Answer Key

Statistics

Course Workbook

2022-23 Edition

Donny Brusca

www.CourseWorkBooks.com

Table of Contents

Chapter 1 Data Collection

1.1 Types of Data

1. (3)	2. a) quantitative b) qualitative c) qualitative d) quantitative
3. (4) The numerical responses would not be used arithmetically, and the order of the categories is arbitrary.	4. (1)
5. (1)	
6. bivariate: the two variables represent the sales quarter (Q1, Q2, Q3, or Q4) and the region (East, West, North, and South); the data values are the sales figures.	

1.2 Sampling

1. The population is all the bolts in the shipment. The sample is the 100 selected bolts.	2. The population is all the mall shoppers. The sample is every sixth person within the 3-hour period.
3. a) Because the average of $350 is based on a sample, this is a statistic. b) Because the average of $425 is based on a population, this is a parameter.	4. a) parameter b) statistic c) statistic
5. a) 3 b) 2 c) 4 d) 1	
6. (3) cluster sampling: if each franchise is a heterogeneous cluster, we can limit the number of stores to which we need to travel. Incorrect responses explained: (1) and (2) may require that we travel to many or all of the 200 franchises (4) is only appropriate when there are notable groups in the population that need to be accounted for	

1.3 Methods

1. (4)	2. (2)
3. (2)	4. double-blinding
5. a) The control group of plants would receive the normal level of CO2 (300 ppm). There should be two experimental groups, one which is exposed to 400 ppm and one which is exposed to 500 ppm. b) The independent variable is level of CO_2 exposure. The dependent variable is the rate of photosynthesis.	

1.4 Bias

1. (3) Seniors or physics students may be biased by aspects of class scheduling specific to their groups. Selecting only students from the cafeteria would omit students who have already chosen not to eat there.	
2. (4) Allowing subjects to self-select their participation can lead to bias. Honors calculus students may tend to spend more (or less) time on homework due to the nature of their courses. Surveying only teenagers at a movie theater would omit other age groups as well as people who don't like to go to movie theaters.	
3. (4) People who attend a football game are more likely to prefer an increase in the sports budget since they are sports fans.	
4. (2)	5. (1)

2.1 Categorical Data

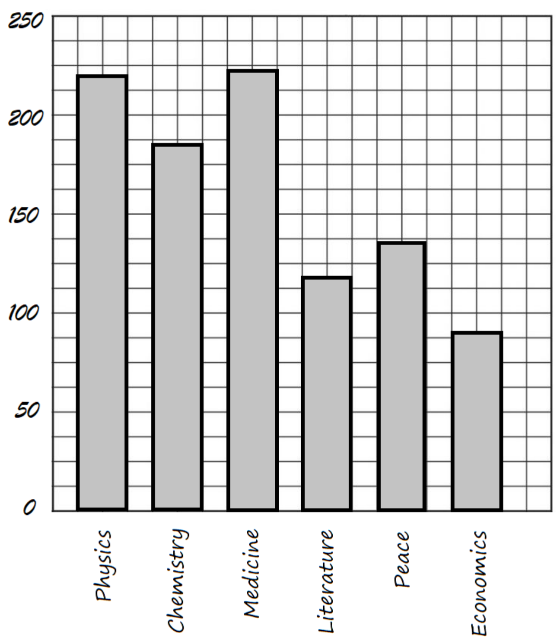
1.

Pet	Frequency
Rabbit	3
Cat	8
Dog	10
Fish	3

2.

Player	Frequency (f)	Relative Frequency (rf)
Able	13	0.19
Baker	18	0.26
Charlie	10	0.14
Daniels	10	0.14
Edwards	19	0.27

3.

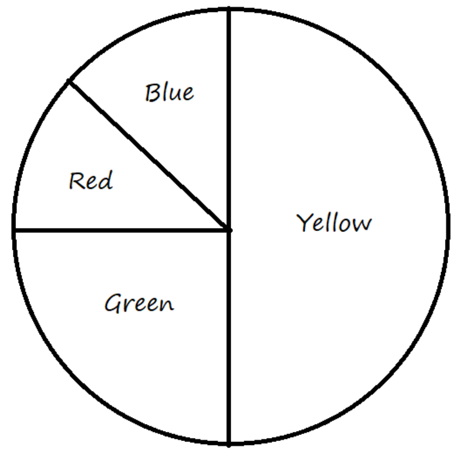


4. $\sum f = 973$

Category	Nobel Prizes	Relative Frequency
Physics	219	0.23
Chemistry	186	0.19
Medicine	224	0.23
Literature	118	0.12
Peace	137	0.14
Economics	89	0.09

5.

Color	Frequency	Relative Frequency
Red	5	$0.125 = \frac{1}{8}$
Yellow	20	$0.50 = \frac{1}{2}$
Green	10	$0.25 = \frac{1}{4}$
Blue	5	$0.125 = \frac{1}{8}$



2.2 Quantitative Data

1.

Result (x)	Frequency (f)	Relative Frequency (rf)
1	5	0.25
2	3	0.15
3	1	0.05
4	5	0.25
5	4	0.20
6	2	0.10

2.

Result (x)	Frequency (f)	Cumulative Frequency (cf)
1	5	5
2	3	8
3	1	9
4	5	14
5	4	18
6	2	20

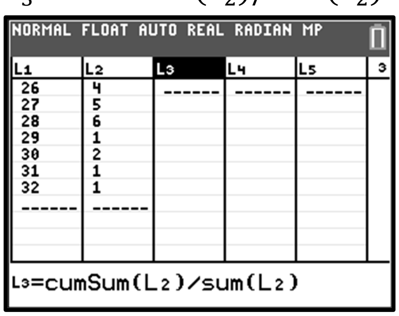
3.

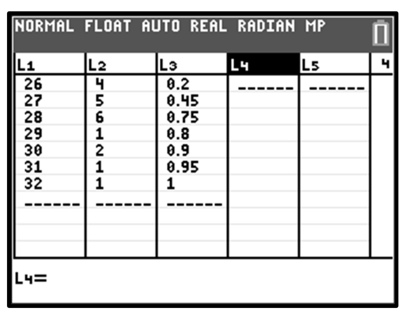
Bases (x)	Frequency (f)	Cumulative Frequency (cf)	Relative Frequency (rf)	Cumulative Relative Frequency (crf)
1	25,006	25,006	0.633	0.633
2	7,863	32,869	0.199	0.832
3	671	33,540	0.017	0.849
4	5,944	39,484	0.151	1.000

4.

Age (x)	Frequency (f)	Relative Frequency (rf)	Cumulative Frequency (cf)
26	4	0.20	4
27	5	0.25	9
28	6	0.30	15
29	1	0.05	16
30	2	0.10	18
31	1	0.05	19
32	1	0.05	20

5. $L_3 = \text{cumSum}(L_2)/\text{sum}(L_2)$





2.3 Graphs of Small Data Sets

1. a) 40 b) $\frac{18}{40} = 45\%$

2. a) 50 b) $\frac{15}{50} = 30\%$

3.

Value	Frequency
0	3
1	3
2	7
3	6
4	6
5	3
6	4
7	8
8	2
9	8

4.

4	9
5	2666
6	77
7	239
8	12779
9	1
10	3

Key: 4|9 = 49

5.

0	2 3 3 4
0	5 5 5 6 6 7 7 7 8 8 8 9
1	0 0 0 1 1 2 2 2 3 3 4 4
1	9
2	1 2 2 3
2	5 7
3	1

Key: 3|1 = 31

6.

Before		*After*
9 8 8 8	6	
8 6 6 4 1 1 0	7	
8 8 6 2	8	6 7 8 8
6 0	9	0 2 2 4 5 8 9 9
4	10	0 4 4
0	11	8
	12	4 4
	13	
	14	6

Key: 2|8|6 represents a pulse of 82 before and a pulse of 86 after

2.4 Classes

1. 25 – 18 = 7

2.

Yards	Field Goals	Cumulative Frequency
0 – 19	1	1
20 – 29	236	237
30 – 39	281	518
40 – 49	236	754
50 – 59	120	874

3. a) class width is $7 - 0 = 7$.
   b) boundaries are 6.5 and 13.5,
   midpoint is $\frac{7+13}{2} = 10$.

4. a) class width is $23 - 15 = 8$.
   b) boundaries are 30.5 and 38.5,
   midpoint is $\frac{31+38}{2} = 34.5$.

5. range $= 135 - 17 = 118$
   class width $= \frac{118}{8} = 14.75 \rightarrow 15$
   classes: 17 – 31, 32 – 46, 47 – 61,
   62 – 76, 77 – 91, 92 – 106, 107 – 121,
   122 – 136

6. range $= 94 - 20 = 74$
   class width $= \frac{74}{7} \approx 10.6 \rightarrow 11$
   classes: 20 – 30, 31 – 41, 42 – 52,
   53 – 63, 64 – 74, 75 – 85, 86 – 96

7. range $= 58 - 2 = 56$
   class width $= \frac{56}{6} \approx 9.3 \rightarrow 10$
   classes: 2 – 11, 12 – 21, 22 – 31,
   32 – 41, 42 – 51, 52 – 61
   Yes, by subtracting 2 from each limit, the new intervals would be
   0 – 9, 10 – 19, 20 – 29, 30 – 39, 40 – 49, 50 – 59
   (The last class includes the maximum.)

8. range $= 60 - 12 = 48$
   class width $= \frac{48}{5} = 9.6 \rightarrow 10$
   classes: 12 – 21, 22 – 31, 32 – 41,
   42 – 51, 52 – 61
   We cannot shift the intervals. To do so would mean subtracting 2 from each limit (10 – 19, etc.), but this would make the last class $50 - 59$, which would exclude the maximum of 60.

