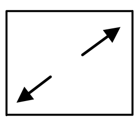
To Enlarge the Document Use the Slide-Bar in Lower Right Hand Corner

Bb213HWK

To enlarge pdf file click  icon. When the oval icon appears at the bottom of the screen click + to enlarge and – to decrease, click the X to eliminate the icon.

“Frequency Distribution”

Identify the Class Width, Class Marks (class midpoints), Class Boundary, and Tally Marks for the Frequency Distribution below.

Classes	Frequency	Class Mark	Class Boundaries	Tally Marks
10 – 13	1			
14 – 17	0			
18 – 21	15			
22 – 25	7			
26 − 29	2			

Solution:

Class Width: subtract the first “lower class limit” which is 10 from the NEXT “lower class limit” which is 14 thus: 14 – 10 = **4 Class Width**.

Class Mark (midpoints): The first class midpoint is found by adding the lower class limit of 10 to the upper class limit of 13 and dividing by 2 thus: (10 + 13) ÷ 2 = 11.5. Note: since the Class Width is consistently 4, just add 4 to 11.5 which are equal to 15.5 for the next Class Mark etc.

Class Boundary: It is the value that is between the upper class limit and the next lower class limit. The first Class Boundary would be the first upper class limit of 13 added to the next lower class limit of 14 and divide by 2 thus: $(13 + 14) \div 2 = 13.5$. Just add the class width of 4 to get the other values.

IMPORTANT: some online homework problems ask for the class boundary that is below the first lower class limit which is 10 in this example. Subtract the class width of **4** from the first class boundary of 13.5 thus: $(13.5 - 4) =$ **9.5**

Note: Class Boundaries are on the online homework, they are NOT required on the Test.

Classes	**Frequency**	**Class Mark**	**Class Boundaries**	**Tally Marks**
10 – 13	1	11.5	13.5	I
14 – 17	0	15.5	17.5	
18 – 21	15	19.5	21.5	卌 卌 卌
22 – 25	7	23.5	25.5	卌 II
26 – 29	2	27.5	29.5	II

“Measures of Central Tendency”

Listed below is a sample of flight times (in hours) for NASA’s Space Shuttle.

Calculate the following: a) mean b) median c) mode

d) midrange e) inter-quartile range IQR

73, 95, 235, 192, 165, 262, 191, 376, 259, 235

381, 331, 221, 244, 0 ← (Challenger Disaster)

Solution:

a) mean $\bar{x}$ = **217.333** Note: take all values out a minimum of 3 decimal places

b) median = **235** Note: this Med on the calculator, it’s the value in the middle

c) mode: (use the “sorting” instructions) = **235 it is unimodal**

Note: some data sets have “no mode”, some are “multi-modal”

d) midrange: $\frac{Max\ x + Min\ x}{2}$ thus: $\frac{381+0}{2}$ = **190.5**

e) IQR: $Q_3 - Q_1$, 262 – 165 = **97**

"Mean of a Frequency Distribution"

Find the mean of the following "Frequency Distribution"

Classes	Frequency (f)	(midpoint) Class Mark (x)	(f • x)
10 – 13	1		
14 – 17	0		
18 – 21	15		
22 – 25	7		
26 – 29	2		

Solution:

Step 1: Calculate the class width. This is the first lower class limit subtracted from the next lower class limit thus: 14 – 10 = **4**

Step 2: Calculate the Class Marks (midpoints). This is the lower class limit plus the upper class limit divided by 2 thus:

$\frac{(10 + 13)}{2}$ = 11.5 Note: since the class widths are the same (4), just add 4 to 11.5 to get the next Class Mark thus: 4 + 11.5 = 15.5 etc.

Step 3: Find the values for the (f • x) column by multiplying the (f) column value by the (x) column value thus:

1 • 11.5 = 11.5 etc.

Classes	(f)	(x)	(f • x)
10 – 13	1	11.5	11.5
14 – 17	0	15.5	0
18 – 21	15	19.5	292.5
22 – 25	7	23.5	164.5
26 – 29	2	27.5	55.0
	25		523.5

Step 4: Calculate the mean:

Formula: $\bar{x} = \dfrac{\sum(f \bullet x)}{\sum f}$, $\dfrac{523.5}{25}$ = **20.94 answer**

Note: class widths are NOT always the same width value.

"Histogram & Empirical Rule"

For the following data set find: a) median b) mean c) sample standard deviation d) midrange e) mode f) sample variance g) range. Draw a histogram then draw a distribution curve labeling it with the empirical rule values and percentages as demonstrated in class.

155, 142, 149, 130, 151, 163, 151, 142, 156, 133,

138, 161, 128, 144, 172, 137, 151, 133, 147, 166

Solution:

a) median (Med) = **148**,

b) mean $\bar{x}$ = **147.45**

c) sample standard deviation S_x = **12.381**

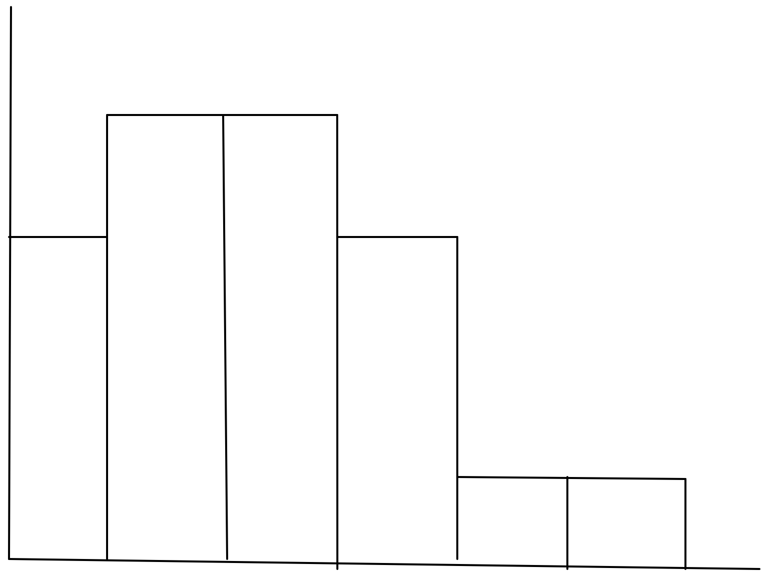
d) midrange $\frac{Max\ x + Min\ x}{2}$, $\frac{172 + 128}{2}$ = **150**

e) mode (sort data) = **151** (it occurred 3 times)

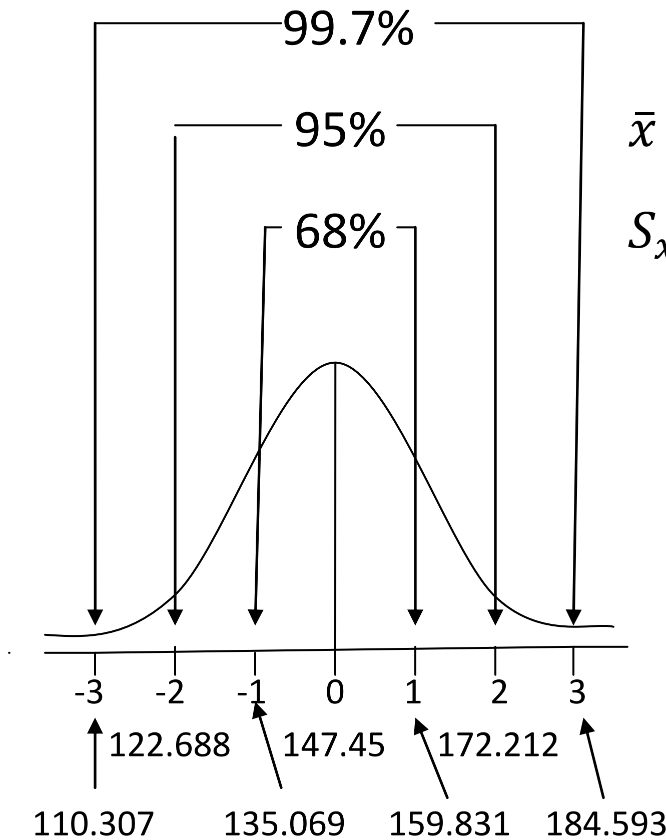
f) sample variance ${S_x}^2$ = 12.381^2 = **153.289**

g) range, 172 – 128 = **44**

Histogram:



Empirical Rule Values& Percentages:



$\bar{x}$ = **147.45**

S_x = **12.381**

"Boxplot and 5 Number Summary"

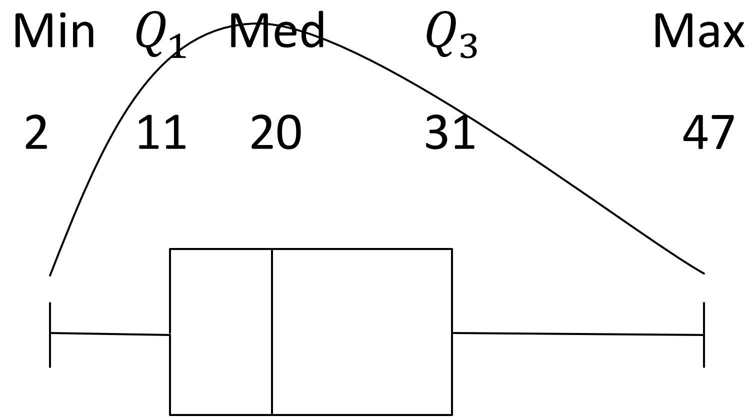
The data below shows costs (in thousands of dollars per megawatt) of 12 nuclear generators. Draw a boxplot for EACH data set (Time & Construction Costs) with the five number summary and calculate the inter-quartile range IQR for each. Finally, describe the shape of the distributions and compare them.

TIME:

2, 3, 10, 12, 17, 19, 21, 26, 30, 32, 41, 47

CONSTRUCTION COSTS:

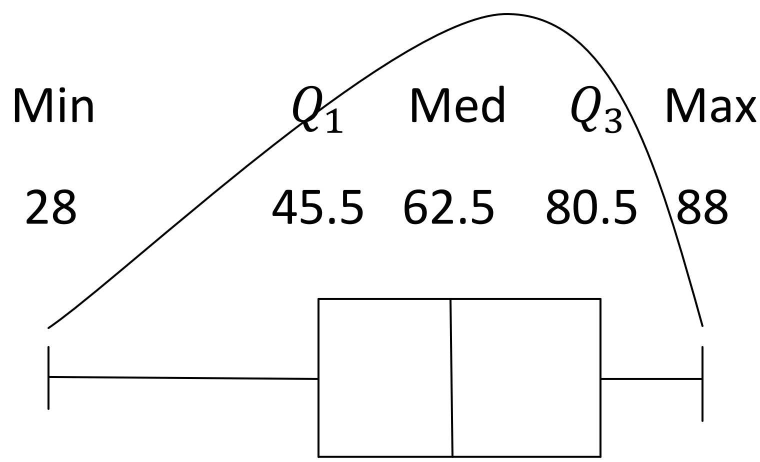
35, 28, 32, 60, 56, 63, 62, 81, 84, 79, 88, 80, TIME:



Positively Skewed

(IQR = 31 – 11 = **20**)

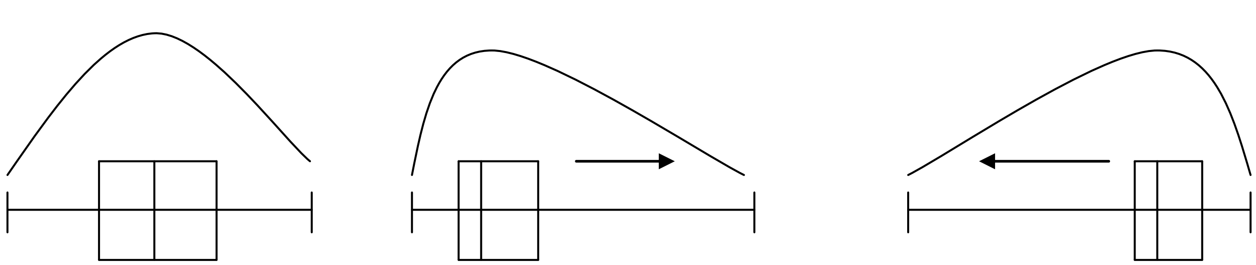
Construction Costs:



Negatively Skewed

(IQR = 80.5 – 45.5 = **35**)

Symmetrical Positively Skewed Negatively Skewed



“Introduction: Locating Z Scores”

The mean weight of yearling Angus steers is 1,152 lbs. Suppose that weights of all such animals can be described by a Normal model with a standard deviation of 84 lbs. Draw a distribution curve for each question, label it with the empirical rule values, and locate the Z scores.

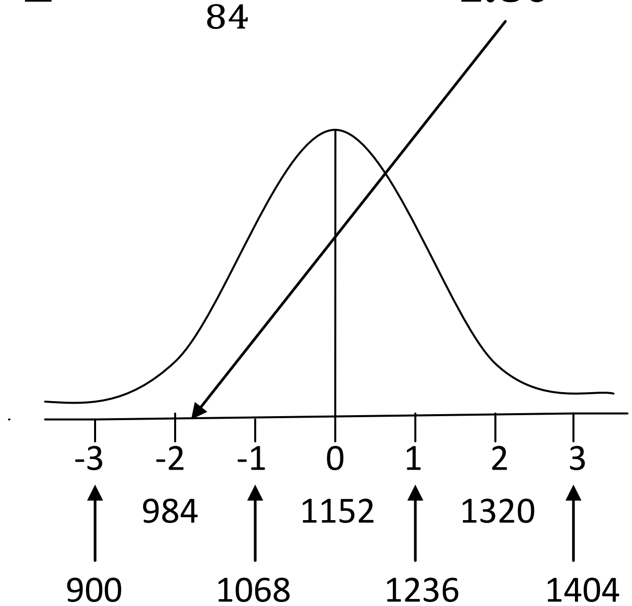
a) How many standard deviations from the mean would a steer weighing 1,000 pounds be?

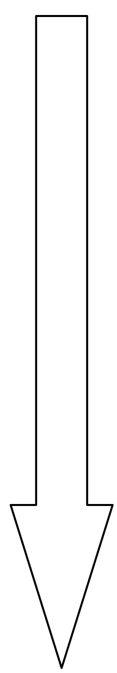
b) Which would be more UNUSUAL, a steer weighing 1,000 lbs. or one weighing 1,250 lbs.?

a) How many standard deviations from the mean would a steer weighing 1,000 pounds be?

$\bar{x}$ = 1,152 ; S_x = 84

$$\boxed{Z = \frac{X - \bar{X}}{S}} \; ; \quad Z = \frac{1{,}000 - 1{,}152}{84} = \mathbf{-\,1.80}$$

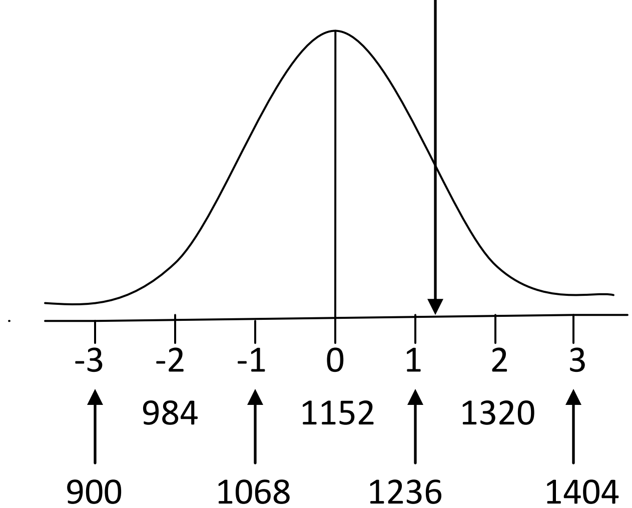




b) Which would be more UNUSUAL, a steer weighing 1,000 lbs. or one weighing 1,250 lbs.?

$\bar{x}$ = 1,152 ; S_x = 84

$$Z = \frac{1{,}250 - 1{,}152}{84} = \mathbf{1.17}$$



Answer: A steer weighing 1,000 pounds is more UNUSUAL than a 1,250 pound steer because of the – 1.8 Z score versus the 1.17 Z score.

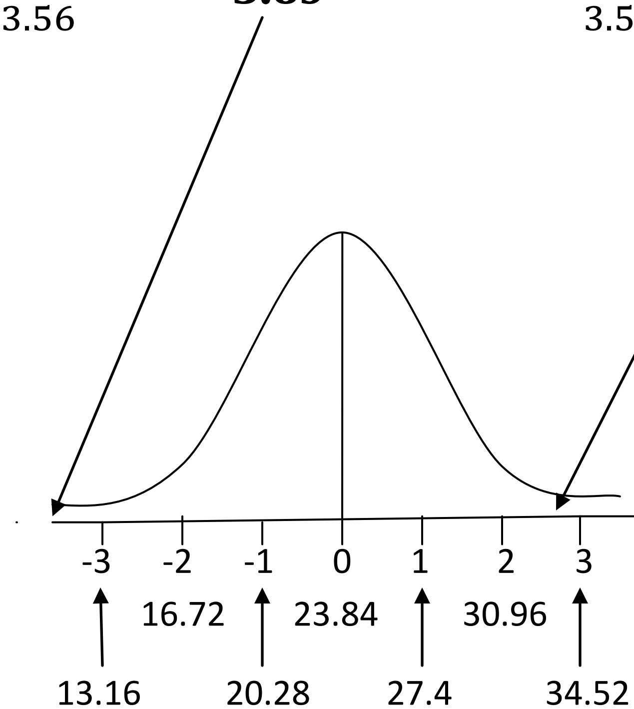
A retiree recorded the speeds of cars driving past his house, where the speed limit read 20 mph. The mean of 100 readings was 23.84 mph, with a standard deviation of 3.56 mph.

a) Which would be more UNUSUAL, a car traveling 34 mph or one going 10 mph? Draw a distribution curve with the empirical rule values and locate the Z scores.

