

Statistics Tutorial, Formula Applications and TI-83 Instructions



Revision 01/16/2018

Table of Contents

	Page
Chapter 1	
✓ Parameter and Statistic	3
✓ Quantitative Data and Categorical Data	3
✓ Discrete vs Continuous Data	3
✓ Identify a Population	3
Chapter 2 and 3	
✓ Edit Feature for the Calculator	4
✓ Descriptive Statistics one set of Data	5
✓ Using the calculator for Statistics for one set of Data	5
✓ Input one set of data with Frequencies	6
✓ Using the calculator for finding the data using frequencies	7
Chapter 4	
✓ Concepts of Probabilities	8
✓ Statistical Thinking	10
✓ Calculating odds	11
Chapter 5	
✓ Parameters for Binominal Distribution	12
✓ Binomial probability Distribution Values	13
Chapter 6	
✓ Standard Normal Z Score Probabilities	14
✓ IQs Z Score Probabilities	14
✓ Probability (finding a Z-score given a shaded area)	16
Chapter 7	
✓ Finding Critical Values	18
✓ Finding Confidence Intervals E for a Proportion	19
✓ Finding Confidence Intervals E for a Mean (Known σ)	20
✓ Finding Confidence Intervals E for a Mean (Unknown σ)	21
✓ Determine Minimum Sample Size	22
Chapter 8	
✓ Statistical Hypothesis	23
✓ Types of Hypothesis Tests	24
✓ Hypothesis Tests	25
✓ Using the T – Chart	26
Chapter 10	
✓ Find the Linear Correlation Coefficient, r and P-value	30

Chapter 1

Statistical Literacy and Critical Thinking

Parameter and Statistic "What is a parameter, and what is a statistic?"

*A parameter is a numerical measurement describing some characteristic of a **population**. (p with p)*

*A statistic is a numerical measurement describing some characteristics of a **sample**. (s with s)*

How do quantitative data and categorical data differ?

*Quantitative data consists of **numbers** representing counts or measurement.*

*Categorical data can be separated into different categories' that are distinguished by some characteristic that is **not numerical**.*

How do discrete and continuous data differ?

The numbers of people surveyed in each of the next several Gallup polls

Discrete because it is quantitative and the answer will be countable

The exact heights of individuals in a sample of several statistics students

Continuous because there can be numerous results. Heights can be anywhere from 1 to 8 feet

The number of Super Bowl football games that must be played before one of the teams scores exactly 75 points

Discrete because it is quantitative.

Identifying the Population

In a Gallup poll of 1010 adults in the United States, 55% of the respondents said that they used local TV stations daily as a source of news. Is the 1010 value a statistic or a parameter? Is the 55% value a statistic or a parameter? Describe the population.

The values of 1010 and 55% are both statistics because they are based on a sample.

The population would consist of all adults in the United States.

Chapter 2 and 3:

EDIT Feature using the Calculator

EDIT – 1: Edit... allows the entry (in columns) and modifications of data to be used with the CALC and TESTS features.

Press **STAT, EDIT, 1: Edit**

Press **ENTER**, and you may delete, enter or modify data that is already in the list columns.

```

2001 CALC TESTS
1: Edit...
2: SortA(
3: SortD(
4: ClrList
5: SetUpEditor
  
```

Clearing lists

Place cursor on top of the list name (L1, etc.) at the top of the column. Press the **CLEAR** key, then press **ENTER**. When you press the **Down arrow** all existing data will be removed from that column.

L1	L2	L3	2	L1	L2	L3	2
10	23				23		
12	21				21		
13	25				25		
20	30				30		
18	32				32		
17	40				40		
-----				---			

Changing existing entry

Change the 18 to 20. Press **STAT, 1: Edit, ENTER**, Move down arrow to the highlight 20, type 18, and hit **ENTER**. The 18 will replace the 20.

L1	L2	L3	1	L1	L2	L3	1
16				16			
12				12			
13				13			
20				20			
-----				-----			
L1(5)=18							

Return to Home Screen

Press **2nd MODE [QUIT]** to end the **EDIT** mode and return to the home screen



Descriptive Statistics for one set of Raw Data

Find the midrange, mean, median, and mode for the given sample data.

Listed below are the ages of motorcyclists when they were fatally injured in 18 traffic crashes.

17 38 27 14 18 34 16 42 28
24 40 20 23 31 37 21 30 25

14 16 17 18 20 21 23 24 **25 27** 28 30 31 34 37 38 40 42

$$\text{Midrange} = \frac{42+14}{2} = 28$$

$$\text{Median} = \frac{25+27}{2} = 26$$

$$\text{Mean} = \frac{\sum n}{n} = 26.9 = 27$$

Mode None

Using the Calculator

Given 5 numbers determine the mean, median, and mode

Press **STAT, EDIT, 1: Edit**,

Press **ENTER**

EDIT CALC TESTS
1. Edit

Clear the home screen. (See Edit instructions)

Enter all numbers into L1. Press **ENTER** after each entry until all the numbers are inputted.

(10,12,13, 20,18)

L1	L2	L3	1
10	-----	-----	
12			
13			
20			
18			

L1(5)=18			

Once all data is entered. Press **STAT**, right arrow to **CALC**.

After 1-Var Stats press **2nd** and **1** to enter L1



Press **ENTER** for the results.

EDIT **CALC** TESTS
1. 1 Var Stats

1 - Var Stats L1

1-Vars Stats
 $\bar{x} = 14.6$
 $\sum x = 73$
 $\sum x^2 = 1137$
 $Sx = 4.219004622$
 $\sigma x = 3.77592453$
 $n = 5$
 MinX = 10
 Q1=11
 Med = 13

Summary Statistics for one set of Raw Data using Frequency Distribution

Statistical Literacy and Critical Thinking

A set of data has a maximum value of 74, a median value of 61, and a minimum value of 48.