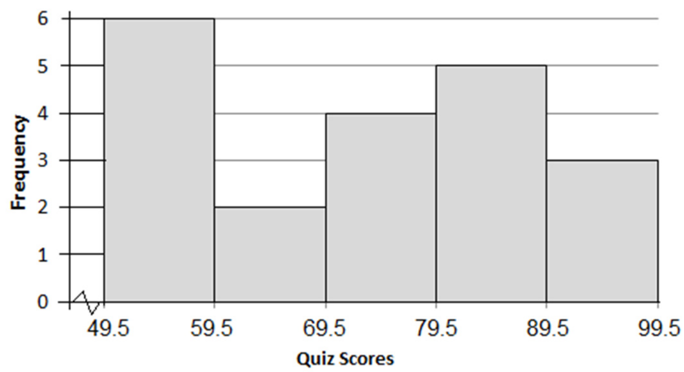
2.5 Histograms

1. Add the frequencies: $2 + 4 + 5 + 4 + 1 = 16$	2. Add the frequencies: $7 + 10 + 3 + 5 = 25$
3. 20	4. a) 3 b) 0 c) 20

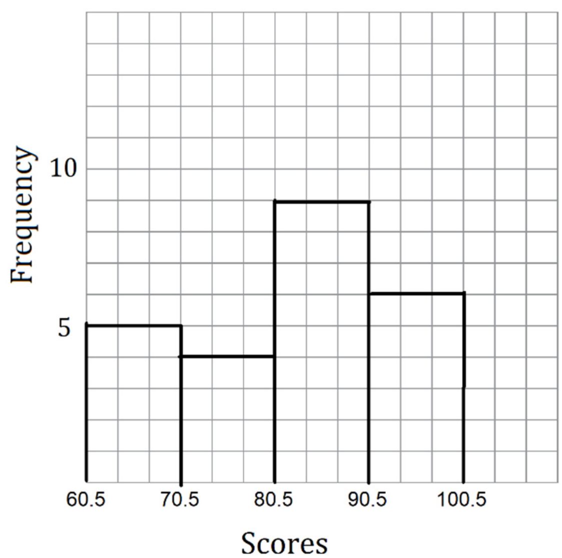
5.

Mathematics Quiz Scores

Interval	Tally	Frequency
50–59	𝍸 I	6
60–69	II	2
70–79	IIII	4
80–89	𝍸	5
90–99	III	3

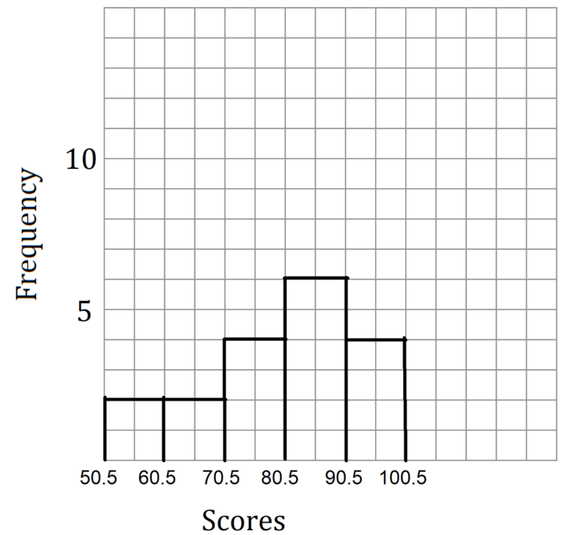


6.



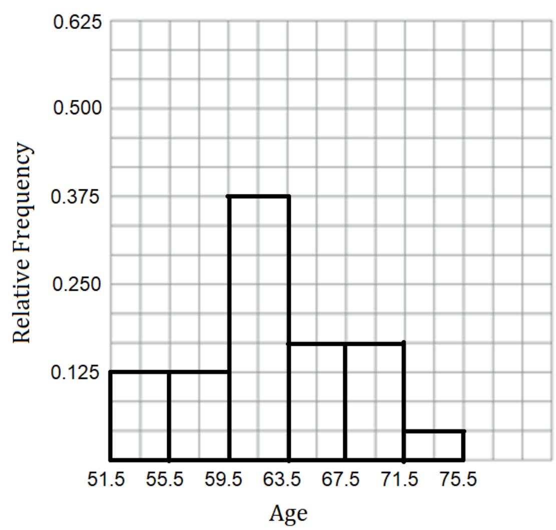
7.

Interval	Tally	Frequency
51–60	II	2
61–70	II	2
71–80	IIII	4
81–90	~~IIII~~ I	6
91–100	IIII	4



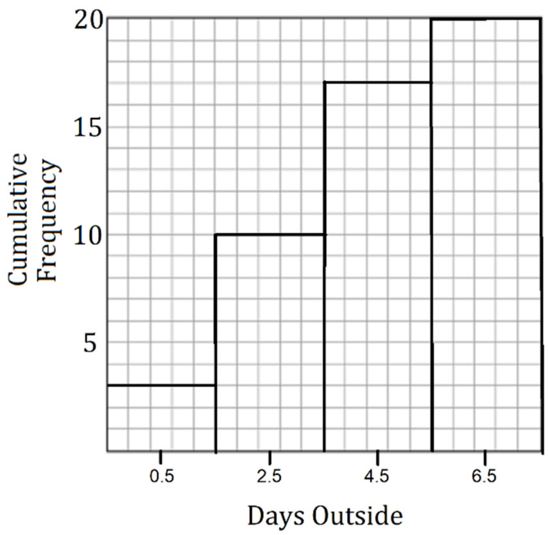
8.

Class	Frequency (f)	Relative Frequency (rf)
52 – 55	3	0.125
56 – 59	3	0.125
60 – 63	9	0.375
64 – 67	4	0.167
68 – 71	4	0.167
72 – 75	1	0.042



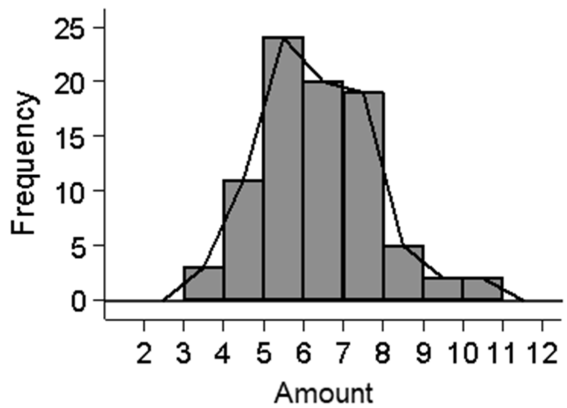
9.

Days	Tally	Frequency	Cumulative Frequency
0 – 1	III	3	3
2 – 3	~~IIII~~ II	7	10
4 – 5	~~IIII~~ II	7	17
6 – 7	III	3	20

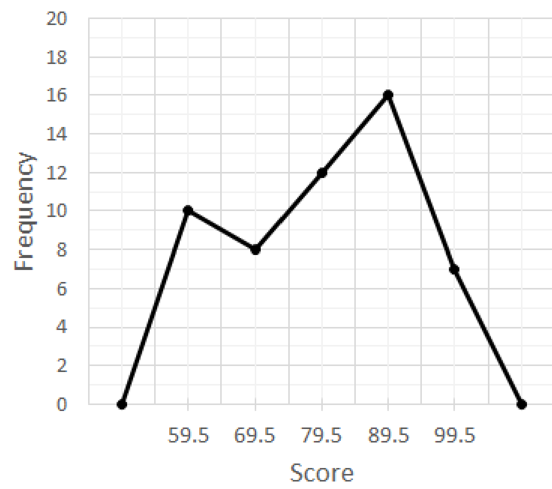


2.6 Frequency Polygons

1.



2.

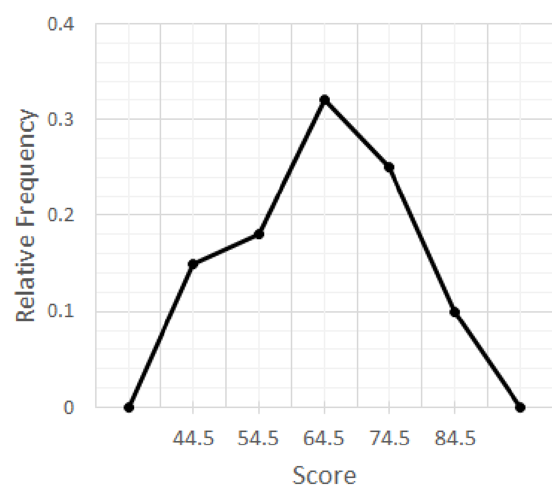


3. a) 3 and 25 b) 100 c) 6 d) 10 – 15

4.

Score (x)	Frequency (f)
50 – 59	5
60 – 69	10
70 – 79	30
80 – 89	40
90 – 99	15

5.



2.7 Ogives

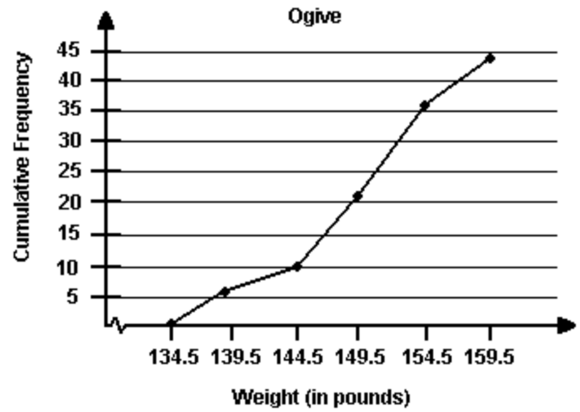
1. (2) 80

2. a) 60 b) 40

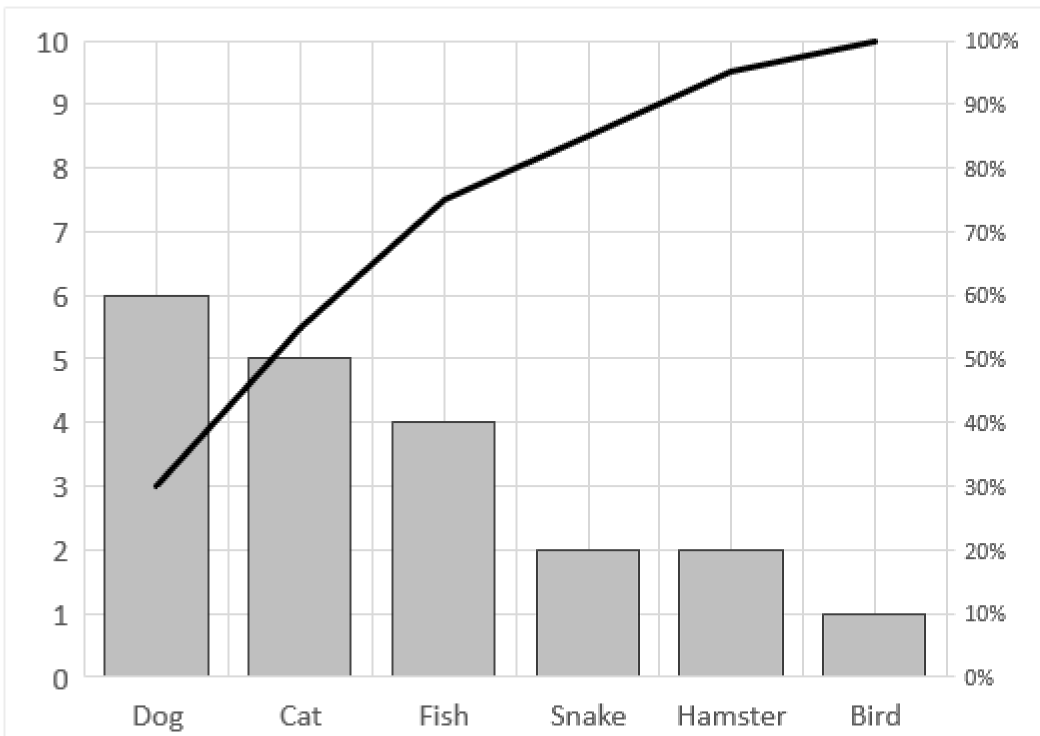
3. (4)

4.

