think this probability is so small?

3. (a) If we have a distribution of *x* values that is **not** normal, do we need to put any restriction on the sample size to claim that the distribution of sample means is approximately normal? Explain.

   (b) If the original distribution of *x* values is known to be normal, do we need to put any restriction on the sample size to claim that the distribution of sample means is normal? Explain.

4. The heights of 18-year-old men are approximately normally distributed, with mean 68 inches and standard deviation 3 inches.

   (a) What is the probability that an 18-year-old man's height is greater than 77 inches?

   (b) Suppose a random sample of nine 18-year-old men is selected. Is this sample size large enough to describe the sampling distribution of $\bar{x}$? If so, describe its distribution and give the mean and standard deviation.

   (c) If a random sample of nine 18-year-old men is selected, what is the probability that the mean height is between 67 and 69 inches?

5. A final exam in Elementary Statistics averages 73 points with standard deviation 7.8 points.

   (a) If 10 students are randomly selected, is it possible to find the probability their mean score is greater than 77? If yes, then find the probability. If no, then explain why not.

   (b) If 35 students are randomly selected, is it possible to find the probability their mean score is greater than 77? If yes, then find the probability. If no, then explain why not.

Confidence Intervals and the TI-83

CONFIDENCE INTERVAL FOR μ (LARGE SAMPLES)

1. STAT
2. Left arrow key once to TESTS menu
3. 7 to get ZInterval
4. Use arrow keys to move cursor over Stats
5. ENTER to select Stats
6. Enter values of s (in place of σ), $\bar{x}$, n, and c (use down arrow once or ENTER to move down line by line)
7. Move cursor over Calculate and hit ENTER
8. Report CI as *lower bound* $< \mu <$ *upper bound* (do not use parentheses format of calculator) and interpret with a complete sentence

CONFIDENCE INTERVAL FOR μ (SMALL SAMPLES)

1. STAT
2. Left arrow key once to TESTS menu
3. 8 to get TInterval
4. Select Stats
5. Enter values of $\bar{x}$, s, n, and c
6. Select Calculate
7. Report CI as *lower bound* $< \mu <$ *upper bound* and interpret

CONFIDENCE INTERVAL FOR p (LARGE SAMPLES)

1. STAT
2. Left arrow key once to TESTS menu
3. ALPHA [A] to get 1-PropZInt
4. Enter values of x (number of successes), n (number of trials), and c
5. Select Calculate
6. Report CI as *lower bound* $< p <$ *upper bound* and interpret

CONFIDENCE INTERVAL FOR $\mu_1 - \mu_2$ (LARGE, INDEPENDENT SAMPLES)

1. STAT
2. Left arrow key once to TESTS menu
3. 9 to get 2-SampZInt
4. Select Stats
5. Enter values of s_1 (in place of σ_1), s_2 (in place of σ_2), $\bar{x}_1$, n_1, $\bar{x}_2$, n_2, and c
6. Select Calculate
7. Report CI as *lower bound* $< \mu_1 - \mu_2 <$ *upper bound* and interpret

CONFIDENCE INTERVAL FOR $\mu_1 - \mu_2$ (SMALL, INDEPENDENT SAMPLES)

1. STAT
2. Left arrow key once to TESTS menu

3. [0] to get 2-SampTInt
4. Select Stats
5. Enter values of $\bar{x}_1$, s_1, n_1, $\bar{x}_2$, s_2, n_2, and c
6. For Pooled select Yes (always leave as Yes)
7. Select Calculate
8. Report CI as $\textit{lower bound} < \mu_1 - \mu_2 < \textit{upper bound}$ and interpret

CONFIDENCE INTERVAL FOR $p_1 - p_2$ (LARGE, INDEPENDENT SAMPLES)

1. [STAT]
2. Left arrow key once to TESTS menu
3. [ALPHA] [B] to get 2-PropZInt
4. Enter values of x_1, n_1, x_2, n_2, and c
5. Select Calculate
6. Report CI as $\textit{lower bound} < p_1 - p_2 < \textit{upper bound}$ and interpret

Chapter 8 – Estimation

Section 8.1 – Estimating μ with Large Samples

1. Using the values of $\bar{x}$ and s you recorded in Section 7.1 (Discovering the Sampling Distribution of $\bar{x}$), find a 95% confidence interval for μ by hand. Does your interval capture the true value of the parameter?

2. A new type of light bulb has been developed which is believed to last longer than ordinary light bulbs. To determine the average life of the new light bulb, a random sample of 100 light bulbs was tested. The sample had a mean life of 1960 hours and a standard deviation of 142 hours. Estimate the true mean life of the new light bulb using a 99% confidence interval. Clearly interpret the meaning of this confidence interval.

3. As part of a Department of Energy survey, 240 families were randomly selected and interviewed concerning the amount of money they spent last year on home heating. The survey indicated that the amount spent by the 240 families had an average of $425 and a standard deviation of $130. Use the sample information to construct a 90% confidence interval for the true mean amount of money spent per family on home heating last year. Clearly interpret the meaning of this confidence interval.