Use the range rule of thumb to estimate the standard deviation, rounding to the nearest tenth.

$$s = \frac{\text{range}}{4} = \frac{13}{4} = 3.25 = 3.3$$

Analysis of Frequency Distribution

Identify, the class width, class midpoints and class boundaries.

Age (years)	Frequency	Mid points	Class boundaries
9-11	3	10	8.5 – 11.5
12-14	2	13	11.5 – 14.5
15-17	5	16	14.5 – 17.5
18-20	3	19	17.5 – 20.5
21-23	6	22	20.5 – 23.5
24-27	4	25	23.5 – 26.5

Class width = 3 (12 - 9)

The difference between the first numbers of two consecutive class intervals.

Class midpoints = $\frac{11+9}{2} = 10$

The midpoint of a class is the sum of the first and last number in the class divided by 2.

Class boundaries

The sum of the last number and first number of two consecutive class intervals divided by 2. $\frac{11+12}{2} = 11.5$

Is it possible to identify the actual age of the 5 people in age from 15 – 17 years?

No! The upper and lower limits include the specific ages. You know 5 people have ages between 15– 17 years, but you do not know their actual age.

Using the Calculator

Given 6 class midpoints and their related frequencies, determine the mean, median, and mode.

Press **STAT**, **EDIT**, 1: **Edit**,

EDIT CALC TESTS
1. Edit

Press **ENTER**,

Clear the home screen (see **EDIT** instructions)
Enter the class midpoints of your data into L1 (10, 13, 16, 19, 22, 25).

L1	L2	L3	Z
10	3		
13	2		
16	5		
19	3		
22	6		
25	4		

L2(7) =			

Enter the frequencies of your class midpoints into L2 (3, 2, 5, 3, 6, 4).

EDIT CALC TESTS
1. 1 Var Stats

Once all data is entered. Press **STAT**, right arrow to **CALC**.

1: 1-Var Stats

1-Var Stats

Press **ENTER**.

With class midpoints in L_1 and corresponding frequencies in L_2 press



1-Var Stats (L_1, L_2)

Press **ENTER** for results

Mean = 18.48

Median = 19.0

Mode = 22

1-Vars Stats
 $\bar{x} = 18.47826087$
 $\sum x = 425$
 $\sum x^2 = 8405$
 $Sx = 5.007898899$
 $\Sigma x = 4.897821740$
 $n = 23$
 Min x = 10
 Q1=16
 Med = 19

Chapter 4:

Basic Concepts of Probability

Criteria for a Probability

Which of the following values may NOT be a probability?

- a) $\frac{1}{3}$ b) -1 c) 0.87 d) 1.01 e) 75%

b and d Probabilities cannot be minus or >1

Relative Frequency Probability

If a person is randomly selected, find the probability that his or her birthday is October 17. Ignore leap years.

$$P(A) = \frac{\text{number of times } A \text{ occurred}}{\text{number of times trial was repeated}} = \frac{1}{365}$$

If a person is randomly selected, find the probability that his or her birthday is in October.

$$P(A) = \frac{\text{number of times } A \text{ occurred}}{\text{number of times trial was repeated}} = \frac{31}{365}$$

Multiplication Rule

A set of digits {0, 1, 2, 3, 4, 5, 6, 7} will be used to create a five-digit code. How many five-digit codes can be created if...

- there are no restrictions? **$8^5 = 32768$**
- the digits may not be repeated? **$8 \times 7 \times 6 \times 5 \times 4 = 6720$**
- the code must begin with a 3 and the digits may be repeated? **$8^4 = 4096$**

Two different cards are drawn from a standard deck of 52 cards. What is the probability that the first card is a king and the second card is a queen?

$$\frac{4}{52} \times \frac{4}{51} = \frac{16}{2652} = .00603$$

A box of marbles has 3 yellow marbles, 2 green marbles, and 5 red marbles. Suppose that two marbles are selected without replacement. What is the probability that neither selected marble is green?

$$\frac{8}{10} \times \frac{7}{9} = \frac{56}{90} = .6222$$

<p>Factorial $n!$</p> <p>Find the value of $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$</p> <p>Using the calculator</p> <p>First enter “n” (5), then select MATH->PRB-then Scroll down to ! and hit ENTER</p>	<p>Math Num CPX PRE 1: r and 2: ${}_n P_r$ 3: ${}_n C_r$ 4: !</p>
<p>Then hit ENTER again and the answer 120 appears</p>	<p>5! 120</p>
<p>Permutations nPr</p> <p>If a president and vice president are to be selected from 5 committee members how many permutations can be selected. In this example order is important so we use Permutations ${}_n P_r$</p> ${}_5 P_2 = \frac{5!}{(5-2)!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1} = \frac{120}{6} = 20$ <p>Using the calculator</p> <p>First enter “n”(5), then select MATH->PRB-then Scroll down to nPr and hit ENTER</p>	<p>Math Num CPX PRE 1: r and 2: ${}_n P_r$</p>
<p>Press 2 and hit ENTER again and the answer 20 appears</p>	<p>${}_5 P_2$ ANS 20</p>
<p>Combinations ${}_n C_r$</p> <p>If 3 committee members are select from 5 employees how many combinations can be selected. In this example order is not important so we use Combinations ${}_n C_r$</p> ${}_5 C_3 = \frac{5!}{(5-3)!(3!)} = \frac{5 \times 4 \times 3 \times 2 \times 1}{2 \times 1 \times 3 \times 2 \times 1} = \frac{120}{12} = 10$ <p>Using the calculator</p> <p>First enter “n”(5), then select MATH->PRB-then Scroll down to ${}_n C_r$ and hit ENTER</p>	<p>Math Num CPX PRE 1: r and 2: ${}_n P_r$ 3: ${}_n C_r$</p>
<p>Press 3 and hit ENTER again and the answer 10 appears</p>	<p>${}_5 C_3$ ANS 10</p>

Statistical Thinking

During a certain week the number of late and on-time flights were counted for three airlines.