Class	Frequency (f)	Cumulative Frequency (cf)
135 – 139	6	6
140 – 144	4	10
145 – 149	11	21
150 – 154	15	36
155 – 159	8	44



5.



Chapter 3 Descriptive Statistics

3.1 Center

1. mode	2. median
3. (1)	4. mean = 79, median = 79, mode = 78
5. (2) mode = median = 6	6. (1) mean = 17, median = 18, mode = 22
7. data values are: 48, 63, 65, 69, 71, 74, 74, 78, 82, 83, 85, 85, 89, 90, 94, 96, 100 mean ≈ 79.2, median = 82, modes are 74 and 85 (bimodal)	8. data values are: 0.8, 1.5, 1.6, 1.8, 2.1, 2.3, 2.4, 2.5, 3.0, 3.4, 3.5, 3.6, 3.9, 4.0, 4.0 mean ≈ 2.7, median = 2.5, mode = 4.0

9. (3) Enter into the calculator as L1 and L2 and find 1-Var Stats, or calculate as below.

Score (x)	Frequency (f)	(xf)
96	2	192
92	5	460
88	3	264
84	2	168
78	4	312
60	1	60
Σ	**17**	**1456**

mean $= \frac{\sum xf}{\sum f} = \frac{1456}{17} \approx 85.6$,

median is value at position $\frac{17+1}{2} = 9$, which is 88

mode is the value with the highest frequency, which is 92

10. mean $=$ 5.625, median $=$ 5, mode $=$ 10

11. data values are 15, 25, 20, 20, 30 mean = 22, median = 20, mode = 20	12. outlier is 2, mean = 23, trimmed mean (without 2 and 36) = 24.6
13. mean = \$44.75, trimmed mean = \$41.50	14. mean = \$14.25, trimmed mean = \$15.33

15. a) mean = \$225,000, median = \$175,000
b) the median because the mean is higher than all but one of the values (an outlier, \$700,000)
c) trimmed mean ≈ \$186,100

16. 131 – 150;
There are 44 total scores, so the median would be the average of the 22nd and 23rd highest scores.

17. 71-80;
Out of 31 students, the 16th lowest value is the median, which is within 71-80 interval.

18.

Class	Midpt ($\boldsymbol{x}$)	Freq ($\boldsymbol{f}$)	($\boldsymbol{xf}$)
13 – 22	17.5	6	105
23 – 32	27.5	8	220
33 – 42	37.5	7	262.5
43 – 52	47.5	5	237.5
53 – 62	57.5	4	230
Σ		**30**	**1055**

mean is $\frac{\sum xf}{\sum f} = \frac{1055}{30} \approx 35.17$

19.

Class	Midpt ($\boldsymbol{x}$)	Rel Freq ($\boldsymbol{rf}$)	($\boldsymbol{x \cdot rf}$)
25 – 29	27	0.15	4.05
30 – 34	32	0.25	8.0
35 – 39	37	0.30	11.1
40 – 44	42	0.25	10.5
45 – 49	47	0.05	2.35
Σ		**1.00**	**36.0**

mean is $\sum(x \cdot rf) = 36.0$

20.

$\boldsymbol{x}$	$\boldsymbol{w}$	$\boldsymbol{xw}$
85	0.05	4.25
80	0.35	28
100	0.20	20
90	0.15	13.5
93	0.25	23.25
Σ	**1.00**	**89.0**

weighted mean is $\sum xw = 89$.

21.

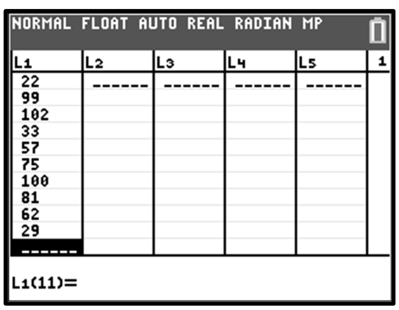
Course	Gr	Pts ($\boldsymbol{x}$)	Creds ($\boldsymbol{w}$)	($\boldsymbol{xw}$)
MAT 210	A	4	4	16
SOC 150	C	2	3	6
BIO 245	B	3	4	12
ENG 110	F	0	3	0
GRK 101	D	1	2	2
ART 205	B	3	1	3
Σ			**17**	**39**

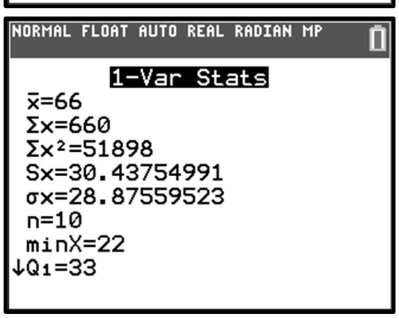
weighted mean is $\frac{\sum xw}{\sum w} = \frac{39}{17} \approx 2.29$.

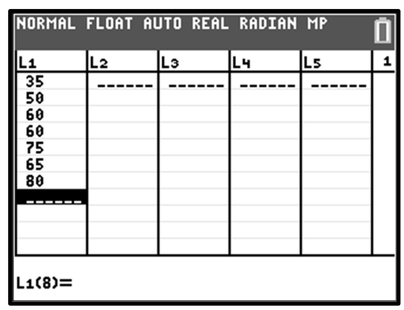
3.2 Shape

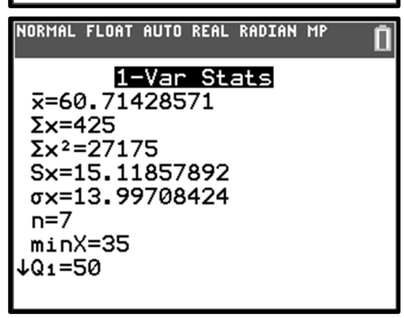
1. skewed to the right	2. skewed to the left
3. symmetrical, but with outliers at 9.45	4. symmetrical, but bimodal
5. a) 5 b) 3 c) 1 d) 6 e) 2 f) 4	
6. (1) because the distribution is skewed right	

3.3 Spread

<table>
<tr><td>1. City A (22)</td><td>2. an outlier such as a very low score could greatly affect the range without affecting the median</td></tr>
<tr><td>3. (2)</td><td>4. (1)</td></tr>
<tr><td>5. The first set, as shown by the smaller SD.</td><td>6. McCrane; a larger SD means more variability</td></tr>
<tr><td>7. Press [STAT]<CALC>[1] for 1-Var Stats to get $\mu = 66$ (listed as $\bar{x}$) and $\sigma \approx 28.9$.


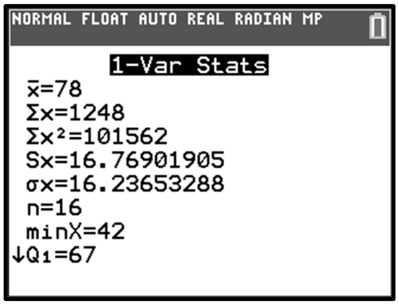


</td><td>8. Press [STAT]<CALC>[1] for 1-Var Stats to get $\bar{x} \approx$ 60.7 and $s \approx 15.1$.


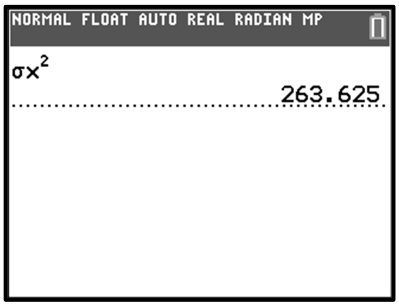


</td></tr>
</table>

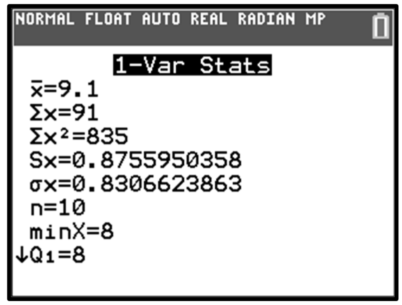
9. $\mu = 78$ and $\sigma \approx 16.8$



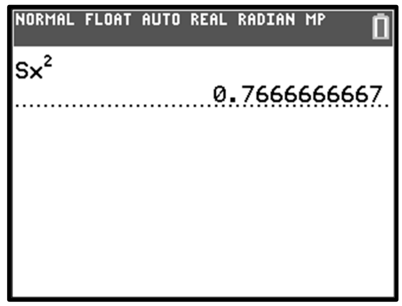
Press [VARS] [5] [4] [x^2] [ENTER] to get $\sigma^2 = 263.6$



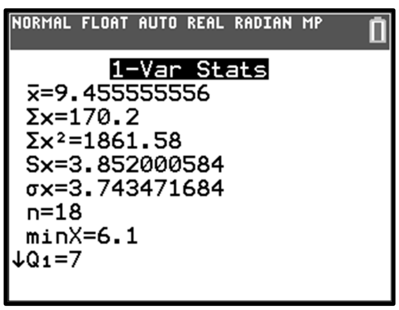
10. $\bar{x} = 9.1$ and $s \approx 0.88$



Press [VARS] [5] [3] [x^2] [ENTER] to get $s^2 \approx 0.77$.