$\bar{x}$ = 23.84 ; S_x = 3.56

$$Z = \frac{X - \bar{X}}{s}$$

$$Z = \frac{10 - 23.84}{3.56} = -\mathbf{3.89} \qquad = \frac{34 - 23.84}{3.56} = \mathbf{2.85}$$



Answer: Both speeds are unusual but 10 mph is more unusual.

"Introduction to Probability"

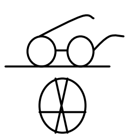
If two dice are rolled <u>ONE</u> time, find the following probabilities:

a) A sum of 6; (Hint: there are <u>36</u> possibilities in the sample space when rolling two dice)

b) Doubles;

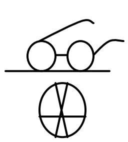
a) A sum of 6; (Hint: there are <u>36</u> possibilities in the sample space when rolling two dice)

Solution: (1,5) (2,4) (3,3) (4,2) (5,1) <u>Note</u>: (3,3) is counted once

$\frac{\mathbf{S}}{N}$;  ; $\frac{5}{36}$ = <u>**.138**</u>

b) Doubles;

Solution: (1,1) (2,2) (3,3) (4,4) (5,5) (6,6)

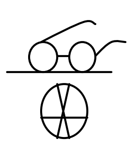
$\frac{\mathbf{S}}{N}$;  ; $\frac{6}{36}$ = <u>**.167**</u>

c) A 2001 survey asked 1,005 adults how the United States should deal with the current energy situation: “by more production”, “more conservation”, “both”, or “no opinion”. The results are the following:

Response	Number
“More production”	332
“More conservation”	563
“Both”	80
“No opinion”	30

a) What is the probability that the person responded “More production”? Note: calculate the total of people surveyed thus: Total = 1,005

a) What is the probability that the person responded “More production”? Note: calculate the total of people surveyed thus: Total = 1,005

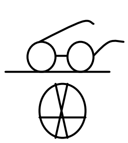
$\frac{S}{N}$;  ; $\frac{332}{1{,}005}$ = **<u>.330</u>**

In a college class of 250 graduating seniors, 50 have jobs waiting, 10 are going to medical school, 20 are going to law school, and 80 are going to various OTHER kinds of graduate schools.

a) What is the probability that a student is going to graduate school?

b) What is the probability that a student is NOT going to any type of graduate school?

a) What is the probability that a student is going to graduate school?

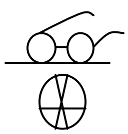
$\frac{S}{N}$;  ; $\frac{(10 + 20 + 80)}{250}$; $\frac{110}{250}$ = **.44**

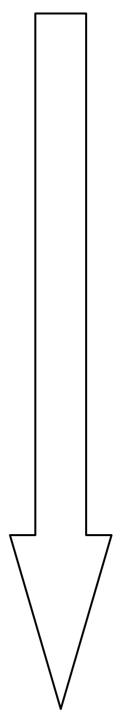
b) What is the probability that a student is NOT going to any type of graduate school?

Note: Use the "Compliment Rule"

250 – (People going to various graduate schools)

250 – (10 + 20 + 80); 250 – 110 = 140

$\frac{S}{N}$;  ; $\frac{140}{250}$ = **.56**



“Addition Rule of Probability”

In a statistics class there are 18 juniors and 10 senior; 6 of the seniors are females, and 12 of the juniors are males. If a student is selected at random, find the probability of selecting the following.

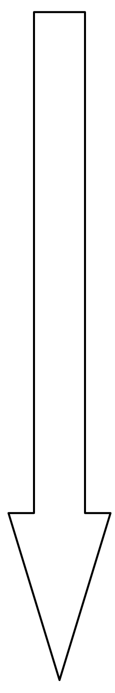
a) A Junior OR a Female,

b) A Senior OR a Female,

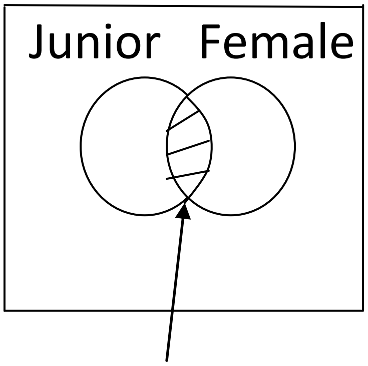
c) A Junior OR a Senior,

Summary:

- 18 juniors + 10 seniors = 28 Total
- 12 male juniors + 6 female juniors
- 4 male seniors + 6 female seniors



a) A Junior OR a Female, Note: when you see "or" you add the fractions ($\frac{s}{n}$), then determine if they can occur at the same time, if they do, subtract that fraction.



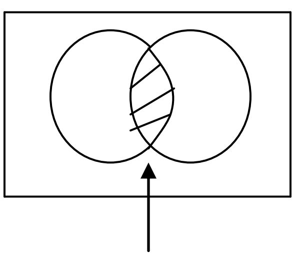
Junior & Female

$$\frac{s}{n} + \frac{s}{n} - \frac{s}{n}$$

$$\frac{18}{28} + \frac{(6+6)}{28} - \frac{6}{28} = \frac{24}{28} = \mathbf{.857}$$

b) A Senior OR a Female, Note: "or" means add fractions, subtract the intersection

Sen. Fem.

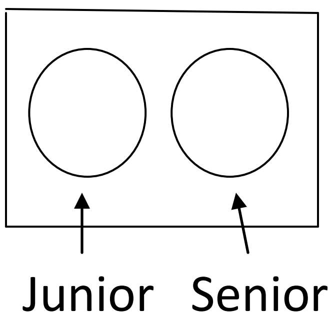


Sen. & Fem.

$$\frac{s}{n} + \frac{s}{n} - \frac{s}{n}$$

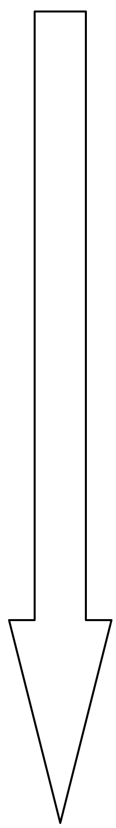
$$\frac{10}{28} + \frac{(6+6)}{28} - \frac{6}{28} = \frac{16}{28} = \mathbf{.571}$$

c) A Junior <u>OR</u> a Senior, <u>Note</u>: there is no intersection therefore just add the fractions



$$\frac{s}{n} + \frac{s}{n}$$

$$\frac{18}{28} + \frac{10}{28} = \frac{28}{28} = \mathbf{1.0\ or\ 100\%}$$



The following table represents a study of people who refused to answer survey questions by age group.

	18 – 21	22 – 29	30 – 39	40 – 49	50 – 59	60 & over	Total
Responded:	**73**	**255**	**245**	**136**	**138**	**202**	
Refused:	**11**	**20**	**33**	**16**	**27**	**49**	
Total							

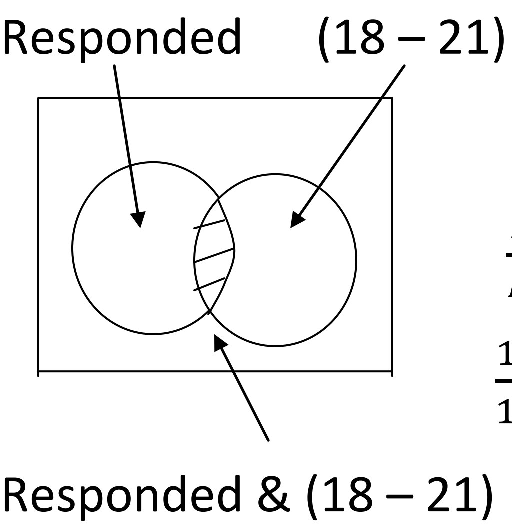
Hint: Calculate the Row & Column totals above

a) What is the probability that the selected person responded "OR" is in the 18 – 21 age bracket?

b) Find the probability that a selected person responds "OR" is between the ages of 22 & 39.

	18 – 21	22 – 29	30 – 39	40 – 49	50 – 59	60 & over	Total
Responded:	**73**	**255**	**245**	**136**	**138**	**202**	**1,049**
Refused:	**11**	**20**	**33**	**16**	**27**	**49**	**156**
Total	**84**	**275**	**278**	**152**	**165**	**251**	**1,205**

a) What is the probability that the selected person responded "OR" is in the 18 – 21 age bracket?

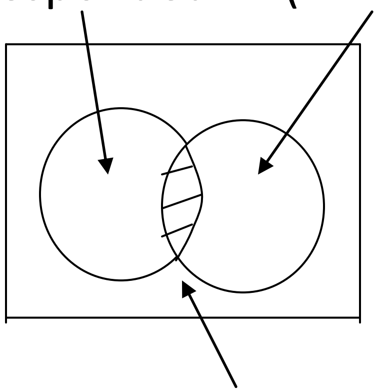


$$\frac{S}{N} + \frac{S}{N} - \frac{S}{N}$$

$$\frac{1049}{1205} + \frac{84}{1205} - \frac{73}{1205} = \frac{1060}{1205} = \textbf{.879 answer}$$

b) Find the probability that a selected person responds "OR" is between the ages of 22 & 39.

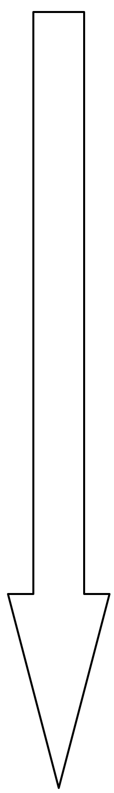
Responded (22 – 39)



$$\frac{S}{N} + \frac{S}{N} - \frac{S}{N}$$

$$\frac{1049}{1205} + \frac{(275+278)}{1205} - \frac{(255+245)}{1205} = \frac{1102}{1205} = \mathbf{.914}$$

Responded & (22 – 39)



“Multiplication of Probabilities”

You roll a fair die three times. What is the probability that:

a) You roll all 6’s

b) You roll all ODD numbers

c) None of your rolls gets a number divisible by 3.

a) You roll all 6’s

Note: no mention or “OR” therefore you multiply the fractions $(\frac{s}{n})$

and the events are “Independent” thus:

$(\frac{s}{n})$ $(\frac{s}{n})$ $(\frac{s}{n})$ which is $(\frac{1}{6})$ $(\frac{1}{6})$ $(\frac{1}{6})$ = $(.167)^3$ = **.0046**

b) You roll all ODD numbers

Note: no mention of "OR" therefore you multiply the fractions $(\frac{s}{N})$

and the events are "Independent" thus:

The possibilities of ODD rolls are (1, 3, 5) therefore:

$(\frac{s}{n})$ $(\frac{s}{n})$ $(\frac{s}{n})$ which is $(\frac{3}{6})^3$ = $.50^3$ = **.125**

c) None of your rolls gets a number divisible by 3.

Note: no mention of "OR" therefore multiply the fractions and the events are independent therefore:

The possibilities of NOT divisible by 3 are (1, 2, 4, 5)

$(\frac{s}{n})$ $(\frac{s}{n})$ $(\frac{s}{n})$ which is $(\frac{4}{6})^3$ = $.667^3$ = **.296**

A certain bowler can bowl a strike 70% of the time. What is the probability that the bowler:

a) Goes <u>three</u> consecutive frames WITHOUT a strike.

b) The bowler makes the first strike in the third frame.

c) The bowler gets a perfect game (12 consecutive strikes)

a) Goes <u>three</u> consecutive frames WITHOUT a strike.

<u>Note</u>: no mention of "OR" therefore you multiply the fractions and the events are independent. Use the "Compliment Rule": $1 - .70 = .30$.

(.30) (.30) (.30) which is $(.30)^3$ = **.027**

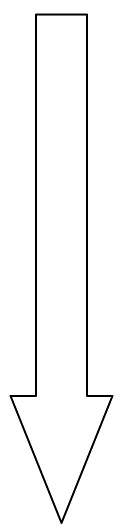
b) The bowler makes the first strike in the third frame.

Note: no mention of "OR" thus:

(.30 no strike) (.30 no strike) (.70 strike) = **.063**

c) The bowler gets a perfect game (12 consecutive strikes)

Note: no mention of "OR" thus: $(.70)^{12}$ = **.014**



A junk box in your room contains a dozen old batteries, five of which are totally dead. You start picking batteries one at a time and testing them. Find the probability of the following:

a) The first two batteries you choose are BOTH good.

b) The first four you pick all work.

c) You have to pick 5 batteries before you find one that works.

a) The first two batteries you choose are BOTH good.

Note: no mention of “OR” BUT “Dependence” is strongly implied.

Summary: there are 5 bad batteries + 7 good batteries = 12 Total

Formula: $\frac{s}{n} \bullet \frac{s-1}{n-1}$ thus: $\frac{7}{12} \bullet \frac{6}{11} = \frac{42}{132} =$ **.318**

b) The first four you pick all work.

Note: no mention of "OR" BUT "Dependence" is strongly implied.

Formula: $\frac{s}{n} \bullet \frac{s-1}{n-1}$ thus: $\frac{7}{12} \bullet \frac{6}{11} \bullet \frac{5}{10} \bullet \frac{4}{9} = \frac{840}{11{,}880}$ = **.071**

c) You have to pick 5 batteries before you find one that works.

Note: no mention of "OR" BUT "Dependence" is strongly implied.

In other words "the first 4 batteries you pick are bad, the last battery is good"

$\frac{5}{12} \bullet \frac{4}{11} \bullet \frac{3}{10} \bullet \frac{2}{9} \bullet \frac{7}{8} = \frac{840}{95{,}040}$ = **.009**

“At Least One Probability”

The probability of a randomly selected car crashing during a year is .0480. If a family has 4 cars, find the probability that AT LEAST ONE of them has a car CRASH during the year.

Formula: P (at least one) = 1 – P(none or compliment)

Formula: P(at least one) = 1 – P(none or “No Crash”)

Thus: 1 − .0480 = .952

P(at least one) = 1 – $(.952)^4$; (4 cars in the household)

P(at least one) = 1 − .8213 = **.1787 answer**

“Probability: Permutations, Combinations”

The Bureau of Fisheries once asked for help in finding the shortest route for getting samples from locations in the Gulf of Mexico. How many routes are possible if samples must be taken in a particular ORDER of 6 locations from a list of 20 locations?

Permutation Rule: “ORDER” is important and you can “NOT REPEAT” numbers.

N = “big” number, r = “little” number

thus: n = 20, r = 6

Thus: ${}_{20}P_{6}$ = **27,907,200 different ways, answer**

When testing for current in a cable with 5 color-coded wires, the technician used a meter to test 2 wires at a time. How many different tests are required for every possible pairing of the two wires?

Note: there is NO mention of ORDER, it is a Combination

Combination Rule: there is NO ORDER and you can REPEAT numbers, n = “big” number, r = “little number”

thus: $n = 5$, $r = 2$

$_5C_2 =$ **10 different ways, answer**

“Probability Distributions”

Compute the mean, variance, and standard deviation of the following “Probability Distribution” of the genetic disorder in a random sample of infants. Create a line graph then calculate $\mu - 2Ó$ and $\mu + 2Ó$. Finally, calculate the Z Scores for 2 genetic disorders and 0 genetic disorders.

x	P(x)	x • P(x)	x^2 • P(x)
0	.125		
1	.375		
2	.375		
3	.125		

<u>Step 1</u>: Fill in the column below “x • P(x)”. The first value would be 0 • .125 = 0, the second value would be 1 • .375 = .375 etc. Finally sum this column.

Step 2: Fill in the values under the "x^2 • P(x)" column. Square the value under the x column namely 0^2 then multiply this times the value under P(x) namely .125 thus 0. For the next value, 1^2 times .375 = .375 etc. Finally, sum this column.

x	P(x)	x • P(x)	x^2 • P(x)
0	.125	0	0
1	.375	.375	.375
2	.375	.75	1.50
3	.125	.375	1.125
		$\sum$ 1.50	$\sum$ 3.00

Step 3: Calculate the mean. Formula: $\mu = \sum X \bullet P(X)$

$\mu = \mathbf{1.50}$

Step 4: Calculate the Variance. Formula: $Ó^2 = \sum X^2 \bullet P(X) - \mu^2$

$Ó^2 = 3.00 - 1.50^2 = \mathbf{.75}$

Step 5: Calculate the Standard Deviation: $Ó = \sqrt{\sum X^2 \bullet P(X) - \mu^2}$

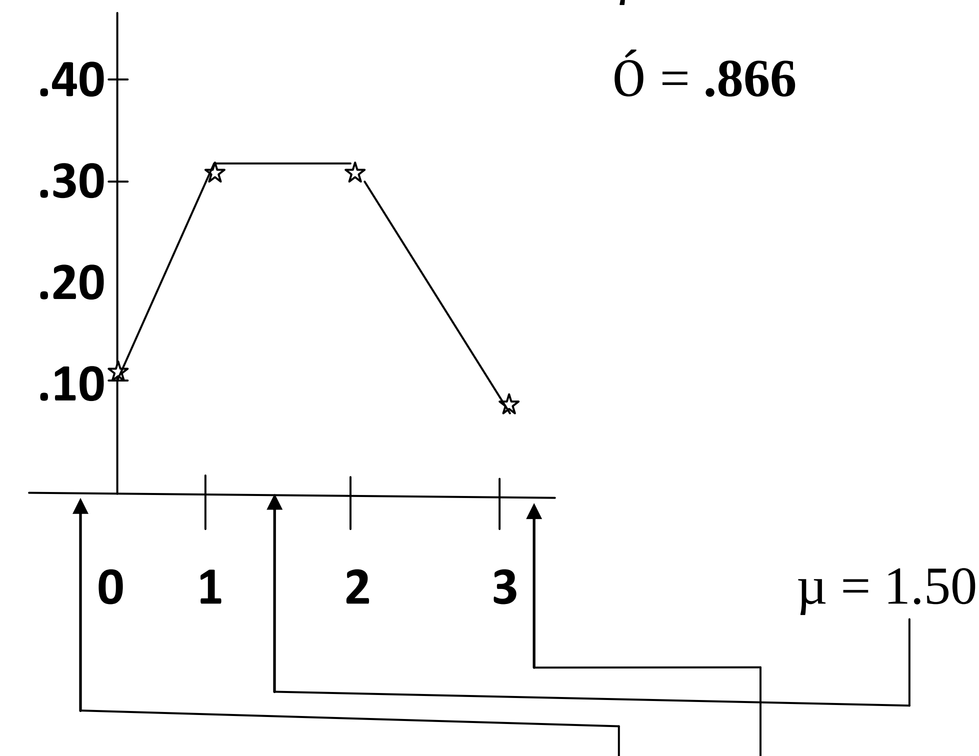
$Ó = \sqrt{.75} = \mathbf{.866}$

Note: enter the x column under L_1 and the P(x) column under L_2

then follow the "Graphing 2 Variable Stats" instructions

μ = **1.50**

Ó = **.866**



μ − 2Ó, 1.50 – 2 (.866) = **− .232**

μ + 2Ó, 1.50 + 2 (.866) = **3.232**

$$Z = \frac{X-\mu}{\sigma}, \quad Z = \frac{2(genetic\ disorders) - 1.50}{.866} = \mathbf{.577}$$

$$Z = \frac{X-\mu}{\sigma}, \quad Z = \frac{0(genetic\ disorders) - 1.50}{.866} = \mathbf{-1.732}$$

“Binomial Probabilities”

An Olympic archer is able to hit the bulls-eye 80% of the time. Assume each shot is independent of the others. If an archer shoots 6 arrows, what is the probability of getting exactly 4 bulls-eyes? Calculate the mean, variance, and standard deviation.