	Number of On-time Flights	Number of Late Flights	Total
Southern Airlines	34	11	45
Statewide Airlines	28	3	31
Westward Airlines	30	8	38
Total	92	22	114

What is the probability that the flight was a Westward Airlines flight and was late?

$$\frac{8}{114} = .0702$$

What is the probability that the flight was late, **given** that the flight was on Westward Airlines?

$$\frac{8}{38} = .2105$$

An experiment has a sample space given by $S = \{1,2,3,4,5,6,7,8,9,10\}$. The events D, E, and F are defined as $D = \{1,2,3,4\}$, $E = \{2,4,6,8,10\}$, and $F = \{3,5,7\}$.

Assume each outcome is equally likely:

Are events D and E disjoint (mutually exclusive)? **No**

Are events D and F disjoint (mutually exclusive)? **No**

Are events E and F disjoint (mutually exclusive)? **Yes**

Determine whether events G and H are independent or dependent.

G: Joe forgets to set his alarm clock. H: Joe is late for work the next day. **Dependent**

G: Milk is on sale. H: Notebook paper is on sale. **Independent**

G: A student earns an "A" in English 1. H: A student in English 1 is a girl. **Independent**

Calculating Odds

A box contains 6 white chips, 3 blue chips, 2 yellow chips, and 1 green chip. A player randomly selects one chip from the box. If the player selects a blue chip, he/she wins \$3. If the player selects any other color chip, he/she loses \$1

Find the player's odds of winning?

$$\text{odds} = \frac{\# \text{ favorable}}{\# \text{ unfavorable}} = \frac{3}{9} = \frac{1}{3} \quad 1:3$$

Find the player's expected value?

$$u = \sum x \cdot P(x) = 3\left(\frac{3}{12}\right) + -1\left(\frac{9}{12}\right) = 0$$

When you give a casino \$5 for a bet on the "pass the line" in the game of craps, there is a 244/495 probability that you will win \$5 and a 251/495 probability that you will lose \$5.

What are your odds of winning?

$$\text{odds} = \frac{\# \text{ favorable}}{\# \text{ unfavorable}} = 244:251$$

What is your expected value?

$$u = \sum x \cdot P(x) = 5\left(\frac{244}{495}\right) + -5\left(\frac{251}{495}\right) = -.0707$$

In the long run, how much do you lose for each dollar bet? $\frac{-0.0707}{5} = -0.014 = 1.4 \text{ cents}$

Chapter 5:

Parameters for Binomial Distribution

Among 945 births, there were 879 girls born to parents. Assuming boys and girls are equally likely, find the mean and standard deviation for the number of girls born in this group of 945 babies.

$$u = n \cdot p \quad \text{mean binomial}$$

$$\sigma = \sqrt{n \cdot p \cdot q} \quad \text{standard deviation binomial}$$

$$\sigma^2 = n \cdot p \cdot q \quad \text{variance binomial}$$

For this example, $n = 945$ and $p = 0.50$ $q = 0.50$

Mean: $u = n \cdot p \quad u = 945 \cdot 0.50 = 472.5$ girls

Standard Deviation: $\sigma = \sqrt{945 \cdot .5 \cdot .5} = 15.3704 = 15.4$ girls

variance binomial $\sigma^2 = n \cdot p \cdot q = 945 \cdot 0.50 \cdot 0.50 = 236.25$

Find the minimum usual number and maximum usual number of girls born in this group of 945.

Minimum usual value: $\mu - 2\sigma = 472.5 - 2(15.4) = 441.7$ girls

Maximum usual value: $\mu + 2\sigma = 472.5 + 2(15.4) = 503.3$ girls

Binomial Probability Distribution Values

<p>Find the probability that 2 successes will occur out of 5 trials if the probability of success is 0.3.</p> $P(x) = \frac{n!}{(n-x)!x!} \cdot p^x q^{n-x} \text{ binomial probability}$ $Q = 1.0 - p = 1.0 - .3 = .7$ $P(2) = \frac{5!}{(5-2)!2!} \cdot .3^2 \cdot .7^3 = .3087$	
<p>Using the calculator:</p> <p>Press 2nd VARS [DISTR].</p> <p>Scroll down to 10:binompdf(</p> <p>Press ENTER</p>	<div style="border: 1px solid black; padding: 5px;"> <p>DISTR Draw 10. binompdf(</p> </div>
<p>Enter n, p, x-value n = 5, p = 0.3, x = 2</p> <p>and press ENTER to get the answer .3087</p>	<p>Binompdf (5, .3, 2)</p>
<p>Find the probability that 2 or less successes will occur out of 5 trials if the probability of success is 0.3. That is, find $P(x \leq 2)$.</p>	
<p>Press 2nd VARS [DISTR].</p> <p>Scroll down to A:binomcdf(</p> <p>Press ENTER.</p>	<div style="border: 1px solid black; padding: 5px;"> <p>DISTR Draw A. binomcdf(</p> </div>
<p>Enter n, p, x-value n = 5, p = 0.3, x = 2</p> <p>and press ENTER to get the answer .83692.</p>	<p>binomcdf (5, .3, 2)</p> <p style="text-align: right;">.83692</p>

Chapter 6:

Standard Normal Z Score Probabilities

Find the probability that a z-score is between -1.5 and 2. That is, find $P(-1.5 \leq z \leq 2)$

Note: $\mu = 0$ because we are talking Standard Normal Distribution
 $\sigma = 1$ because we are talking Standard Normal Distribution

$$z = \frac{x - \mu}{\sigma} \text{ standard score } z = \frac{-1.5 - 0}{1} = -1.5 \quad z = \frac{2 - 0}{1} = 2.0$$

Probability for $Z = -1.5$ is .0668 Probability for $Z = 2.0$ is .9772 $P(-1.5 \leq z \leq 2) = .9104$

Using the Calculator

Press **2nd VARS [DISTR]**.