11. $\bar{x} \approx 9.46$ and $s \approx 3.85$



12. $\mu = \frac{90}{10} = 9$

x	$x-\mu$	$(x-\mu)^2$
5	−4	16
7	−2	4
7	−2	4
8	−1	1
9	0	0
9	0	0
10	1	1
11	2	4
12	3	9
12	3	9
Σ		**48**

$$\sigma^2 = \frac{\sum(x-\mu)^2}{n-1} = \frac{48}{10} = 4.8$$

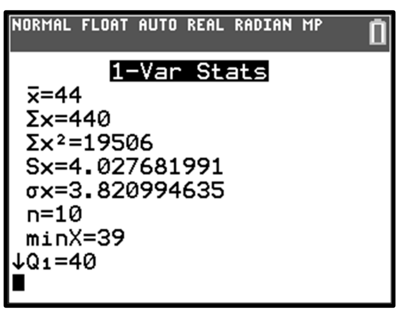
$$\sigma = \sqrt{4.8} \approx 2.2$$

13. $\bar{x} = \frac{440}{10} = 44$

x	$x - \bar{x}$	$(x - \bar{x})^2$
51	7	49
48	4	16
47	3	9
46	2	4
45	1	1
43	–1	1
41	–3	9
40	–4	16
40	–4	16
39	–5	25
Σ		**146**

$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{146}{9} = 16.\overline{2}$

$s = \sqrt{16.\overline{2}} \approx 4.0$

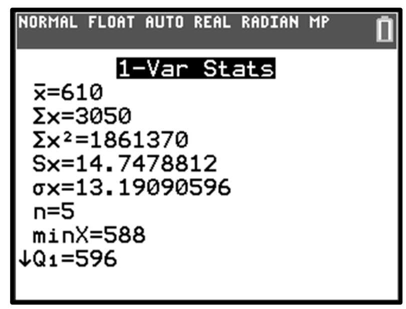


14. $\bar{x} = \frac{\sum x}{n} = \frac{3050}{5} = 610$

x	$x - \bar{x}$	$(x - \bar{x})^2$
612	2	4
588	–22	484
604	–6	36
625	15	225
621	11	121
Σ		**870**

$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{870}{4} = 217.5$

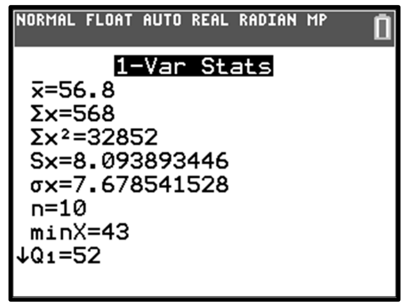
$s = \sqrt{217.5} \approx 14.7$



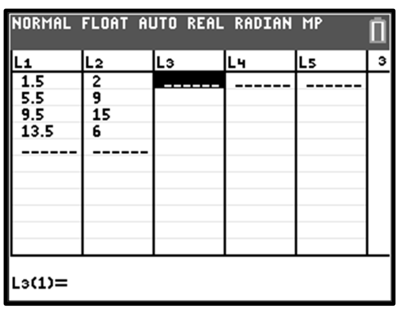
15. $\bar{x} = \frac{\sum x}{n} = \frac{568}{10} = 56.8$

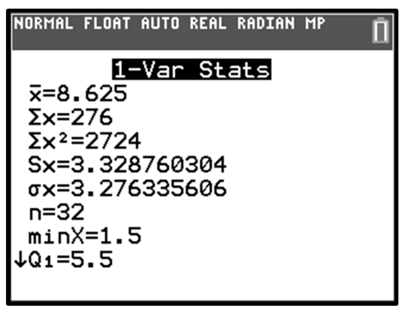
x	$x - \bar{x}$	$(x - \bar{x})^2$
60	3.2	10.24
62	5.2	27.04
43	−13.8	190.44
55	−1.8	3.24
56	−0.8	0.64
61	4.2	17.64
52	−4.8	23.04
69	12.2	148.84
64	7.2	51.84
46	−10.8	116.64
Σ		**589.6**

$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{589.6}{9} = 65.5\overline{1}$ $s = \sqrt{65.5\overline{1}} \approx 8.1$



16. Use the midpoints of the classes as the *x* values.





$\mu = 8.625$ and $\sigma \approx 3.3$, approximately.

17. They are all divided by two as well.

18. The mean increased by five and the range remained the same.

19. a) mean ≈ 11.4, median = 12, mode = 7, range = 15, SD ≈ 5.38
b) the mean, median and mode increase by 5; the range and SD remain the same.

3.4 Position

1. 75% of 40 = 30 students weigh below 150 pounds, so $40 - 30 = 10$ students weigh at least 150 pounds

2. $100\left(\frac{95{,}000+0.5}{125{,}000}\right) = 76$, so the 76th percentile.

3. $100\left(\frac{22+0.5}{30}\right) = 75$, so the 75th percentile.

4. 25, 39, 42, 58, 64, 70, 75, 87, 90, 95
   $p = \frac{b+0.5}{n} = \frac{5}{10} = 0.55$, so 70 is the 55th percentile.

5. second quartile = median = $\frac{35+45}{2}$ = 40

6. 5, 6, 7, 8, 12, 14, 17, 17, 18, 19, 19

   $Q_1 = 7$, $Q_2 = 14$, $Q_3 = 18$

7. 3, 6, 7, | 7, 8, 9, | 9, 9, 10, | 12, 13, 15

   $Q_1 = 7$, $Q_2 = 9$, $Q_3 = 11$

8. 21, 28, | 28, 32, 33, 41, 45, | 50, 53

   $Q_1 = 28$, $Q_2 = 33$, $Q_3 = 47.5$, IQR = 19.5

9. 71, 71, 72, 74, 74, | 75, 78, 79, 79, 83

   $Q_3 = 79$ and $Q_1 = 72$, so IQR = 7

10. $Q_1 = 70$, $Q_2 = 80$, $Q_3 = 90$

11. The corresponding frequency table would show:

Minutes Used	Frequency
31–40	2
41–50	3
51–60	5
61–70	9
71–80	11

25% of 30 is 7.5, so the first quartile would be between the 7th and 8th smallest values out of 30. This falls within the 51–60 interval.

12. $74 + 6 = 80$

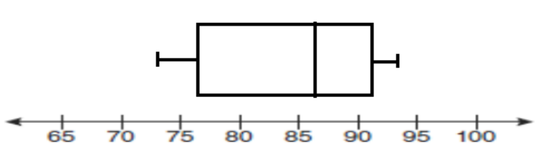
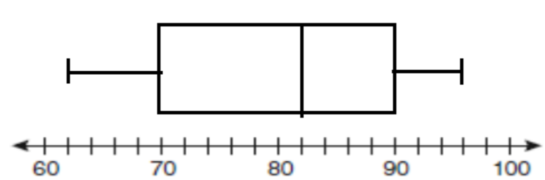
13. $85 - 2(4) = 77$

14. $z = \frac{99.5 - 98.6}{0.62} = 1.45$

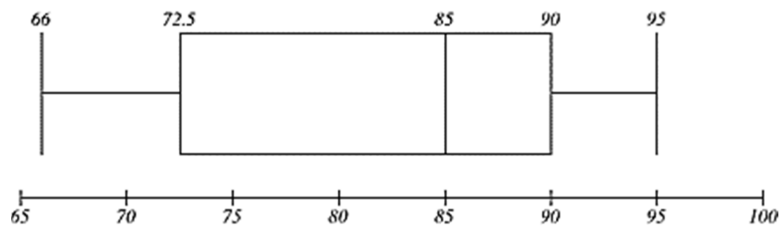
15. $z = \frac{2.67 - 3.0}{0.2} = -1.65$

16. Jason: $z = \frac{1150 - 1000}{100} = 1.50$ Mary: $z = \frac{26 - 22}{2} = 2.00$ Mary performed better.	17. East Point: $z = \frac{95 - 80}{10} = 1.5$ West Point: $z = \frac{95 - 84}{4} = 2.75$ The temperature at West Point was more unusual.
18. Lion: $z = \frac{70 - 60}{10} = 1$ Rhino: $z = \frac{56 - 40}{8} = 2$ The rhino is running relatively faster.	19. $1.5 = \frac{x - 80}{10}$ $x = 1.5(10) + 80 = 95$
20. $-0.5 = \frac{x - 5.1}{0.9}$ $x = -0.5(0.9) + 5.1 = 4.65$	21. $3 = \frac{7.5 - \mu}{0.5}$ $3(0.5) - 7.5 = -\mu$ $\mu = -3(0.5) + 7.5 = 6$
22. $2.78 = \frac{125 - 115}{\sigma}$ $\sigma = \frac{10}{2.78} \approx 3.6$	23. $-0.86 = \frac{1241 - 1509}{\sigma}$ $\sigma = \frac{263}{0.86} \approx 312$

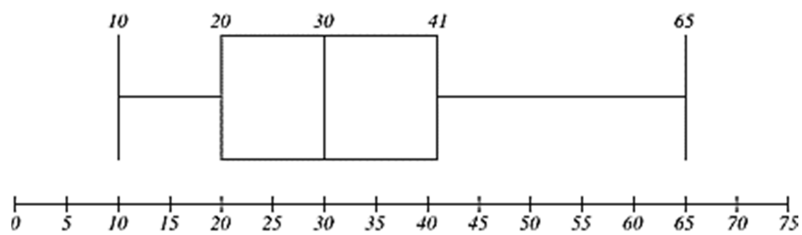
3.5 Boxplots

1. 81	2. 75
3. 10	4. 84
5. 30	6. 4
7. 75 – 15 = 60	8. 25%
9. (4) 75–88	
10. (a) = (2) right skewed; (b) = (3) no skew; (c) = (1) left skewed	
11. $Q_1 = 77$, $Q_2 = 87$, $Q_3 = 91$  	12. $Q_1 = 70$, $Q_2 = 82$, $Q_3 = 90$ 

13.


66
72.5
85
90
95
65
70
75
80
85
90
95
100

14.


10
20
30
41
65
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75

Chapter 4 Bivariate Data

4.1 Contingency Tables

1. a) $\frac{15}{113} \approx 13.3\%$ of the students are undecided.

   b) $\frac{31}{60} \approx 51.7\%$ of the 9th graders are watching.

2.

	Fiction	Nonfiction	Total
Hardcover	28	52	80
Paperback	94	36	130
Total	122	88	210

	Fiction	Nonfiction	Total
Hardcover	13.3%	24.8%	38.1%
Paperback	44.8%	17.1%	61.9%
Total	58.1%	41.9%	100%

3. Given data in bold below.

	Coca-Cola	Sprite	Total
Table	16	**14**	30
Garbage	34	8	**42**
Total	**50**	22	**72**

4. $\frac{(48)(13)}{100} = 6.24 > 4$; negative association

5.