Note: This is a binomial probability because an archer either hits a bulls-eye or doesn’t.

Summary: p = .80

q = 1 − .80 = .20 (you have to calculate this)

n = 6 (it’s the BIG number)

x = 4 (it’s the LITTLE number)

IMPORTANT: 0! = 1

Review of storing “scientific notation”

$.04^8 = 6.5536_e - 12$, always use the “store instructions”

$$P(X) = \frac{n!}{(n-x)!\,x!} \bullet p^x \bullet q^{n-x}$$

$$P(X) = \frac{6!}{(6-4)!\,4!} \bullet .80^4 \bullet .20^{6-4}$$

$$P(X) = \frac{6!}{2!\,4!} \bullet .80^4 \bullet .20^2$$

$$P(X) = \frac{720}{2 \bullet 24} \bullet .409 \bullet .04$$

$P(X) = \frac{720}{48} \bullet .409 \bullet .04$, $P(X) = 15 \bullet .409 \bullet .04 =$ **.245**

µ = np, µ = 6 • .80 = **4.8**

Ó^2 = npq, Ó^2 = 6 • .80 • .20 = **.96**

Ó = $\sqrt{npq}$, Ó = $\sqrt{6 \bullet .80 \bullet .20}$ = **.979**

REVIEW: $.04^8$ = 6.5536_E −12 is scientific notation: (**store it**)

“Areas Under the Curve & Tail Area”

For the following problems, draw a distribution curve for <u>each</u> Z score, find the Z Table area then shade the area on the distribution curve.

a) $Z > -2.05$ b) $Z < +.33$ c) $1.2 < Z < 1.8$

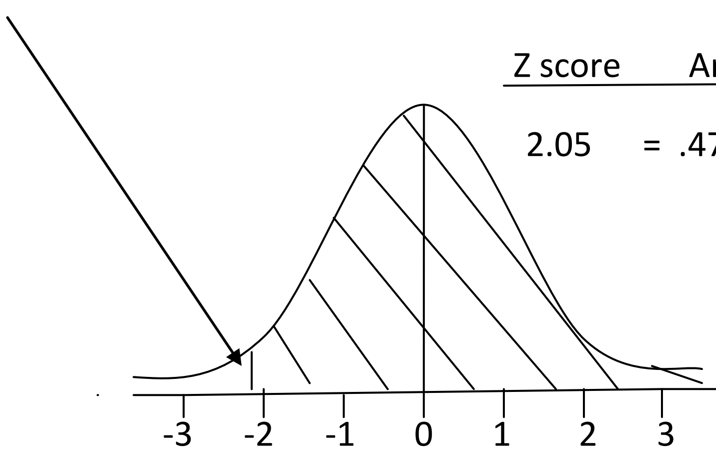
d) Between Z = 2.47 and Z = −1.03

e) To the RIGHT of Z = 1.92 and to the LEFT of Z = −.44.

AREA, PERCENTAGE, AND PROBABILITY ARE INTERCHANGEABLE TERMS

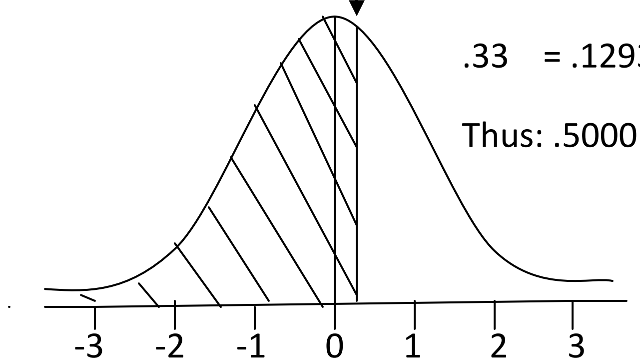
<u>IMPORTANT</u>: Z Score areas are <u>ALWAYS</u> to the middle, <u>NOT</u> to the tail.

a) $Z > -2.05$



Z score	Area
2.05	= .4798 + .5000 = **.9798 area**

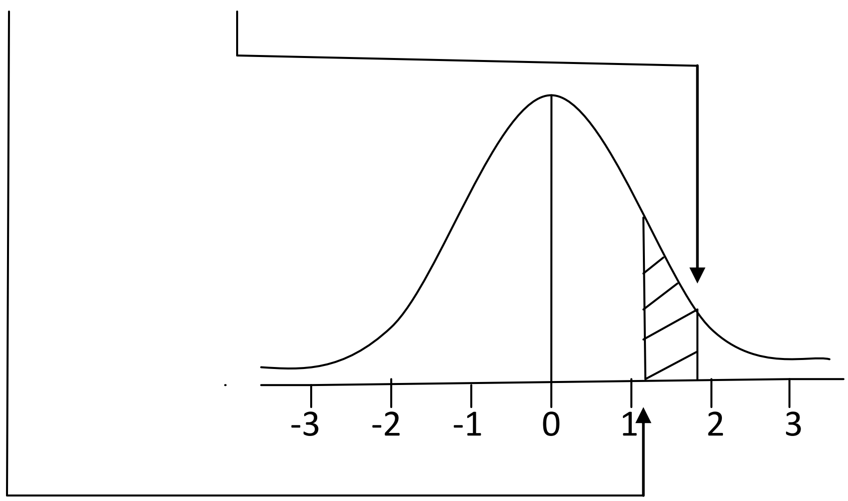
b) $Z < +.33$



Z score	Area
.33	= .1293 area

Thus: .5000 + .1293 = **.6293 area**

c) $1.2 < Z < 1.8$ (Z is between 1.2 & 1.8)

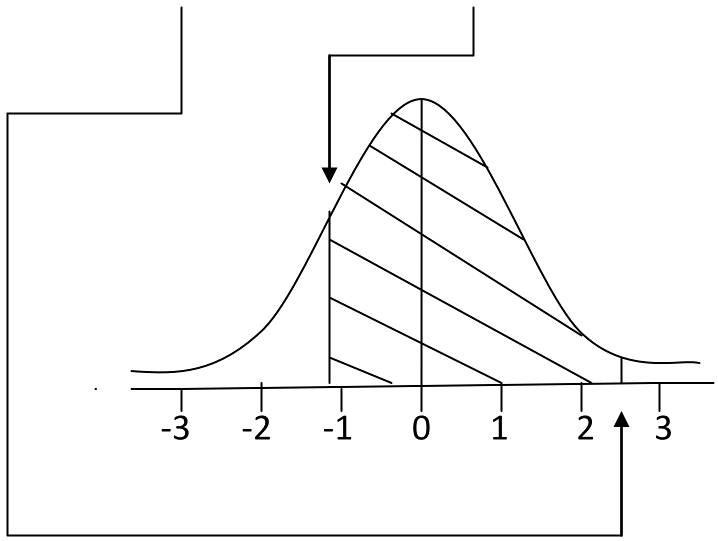


<u>Note</u>: when there is area on one side of the distribution curve and there is UNSHADED area to the left and right, <u>SUBTRACT</u> the smaller area from the larger area.

Z score		Area
1.8	=	.4641
1.2	=	.3849

thus: .4641 − .3849 = **.0792 area**

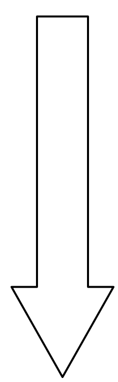
d) Between Z = 2.47 and Z = −1.03



Note: Because the Z score areas are to the middle, ADD the Z table values.

Z score		Area
1.03	=	.3485
2.47	=	.4932

thus: .3485 + .4932 = **.8417 area**

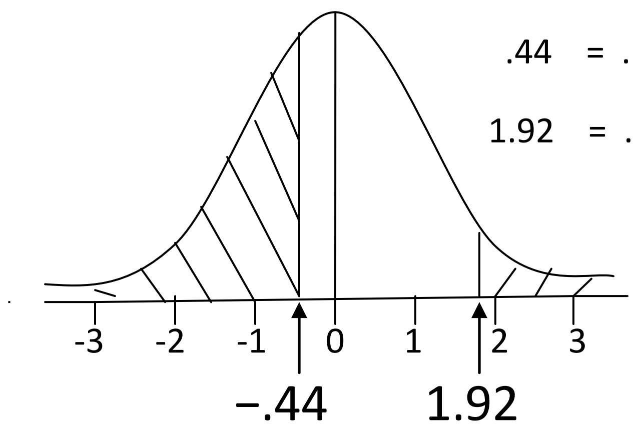


e) To the RIGHT of Z = 1.92 and to the LEFT of Z = −.44.

Step 1: Draw the distribution and locate the two Z scores

Step 2: Shade the desired (area, percentage, probability)

Step 3: Subtract the 4 digit Z table area from .5000



Z score	Area
.44	= .1700
1.92	= .4726

.5000 − .1700 = .3300

.5000 − .4726 = + .0274

.3574 (area, percentage, probability)

"Area Under the Curve Application Problems"

Some IQ tests are standardized to a normal model, with a mean of 100 and a standard deviation of 16. For each question, draw a distribution curve, label it with the empirical rule values, locate the Z score, and shade the (area, percentage, probability) in question.

a) What percent of people have IQ scores ABOVE 116?

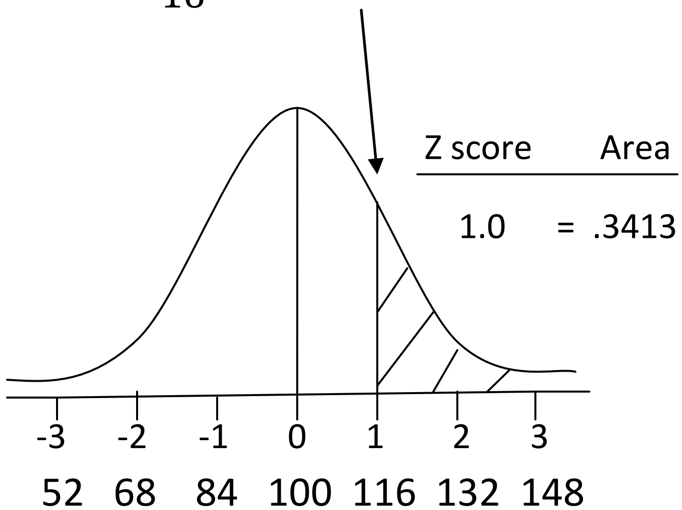
b) What is the probability that people would have IQ scores between 68 & 84?

c) What percent of people have IQ scores ABOVE 132?

a) What percent of people have IQ scores ABOVE 116?

$$Z = \frac{X - \bar{X}}{S}$$

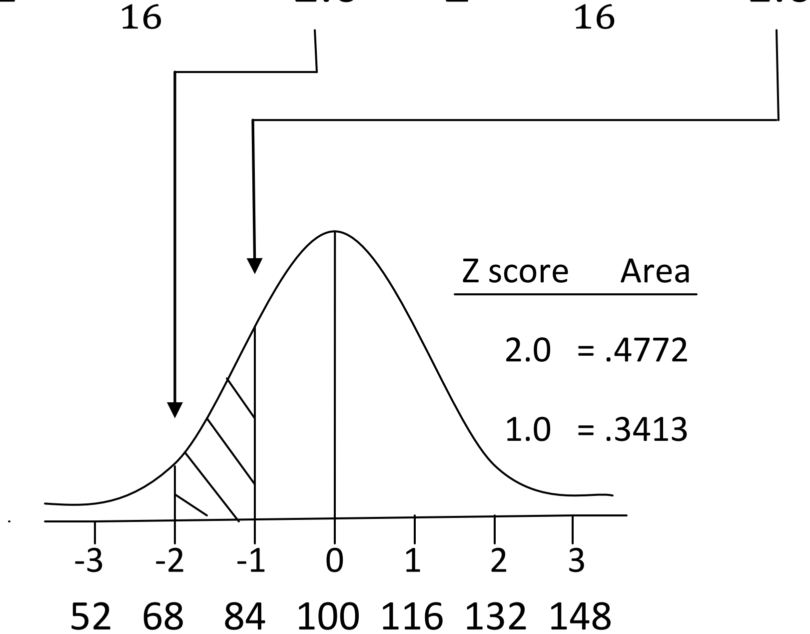
$$Z = \frac{116 - 100}{16} = 1.0 \text{ Z score}$$



Because it is “tail area” subtract: .5000 − .3413 = .1587 or **15.87%**

b) What is the probability that people would have IQ scores between 68 & 84?

$$Z = \frac{68 - 100}{16} = -2.0 \qquad Z = \frac{84 - 100}{16} = -1.0$$

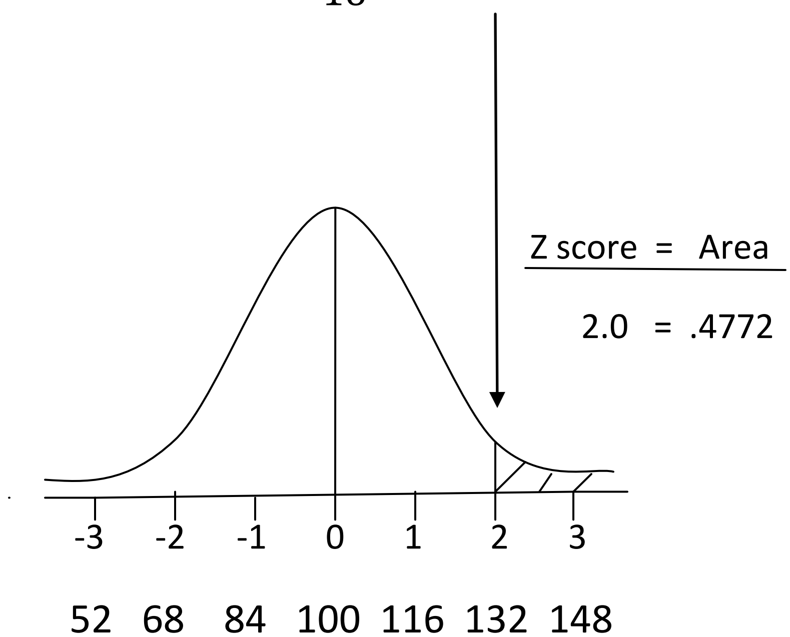


Because the area is on one side of the distribution with unshaded on both sides, subtract the smaller from the bigger area.

.4772 − .3413 = .1359 or **13.59%**

c) What percent of people have IQ scores ABOVE 132?

$$Z = \frac{132 - 100}{16} = 2.0$$



Because this is "tail area" subtract:

$.5000 - .4772 = .0228$ or **2.28%**

"Working Problems Backwards"

Since the 19th century it was believed that normal body temperature was 98.6°; this has recently been challenged. In a 1992 medical study, researchers reported that a more accurate figure is 98.2° F. The standard deviation was $.7^\circ$F. For each question, draw a distribution curve, label it with the empirical rule values, locate the Z score, and shade the (area, percentage, probability). Note: use 98.2° F as the sample mean ($\bar{x}$).