Scroll down to **2: normal cdf** Press **ENTER**

Since z-scores are for the Standard Normal Distribution, the sequence is normalcdf (smaller z, larger z).

Enter (smaller IQ, larger IQ,)

Press **ENTER** to get the answer **.91044**.

DISTR Draw
2. normalcdf(

normalcdf (-1.5, 2)

IQs Z Score Probabilities

Adult IQs are normally distributed with $\mu = 100$ and $\sigma = 15$. Find the probability that a randomly selected IQ is less than 112. That is, find $P(x < 112)$.

$$z = \frac{x - \mu}{\sigma} \text{ standard score } z = \frac{112 - 100}{15} = 0.8 \text{ The probability of a Z score of 0.8 is .7881}$$

Using the Calculator

Press **2nd VARS [DISTR]**.

Scroll down to **2: normal cdf** Press **ENTER**

Note: Use -9999 for the smaller IQ score Make sure to use (-) key Enter (smaller IQ, larger IQ, μ , σ)

Press **ENTER** to get the answer **.7881**

DISTR Draw
2. normalcdf(

normalcdf (-9999, 112, 100, 15)

IQs Z Score Probabilities

Adult IQs are normally distributed with $\mu = 100$ and $\sigma = 15$. Find the probability that a randomly selected IQ is at least 122. That is, find $P(x \geq 122)$.

$$z = \frac{x - \mu}{\sigma} \text{ standard score } z = \frac{122 - 100}{15} = 1.47 \text{ The probability of a Z score of 1.47 is .9292}$$

Because we need at least, the $P = 1 - .9292 = .0708$

Using the Calculator

Press **2nd VARS [DISTR]**.

Scroll down to **2: normal cdf** Press **ENTER**

DISTR Draw 2. normalcdf(

Note: Use 9999 for the larger IQ score

Enter (smaller IQ, larger IQ, μ , σ)

Press **ENTER** to get the answer **.0712334139**

nomalcdf 122, 9999,
100, 15)

IQs Z Score Probabilities

Adult IQs are normally distributed with $\mu = 100$ and $\sigma = 15$. Find the probability that a randomly selected IQ is between 112 and 122. That is, find $P(112 \leq x \leq 122)$.

$$z = \frac{x - \mu}{\sigma} = \frac{122 - 100}{15} = 1.47 \text{ The probability of a Z score of 1.47 is .9292}$$

$$z = \frac{x - \mu}{\sigma} = \frac{112 - 100}{15} = 0.8 \text{ The probability of a Z score of 0.8 is .7881}$$

$$P(112 \leq x \leq 122) = .9292 - .7881 = .1411$$

Using the Calculator

Press **2nd VARS [DISTR]**.

Scroll down to **2: normal cdf** Press **ENTER**

DISTR Draw 2. normalcdf(

Enter (smaller IQ, larger IQ, μ , σ)

Press **ENTER** to get the answer **.1406**

nomalcdf (112, 122,
100, 15)

IQs Z Score Sample Probabilities

Adult IQs are normally distributed with $\mu = 100$ and $\sigma = 15$. Find the probability if 9 students are randomly selected that one of those selected has an IQ of at least 122. That is, find $P(x \geq 122)$.

$$\sigma_x = \frac{\sigma}{\sqrt{n}} \text{ central limit theorem (standard error)} \quad \sigma_x = \frac{15}{\sqrt{9}} = 5 \quad z = \frac{x - \mu}{\sigma_x}$$

$$z = \frac{122 - 100}{5} = 4.4 \quad \text{Probability of .9999} \quad P(x \geq 122) = 1 - .9999 = .0001$$

Using the Calculator

Press **2nd VARS [DISTR]**.

Scroll down to **2: normal cdf** Press **ENTER**

DISTR Draw
2. normalcdf(

The sequence is **normalcdf (smaller, larger, μ , σ_x)**

Enter 122,9999,100,5)

Press **ENTER** to get the answer **.00000416**

normalcdf (122, 9999,
100, 5)

Probability (finding a Z-score given a shaded area)

Given the shaded area is .75. Find the z-score for this area using Positive Z score chart or technology. Z chart shows a Z score of 0.67 for a probability of .75



Press **2nd VARS [DISTR]**.

Scroll down to **3:invNorm** Press **ENTER**.

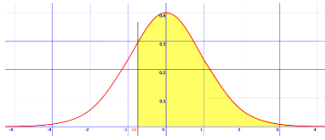
DISTR Draw
3. invNorm

Enter .75) and press **ENTER** to get the answer **Z = .6744897495**

InvNorm (.75)

Probability (finding a Z-score given a shaded area)

Given the shaded area is .75. Find the z-score for area to the left of this shaded area using Negative Z score chart or technology. Z chart shows a Z score of -0.67 for a probability of .25



Subtract .75 from 1 to get the unshaded area. $1 - .75 = .25$

Press **2nd VARS [DISTR]**.

Scroll down to **3:invNorm**

Press **ENTER**.

Enter .25) and press **ENTER** to get $Z = -.6744897495$

DISTR Draw
3. invNorm

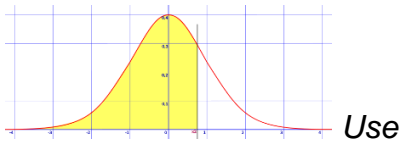
InvNorm (.25)

-.

Probability (finding a Q-score given a shaded area)

Adult IQs are normally distributed with $\mu = 100$ and $\sigma = 15$. Find the IQ scores for those individuals whose IQ exceed 75% of the population. Z chart shows a Z score of 0.67 for a probability of .75

$$z = \frac{x - \mu}{\sigma} \quad .67 = \frac{x - 100}{15} \quad x = .67(15) + 100 \quad x = 110.05$$



Press **2nd VARS [DISTR]**.

Scroll down to **3:invNorm** Press **ENTER**.