	In Favor	**Opposed**	**Total**
Liberal	74	74	148
Moderate	15	15	30
Conservative	11	11	22
Total	100	100	200

4.2 Bivariate Bar Graphs

1.

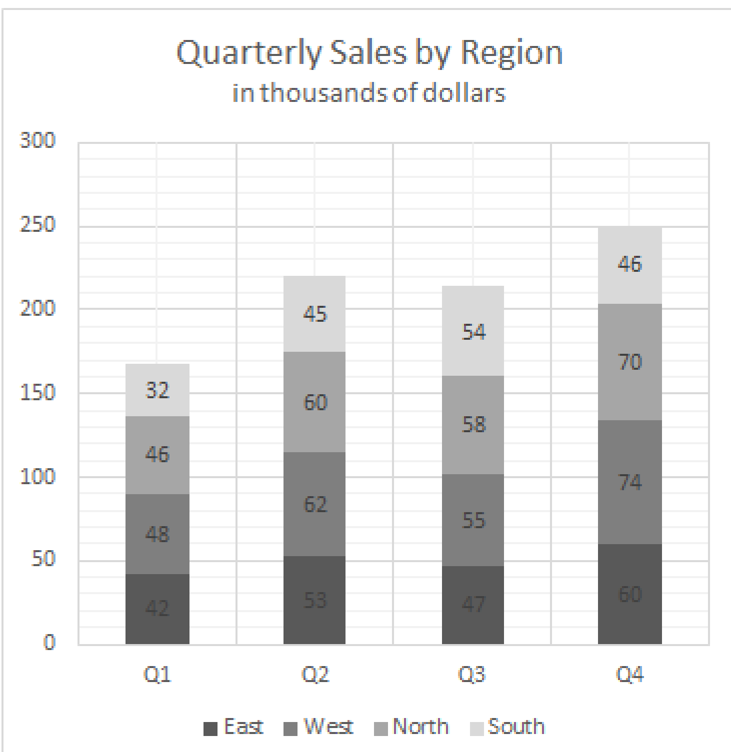
P2P Payments in Billions of Dollars

	Q1	Q2	Q3	Q4	Total
PayPal	27	30	33	36	126
Venmo	10	12	14	17	53
Zelle	21	25	28	32	106
Total	58	67	75	85	285

2.

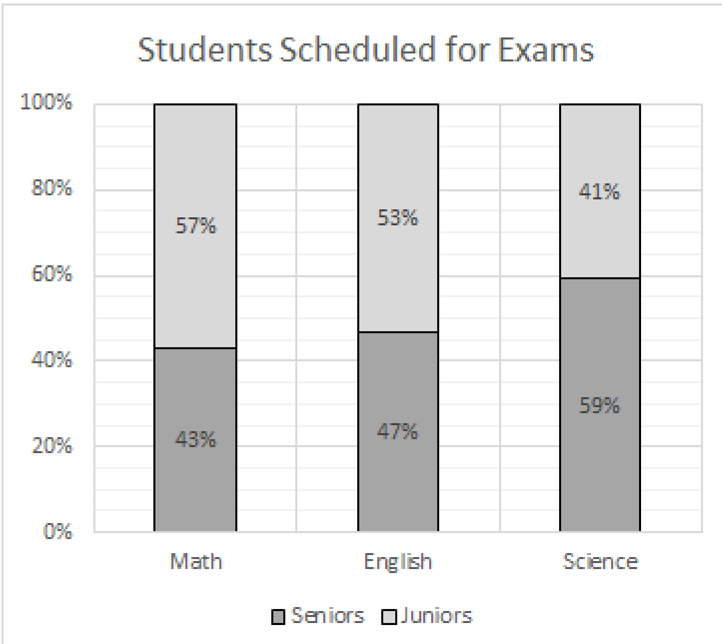
	Plan A	Plan B	Plan C	Plan D	Total
Year 1	3	8	6	4	21
Year 2	10	2	4	3	19
Total	13	10	10	7	40

3.



4.

	Math	English	Science
Seniors	43%	47%	59%
Juniors	57%	53%	41%
Total	100%	100%	100%



5. a)

	White	Blue	Total
Captured	15%	40%	65%
Not Captured	5%	40%	45%
Total	20%	80%	100%

b)

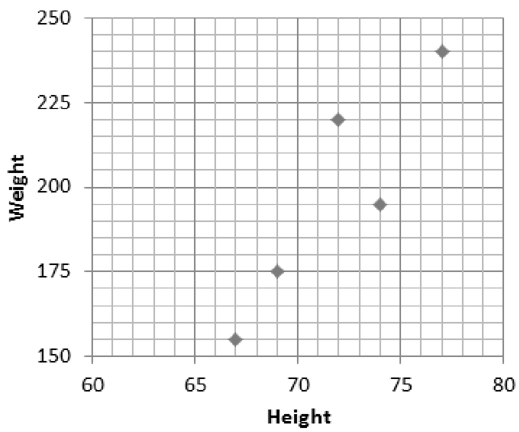
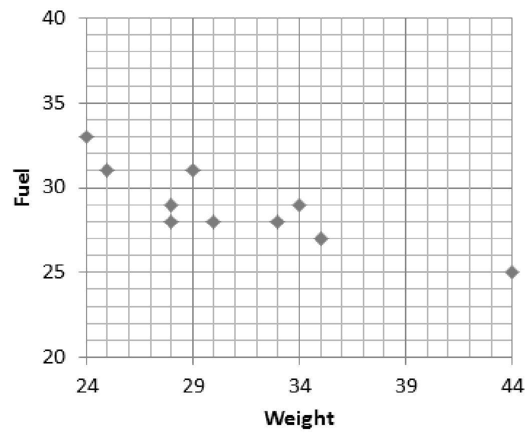
	White	Blue	Total
Captured	150	400	650
Not Captured	50	400	450
Total	200	800	1000

a) Calculate the area of each rectangle.

For example, for the captured white pigeons, $0.75 \times 0.2 = 0.15 = 15\%$.

b) Multiply each percentage by 1000.

4.3 Scatter Plots

<table>
<tr><td>1. (2)</td><td>2. (2)</td></tr>
<tr><td colspan="2">3. (3)</td></tr>
<tr><td>4.
</td><td>5.
</td></tr>
</table>

4.4 Correlation

1. (3)	2. (2)
3. a) positive: children usually gain weight as they age and grow b) negative: as the volume of water increases, the remaining space decreases c) none: shoe size and hair length are unrelated d) positive: more people go to the beach when the temperature is higher	4. a) positive, causal b) positive, not causal; hot temperatures lead to higher sales and more fires c) negative, causal d) positive, not causal; the size and severity of the fire, which results in more firefighters being called e) negative, not causal; the degree of civilization and industrialization f) negative, not causal; higher temperatures may lead to less demand for snow shovels and may also lead to more ocean swimmers, resulting in more opportunity for shark attacks
5. (1)	
6. positive correlation	7. negative correlation
8. negative correlation	9. positive correlation
10. no correlation	11. positive correlation

Chapter 5 Probability

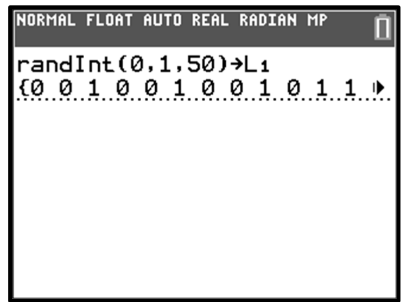
5.1 Theoretical and Empirical Probability

1. $\frac{1}{4}$	2. $\frac{6}{22}=\frac{3}{11}$
3. $\frac{1}{6}$	4. $\frac{23}{29}$
5. $\frac{6}{20}=\frac{3}{10}$	6. $\frac{13}{52}=\frac{1}{4}$
7. $\frac{5}{8}$	8. $P(red)=\frac{30}{90}$ $P(white)=\frac{31}{90}$ $P(blue)=\frac{29}{90}$ White is the most likely to be picked.
9. $\frac{2{,}000}{80{,}000}=\frac{1}{40}$	10. $\frac{8}{20}=\frac{2}{5}$
11. The trials in this case are 100 products per month for 10 months, or 1,000. The empirical probability of a faulty bulb is $\frac{20}{1000}=\frac{1}{50}$.	
12. $\frac{\frac{4}{9}}{\frac{5}{9}}=\frac{4}{5}=4:5$	13. a) $\frac{4}{100}=\frac{1}{25}$ b) $4:96$
14. a) $4:1$ b) $\frac{4}{5}$	15. $6+2+2+2+2+1=15$ a) $\frac{6}{15}=\frac{2}{5}$ b) $2:3$

5.2 Simulation of Random Trials

1. $20\% = \frac{2}{10}$.
   There are 10 numbers from 0 to 9, so any two numbers (such as 0 and 1) can represent the event occurring.

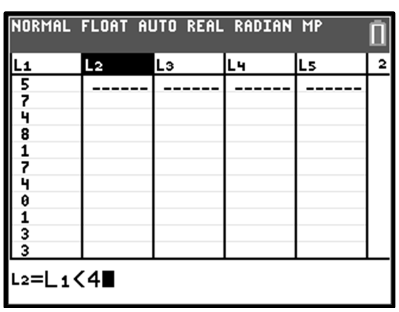
2. Let 0 represent "heads" and 1 represent "tails." Then, enter the function randInt(0,1,50) $\rightarrow L_1$ for the results to be stored in list L1.

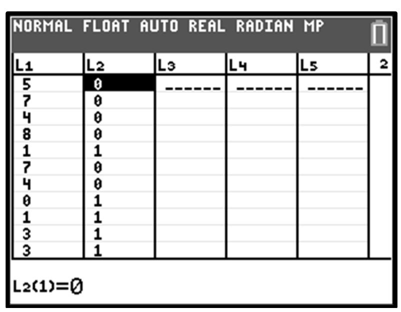


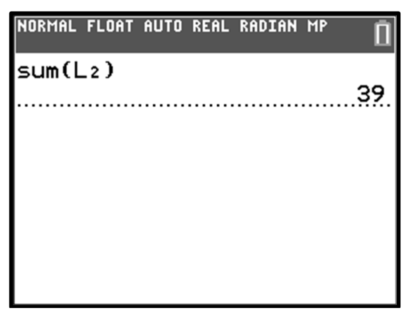
3. The random numbers that are generated may include duplicates.