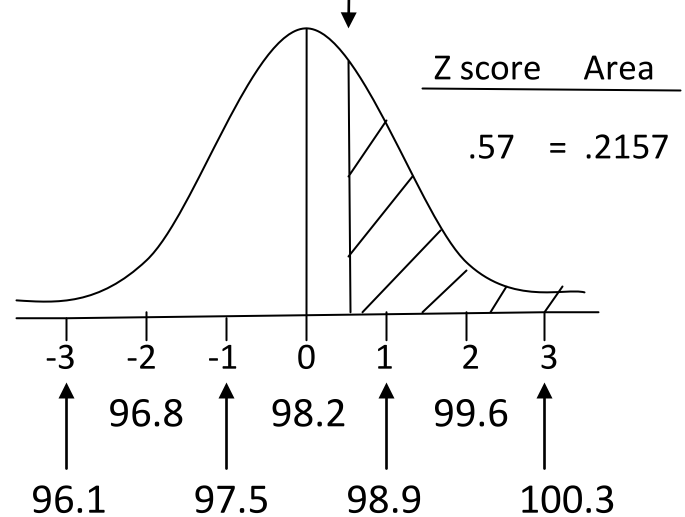
a) What percent of people have body temperatures ABOVE 98.6°?

b) Below what body temperature are the COOLEST 20% of all people?

a) What percent of people have body temperatures ABOVE 98.6°?

$$Z = \frac{X - \bar{X}}{S}$$

$$Z = \frac{98.6 - 98.2}{.7} = .57 \text{ Z score}$$



Because this is "tail area" subtract:

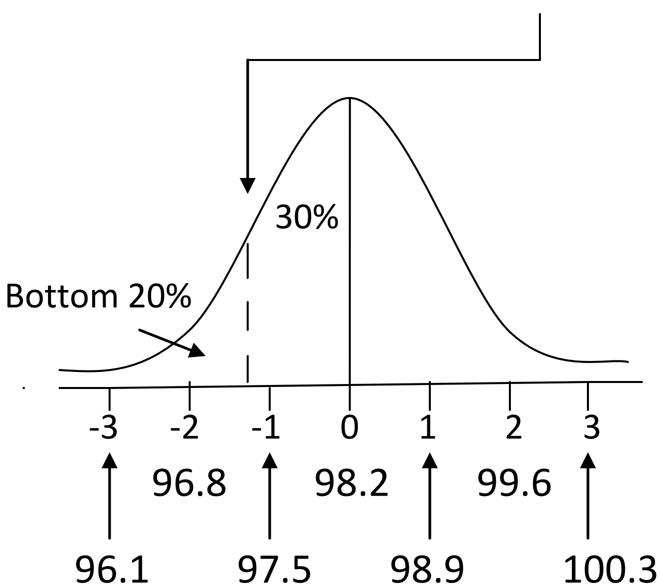
.5000 − .2157 = .2843 or **28.43%**

b) Below what body temperature are the <u>COOLEST 20%</u> of all people?

<u>Note:</u> The percentage was given and we have to find the data value that corresponds to it. In this type of question you have to work the problem BACKWARDS.

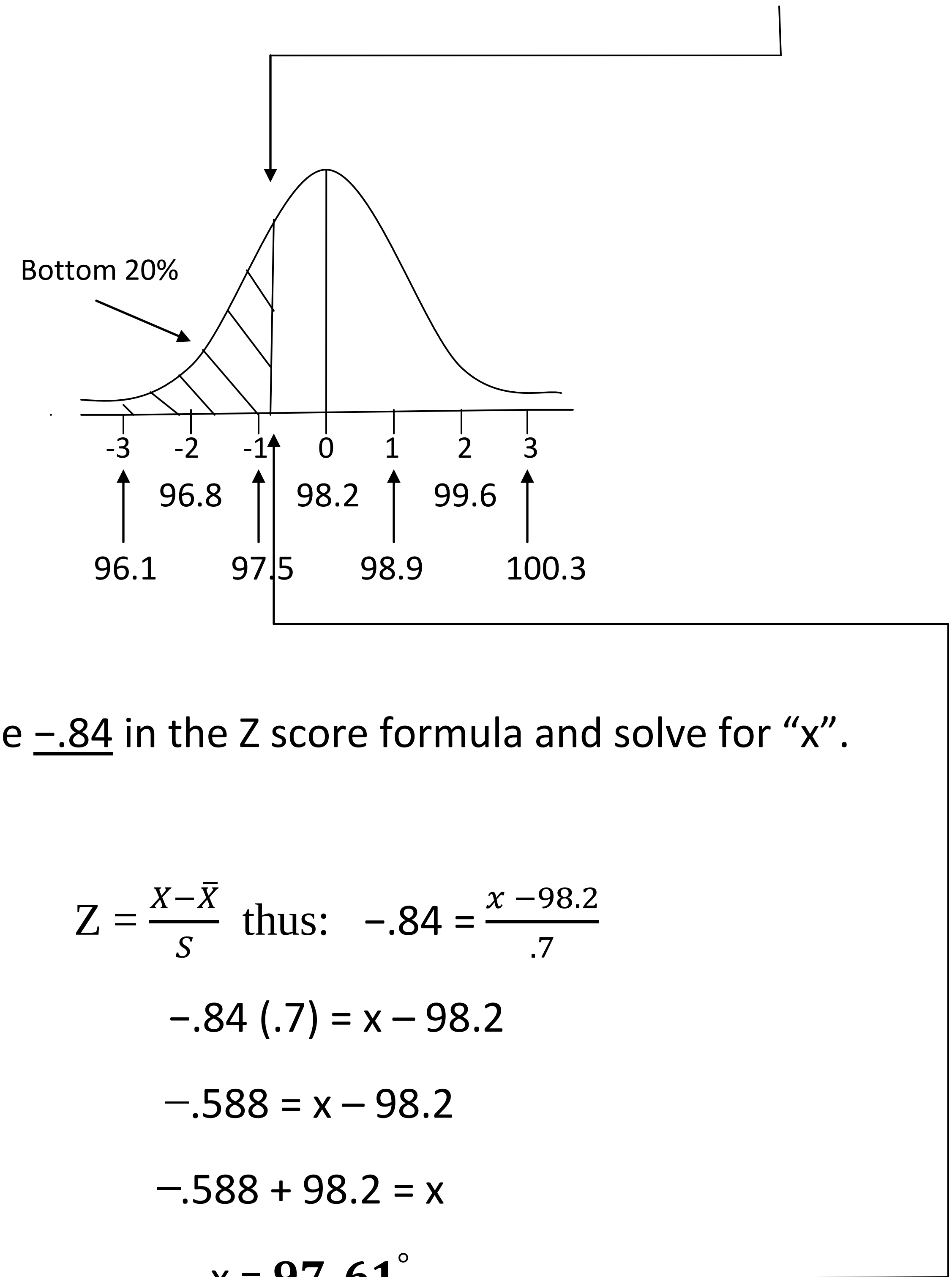
<u>Step 1</u>: The "coolest" 20% refers to the bottom left side of the distribution curve.

<u>Step 2</u>: Draw the distribution curve and "estimate" where this area would occur with a <u>dashed line</u>.



<u>Step 3</u>: Find the Z score associated with .3000; the closest numbers are: .2995 & .3023. The closest number is .2995 which has a Z score of .84 or <u>−.84</u> because it's on the left side of the distribution.

Step 4: Draw the distribution curve and locate the −.84 Z score.



Step 5: Use −.84 in the Z score formula and solve for "x".

$$Z = \frac{X - \bar{X}}{S} \quad \text{thus:} \quad -.84 = \frac{x - 98.2}{.7}$$

$$-.84\,(.7) = x - 98.2$$

$$-.588 = x - 98.2$$

$$-.588 + 98.2 = x$$

$$x = \mathbf{97.61}^{\circ}$$

"Central Limit Theorem"

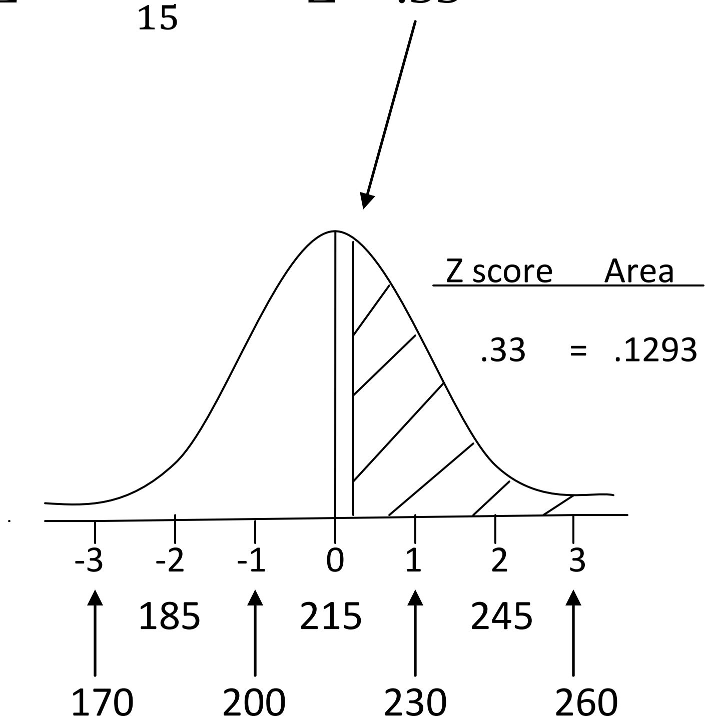
The average cholesterol content of a certain brand of egg is 215 milligrams, and the standard deviation is 15 milligrams. Assume the variable is normally distributed. Draw a distribution curve for the following probabilities showing the empirical rule values then shade the appropriate area. Assume: $\mu = 215$, $\sigma = 15$

a) If a single egg is selected, find the probability that the cholesterol content will be GREATER than 220 milligrams.

b) If a SAMPLE of 25 eggs is selected, find the probability that the mean of the sample will be LARGER than 220 milligrams.

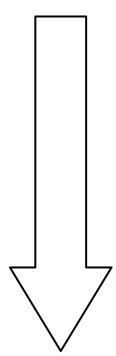
a) If a single egg is selected, find the probability that the cholesterol content will be GREATER than <u>220 milligrams</u>.

$$Z = \frac{X - \mu}{\sigma} \qquad Z = \frac{220 - 215}{15} \qquad Z = .33$$



Because this is "tail area" subtract:

$$.5000 - .1293 = .3707 \text{ or } \mathbf{37.07\%}$$



b) If a SAMPLE of 25 eggs is selected, find the probability that the mean of the sample will be LARGER than 220 milligrams.

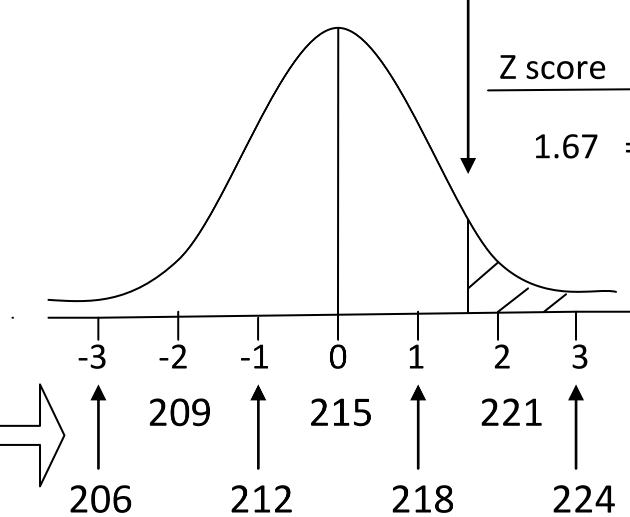
Note: The problem is based on a SAMPLE of 25, therefore use the "Central Limit Theorem". Keep the mean the same (215 milligrams) but TWEAK the standard deviation.

$$Ó_{\bar{x}} = \frac{ó}{\sqrt{n}} \text{ thus: } Ó_{\bar{x}} = \frac{15}{\sqrt{25}} = \underline{3} \qquad \text{Note: } \mu_x = 215 \text{ mlg.}$$

$$Z = \frac{\bar{x} - \mu_x}{Ó_{\bar{x}}}$$

$$Z = \frac{220 - 215}{3} = 1.67$$

Z score	Area
1.67 =	.4525



Note: new Empirical Rule Values

Because it is "tail area" subtract: .5000 – .4525 = .0475 or **4.75%**

“Binomial Problems Using the Z-Table”

When Mendel conducted his famous hybridization experiments; he used peas with green pods. One experiment involved crossing peas in such a way that 25% (or 145) of the 580 offspring peas were expected to have yellow pods. Instead of getting 145 peas with yellow pods, he obtained 152. Assume that Mendel’s 25% rate is correct.

Find the probability that among the 580 offspring peas, EXACTLY 152 have yellow pods. Draw a distribution curve showing the empirical rule values, Z Scores, and shaded area.

Formulas: $\mu = np$ Ó = $\sqrt{npq}$ $q = 1 - p$

Note: p is always given in the problem thus: p = **.25**

n = Big Number thus: n = **580**

Step 1:

μ = np, μ= 580 (.25) , μ = **145**, q = 1 – p, q = 1 − .25, q = **.75**

Ó = $\sqrt{npq}$, Ó = $\sqrt{(580)(.25)(.75)}$, Ó **= 10.428**

Step 2: Calculate the "Correction Factor" x − .5 to x + .5

Note: The question asks for "EXACTLY" 152 yellow pods

thus: x = 152

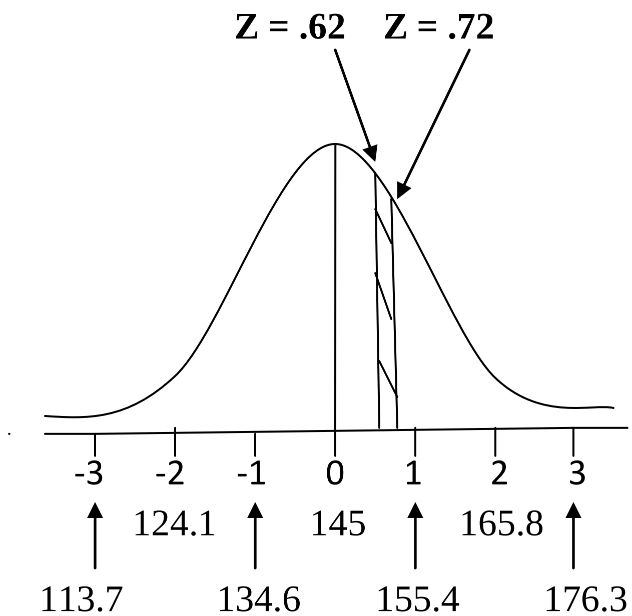
152 − .5 = **151.5** 152 + .5 = **152.5**

Step 3: Calculate a Z-Score for both "Correction Factor Values"

Formula: $Z = \frac{X - \mu}{\sigma}$ where: μ = **145**, σ = **10.428**

$Z = \frac{151.5 - 145}{10.428}$, Z = .62 which has an Area of **.2324**

$Z = \frac{152.5 - 145}{10.428}$, Z = .72 which has an Area of **.2642**



<u>Step 4</u>: Because there is unshaded to the "middle" and unshaded to the tail, subtract the smaller area from the bigger area.

$$
\begin{array}{r}
.2642 \\
-\ .2324 \\
\hline
\end{array}
$$

.0318 or 3.18% probability answer

"Introduction to Confidence Intervals"

Draw a distribution curve for each Z value, shade the appropriate area, and locate $Z_{\alpha/2}$ for the following:

a) $Z_{\alpha/2}$ for the 99% Confidence Level

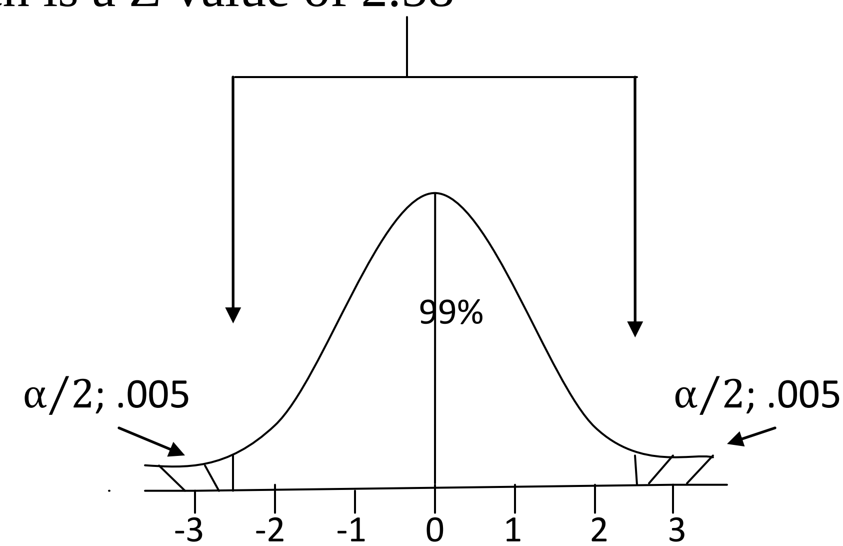
b) $Z_{\alpha/2}$ for a 98% Confidence Level

c) $Z_{\alpha/2}$ for a 95% Confidence Level

a) $Z_{\alpha/2}$ for the 99% Confidence Level,

$1 - .99 = .01$, this is α (alpha or uncertainty), $Z_{.01/2}$, $Z_{.005}$

Because it is "tail area" Subtract: $.5000 - .005 = .4950$. The Z value for .4950 is between .4949 & .4951. Since they are of equal distance you can use either value, I will use .4951 which is a Z value of 2.58

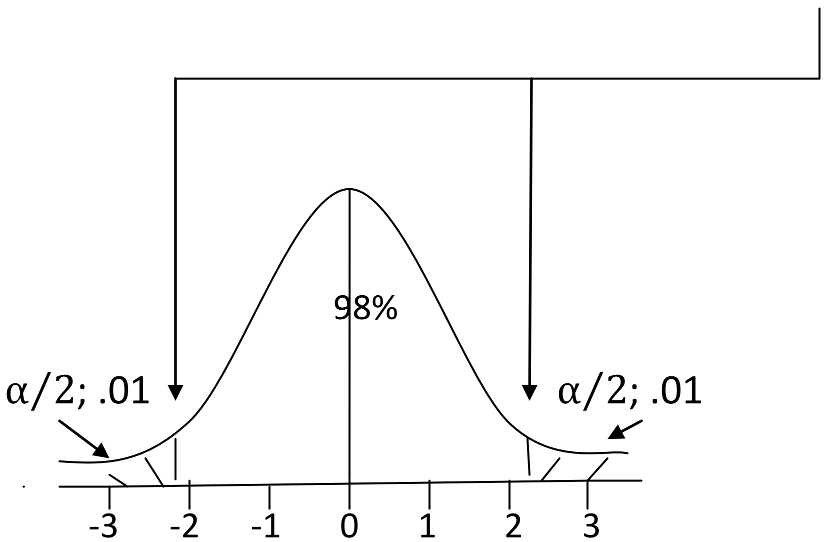


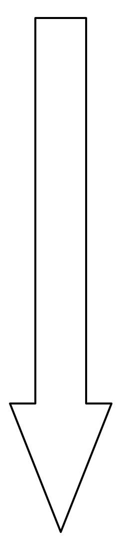
b) $Z_{\alpha/2}$ for a 98% Confidence Level