Enter .75) and press **ENTER** to get the answer $Z = .6744897495$

The IQ score for those that exceed 75% of the population is equal to: $IQ > 100 + (.67 \times 15)$

$$IQ > 100 + 10.1175 = 110.05$$

DISTR Draw
3. invNorm

InvNorm (.75)

Chapter 7

Finding the “Critical Value” for a specific “Confidence Level”

Example: Find the “Critical Value” for a “Confidence Level of 90%”.

$z_{\alpha/2}$ is also known as the *critical value for two tailed data*.

1. $\alpha = 1 - \text{confidence level} = 1 - .90 = .10$
2. $\frac{\alpha}{2} = \frac{.10}{2} = .05$
3. $1 - .05 = .95$
4. After looking up a .95 on the Z-score chart (Table A-2), the Critical value Z-score. was determined to be 1.645

Therefore the *critical value* of a 90% Confidence Level is 1.645

The table below was developed using the procedure outlined above.

Confidence Level	Critical Value (Z-score)
0.90	1.645
0.91	1.70
0.92	1.75
0.93	1.81
0.94	1.88
0.95	1.96
0.96	2.05
0.97	2.17
0.98	2.33
0.99	2.575

Finding Confidence Intervals E for a Proportion

$$E = z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} \quad \text{proportion } (\hat{p} \text{ and } \hat{q} \text{ are known})$$

$$E = z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} \quad \text{proportion (use 0.25 when p and q are not known)}$$

Sample: A research poll 200 respondents were asked if they felt vulnerable to identity theft. Thirty said yes. Find the confidence interval given a confidence level of 95%.

$N = 200$ $x = 30$ confidence interval of 95% $z_{\alpha/2} = 1.96$ $p = \frac{30}{200} = .15$ $q = 1 - .15 = .85$

$$E = z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.96 \sqrt{\frac{.15 \times .85}{200}} = 0.4949$$

Using the calculator:

Press **STAT** and the right arrow twice to select **TESTS**.

Use the down arrow to select **A:1-PropZInt...**
Press **ENTER**.

EDIT CALC TESTS
 A: 1-PropZInt

Enter number of favorable outcomes, x , sample size, n , and confidence level.

Use left arrow to select **Calculate** and press **ENTER**.

1-PropZInt
 x:30
 n:200
 C-Level:.95
 Calculate

Results:

95% confidence interval is 0.10051 to 0.19949 when $P = 0.15$ and sample size is 200.

$$E = \frac{.19949 - .10051}{2} = \frac{.09898}{2} = 0.4949$$

```

1 - PropZInt
(.10051,.19949)
^
p = .15
n = 200
  
```

To find a confidence interval for Mean (known σ)

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Example: A sample of 5 students were asked if they felt their test was fair. Three said yes. Find the confidence interval given a confidence level of 95% and a σ of 5.

$\sigma = 5$, mean = 3, $n = 5$, Confidence level = .95 $z_{\alpha/2} = 1.96$

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 1.96 \frac{5}{\sqrt{5}} = 4.3827 = 4.383$$

Or use the calculator:

Press **STAT** and the right arrow twice to select **TESTS**.

EDIT CALC TESTS
7: ZInterval

Use the down arrow to select **7:Zinterval...**

Press **ENTER**.

Use right arrow to select **Stats**

```
ZInterval
Inpt:Data Stats
σ:5
x̄:3
n:5
C-Level:.95
Calculate
```

Enter population standard deviation, mean, sample size, and confidence level.

Press down arrow to select Calculate and press **ENTER**.

Results:

The 95% confidence interval for Distribution is -1.38 to 7.83 when the mean is 3 the population standard deviation is 5.0, and the sample size is 5.

Z Interval
(-1.3837, 7.3826)
 $\bar{x} = 3$
 $n = 5$

$$E = \frac{7.3826 - (-1.3837)}{2} = 4.383$$

To find a confidence interval for Mean (unknown σ)

Example: Find the confidence interval for a sample of 4 students given a sample mean of 6.8, a sample standard deviation of 1 and a confidence level of 95%

Use the Table A3.

Go to row on the chart for a degree of freedom of 3 ($n - 1$).

This is a two tail.

Under two tail the values are :

0.01	0.02	0.05	0.10	0.20
5.841	4.541	3.182	2.353	1.638

In the row for degree of freedom of 3, the h number on this line for 95% confidence is 3.182 corresponds to a one tail of 0.005.

$$t_{\alpha/2} = 3.182 \quad E = t_{\alpha/2} \frac{s}{\sqrt{n}} = 3.182 \frac{1}{\sqrt{4}} = 1.591$$

Using the calculator:

Press **STAT** and the right arrow twice to select **TESTS**.

EDIT CALC TESTS
8: TInterval

Use the down arrow to select **8:TInterval...**

Press **ENTER**.

Use right arrow to select **Stats**

Enter mean, sample standard deviation, sample size, and confidence level.

Press down arrow to select Calculate and press **ENTER**.

```
TInterval
Inpt:Data Stats
x̄:6.8
Sx:1
n:4
C-Level:.95
Calculate
```

Results:

The 95% confidence level for the Student's t Distribution is 5.21 to 8.39 when the mean is 6.8, the sample standard deviation is 1.0, and the sample size is 4

$$E = \frac{8.3912 - 5.2088}{2} = 1.591$$

T Interval
(5.2088, 8.3912)
x = 6.8
Sx = 1.0
n = 4.0

Determine Minimum sample size

(Note: When finding the minimum sample size, always round up.)