4. Let each *pair* of digits represent one of our selected numbers.
   For example, if the list contains 927946296491601763 01…,
   then our selected numbers are 92, 79, 46, 29, 64, 91, 60, 17, 63, 01, etc.

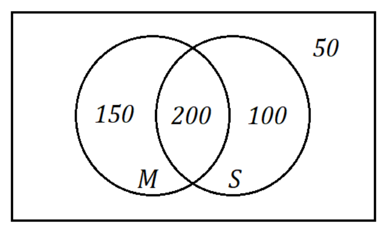
5. Fill L_2 using the test formula $L_1 < 4$. Then, calculate sum(L_2), as shown below.







5.3 Probability Involving And or Or

<table>
<tr>
<td>1. $\frac{6}{11}$</td>
<td>2. $\frac{4}{5}$</td>
</tr>
<tr>
<td>3. $\frac{3}{8}+\frac{2}{8}=\frac{5}{8}$</td>
<td>4. $P(pen\ or\ red) =$
$P(pen)+P(red)-P(red\ pen) =$
$\frac{6}{14}+\frac{9}{14}-\frac{4}{14}=\frac{11}{14}$</td>
</tr>
<tr>
<td>5. a) $P(A\ and\ B) = P(A \cap B) =$
$P(\{5,8\}) = \frac{2}{10} = \frac{1}{5}$
b) $P(A\ or\ B) = P(A \cup B) =$
$P(\{2,3,4,5,7,8,9\}) = \frac{7}{10}$</td>
<td>6.
$P(A\ or\ B) = P(A)+P(B)-P(A\ and\ B)$
$P(A\ or\ B)+P(A\ and\ B) = P(A)+P(B)$
[add $P(A\ and\ B)$ to both sides]
$P(A\ and\ B) = P(A)+P(B)-P(A\ or\ B)$
[subtract $P(A\ or\ B)$ from both sides]</td>
</tr>
<tr>
<td>7. $P(G\ or\ A)$
$= P(G)+P(A)-P(G\ and\ A)$
$= \frac{11}{20}+\frac{9}{20}-\frac{5}{20}=\frac{15}{20}=\frac{3}{4}$</td>
<td>8.


$\frac{50}{500}=\frac{1}{10}$</td>
</tr>
<tr>
<td>9. $P(A) = 0.05$, $P(B) = 0.08$, and
$P(A\ and\ B) = 0.004$
a) not mutually exclusive because
$P(A\ and\ B) \neq 0$
b) $P(A\ or\ B) = 0.05+0.08-0.004 =$
0.126</td>
<td>10. a) $P(C\ or\ B) = P(C)+P(B)$
$= \frac{56}{100}+\frac{26}{100}=\frac{82}{100}=0.82$
b) $P(C\ or\ H)$
$= P(C)+P(H)-P(C\ and\ H)$
$= \frac{56}{100}+\frac{58}{100}-\frac{38}{100}=\frac{76}{100}=0.76$</td>
</tr>
</table>

5.4 Conditional Probability

1.

	Online	TV	Radio	Total
Car Ads	18	16	*11*	45
Insurance Ads	*16*	*25*	14	55
Total	*34*	41	*25*	*100*

2. a) $\frac{10}{50} = \frac{1}{5} = 0.2$ b) $\frac{16}{50} = \frac{8}{25} = 0.32$ c) $\frac{18+16}{50} = \frac{34}{50} = \frac{17}{25} = 0.68$

$OR\ 1 - \frac{8}{25} = \frac{17}{25}\ OR\ 1 - 0.32 = 0.68$

3. a) $P(F) = \frac{72}{240} = \frac{3}{10} = 0.3$ *[from the Total row]*

b) $P(C) = \frac{80}{240} = \frac{1}{3} = 0.\overline{3}$ *[from the Total column]*

c) $P(F|C) = \frac{24}{80} = \frac{3}{10} = 0.3$ *[from the first row]*

d) $P(C|F) = \frac{24}{72} = \frac{1}{3} = 0.\overline{3}$ *[from the first column]*

e) $P(C\ and\ F) = \frac{24}{240} = \frac{1}{10} = 0.1$ *[from the one cell and the grand total]*

4. It is helpful to calculate the totals first:

	Dogs	Cats	Rabbits	Total
Girls	53	72	25	150
Boys	62	28	40	130
Total	115	100	65	280

a) $P(G|R) = \frac{25}{65} = \frac{5}{13} \approx 0.385$

b) $P(R|G) = \frac{25}{150} = \frac{1}{6} = 0.1\overline{6}$

c) $P(B|D\ or\ C) = \frac{62+28}{115+100} = \frac{90}{215} = \frac{18}{43} \approx 0.419$ *[from the first two columns]*

5.

	Female	Male	Total
Positive	50	30	80
Negative	70	60	130
Total	120	90	210

a) $P(P|F) = \frac{50}{120} = \frac{5}{12} = 0.41\overline{6}$

b) $P(N|M) = \frac{60}{90} = \frac{2}{3} = 0.\overline{6}$

c) $P(M|P) = \frac{30}{80} = \frac{3}{8} = 0.375$

6. a)

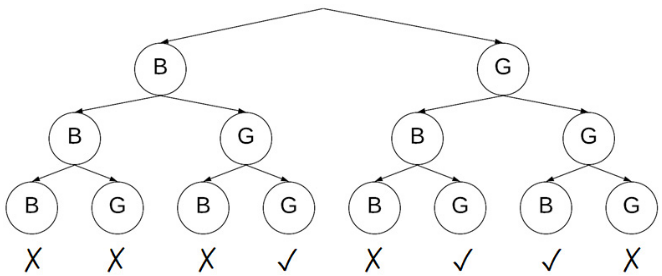
	Defective	Not Defective	Total
Machine A	10	190	200
Machine B	9	291	300
Machine C	5	495	500
Total	24	976	1000

b) Let C represent the item is produced by Machine C and let D represent the item is defective. $P(C|D) = \frac{5}{24} = 0.208\overline{3}$

5.5 Sequence of Events

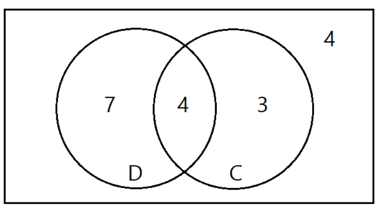
1. $\frac{1}{3} \times \frac{3}{7} = \frac{1}{7}$	2. $\left(\frac{1}{2}\right)^5 = \frac{1}{32}$
3. $\frac{3}{4} \times \frac{1}{2} = \frac{3}{8}$	4. $0.95 \times 0.93 \times 0.98 \approx 87\%$
5. $6 \times 20 = 120$	6 $4 \times 2 = 8$
7. a) $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^{10} = 1{,}024$ b) $\frac{1}{1{,}024}$	8. a) $9 \times 10 \times 10 \times 2 = 1{,}800$ b) $\frac{1}{1{,}800}$

9.



a) $\frac{3}{8}$ (see check marks above) b) $\frac{7}{8}$ (all except the first leaf)

10.



a) 3 b) 4

11. a) dependent (due to genetics) b) dependent (due to growth with age) c) independent d) dependent (without replacement) e) independent

12. $P(at\ least\ one\ blue) =$

$1 - P(red\ or\ white\ on\ all\ 5\ picks) =$

$1 - \left(\frac{2}{3}\right)^5 = 1 - \frac{32}{243} = \frac{211}{243} \approx 87\%$

13. $\frac{1}{20} \times \frac{1}{19} = \frac{1}{380}$

14. $\frac{5}{7} \times \frac{2}{6} = \frac{5}{21}$

15. $\frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$

16. $\frac{3}{4} \times \frac{2}{3} = \frac{1}{2}$

17. $P(M|S) = \frac{P(S\ and\ M)}{P(S)} = \frac{\frac{3}{10}}{\frac{3}{5}} = \frac{15}{30} = \frac{1}{2}$

18. $P(H_1\ and\ H_2) = P(H_1) \cdot P(H_2|H_1) =$

$\frac{13}{52} \times \frac{12}{51} = \frac{156}{2652} = \frac{1}{17}$

19. $P(same\ suit) =$

$P(2Hs\ or\ 2Ds\ or\ 2Cs\ or\ 2Ss) =$

$\frac{1}{17} + \frac{1}{17} + \frac{1}{17} + \frac{1}{17} = \frac{4}{17}$

20.	a) $\frac{10}{25}\times\frac{10}{25}\times\frac{10}{25}\times\frac{10}{25}\times\frac{10}{25}=\left(\frac{10}{25}\right)^5=\frac{32}{3125}$ b) $\frac{10}{25}\times\frac{9}{24}\times\frac{8}{23}\times\frac{7}{22}\times\frac{6}{21}=\frac{6}{1265}$
21.	$P(A')=\frac{3}{4}$, $P(B')=\frac{2}{3}$, and $P(C')=\frac{1}{2}$, so $P(A'\,and\,B'\,and\,C')=\frac{3}{4}\times\frac{2}{3}\times\frac{1}{2}=\frac{6}{24}=\frac{1}{3}$

5.6 Bayes' Theorem

1.	Let A = the patient has arthritis and H = the patient has hay fever. We want to find $P(A\|H)$. $P(A)=0.10$, $P(H)=0.05$, and $P(H\|A)=0.07$ $P(A\|H)=\frac{P(A\ and\ H)}{P(H)}=\frac{P(A)\times P(H\|A)}{P(H)}=\frac{(0.10)(0.07)}{(0.05)}=0.14$
2.	Cards are RR, GG, and RG. There are 6 sides of the cards to choose from. a) Let O_G = the other side is green and T_G = this side is green. $P(O_G\|T_G)=\frac{P(O_G\ and\ T_G)}{P(T_G)}=\frac{\frac{2}{6}}{\frac{3}{6}}=\frac{2}{3}$ b) Either take the complement of the answer to part a), $1-\frac{2}{3}=\frac{1}{3}$ *OR* Let O_R = the other side is red and T_G = this side is green. $P(O_R\|T_G)=\frac{P(O_R\ and\ T_G)}{P(T_G)}=\frac{\frac{1}{6}}{\frac{3}{6}}=\frac{1}{3}$
3.	Let A = deGrom is the starting pitcher and B = the Mets win. Given: $P(A)=0.20$, $P(B\|A)=0.60$, and $P(B\|A')=0.45$ $P(A')=1-P(A)=1-0.20=0.80$ We want to find $P(A\|B)$. $P(A\|B)=\frac{P(A)\cdot P(B\|A)}{P(A)\cdot P(B\|A)+P(A')\cdot P(B\|A')}=\frac{(0.20)(0.60)}{(0.20)(0.60)+(0.80)(0.45)}=0.25$
4.	Let D = the email is detected as spam and S = the email is spam. We want to find $P(S'\|D)$. $P(S)=P(S')=0.5$, $P(D\|S)=0.99$, and $P(D\|S')=0.05$ $P(S'\|D)=\frac{P(S')\cdot P(D\|S')}{P(S')\cdot P(D\|S')+P(S)\cdot P(D\|S)}=\frac{(0.5)(0.05)}{(0.5)(0.05)+(0.5)(0.99)}=\frac{5}{104}\approx 0.048$

5. $P(M|T) = \frac{P(M \text{ } and \text{ } T)}{P(T)}$, where M is the event that you have the meta-gene and T is the event that you test positive.