1 − .98 = .02 which is (α or uncertainty)

$Z_{\alpha/2}$, $Z_{.02/2}$, $Z_{.01}$, Subtract: $.5000 - .01 = .4900$

.4900 falls between .4898 & .4901 therefore use .4901 because it is closer, .4901 has a Z score of 2.33



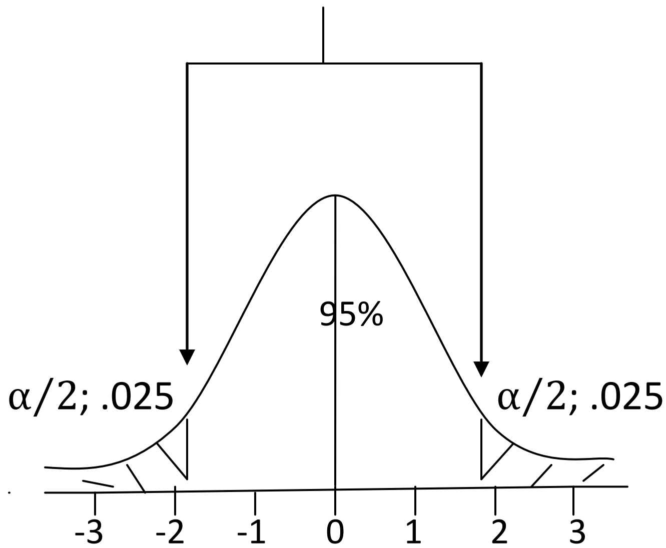


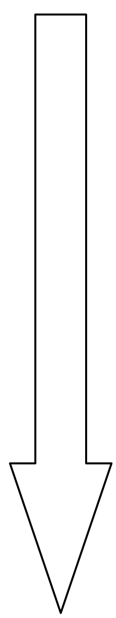
c) $Z_{\alpha/2}$ for a 95% Confidence Level

$1 - .95 = .05$, $Z_{.05/2}$, $Z_{.025}$

Subtract: $.5000 - .025 = .4750$

.4750 has a Z score of 1.96





A sample of reading scores of 35 fifth graders has a mean of 82. The standard deviation of the sample is 15. Draw a distribution curve for each Z score, label $Z_{\alpha/2}$, and shade the appropriate area.

a) Calculate the 95% Confidence Interval of the mean reading scores of all fifth-graders.

b) Calculate the 99% Confidence Interval of the mean reading scores of all the fifth graders previously mentioned. Follow all directions

Note: When you see the word “Interval” you calculate two numbers, a “minimum and a maximum”.

$$\bar{x} - Z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) \quad \bar{x} + Z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

1 − .95 = .05 which is α (uncertainty), $Z_{\alpha/2}$, $Z_{.05/2}$, $Z_{.025}$

Subtract: .5000 − .025 = .4750 which has a Z score of 1.96

$$82 - 1.96\left(\frac{15}{\sqrt{35}}\right), \quad 82 + 1.96\left(\frac{15}{\sqrt{35}}\right)$$

$$82 - 1.96\left(\frac{15}{5.916}\right), \quad 82 + 1.96\left(\frac{15}{5.916}\right)$$

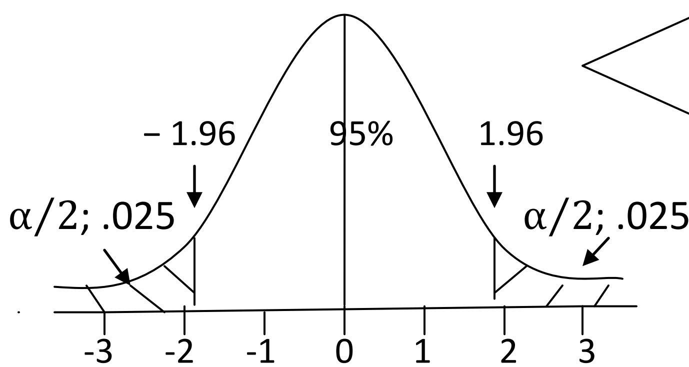
$$82 - 1.96\,(2.535), \quad 82 + 1.96\,(2.535)$$

$$82 - 4.968 \quad , \quad 82 + 4.968$$

Minimum Maximum

77.03 86.968 "Confidence Interval"

<u>NOTE</u>: Show this on the Test.



b) $1 - .99 = .01$ which is α (uncertainty), $Z_{\alpha/2}$, $Z_{.01/2}$

$Z_{.005}$, Subtract: $.5000 - .005 = .4950$ which is a

2.58 Z score

$$\bar{x} - Z\alpha/2\ (\frac{s}{\sqrt{n}}) \quad \bar{x} + Z\alpha/2\ (\frac{s}{\sqrt{n}})$$

$82 - 2.58\ (\frac{15}{\sqrt{35}})$ $\quad$ $82 + 2.58\ (\frac{15}{\sqrt{35}})$

$82 - 2.58\ (\frac{15}{5.916})$ $\quad$ $82 + 2.58\ (\frac{15}{5.916})$

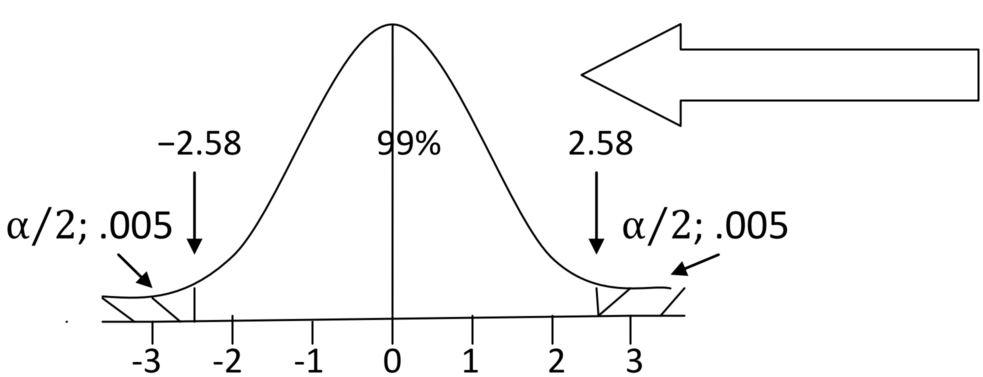
82 – 2.58 (2.535) $\quad$ 82+ 2.58 (2.535)

82 – 6.540 $\quad$ 82 + 6.540

Minimum $\quad$ Maximum

75.46 $\quad$ **88.54 "Confidence Interval"**

Note: Show this on the Test



“Confidence Intervals for Proportions”

For parts a) and b) find $\hat{p}$ and $\hat{q}$:

$\hat{p} = \frac{x}{n} \quad \hat{q} = 1 - \hat{p}$

Note: n = the BIG number

x = the LITTLE number

a) n = 60 , x =35

b) n = 95, x = 43

a) n = 60 , x =35

$\hat{p} = \frac{35}{60}$ = **.58**3 , $\hat{q} = 1 - .583$ = **.417**

b) n = 95, x = 43 ,

$\hat{p} = \frac{43}{95}$ = **.452**, $\hat{q} = 1 - .452$ = **.548**

In 1998 a reproductive clinic reported 49 live births to 207 women under the age of 40 who had previously been unable to conceive.

a) Find the 90% Confidence Interval for the success rate at this clinic. Draw a distribution curve with the $Z_{\alpha/2}$ values and shade the appropriate areas.

Step 1: Calculate $\hat{p}$ & $\hat{q}$

$\hat{p} = \frac{49}{207}$ = **.236**, $\hat{q} = 1 - .236 =$ **.764**

Step 2: Calculate $Z_{\alpha/2}$

1 − .90 = .10 which is α (uncertainty) , $Z_{\alpha/2}$, $Z_{.10/2}$,

$Z_{.05}$, Subtract: $.5000 - .05 = .4500$ which is a

Z score of 1.65

$$\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}},\ \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$.236 - 1.65\sqrt{\frac{.236\ (.764)}{207}}\ , \quad .236 + 1.65\sqrt{\frac{.236\ (.764)}{207}}$$

$$.236 - 1.65\sqrt{\frac{.180}{207}}\ , \quad .236 + 1.65\sqrt{\frac{.180}{207}}$$

$$.236 - 1.65\sqrt{8.695\ \ldots_{E\ -4}}\ , \quad .236 + 1.65\sqrt{8.695\ \ldots_{E\ -4}}$$

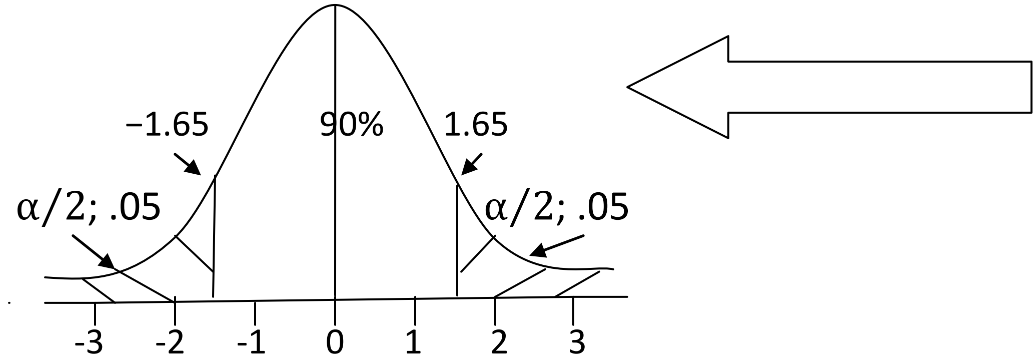
$$.236 - 1.65(.029)\ , \quad .236 + 1.65(.029)$$

$$.236 - .047 \qquad .236 + .047$$

Minimum Maximum

.189 **.283**

Note: Show This on Test



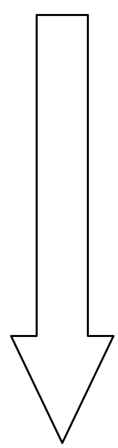
"Confidence Interval When $\hat{p}$ is Given"

In 2004 ACT, Inc. reported that 74% of 1,644 randomly selected college freshman returned to college the next year. Estimate the national freshman to sophomore retention rate by constructing a 98% Confidence Interval. Draw a distribution curve with the $Z_{\alpha/2}$ values and shade the appropriate areas.

Note: $\hat{p}$ = .74 , $\hat{q}$ = 1 − .74 = .26 , n = 1,644 , 1 − .98 = .02 α

$Z_{\alpha/2}$, $Z_{.02/2}$, $Z_{.01}$, Subtract: $.5000 - .01 = .4900$

which has a Z score of 2.33



$$.74 - 2.33\sqrt{\frac{.74\,(.26)}{1{,}644}}\,, \quad .74 + 2.33\sqrt{\frac{.74\,(.26)}{1{,}644}}$$

$$.74 - 2.33\sqrt{\frac{.192}{1{,}644}}\,, \quad .74 + 2.33\sqrt{\frac{.192}{1{,}644}}$$

$$.74 - 2.33\sqrt{1.167\ldots_{E\,-4}}\,, \quad .74 + 2.33\sqrt{1.167\ldots_{E\,-4}}$$

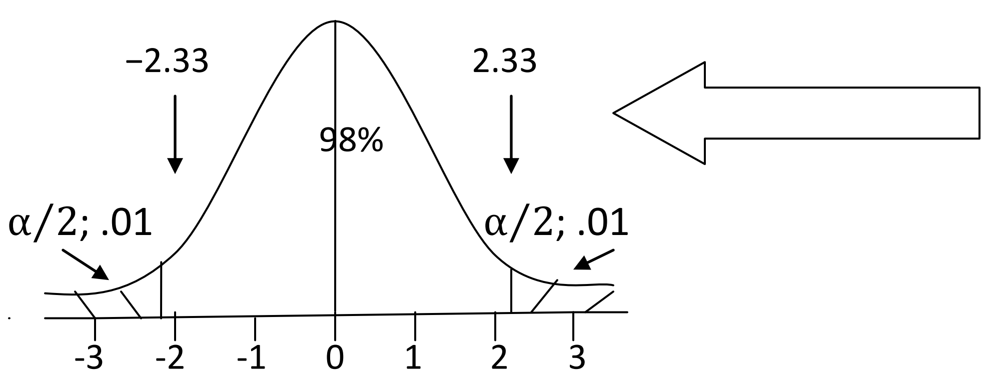
$$.74 - 2.33(.010)\,, \quad .74 + 2.33(.010)$$

$$.74 - .023\,, \quad .74 + .023$$

Minimum Maximum

.717 **.763**

<u>Note</u>: Show This on Test



"Confidence Intervals Using T-Table"

Find the T-Table values for the following. Draw a distribution curve for each, show $T_{\alpha/2}$, and shade the appropriate area.

a) $T_{\alpha/2}$ and n = 18 for the 99% Confidence Level

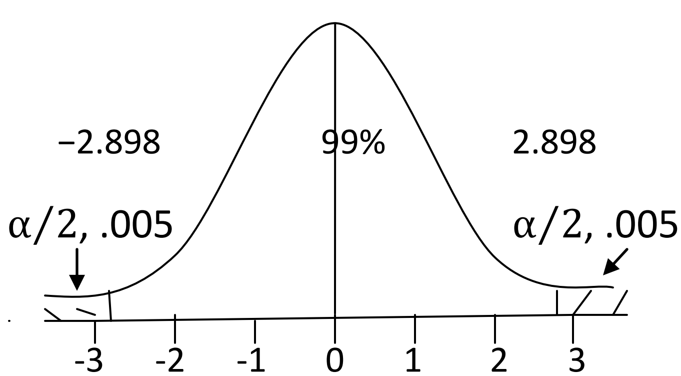
b) $T_{\alpha/2}$ and n = 23 for the 95% Confidence Level

a) $T_{\alpha/2}$ and n = 18 for the 99% Confidence Level

Note: degrees of freedom 18 – 1 = 17

1 − .99 = .01 α, $T_{\alpha/2}$, $T_{.01/2}$, $T_{.005}$, $T_{17\ .005}$ = **2.898**

Note: n < 30 use the T-Table, "Do Z's Don't T's"

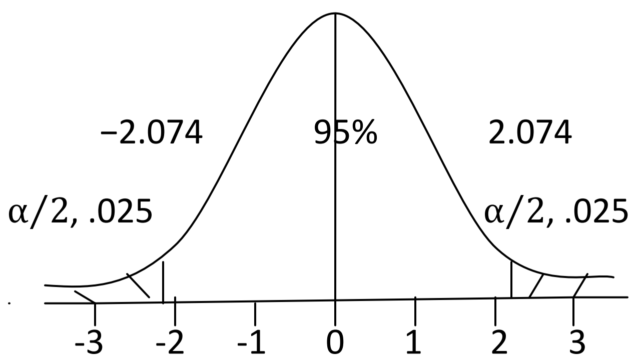


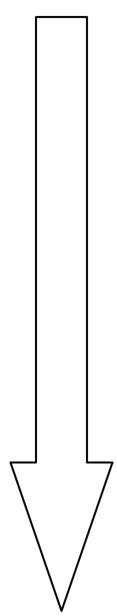
b) $T_{\alpha/2}$ and n = 23 for the 95% Confidence Level

<u>Note</u>: degrees of freedom 23 – 1 = 22

$1 - .95 = .05\ \alpha$, $T_{\alpha/2}$, $T_{.05/2}$, $T_{.025}$, $T_{22\ .025}$ = **2.074**

<u>Note</u>: n < 30 use the T-Table, "Do Z's Don't T's"





The number of grams of carbohydrates in a 12-ounce serving of a regular soft drink is listed below. Estimate the mean number of carbohydrates in all brands of soda with a 95% Confidence Interval. Draw a distribution curve with the $T_{\alpha/2}$ values and shade the appropriate areas.