Proportion (\hat{p} and \hat{q} are known)

Sample question: 41% of Jacksonville residents said that they had been in a hurricane. How many adults should be surveyed to estimate the true proportion of adults who have been in a hurricane, with a 95% confidence level and an interval of 6%?

$$E = \frac{\text{interval}}{2} = \frac{.06}{2} = .03 \quad q = 1 - P = .59 \quad \text{Critical value for 95\% confidence level} = 1.96$$

$$n = \frac{[z_{\alpha/2}]^2 \hat{p} \hat{q}}{E^2} \quad \text{proportion } (\hat{p} \text{ and } \hat{q} \text{ are known})$$

$$n = \frac{[1.96]^2 (.41)(.59)}{.03^2} = 1032.536711 = 1032.536711 = 1033$$

Proportion (\hat{p} and \hat{q} are not known)

Sample question: Jacksonville residents said that they had been in a hurricane. How many adults should be surveyed to estimate the true proportion of adults who have been in a hurricane, with a 95% confidence level and an interval of 6%?

$$n = \frac{[z_{\alpha/2}]^2 \cdot 0.25}{E^2} \quad \text{proportion } (\hat{p} \text{ and } \hat{q} \text{ are not known})$$

$$n = \frac{[1.96]^2 (.25)}{.03^2} = 1067.1111 = 1067.1111 = 1068$$

Proportion (known standard deviation of the population σ)

Sample question: Suppose we want to know the average age of a Florida State College student, plus or minus 0.5 years. We'd like to be 99% confident about our result. From a previous study, we know that the standard deviation for the population σ is 2.9.

$$n = \left[\frac{z_{\alpha/2} \sigma}{E} \right]^2$$

$$n = \left[\frac{2.575(2.9)}{.5} \right]^2 = 223.054225 = 223.054225 = 224$$

Chapter 8:

Statistical Hypotheses

The best way to determine whether a statistical hypothesis is true would be to examine the entire population. Since that is often impractical, researchers typically examine a random sample from the population. If sample data are not consistent with the statistical hypothesis, the hypothesis is rejected.

There are two types of statistical hypotheses.

- **Null hypothesis.** The null hypothesis, denoted by H_0 , is usually the hypothesis that sample observations result purely from chance.

✓ H_0 is always =.

- **Alternative hypothesis.** The alternative hypothesis, denoted by H_1 , is the hypothesis that sample observations are influenced by some non-random cause.

✓ H_1 is never =. Options are <, >, ≠.

Chart for establishing $H_0: P$ and $H_1: P$

Claim	$H_0: P$	$H_1: P$	Comments
$P > 5$	$H_0: P = 5$	$H_1: P > 5$	
$P < 5$	$H_0: P = 5$	$H_1: P < 5$	
$P = 5$	$H_0: P = 5$	$H_1: P \neq 5$	If difference between the assumption and the claim is close. For instance, Claim 20% of adults smoke. Survey showed 21% of respondents smoke.
$P \neq 5$	$H_0: P = 5$	$H_1: P \neq 5$	If difference between the assumption and the claim is close. For instance, Claim 20% of adults smoke. Survey showed 21% of respondents smoke.
$p \geq 5$	$H_0: P = 5$	$H_1: P < 5$	H_1 is opposite the claim
$p \leq 5$	$H_0: P = 5$	$H_1: P > 5$	H_1 is opposite the claim

Types of Hypothesis Tests - P Value One tail vs two tail

$H_1: P <$ Left Tailed test Actual P value

$H_1: P >$ Right-tailed test $1 -$ actual P value

$H_1: P \neq$ Two tailed test. If test is a 2 tail, you have to double your calculated P value

Acceptance/Rejection

Condition	Conclusion
Original claim does not include equality and you reject H_0 because the p value is $<$ or $=$ to given Critical value α .	There is sufficient evidence to support the original claim
Original claim does not include equality and you fail to reject H_0 because the p value is $>$ than the given Critical value α .	There is not sufficient evidence to support the original claim
Original claim includes equality and you reject H_0 because the p value is $<$ or $=$ to the given Critical value α .	There is sufficient evidence to warrant rejection of the original claim
Original claim includes equality and you fail to reject H_0 because the p value is $>$ than the given Critical value α .	There is not sufficient evidence to warrant rejection of the original claim

Hypothesis Tests for Proportions (Test Statistic, P value and sample proportion).

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Example: In sampling 200 people, we found that 30% of them favored a certain candy. Use $\alpha = 10\%$ to test the hypothesis that the proportion of people who favored that candy is less than 35%.

Solution: This represents a one-sample test of proportion. So we use the "**1-PropZTest**" function. The sample proportion is 30% or $p = 0.30$, and the hypothesized claim is $p \geq 0.35$. Therefore, the hypotheses are $H_0: p = 0.35$ and $H_1: p < 0.35$. This is a two tail test because the claim includes an equal sign.

[NOTE: If the p-value $< \alpha$, reject the null hypothesis; otherwise, do not reject the null hypothesis].

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{.30 - .35}{\sqrt{(.35)(.65) / 200}} = -1.48 \quad P = .0694$$

This is a left tail test because $H_1: P <$ so $P = .0694$

OR Use the Calculator:

Press **STAT** and the right arrow to select **TESTS**.

Use the down arrow to select **5.1-PropZTest**

Press **ENTER**

Enter: $H_0 p = 0.35$, number of favorable outcomes (x), sample size (n), and select the alternate hypothesis $p < 0.35$

Use down arrow to select Calculate and press **ENTER**

Results:

The Test Statistic is -1.48

Since the p value = 0.069 is less than $\alpha = 0.10$, reject the null hypothesis $H_0 p = 0.35$

The hypothesized claim is $p \geq 0.35$ includes equality. We rejected H_0 and conclude that there is sufficient evidence to warrant rejection of the original claim.