The numerator, $P(M \text{ } and \text{ } T)$, is the probability that you have the meta-gene and test positive, which can be written as $P(M) \cdot P(T|M)$.

The denominator, $P(T)$, is the probability that you test positive, which can be broken down into the probability that you test positive and *have* the meta-gene OR you test positive and *don't have* the meta-gene, which can be written as $P(M) \cdot P(T|M) + P(M') \cdot P(T|M')$.

$$P(M|T) = \frac{P(M \text{ } and \text{ } T)}{P(T)} = \frac{P(M) \cdot P(T|M)}{P(M) \cdot P(T|M) + P(M') \cdot P(T|M')} = \frac{0.0001 \cdot 0.99}{0.0001 \cdot 0.99 + 0.9999 \cdot 0.01} \approx 0.01$$

Even though you tested positive, you have only a 1% chance of having the meta-gene.

6.1 Discrete Random Variables

1. $P(x > 4) = 0.11 + 0.16 + 0.32 = 0.59$

2. $P(7 \leq x \leq 9)$
$= 0.09 + 0.13 + 0.13 = 0.35$

3. $E(X) = \sum xP(x) = 4.9$

x	$P(x)$	$xP(x)$
2	0.19	0.38
3	0.11	0.33
4	0.11	0.44
5	0.11	0.55
6	0.16	0.96
7	0.32	2.24
Σ	**1.00**	**4.90**

4. $E(X) = \sum xP(x) = 7.61$

x	$P(x)$	$xP(x)$
5	0.30	1.50
6	0.09	0.54
7	0.09	0.63
8	0.13	1.04
9	0.13	1.17
10	0.13	1.30
11	0.13	1.43
Σ	**1.00**	**7.61**

5. $\frac{1}{11}$ The sum of the probabilities must equal 1.

6.

x	$P(x)$	$xP(x)$
1	0.105	0.105
2	0.211	0.422
3	0.281	0.843
4	0.175	0.700
5	0.105	0.525
6	0.070	0.420
7	0.035	0.245
8	0.018	0.144
Σ	**1.000**	**3.404**

$E(X) = \sum xP(x) \approx 3.4$

7.

x	f	$P(x)$	$xP(x)$
0	6	0.03	0.00
1	12	0.06	0.06
2	29	0.15	0.30
3	57	0.30	0.90
4	42	0.22	0.88
5	30	0.16	0.80
6	16	0.08	0.48
Σ	**192**	**1.00**	**3.42**

$E(X) = \sum xP(x) \approx 3.4$

8. $E(X) = \sum xP(x) = -1.2$

You should expect a net loss of $1.20.

Prize (\$)	**Gain (x)**	$P(x)$	$xP(x)$
500	495	0.005	2.475
100	95	0.005	0.475
20	15	0.04	0.6
0	-5	0.95	-4.75
Σ			$\mathbf{-1.2}$

9. $E(X) = \sum xP(x) = -\frac{20}{38} \approx -0.53$

You should expect to lose $0.53.

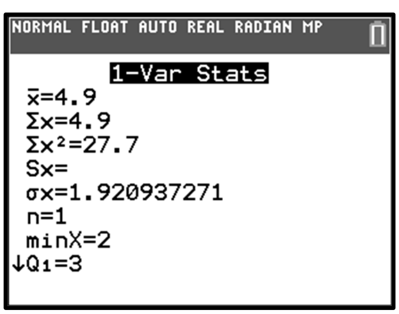
Result	**Net (x)**	$P(x)$	$xP(x)$
Win	350	$\frac{1}{38}$	$\frac{350}{38}$
Loss	-10	$\frac{37}{38}$	$-\frac{370}{38}$
	Σ		$\mathbf{-\frac{20}{38}}$

10.

x	$P(x)$	$xP(x)$	x^2	$x^2P(x)$
2	0.19	0.38	4	0.76
3	0.11	0.33	9	0.99
4	0.11	0.44	16	1.76
5	0.11	0.55	25	2.75
6	0.16	0.96	36	5.76
7	0.32	2.24	49	15.68
Σ	**1.00**	**4.90**		**27.70**

$$\sigma^2 = E(X^2) - E(X)^2$$
$$= \sum x^2P(x) - [\sum xP(x)]^2$$
$$= 27.7 - 4.9^2 = 3.69$$
$$\sigma = \sqrt{3.69} \approx 1.92$$

Check:

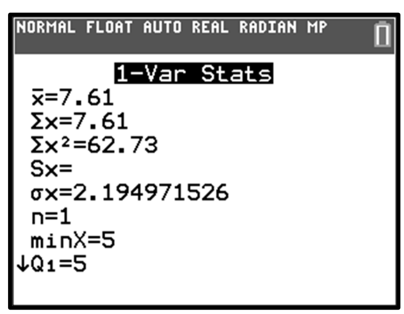


11.

x	$P(x)$	$xP(x)$	x^2	$x^2P(x)$
5	0.30	1.50	25	7.50
6	0.09	0.54	36	3.24
7	0.09	0.63	49	4.41
8	0.13	1.04	64	8.32
9	0.13	1.17	81	10.53
10	0.13	1.30	100	13.00
11	0.13	1.43	121	15.73
Σ	**1.00**	**7.61**		**62.73**

$$\sigma^2 = E(X^2) - E(X)^2$$
$$= \sum x^2P(x) - [\sum xP(x)]^2$$
$$= 62.73 - 7.61^2 = 4.8179 \approx 4.82$$
$$\sigma = \sqrt{4.8179} \approx 2.19$$

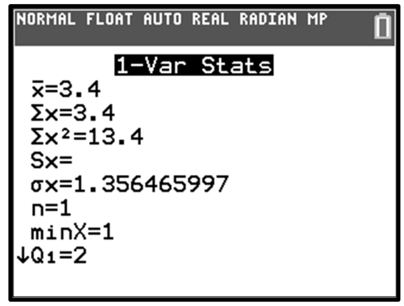
Check:



12.

x	$P(x)$	$xP(x)$	x^2	$x^2P(x)$
1	$\frac{3}{25}$	$\frac{3}{25}$	1	$\frac{3}{25}$
2	$\frac{4}{25}$	$\frac{8}{25}$	4	$\frac{16}{25}$
3	$\frac{5}{25}$	$\frac{15}{25}$	9	$\frac{45}{25}$
4	$\frac{6}{25}$	$\frac{24}{25}$	16	$\frac{96}{25}$
5	$\frac{7}{25}$	$\frac{35}{25}$	25	$\frac{175}{25}$
Σ	**1**	$\mathbf{\frac{85}{25} = \frac{17}{5}}$		$\mathbf{\frac{335}{25} = \frac{67}{5}}$

Check:



$$\sigma^2 = E(X^2) - E(X)^2 = \sum x^2P(x) - \left[\sum xP(x)\right]^2 = \frac{67}{5} - \left(\frac{17}{5}\right)^2 = \frac{46}{25}$$

$$\sigma = \sqrt{\frac{46}{25}} \approx 1.36$$

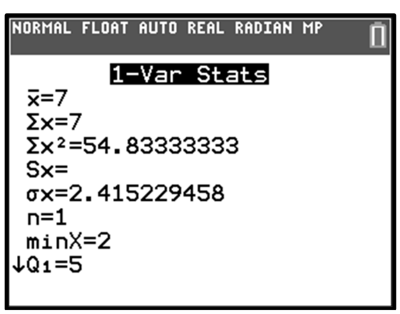
13.

x	2	3	4	5	6	7	8	9	10	11	12	Σ
$P(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	
$xP(x)$	$\frac{2}{36}$	$\frac{6}{36}$	$\frac{12}{36}$	$\frac{20}{36}$	$\frac{30}{36}$	$\frac{42}{36}$	$\frac{40}{36}$	$\frac{36}{36}$	$\frac{30}{36}$	$\frac{22}{36}$	$\frac{12}{36}$	$\frac{252}{36} = 7$
x^2	4	9	16	25	36	49	64	81	100	121	144	
$x^2P(x)$	$\frac{4}{36}$	$\frac{18}{36}$	$\frac{48}{36}$	$\frac{100}{36}$	$\frac{180}{36}$	$\frac{294}{36}$	$\frac{320}{36}$	$\frac{324}{36}$	$\frac{300}{36}$	$\frac{242}{36}$	$\frac{144}{36}$	$\frac{1974}{36} = \frac{329}{6}$

$$\sigma^2 = E(X^2) - E(X)^2 = \sum x^2P(x) - \left[\sum xP(x)\right]^2 = \frac{329}{6} - 7^2 = \frac{329}{6} - \frac{294}{6} = \frac{35}{6}$$

$$\sigma = \sqrt{\frac{35}{6}} \approx 2.415$$

Check:



6.2 Binomial Distributions

1. $P(X=7) = {}_8C_7(0.9)^7(0.1)^1 \approx 0.383$ OR binompdf $(8, 0.9, 7) \approx 0.383$	2. $P(X=1) = {}_5C_1(0.85)^1(0.15)^4 \approx 0.002$ OR binompdf $(5, 0.85, 1) \approx 0.002$
3. $P(X \le 7) = 1 - P(X=8) =$ $1 - 0.9^8 \approx 0.570$ OR $1 -$ binompdf $(8, 0.9, 8) \approx 0.570$ OR binomcdf $(8, 0.9, 7) \approx 0.570$	4. $P(X \ge 8) = 1 - P(X \le 7) =$ $1 -$ binomcdf $(10, 0.85, 7) \approx 0.820$ OR binomcdf $(10, 0.15, 2) \approx 0.820$
5. $P(X \ge 13) = 1 - P(X \le 12) =$ $1 -$ binomcdf $(20, 0.25, 12) \approx 0.00018$ OR binomcdf $(20, 0.75, 7) \approx 0.00018$	6. $P(X \ge 3) = 1 - P(X \le 2) =$ $1 -$ binomcdf $(5, 0.3, 2) \approx 0.163$ OR binomcdf $(5, 0.7, 2) \approx 0.163$
7. binomcdf $(7, 0.2, 5) -$ binomcdf $(7, 0.2, 2) \approx 0.148$	8. binomcdf $(10, 0.85, 7) -$ binomcdf $(10, 0.85, 3) \approx 0.180$
9. $\mu = np = 100 \times 0.85 = 85$ $\sigma = \sqrt{npq} = \sqrt{(100)(0.85)(0.15)} =$ $\sqrt{12.75} \approx 3.571$	10. For $x = 1$, ${}_nC_x = {}_nC_1 = n$. *[Consider the 10 white/gray cards example; there are 10 ways to have exactly one card gray side up.]* So, for $x = 1$, $P = npq^{n-1}$.
11. (2) $p = 0.25$. The graph is skewed right, so $p < 0.5$. Also, $\mu = np$. So, if $p = 0.1$, then the mean would be $(20)(0.1) = 2$, which doesn't match the graph. For $p = 0.25$, the mean is $(20)(0.25) = 5$, which matches the graph.	