48, 41, 30, 37, 45, 34, 52, 45, 46, 40,

33, 40, 43, 35, 46, 52, 41, 45, 38, 41

Step 1: Calculate $\bar{x}$, S_x and n: $\bar{x}$ = 41.6 , S_x = 5.995 n = 20

Note: n $<$ 30 use the T-Table, “Do Z’s Don’t T’s”

Step 2: Find $T_{\alpha/2}$: 1 − .95 = .05 α , $T_{\alpha/2}$, $T_{.05/2}$, $T_{.025}$

Note: degrees of freedom 20 – 1 = 19 , $T_{19\,.025}$ = **2.093**

$$\bar{x} - T_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) \quad \bar{x} + T_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$41.6 - 2.093\left(\frac{5.995}{\sqrt{20}}\right), \quad 41.6 + 2.093\left(\frac{5.995}{\sqrt{20}}\right)$$

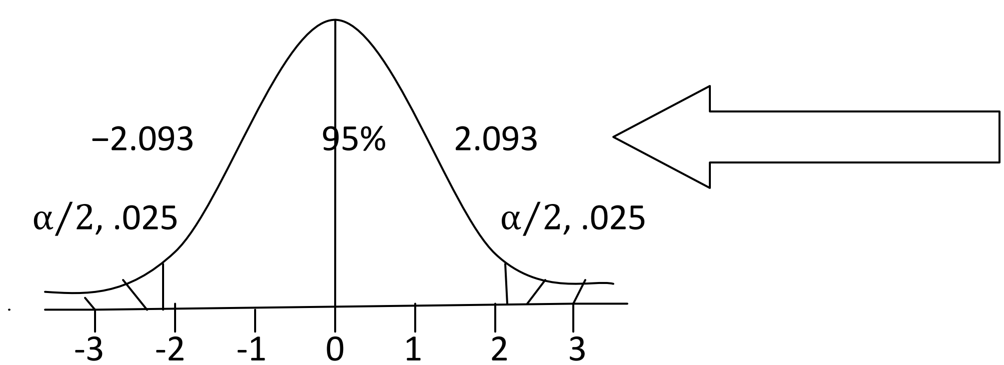
$$41.6 - 2.093\left(\frac{5.995}{4.472}\right), \quad 41.6 + 2.093\left(\frac{5.995}{4.472}\right)$$

$$41.6 - 2.093\,(1.341), \quad 41.6 + 2.093\,(1.341)$$

$$41.6 - 2.806 \qquad 41.6 + 2.806$$

Minimum = **38.794** Maximum = **44.406**

Note: Show This on Test



"Sample Size for Proportions"

Use the given data to find the minimum SAMPLE size required to estimate a population proportion or percentage.

a) Margin of Error = .045, Confidence Level = 95%, assume p = .5 and q = .5. Note: This is for Online Homework Only.

$$n = \frac{(Z_{\alpha/2})^2 \bullet (pq)}{E^2}$$

$1 - .95 = .05\ \alpha\ ,\quad Z_{.05/2}\ ,\quad Z_{.025}$

Note: "Do Z's Don't T's" thus: .5000 − .025 = .4750 which is a Z score of 1.96

$$n = \frac{1.96^2 \bullet (.5)(.5)}{.045^2}\ ,\qquad \frac{3.8416 \bullet .25}{.002025}\ ,\qquad \frac{.9604}{.002025}$$

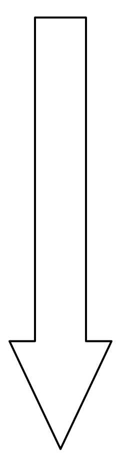
n = 474.27 or **475 sample size**

(Round ALL decimals up to the next whole number)

“Sample Sizes For Means”

A university wishes to estimate the average number of hours students spend doing homework per week. The standard deviation from a previous study is 6.2 hours. How large a SAMPLE must be selected if he wants to be 99% confident of finding whether the true mean DIFFERS from the sample mean by 1.5 hours?

Note: The question is asking for the Sample Size “n” and the Error “E” is given as 1.5 hours. Choose the formula below because it provides the population standard deviation “Ó" from a “previous study”, 6.2 hours.



$$n = \frac{(Z\alpha/2 \bullet \acute{O})^2}{E^2}$$

$1- .99 = .01\ \alpha$, $Z_{\alpha/2}$, $Z_{.01/2}$, $Z_{.005}$,

Note: "Do Z's Don't T's" thus .5000 − .005 = .4950 which is a Z score of 2.58

$$n = \frac{(2.58 \bullet 6.2)^2}{1.5^2}$$

n = $\frac{(15.996)^2}{1.5^2}$, n = $(10.664)^2$, n = 113.72 ,

n = **114 sample size**

Note: always round up ANY decimal to the next whole number.

IMPORTANT: the formula for sample sizes for Proportions has the .25 (decimal), the formula for sample sizes for means does NOT.

Introduction to Hypothesis Tests”

Identify H_0 and H_1 for the following problems. Draw a “generic” distribution curve without the Z values and show where the shaded area occurs.

a) The mean age of accountants is GREATER than 30 years

b) The mean amount charged by credit card users each month is AT LEAST $250.

c) The mean length of time managers spend on paperwork each day is LESS THAN 3 hours.

d) The mean monthly maintenance cost of an aircraft IS $3,271.

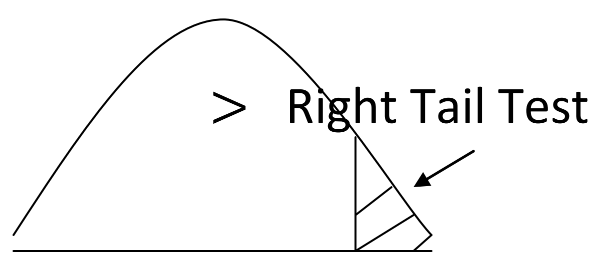
Solution:

a) The mean age of accountants is GREATER than 30 years

Note: “greater than” does not include the 30 years, it’s H_1

H_0: $\mu \leq 30$ years

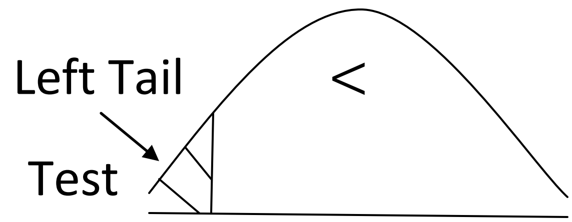
H_1: $\mu > 30$ years



b) The mean amount charged by credit card users each month is AT LEAST \$250. Note: "At Least" includes \$250, it's H_0

H_0: $\mu \geq \$250$

H_1: $\mu < \$250$

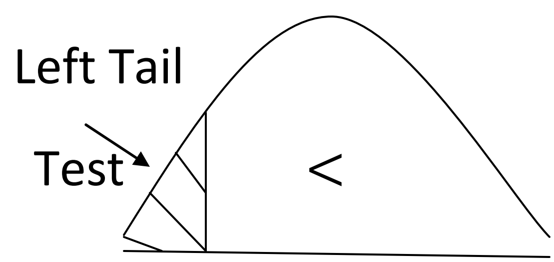


c) The mean length of time managers spend on paperwork each day is LESS THAN 3 hours.

Note: "Less Than" does not include 3 hours, it's H_1

H_0: $\mu \geq$ 3 hours

H_1: $\mu <$ 3 hours

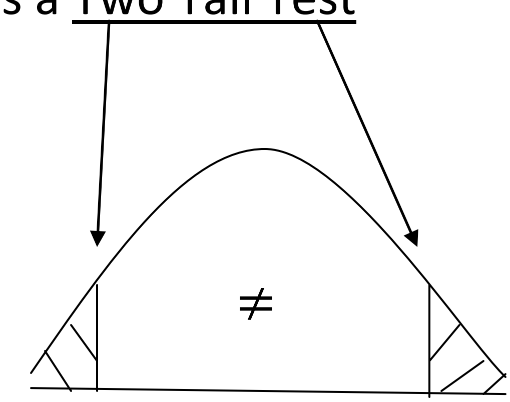


d) The mean monthly maintenance cost of an aircraft IS \$3,271.

Note: "is" means = \$3,271, it's a Two Tail Test

H_0: μ = \$3,271

H_1: $\mu \neq$ \$3,271



“Critical Value / Critical Region”

For the following problems, find the Critical Value “CV” and Critical Region “CR”. Draw a distribution curve and locate the appropriate shaded region.

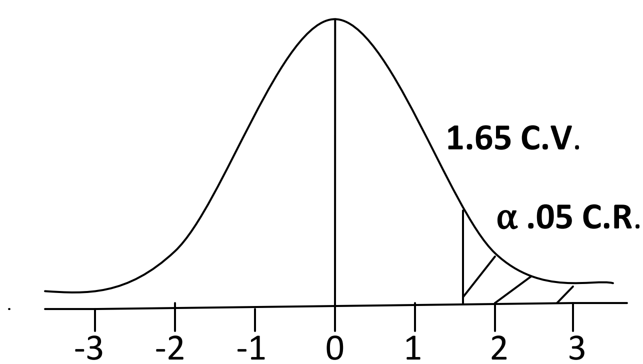
a) Right Tailed Test, $\alpha = .05$

b) Two Tailed Test, $\alpha = .05$

c) Left Tailed Test, $\alpha = .01$

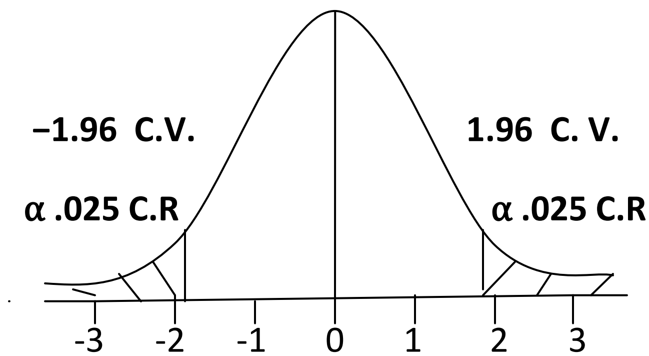
a) Right Tailed Test, $\alpha = .05$

<u>Note</u>: Because all of the shaded area is on the right side, “Do Z’s Don’t T’s”, subtract: $.5000 - .05 = .4500$ which is a Z score of 1.65.



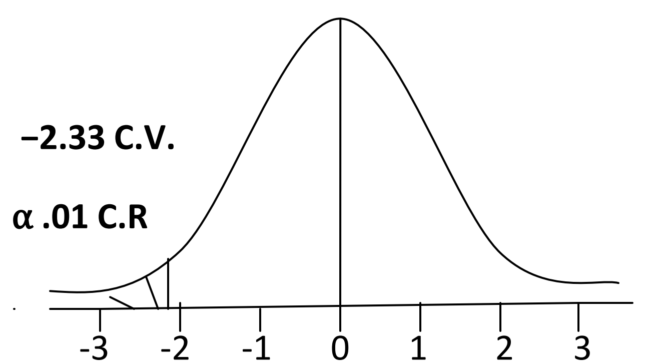
b) Two Tailed Test, $\alpha = .05$

<u>Note</u>: Because this is a "two tailed" test you have to divide α by 2 , $Z_{\alpha/2}$, $Z_{.05/2}$, $Z_{.025}$, Subtract: $.5000 - .025 = .4750$ which has a Z score of 1.96.



c) Left Tailed Test, $\alpha = .01$

<u>Note</u>: Because all of the shaded area is on the left side, subtract all of α: $.5000 - .01 = .4900$ which has a Z score of -2.33.



"6 Step Hypothesis Test"

Test the given hypothesis by following the "6 – Step" Method. Draw the distribution curve and show the C.V. & C.R. then label the distribution curve with the empirical rule values as demonstrated in class.

Based on a random sample of 100 drivers, test the claim that the mean number of people failing to use a seatbelt while driving is 20. Recently a random sample of 100 drivers were checked by the State Highway Patrol for failing to use a seatbelt; the results were: $\bar{x}$ = 18.7 failing to use a seat belt and S_x = 3. Test at the $\alpha = .05$ significance level.

Step 1: (Null Hypothesis), H_0: μ = 20

Step 2: (Alternative Hypothesis), H_1: μ ≠ 20

Step 3: (State the Alpha Value) , $\alpha = .05$

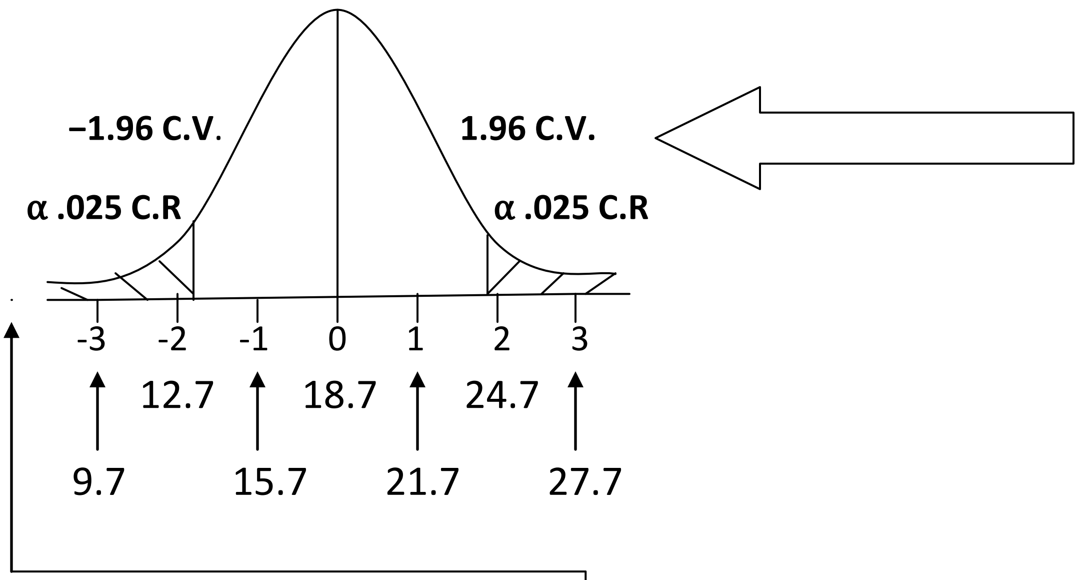
Step 4: (Draw the Distribution Curve and Locate C.V. & C.R.)

Note: It is a "two tail test", because H_1: $\mu \neq 20$

$Z_{\alpha/2}$, $Z_{.05/2}$, $Z_{.025}$, "Do Z's Don't T's"

Subtract: .5000 − .025 = .4750 which is a Z score = **1.96**

NOTE: Show This on Test



Step 5: (Test Statistic)

$$Z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$$

Note: $\mu_{\bar{x}}$ is from H_0: $\mu = 20$

$$Z = \frac{18.7 - 20}{\frac{3}{\sqrt{100}}}, \quad Z = \frac{-1.3}{\frac{3}{10}}, \quad \frac{-1.3}{.3} \quad = -\mathbf{4.33\ test\ statistic}$$

Step 6: **Reject H_0** (because the Test Statistic is in rejection area)

In a study of distances traveled by buses before the first major engine failure, a sample of 191 buses results in a mean of 96,700 miles and a standard deviation of 37,500 miles. At the .05 significance level, test the claim that the mean distance traveled before a major engine failure is MORE THAN 90,000 miles. Assume that the sample standard deviation can be used for Ó. Draw the distribution curve and show the C.V. & C.R. then label the distribution curve with the empirical rule values as demonstrated in class. *Calculate a P-Value*

Step 1: H_0: $\mu \leq 90{,}000$ miles   Note: "more than" does not include 90,000 miles

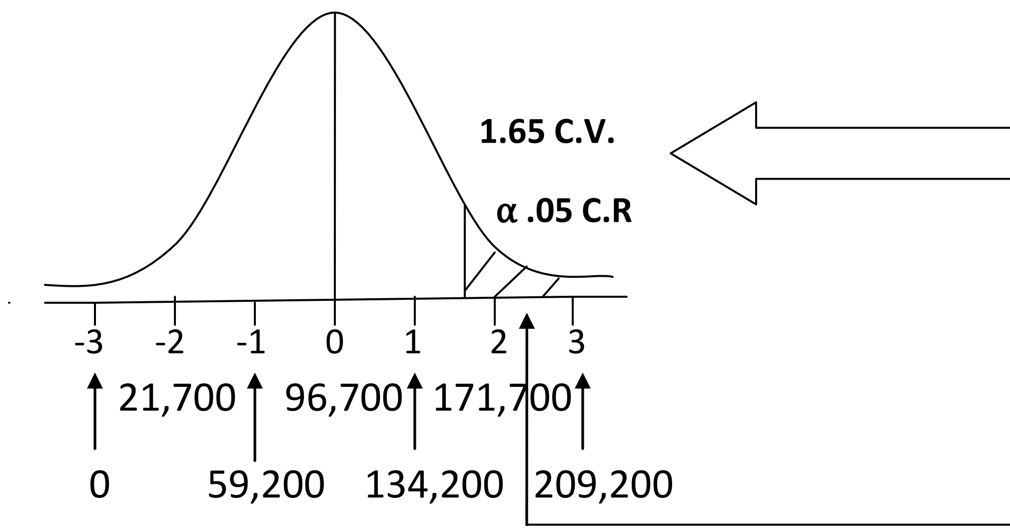
Step 2: H_1: $\mu > 90{,}000$ miles

Step 3: $\alpha = .05$

Step 4: (C.V. & C.R.) Note: because H_1: $\mu > 90{,}000$, this is a "right tail" test therefore DO NOT divide α by 2.

"Do Z's Don't T's" Subtract: .5000 − .05 = .4500 which is a Z score of 1.65

Note: Show This on Test



Step 5: Note: this is from H_0: $\mu \leq 90{,}000$

$$Z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} \qquad Z = \frac{96{,}700 - 90{,}000}{\frac{37{,}500}{\sqrt{191}}}$$

$$Z = \frac{6{,}700}{\frac{37{,}500}{\sqrt{13.820}}}, \quad \frac{6{,}700}{2{,}713.458}, \quad Z = \mathbf{2.469\ test\ statistic}$$

P-Value: the area for 2.469 (round to 2.47) = .4932

Because a P-Value is the "tail area" of the test statistic, subtract from .5000 thus: .5000 − .4932 = **.0068 P-Value**

Compare P-Value: $.0068 < .05$ alpha

Step 6: **Reject H_0** "If the P is low the null must GO!"