EDIT CALC TESTS

5: 1-PropZTest

```
1-PropZTest
P0: .35
x: 60
n: 200
PROP<P0 [ ] >P0
Calculate Draw
```

```
1-PropZTest
PROP<.35
z=-1.482498633
P=.0691038658
P=.3
n=200
```

Hypothesis Tests for Proportions for a Mean (known σ)

$$z = \frac{\bar{x} - u}{\sigma / \sqrt{n}}$$

Example: A sample of size 200 has a mean of 20 and a standard deviation of 6 and $\sigma = 6$. Test the hypothesis claim that the population mean is not equal to 19.2 ($\alpha = 10\%$ level of significance).

$$z = \frac{\bar{x} - u}{\sigma / \sqrt{n}} = \frac{19.2 - 20}{6 / \sqrt{200}} = -1.8856 \quad P = .02968 \text{ from chart}$$

This is a two tail test ($H_1: P \neq$) so $P = 2 \times .02968 = .05936$

OR Use the Calculator:

Solution: "The population mean is $\neq 19.2$ means the null hypothesis is $H_0 = 19.2$, and the alternate hypothesis is $H_1 \neq 19.2$.

[NOTE: If the p-value $< \alpha$, reject the null hypothesis; otherwise, do not reject the null hypothesis].

Press **STAT** and the right arrow twice to select **TESTS**.

Use the down arrow to select **1:Z-Test**

Press **ENTER**.

EDIT CALC TESTS
1: Z-Test

Use right arrow to select Stats

Enter given hypothesized mean, standard deviation, mean, and sample size.

Select alternate hypothesis.

Press down arrow to select Calculate and press **ENTER**.

```
Z-Test
Inpt:Data Stats
μ₀:19.2
σ:6
x̄:20
n:200
μ:≠μ₀ <μ₀ >μ₀
Calculate Draw
```

Results:

Since the p-value is 0.059346 is less than $\alpha = 0.10$, reject the null hypothesis $H_0 = 19.2$.

```
Z Test
u ≠ 19.2
Z = 1.885618083
P = .059346
x̄ = 20.0
n = 200
```

Using the T-Chart

Example: A researcher claims that the mean speed of a Miss Williams tennis serve is 125 mph. Therefore, $H_0 = 125$ and the alternate hypothesis is $H_1 \neq 125$. Assume that data is normally distributed for a sample test of 40 serves and the value of α is unknown. Find the critical values for the hypothesis test with a significance level of 0.05. This is a two tail test because equal is included in the claim.

Go to the Table 3 Distribution Critical t values. Look in the row for a degree of freedom (df) of 39 (40-1) and the column for two tail significance level of 0.05. The answer is ± 2.023 .

This is a two tail.

Under two tail the values are:

Two Tail	0.01	0.02	0.05	0.10	0.20
df 39	2.708	2.426	2.023	1.685	1.304

In the row for degree of freedom of 39, the number on the line for two tail of 0.05.

Example: A researcher claims that the mean speed of a Miss Williams tennis serve is > 125 mph. Therefore, $H_0 = 125$ and the alternate hypothesis is $H_1 > 125$. Assume that data is normally distributed for a sample test of 40 serves and the value of α is unknown. Find the critical values for the hypothesis test with a significance level of 0.05. This is a one tail test because equal is not included in the claim.

Go to the Table 3 Distribution Critical t values. Look in the row for df is 39 (40-1) and the column for one tail significance level of 0.05. The answer is 1.685.

Under one tail the values are:

OneTail	0.005	0.01	0.025	0.05	0.10
df 39	2.708	2.426	2.023	1.685	1.304

The P-value for a right-tailed test when the given test statistic is not on the chart.

Example with $n = 28$ and test statistic $t = 3.026$.

The degree of freedom is $28 - 1 = 27$. This is a one tail test.

On the t-Distribution chart go to the row for degree of freedom of 27, the highest number on this line, 2.771 corresponds to a one tail of 0.005.

One Tail	0.005	0.01	0.025	0.05	0.10
df 27	2.771	2.473	2.052	1.703	1.314

The $t = 3.026$ given in the problem is higher than the highest value of 2.771 for .005, therefore the probability must be less than 0.005.

Go to 2nd – VARS – Tcdf hit ENTER

Tcdf (lower, highest, df)

After Tcdf enter (3.026, 9999, 27)

Hit ENTER P value = .0026951941

Hypothesis Tests for a Mean (unknown σ)

Find test statistic and p value of a T-Test

Example: A sample of size 20 has a mean of 110 and a standard deviation of 16. Use the TI-83 calculator to test the hypothesis claim that the population mean is greater than 100 with a level of significance of $\alpha = 5\%$.

$$t = \frac{\bar{x} - u}{\frac{s}{\sqrt{n}}} = \frac{100 - 110}{\frac{16}{\sqrt{20}}} = -2.7951$$

Use technology to find the P-value for a one-tailed test with $n = 20$ and test statistic $t = -2.7951$. This is a T test. On the t chart for df of 24 the value 2.7951 does not exist.

Go to 2nd – VARS – Tcdf hit ENTER

Tcdf (lower, highest, df)

Tcdf (-9999, -2.7951, 19)

Enter you data , scroll down to Paste and press **ENTER**

Press **ENTER** again to see your answer $P = .0057731153 = .0058$

Solution: “The population mean is greater than 100 means the alternate hypothesis is $H_1: \mu > 100$, and the null hypothesis is $H_0: \mu = 100$.

Press **STAT** and the right arrow twice to select **TESTS**.

To select the highlighted
2: T-Test...
Press **ENTER**.

EDIT CALC TESTS
2: T-Test

Use right arrow to select Stats (summary values rather than raw data) and Use the down arrow to Enter the hypothesized mean, sample mean, standard deviation, and sample size. Select alternate hypothesis. Press down arrow to select Calculate and press **ENTER**.

```
T-Test
Inpt:Data  Stats
μ₀:100
x̄:110
Sx:16
n:20
μ:≠μ₀ <μ₀ >μ₀
Draw
```

Results for both methods:
The test statistic $t = 2.7951$.

Since the p-value of 0.0058, is less than $\alpha = 0.05$, reject the null hypothesis $H_0 \mu = 100$.

```
T-Test
μ>100
t=2.795084972
p=.0057733042
x̄=110
Sx=16
n=20
```

Original hypothesis claim that the population mean is greater than 100 does not include equality and you reject H_0 because the p value is $<$ or $=$ to given Critical value α . There is sufficient evidence to support the original claim.