"T-Table Hypothesis Test"

For the following problems, find the Critical Values, draw a distribution curve and shade the affected areas.

a) H_0: μ = 12, n = 27, α = .05

b) H_0: μ ≤ 50, n = 17, α = .10

c) H_0: μ ≥ 1.36, n= 6, α = .01

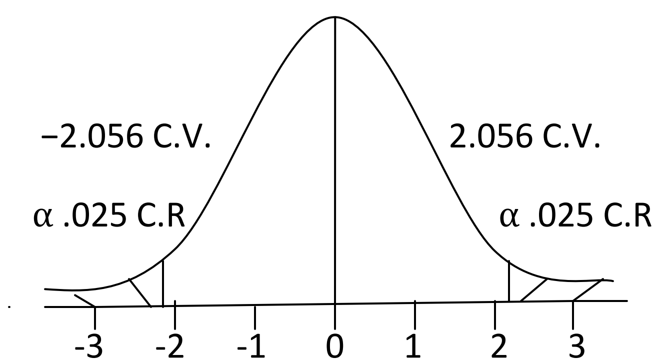
a) H_0: μ = 12, n = 27, α = .05

H_0: μ = 12

H_1: μ ≠ 12 Note: this is a "two tail" test, therefore divide alpha by two

$T_{\alpha/2}$, $T_{.05/2}$, $T_{.025}$ Note: Degrees of Freedom are 27 – 1 = 26

Therefore: $T_{26df.025}$ "Do Z's Don't T's" T-Score = 2.056

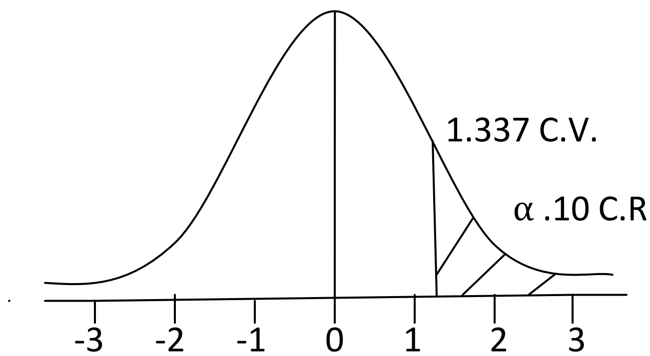


b) H_0: μ ≤ 50, n = 17, α = .10

H_0: μ ≤ 50

H_1: μ > 50, Note: this is a "right tail" test therefore do NOT divide alpha by 2, d.f. = 17 – 1 = 16

$T_{16\,df\,.10}$ = 1.337 T-Score

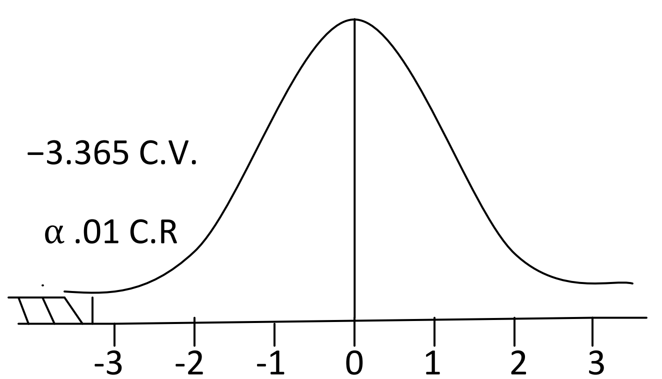


c) H_0: μ ≥ 1.36, n= 6, α = .01

H_0: μ ≥ 1.36

H_1: μ < 1.36 Note: this is a "left tail test" do NOT divide alpha by 2, df = 6 – 1 = 5

$T_{5\,df\,.01}$ = −3.365



A sample of beer cans labeled 16 oz. is randomly selected and the actual contents measured in ounces are as follows:

15.8,   16.2,   16.3   15.9,   15.5,   15.9,   16.0,   15.6,   15.8

Test the claim that the consumer is being CHEATED.
Assume $\alpha = .05$ Use the 6-Step process, draw the distribution curve with the empirical rule values, and show the C.V. and C.R.

Step 1: H_0: $\mu \geq 16$ oz.   Note: "Cheated" means getting less than 16 oz., it's a "left tail test"

Step 2: H_1: $\mu < 16$ oz.

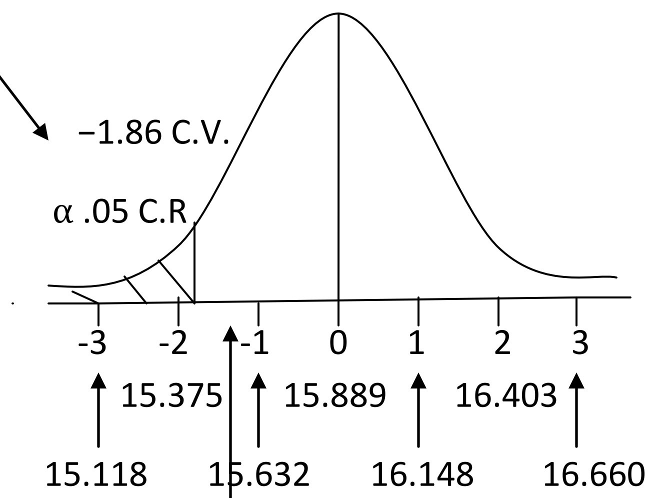
Step 3: $\alpha = .05$

Step 4: C.V. / C.R.   Note: Use the T-Table because n = 9 which is less than 30, df = 9 – 1 = 8

Do NOT divide alpha by 2 thus: $T_{8\,df\ .05}$ = −1.86 T-Score

TI 83 / 84 Summary: $\bar{x}$ = 15.889, S_x = .257, n = 9

Show C.V. / C.R.



Note: $\mu_{\bar{x}}$ is H_0: $\mu \geq 16$ oz.

<u>Step 5</u>: Formula: $T = \dfrac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$ $\quad$ $T = \dfrac{15.889 - 16}{\frac{.257}{\sqrt{9}}}$

$\dfrac{-.111}{\frac{.257}{3}}$, $\dfrac{-.111}{.085}$ = **−1.30 test statistic**

<u>Step 6</u>: **Accept H_0**

"Proportion Hypothesis Test"

In a survey, 813 of the 1,084 respondents indicated support for a ban on household aerosols. At the .01 significance level, test the claim that MORE THAN 70% of the population supports the ban.

Use the 6-Step process, draw the distribution curve, and show the C.V. and C.R. Do NOT show the empirical rule values.

Note: The key number in the "claim" is a percentage, therefore it is "proportion" problem.

Note: use "p" because it is a proportion

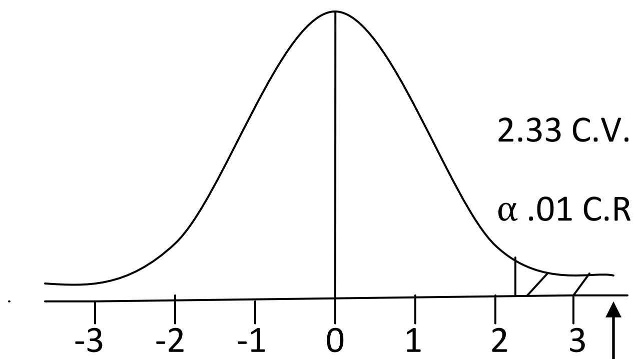
Step 1: H_0: p $\leq$ .70 Note: "More Than" means greater than 70%, it is a "right tail test"

Step 2: H_1: p $>$ .70

Step 3: $\alpha = .01$

Step 4: C.V. / C.R. Note: Use the Z-Table because n = 1,084

Do NOT divide alpha by 2, “Do Z’s Don’t T’s” you have to subtract thus: .5000 − .01 = .4900 which is a 2.33 Z-Score.



Step 5: Formula: $z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$ $\hat{p} = \dfrac{x}{n}$ Note: x = “little” number

n = “big” number

$\hat{p} = \dfrac{813}{1{,}084}$ = .75, p = .70 (it is the H_0 value), q = 1 − .70 = .30

$$z = \frac{.75 - .70}{\sqrt{\frac{(.70)(.30)}{1{,}084}}},\quad \frac{.05}{\sqrt{\frac{.21}{1{,}084}}},\quad \frac{.05}{\sqrt{1.937\ldots_{E-4}}},\quad \frac{.05}{.0139}$$ = **3.597 test statistic**

Step 6: **Reject H_0**

"Chi-Square Hypothesis Test for Standard Deviation"

A random sample of <u>40</u> men results in a standard deviation of <u>9.3</u> beats per minute (sample standard deviation). The standard deviation of the population of adult's heart beats is <u>10</u> beats per minute (population standard deviation). At the .05 significance level test the claim: Pulse rates of men have less VARIATION in their heart beats than the <u>10 beats</u> per minute for all adults. <u>Note</u>: Draw the distribution curve in Step 4.

Summary: Ó = 10   s = 9.3

<u>Note</u>: Because the test is concerned with VARIATION which is the standard deviation Ó, use the "Chi-Square" distribution: χ^2

IMPORTANT: We will only use "Less Than" hypotheses for the population standard deviation Ó on the TEST. We will NOT calculate P-Values for this on the TEST and I will provide the "Chi-Square" table.

Step 1: H_0: Ó $\geq$ 10 beats per minute

Step 2: H_1: Ó $<$ 10 beats per minute

Step 3: α = .05

Step 4: C.V. / C.R. Note: n = 40, Ó = 10, s = 9.3

Find the Chi-Square table value which is the Critical Value where n = 40 therefore the degrees of freedom are: d.f. = 40 – 1 = **39**

For the Chi-Square table use d.f. = 39, then because it is a "Less Than" or "Left Tailed Test" you are concerned with the area to the RIGHT of the .05significance level. The result is 1 − .05 = **.95**

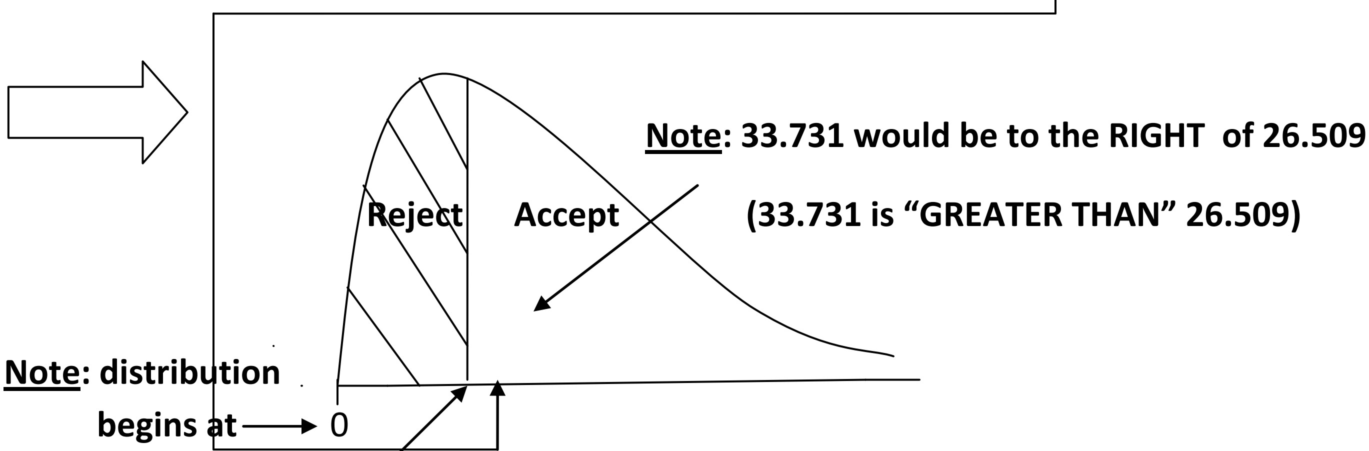
The INTERSECTION in the table is: **26.509 C.V.**

Step 5: Formula: $\chi^2 = \frac{(N-1)\bullet S^2}{Ó^2}$

Note: the sample standard deviation is s

$\chi^2 = \frac{(40-1)\bullet 9.3^2}{10^2}$, $\chi^2 = \frac{39 \bullet 86.49}{100}$, $\frac{3373.11}{100}$ = **33.731 Chi Square**

Show the Distribution Curve on the Test **Test Statistic**



26.509 Chi-Square Table Value

Rule: if the χ^2 value is < C.V. Reject H_0

if the χ^2 value is > C.V. Accept H_0

Step 6: **Accept H_0 because 33.731 > 26.509**

"Regression"

A survey was conducted in the United States and 10 countries of Western Europe to determine the percentage of teenagers who had used marijuana and other drugs. The results are summarized in the table below.

L_1 Independent x	Marj	22	17	40	5	37	19	23	6	7	53	34
L_2 Dependent y	Other	4	3	21	1	16	8	14	3	3	31	24

a) Create a scatterplot and regression line

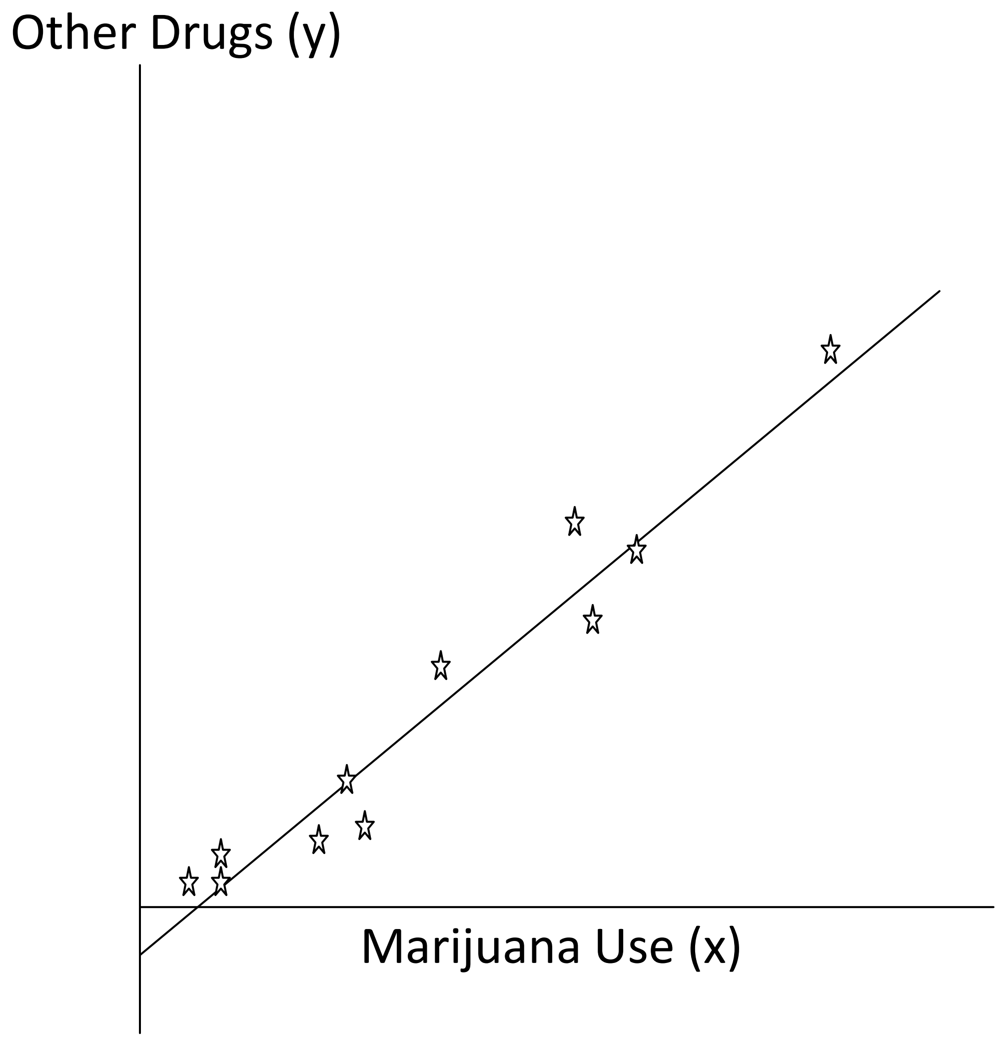
b) Calculate the correlation coefficient (r) and describe its significance.

c) Find the regression equation

d) What is the value of the slope and y-intercept?

e) What percent of "Other Drugs" would you estimate for 30% of teenagers using marijuana. Note: enter 30% as 30. DO NOT convert it to a decimal.

a) Create a scatterplot and regression line



b) Calculate the correlation coefficient (r) and describe its significance.

r = .9341 (Note: 4 decimal places)

“There is a Strong Positive Relationship”

c) Find the regression equation

$y = .615x + -3.067$

d) What is the value of the slope and y-intercept?

slope = .615, y– intercept = −3.067

e) What percent of "Other Drugs" would you estimate for 30% of teenagers using marijuana. <u>Note</u>: enter 30% as 30. DO NOT convert it to a decimal.

$y = .615(30) + -3.067$

$y = 18.45 + -3.067 =$ **15.383 or 15.383%**

(Demonstrate 2 Variable Stats on Calculator Instruction Sheet)

“Hypothesis Test for Correlation Coefficient”

Listed below are systolic blood pressure measurements obtained from the same woman. Is there sufficient evidence to conclude that there is a linear correlation between right and left arm systolic blood pressure measurements? Create a scatter plot & regression line. Assume α = .05

L_1	Right Arm	x	102	101	94	79	79
L_2	Left Arm	y	175	169	182	146	144

<u>Note</u>: p is the population correlation coefficient

<u>Step 1</u>: H_0: p = 0 (zero implies NO correlation)

<u>Step 2</u>: H_1: p ≠ 0 (Not zero implies there IS a correlation)

<u>Step 3</u>: α = .05

<u>Step 4</u>: Go to the Appendix Table 11 “Critical Values for the Pearson Correlation Coefficient” in the text or use the A-6 Table attached to the Calculator Instructions. In the left hand column under ‘n’ is the number of PAIRS of data. For this example there are 5 PAIRS of data. Find the intersection of n = 5 and α = .05, thus: **.878**

(Table A-6 is below)

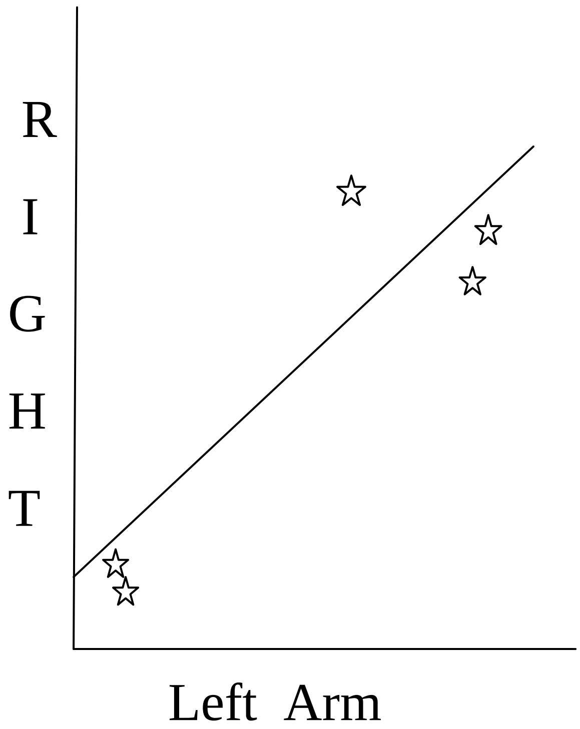
A-6

TABLE A-6 Critical Values of the Pearson Correlation Coefficient *r*

n	$\alpha = .05$	$\alpha = .01$
4	.950	.990
5	.878	.959
6	.811	.917
7	.754	.875
8	.707	.834
9	.666	.798
10	.632	.765
11	.602	.735
12	.576	.708
13	.553	.684
14	.532	.661
15	.514	.641
16	.497	.623
17	.482	.606
18	.468	.590
19	.456	.575
20	.444	.561
25	.396	.505
30	.361	.463
35	.335	.430
40	.312	.402
45	.294	.378
50	.279	.361
60	.254	.330
70	.236	.305
80	.220	.286
90	.207	.269
100	.196	.256

NOTE: To test $H_0: \rho = 0$ against $H_1: \rho \neq 0$, reject H_0 if the absolute value of r is greater than the critical value in the table.

Create a scatter plot and regression line then calculate the correlation coefficient r. thus:



r = .8670 ← correlation coefficient

Step 5: Compare the ABSOLUTE VALUE of r to the Table Value

thus: $|.8670| < .878$ (table value)

RULE: If the r value is Greater Than the Table value, reject the Null Hypothesis. If the r value is Less Than the Table value, accept the Null Hypothesis. Note: the Rule is written on the bottom of Table A-6

Step 6: $.8670 < .878$ therefore **Accept the Null Hypothesis**, there is no strong linear correlation.

“Hypothesis Tests for Two Proportions”

In 2001, one county reported that among 3,132 white women who had babies, 94 were multiple births. There were also 20 multiple births to 606 black women. Does this indicate any racial DIFFERENCE in the likelihood of multiple births? Assume α = .05

Note: let P_1 = white women and P_2 = black women

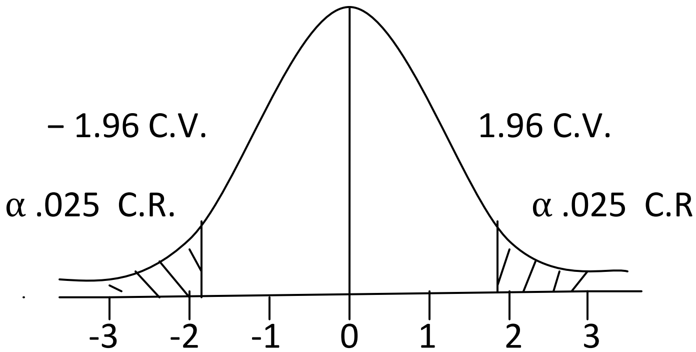
Step 1: H_0: $P_1 = P_2$

Step 2: H_1: $P_1 \neq P_2$

Step 3: α = .05

Step 4: C.V. / C.R. Note: this is a “two tail test” therefore divide α by 2 thus: $Z_{\alpha/2}$, $Z_{.05/2}$, $Z_{.025}$, “Do Z’s Don’t T’s”

subtract: .5000 − .025 = .4750 = 1.96 Z Score.



Step 5: Test Statistic

Formula: $$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}\text{ pooled }\hat{q}\text{ pooled}}{N_1} + \frac{\hat{p}\text{ pooled }\hat{q}\text{ pooled}}{N_2}}}$$

WHERE: $\hat{p} = \frac{x}{n}$ $\quad \hat{p}$ pooled $= \frac{x_1 + x_2}{n_1 + n_2}$

$$\hat{p}_1 = \frac{x}{n}, \frac{94\ white\ women}{3{,}132\ white\ women} = \mathbf{.03} \quad \hat{p}_2 = \frac{x}{n}, \frac{20\ black\ women}{606\ black\ women} = \mathbf{.033}$$

$\hat{p}$ pooled: "Everyone in the Pool" $= \frac{x_1 + x_2}{n_1 + n_2}$

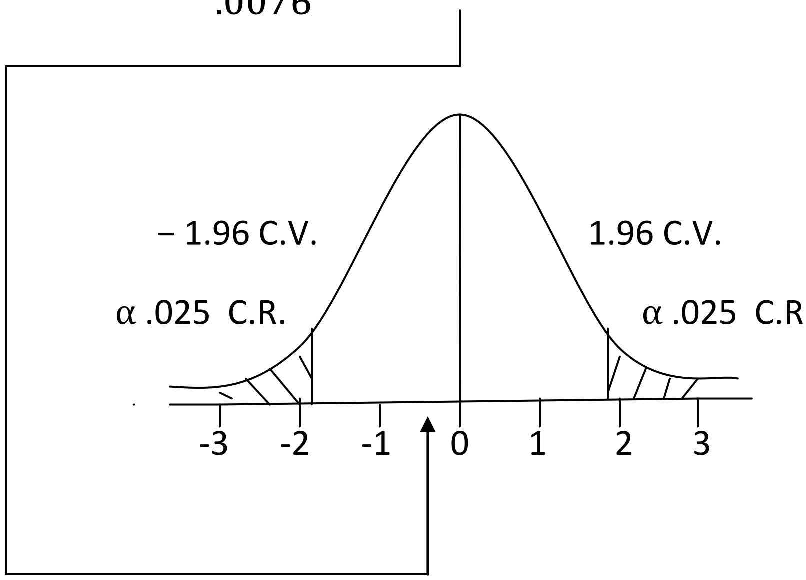
$$\hat{p}\text{ pooled} = \frac{94\ white + 20\ black}{3{,}132\ white + 606\ black} = \frac{114}{3{,}738} = \mathbf{.03}$$

$\hat{q}$ pooled (compliment) = 1 − .03 = **.97**

Note: zero is assumed here because their difference is 0

$$Z = \frac{(.03 - .033) - 0}{\sqrt{\frac{(.03)(.97)}{3{,}132} + \frac{(.03)(.97)}{606}}}, \quad \frac{-.003}{\sqrt{(9.25..._{e-6}) + (4.78..._{e-5})}}, \quad \frac{-.003}{\sqrt{.000059}},$$

$$\frac{-.003}{.0076}, = \mathbf{-.39\ test\ statistic}$$



Step 6: **Accept H_0**

NOTE: THESE PROBLEMS ARE NOT ALWAYS "TWO TAIL TESTS"

"Difference Between Two Means"

The following table represents Heats #2 and #5 of an Olympic swimming event. Are the times for Heats #2 and #5 the SAME given a significance level of .10? Follow the 6-Step procedure, draw the distribution curve and show the C.V. & C.R. DO NOT show the empirical rule values on the distribution curve.

<u>Note</u>: use "Two Variable Stats"

		L_1			L_2
Country	Heat	Time	Country	Heat	Time
ARG	2	256.42	FRA	5	246.76
SLO	2	257.79	JPN	5	249.10
CHI	2	258.68	R0M	5	250.39
MKD	2	259.39	GER	5	250.46
JAM	2	260.00	AUS	5	251.67
NZL	2	261.58	CHN	5	251.81
K0R	2	261.65	CAN	5	252.33
UKR	2	266.30	BRA	5	253.75

Summary: $\bar{x}_1$ = 260.226, S_x (S_1) = 3.031

$\bar{x}_2$ ($\bar{y}$) = 250.783, S_y (S_2) = 2.148 n = 8

<u>Step 1</u>: H_0: $\mu_2 = \mu_5$

<u>Step 2</u>: H_1: $\mu_2 \neq \mu_5$

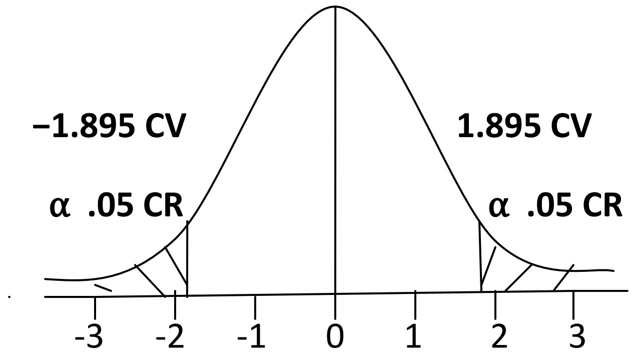
<u>Step 3</u>: $\alpha = .10$

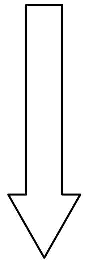
<u>Step 4</u>: C.V. / C.R.

<u>Note</u>: n = 8 (pairs of data) use the T-Table, this is a "two-tailed" test. Divide alpha by 2 "Do Z's Don't T's"

Thus: n – 1 = 7 "degrees of freedom",

$T_{7\ .05} = 1.895$ T Score





Step 5: Test Statistic

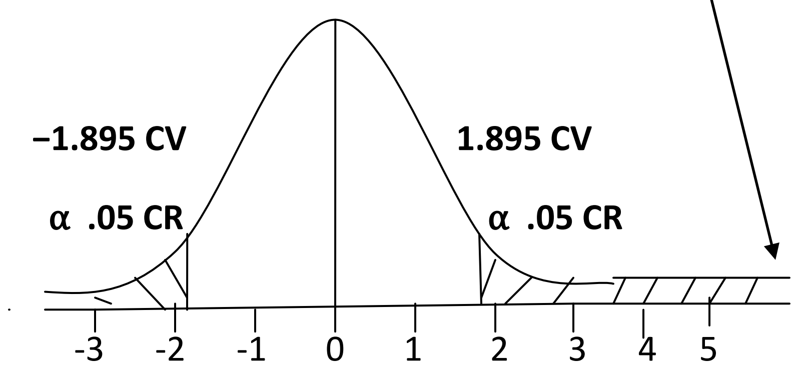
Formula:

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(S_1)^2}{N_1} + \frac{(S_2)^2}{N_2}}}$$

Note: this is 0, if the values are the same, their difference is 0

$$T = \frac{(260.226 - 250.783) - (0)}{\sqrt{\frac{(3.031)^2}{8} + \frac{(2.148)^2}{8}}}, \quad \frac{9.443}{\sqrt{\frac{9.186}{8} + \frac{4.613}{8}}}, \quad \frac{9.443}{\sqrt{1.148 + .576}}$$

$$\frac{9.443}{\sqrt{1.724}}, \quad \frac{9.443}{1.313}$$ = **7.191 Test Statistic**



Step 6: **Reject H_0**

NOTE: THESE PROBLEMS ARE NOT ALWAYS "2 TAIL TESTS"

(End of Course Material)

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(S_1)^2}{N_1} + \frac{(S_2)^2}{N_2}}}$$

X^2cdf(;

$T = \frac{.027}{.784}$; Answer = **.034** $(Y_2 - Y_1) \div (X_2 - X_1)$

$\bar{x} = \frac{\sum(f)X}{\sum f}$ $Q_3 - Q_1$

Midrange = $\frac{highest\ score + lowest\ score}{2}$

$Z = \frac{X - \bar{X}}{S}$ $Z = \frac{X - \mu}{\sigma}$

$\mu = \sum X \bullet P(X)$ $Ó^2 = \sum X^2 \bullet P(X) - \mu^2$ $Ó = \sqrt{\sum X^2 \bullet P(X) - \mu^2}$

$P(X) = \frac{n!}{(n-x)!\,x!} \bullet p^x \bullet q^{n-x}$ $\mu = np$ $Ó^2 = npq$ $Ó = \sqrt{npq}$

$Z = \frac{\bar{x} - \mu_{\bar{x}}}{Ó_{\bar{x}}}$ where: $Ó_{\bar{x}} = \frac{Ó}{\sqrt{n}}$

$Ó^2 = \sum(x - \mu)^2 \bullet P(X)$ $Ó = \sqrt{\sum(x - \mu)^2 \bullet P(X)}$ $Z_{\alpha/2}$ $Z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$

$\bar{x} - \mu$ $\quad s \div \sqrt{n}$ $\quad \alpha$ $\quad \hat{p}$ $\quad \hat{q}$ $\quad \bar{x}$ $\quad \frac{s}{\sqrt{n}}$ $\quad$ Ó $\quad \mu$ $\quad X_1$ $\quad X_2$ $\quad$ Ó^2

$$T = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} \qquad \hat{p} = \frac{x}{n} \quad \hat{q} = 1 - \hat{p} \quad S_1 \quad S_2 \quad \bar{X}_1 \quad \bar{X}_2$$

$$N = \frac{(Z\alpha/2)^2 \bullet .25}{E^2} \qquad \bar{x} - Z\alpha/2 \left(\frac{s}{\sqrt{n}}\right) \quad \bar{x} + Z\alpha/2 \left(\frac{s}{\sqrt{n}}\right) \quad q = 1 - p$$

$$Z = \frac{\bar{x} - \mu_{\bar{x}}}{Ó_{\bar{x}}} \quad \text{where: } Ó_{\bar{x}} = \frac{Ó}{\sqrt{n}} \qquad n = \frac{(Z\alpha/2 \bullet Ó)^2}{E^2} \quad \hat{p} = \frac{x}{n} \quad \hat{q} = 1 - \hat{p}$$

$$\hat{p} - Z\alpha/2 \sqrt{\frac{\hat{p}\hat{q}}{n}}, \quad \hat{p} + Z\alpha/2 \sqrt{\frac{\hat{p}\hat{q}}{n}} \qquad \bar{x} - T\alpha/2 \left(\frac{s}{\sqrt{n}}\right) \quad \bar{x} + T\alpha/2 \left(\frac{s}{\sqrt{n}}\right)$$

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{{S_1}^2}{N_1} + \frac{{S_2}^2}{N_2}}} \qquad z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} \qquad z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \qquad T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(S_1)^2}{N_1} + \frac{(S_2)^2}{N_2}}}$$

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}\text{ pooled } \hat{q}\text{ pooled}}{N_1} + \frac{\hat{p}\text{ pooled } \hat{q}\text{ pooled}}{N_2}}}$$

WHERE: $\hat{p} = \frac{x}{n}$ $\quad \hat{p}$ pooled $= \frac{x_1 + x_2}{n_1 + n_2}$

P_0 $\quad$ ŷ $\quad$ X $\quad$ n $\quad \neq P_0$ $\quad < P_0$ $\quad > P_0$ $\quad$ Ó $\quad \bar{x}$ $\quad$ N $\quad Y_1$

$\neq \mu_0$ $\quad < \mu_0$ $\quad > \mu_0$ $\quad H_0$ $\quad H_1$ $\quad \mu_0$ Ó $\quad \bar{x}$ $\quad$ n $\quad \mu_0$ $\quad \bar{x}$ $\quad S_x$ $\quad$ n

$\bar{x}_1$ $\quad S_{x_1}$ $\quad N_1$ $\quad \bar{x}_2$ $\quad S_{x_2}$ $\quad N_2$ $\quad \neq \mu_2$ $\quad < \mu_2$ $\quad > \mu_2$

$\neq P_2$ $\quad < P_2$ $\quad > P_2$

$Z = \dfrac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$ $T = \dfrac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$ $\chi^2 = \dfrac{(N-1) \bullet S^2}{\text{Ó}^2}$ $Z = \dfrac{(x + .5) - \frac{n}{2}}{\frac{\sqrt{n}}{2}}$

$\geq$ $\pm$ $\leq$ X^{-1} $-$ X^2 $+$ $-$ $\bullet$ X^{-1} $(-)$ X T Θ N $\backslash Y_1$ $\backslash Y_2$ $\backslash Y_3$ $\frac{S}{N}$

$([A] - [B])^{-1} [C]$ $\cup$ $\cap$ A' $A \bullet B$ $(A' \bullet B) + (A \bullet B)$ $\dfrac{A \bullet B}{(A' \bullet B) + (A \bullet B)}$

P_0 ŷ X n $\neq P_0$ $< P_0$ $> P_0$ Ó $\bar{x}$ N Y_1

$\neq \mu_0$ $< \mu_0$ $> \mu_0$ H_0 H_1 μ_0 Ó $\bar{x}$ n μ_0 $\bar{x}$ S_x n

$\bar{x}_1$ S_{x_1} N_1 $\bar{x}_2$ S_{x_2} N_2 $\neq \mu_2$ $< \mu_2$ $> \mu_2$

$\neq P_2$ $< P_2$ $> P_2$

where $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $P(X) = \dfrac{n!}{(n-x)!\,x!} \bullet p^x \bullet q^{n-x}$ nCr nPr

$\mu = \sum X \bullet P(X)$ $Y = \frac{1}{4} x^2 + x + 2$

$P = R\left[\dfrac{1-(1+i)^{-n}}{i}\right]$ $R = \dfrac{S\,i}{(1+i)^n - 1}$ $S = R\left[\dfrac{(1+i)^n - 1}{i}\right]$

$r_e = (1 + \frac{r}{m})^m - 1$ $A = P(1+i)^n$ $I = Prt$ $A = P(1 + rt)$

For Vertex use: $X = \dfrac{-b}{2a}$ For X intercepts use the quadratic: $\dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

$T = \frac{Ln(multiple)}{Ln(1+r)}$ $A = Pe^{rt}$ $r_e = \left(1 + \frac{r}{m}\right)^m - 1$ $A = P\left(1 + \frac{r}{m}\right)^{tm}$

$r_e = e^r - 1$