Chapter 10:

To find the linear correlation coefficient, r , p -value

Example: Given sample of size 5 and a level of significance of $\alpha = 5\%$. find the linear correlation, r and p -value with a hypothesis of >0 .

x	230	265	358	480	530
y	15.9	15.7	15.4	15.3	14.9

Make sure the data for the x data is entered in L1 and y data is entered in L2.

Press **STAT, TESTS**, and press the down arrow to **E:LinRegTTest**. Press **ENTER**.

Select the alternate hypothesis > 0 .

Use the down arrow to get to calculate and press **ENTER**.

Results:

If linear correlation r is to the left or right of the Table 5 – values, the r lies within a critical region.

Since the linear correlation $r = -.959$ is to the left of Table 5 values of $\pm .878$ and the p -value is $0.005 < \alpha = (0.05)$ support the null hypothesis.

Conclude that there is a positive relationship between the x and y values.

EDIT CALC TESTS
E: LinRegTTest

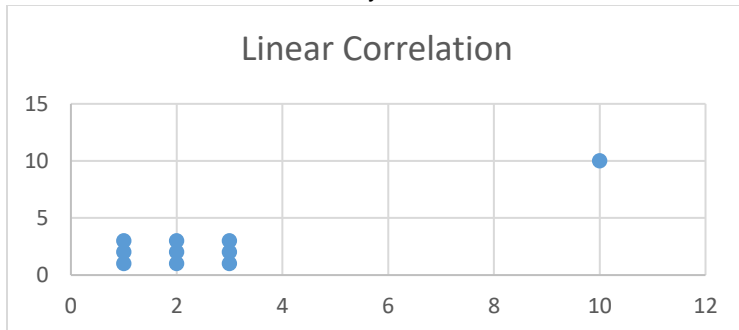
```
LinRegTTest
Xlist:L1
Ylist:L2
Freq:1
 $\beta$  &  $p$ :  $\neq 0 < 0$ 
RegEQ:Y1
Calculate
```

```
LinRegTTest
y = a + bx
 $\beta > 0$  and  $p > 0$ 
t = -5.858137033
p = .9950411445
a = 16.49091023
b = -.0028204784
s = .1259514044
 $r^2 = .9196093432$ 
r = -.9589626391
```

Find the effects of an outlier

Review the scatterplot below:

- a. Examine the pattern of all points and subjectively determine whether there appears to be a correlation between x and y .



Given sample of size 10 and a level of significance of $\alpha = 5\%$.
The ten points on the scatterplot are shown in the table below:

x	1	1	1	2	2	2	3	3	3	10
y	1	2	3	1	2	3	1	2	3	10

- b. **Enter Data**

STAT 1:Edit **Clear Old Data** [Omit this step if there is no old data in L1 and/or L2.
Enter **◀STAT 1:Edit** (Enter input (x) values in L1 and output (y) values in L2.) Make sure every entry is correct.

- c. Using technology find the value of the correlation coefficient r and determine whether there is a linear correlation.

Press **STAT, TESTS**, and press the down arrow to **E:LinRegTTest**. Press **ENTER**.

EDIT CALC TESTS
E: LinRegTTest

The Freq List should be 1 and StoreRegEq should be blank

Scroll down to Calculate and hit **ENTER**

LinReg
x List: L1
y List: L2
Frequency:List:1
Store RegEQ:
Calculate

Table 5 Linear Correlation for .05 confidence level and $n=10$ is ± 0.632 .

So $r = .91$ falls in the critical area to the right outside of ± 0.632 and there does appear to be linear correlation.

LinReg
 $y = a+bx$
 $a = .2641509434$
 $b = .9056603774$
 $s = 1.19551047$
 $r^2 = .8202207191$
 $r = .9056603774$

Find the effects of an outlier (continued)

d Now remove the point (10,10) from L1 and L2 and repeat parts a and b

x	1	1	1	2	2	2	3	3	3
y	1	2	3	1	2	3	1	2	3

e Using technology find the value of the correlation coefficient r and determine whether there is a linear correlation.

Choose **STAT, TESTS, E:LinRegTTest**

EDIT CALC TESTS
E: LinRegTTest

Table 5 Linear Correlation for .05 confidence level and $n=9$ is ± 0.666 .

So $r = 0$ falls inside ± 0.666 , so there does not appear to be linear correlation.

LinReg
 $y = a+bx$
 $a = 2$
 $b = 0$
 $s = .9258200998$
 $r^2 = 0$
 $r = 0$

Finding the equation of the Regression Line

Example: The data below shows the chest size and weight of 5 bears. Find the regression equation, letting chest size be the variable x and weight be the variable y .

x chest size	40	53	38	43	44	58
y weight	227	360	153	206	234	414

1. **Clear Old Data** [Omit Step 1 if there is no old data in L1 and/or L2.]

2. **Enter Data for above chart**

STAT 1:Edit (Enter input (x) values in L1 and output (y) values in L2.)

Make sure every entry is correct.

3. Choose **STAT, CALC, 4:LinReg(ax + b)**.

EDIT CALC TESTS
4: LinRegT(ax+b)

4. **Leave Freq List And Store Req blank. Use the down arrow
To move to Calculate. Press ENTER**

LinReg
y = ax+b
x List: L1
y List: L2
Frequency:List
Store RegEQ
Calculate

5. You will see the estimated slope (the value for a) and the estimated y intercept (the value for b).

LinReg
y = ax+b
a = 12.47058824
b = -307.9803922

So the regression equation $y = ax + b$ is $y = 12.47x - 308$

What is the best predicted weight of a bear with a chest size of 41 inches?

$$y = 12.47(41) - 308 = 240.5027 = 203 \text{ lbs.}$$