12. a) $\mu = np = (25)(0.4) = 10$,
$\sigma^2 = npq = (25)(0.4)(0.6) = 6$,
$\sigma = \sqrt{6} \approx 2.45$
b) $\mu = (100)(0.75) = 75$,
$\sigma^2 = (100)(0.75)(0.25) = 18.75$,
$\sigma = \sqrt{18.75} \approx 4.33$
c) $\mu = (320)(0.92) = 294.4$,
$\sigma^2 = (320)(0.92)(0.08) = 23.552$,
$\sigma = \sqrt{23.552} \approx 4.85$

13. To calculate each $P(x)$, use binompdf$(6, 0.63, x)$.

$\boldsymbol{x}$	$\boldsymbol{P(x)}$
0	0.003
1	0.026
2	0.112
3	0.253
4	0.323
5	0.220
6	0.063

$\mu = (6)(0.63) = 3.78$
$\sigma^2 = (6)(0.63)(0.37) = 1.3986$
$\sigma = \sqrt{1.3986} \approx 1.18$

6.3 Geometric Distributions

1. $P = q^{x-1}p = (0.25)^1(0.75) = 0.1875$
OR geometpdf$(0.75, 2) = 0.1875$

2. $P = q^{x-1}p = (0.55)^4(0.45) \approx 0.04$
OR geometpdf$(0.45, 5) \approx 0.04$

3. (1) $P = \left(\frac{3}{4}\right)^4 \left(\frac{1}{4}\right) \approx 0.0791$ *OR* geometpdf$(1/4, 5) \approx 0.0791$
(2) $P = \left(\frac{5}{6}\right)^4 \left(\frac{1}{6}\right) \approx 0.0804$ *OR* geometpdf$(1/6, 5) \approx 0.0804$
Answer is (2)

4. a) $P = (0.3)^2(0.7) = 0.063$ *OR* geometpdf$(0.7, 3) = 0.063$
b) $P = (0.3)^0(0.7) + (0.3)^1(0.7) + (0.3)^2(0.7) = (1 + 0.3 + 0.3^2)(0.7) = 0.973$
OR geometcdf$(0.7, 3) = 0.973$
c) $P = 1 - 0.973 = 0.027$ *[complement of the solution to part b]*

5. a) $P = (0.323)^1(0.677) = 0.219$ *OR* geometpdf$(0.677, 2) = 0.219$
b) $P = (0.323)^0(0.677) + (0.323)^1(0.677) = (1 + 0.323)(0.677) = 0.896$
OR geometcdf$(0.677, 2) = 0.896$
c) $P = 1 - 0.896 = 0.104$ *[complement of the solution to part b]*

6. a) $\mu = \frac{1}{p} = \frac{1}{0.01} = 100$ $\sigma^2 = \frac{q}{p^2} = \frac{0.99}{0.01^2} = 9900$ $\sigma = \sqrt{9900} \approx 99.5$
b) 100

6.4 Poisson Distributions

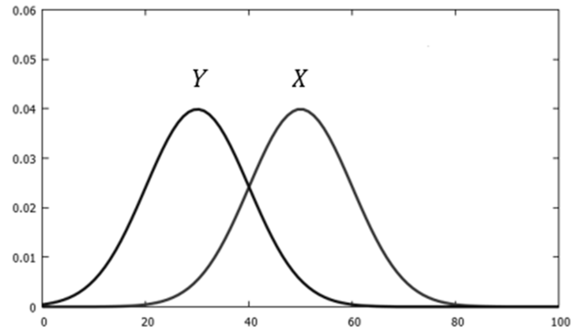
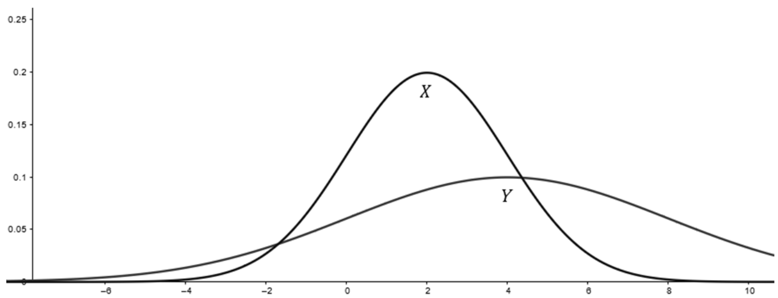
1. $P = \frac{\mu^x e^{-\mu}}{x!} = \frac{6^2 e^{-6}}{2!} \approx 0.04$ OR poissonpdf(6, 2) ≈ 0.04	2. $P = \frac{\mu^x e^{-\mu}}{x!} = \frac{9.8^{10} e^{-9.8}}{10!} \approx 0.12$ OR poissonpdf(9.8, 10) ≈ 0.12
3. a) poissonpdf(8, 7) ≈ 0.140 b) poissoncdf(8,5) ≈ 0.191 c) $1 - 0.191 = 0.809$	
4. a) $P = \frac{5^4 e^{-5}}{4!} = 0.175$ OR poissonpdf(5, 4) $= 0.175$ b) $P = \frac{5^0 e^{-5}}{0!} + \frac{5^1 e^{-5}}{1!} + \frac{5^2 e^{-5}}{2!} + \frac{5^3 e^{-5}}{3!} = 0.007 + 0.034 + 0.084 + 0.140 = 0.265$ OR poissoncdf(5, 3) $= 0.265$ c) $p = 1 - (0.265 + 0.175) = 0.560$ *[subtract the solutions to parts a and b from 1]* OR $1 -$ poissoncdf(5, 4) $= 0.560$	
5. a) $P = \frac{8^8 e^{-8}}{8!} \approx 0.14$ OR poissonpdf(8, 8) ≈ 0.14 b) $P = \frac{8^0 e^{-8}}{0!} = e^{-8} \approx 0.00$ OR poissonpdf(8, 0) ≈ 0.00 c) $P = \frac{8^7 e^{-8}}{7!} + \frac{8^8 e^{-8}}{8!} + \frac{8^9 e^{-8}}{9!} = 0.14 + 0.14 + 0.12 = 0.40$ OR poissonpdf(8, 7) + poissonpdf(8, 8) + poissonpdf(8, 9) $= 0.40$ OR poissoncdf(8, 9) $-$ poissoncdf(8, 6) $= 0.40$	
6. a) $\sigma^2 = \mu = 42.5$ $\sigma = \sqrt{42.5} = 6.5$ b) poissonpdf(42.5, 40) $= 0.058$ c) poissoncdf(42.5, 39) $= 0.330$ d) $1 -$ poissoncdf(42.5, 50) $= 0.112$	

7.1 Continuous Random Variables

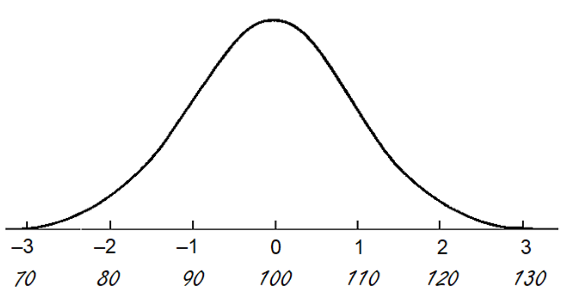
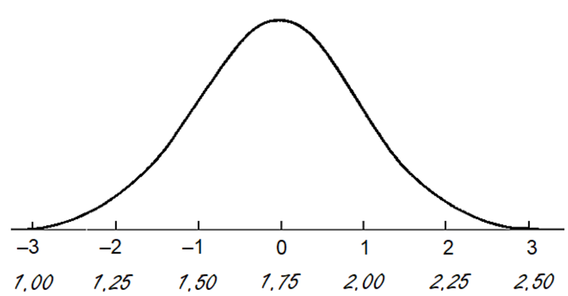
1. a) $\mu = \frac{0+23}{2} = 11.5$ $s = \sqrt{\frac{(23-0)^2}{12}} \approx 6.64$ b) $h = \frac{1}{23}$ c) $A = (18-2)\left(\frac{1}{23}\right) = \frac{16}{23} \approx 0.696$	2. a) $\mu = \frac{8+23}{2} = 15.5$ $s = \sqrt{\frac{(23-8)^2}{12}} \approx 4.33$ b) $h = \frac{1}{15}$ c) $A = (23-12)\left(\frac{1}{15}\right) = \frac{11}{15} \approx 0.733$
3. $A = (7.5-2.5)\left(\frac{1}{8}\right) = \frac{5}{8} = 0.625$	4. $A = (5-2)\left(\frac{1}{7}\right) = \frac{3}{7} = 0.429$
5. a) 0.10 b) 0.53 c) $0.53 + 0.30 + 0.10 = 0.93$ d) $0.30 + 0.10 = 0.40$ e) $1 - 0.93 = 0.07$ f) 0	

7.2 Transform Random Variables

1. a) $\mu_Y = a + b\mu_X = 1000 + 1.05(65000) = \$69.250$ b) $\sigma_Y = \|b\|\sigma_X = 1.05(10000) = \$10,500$ c) $\sigma_Y^2 = b^2\sigma_X^2 = (1.05)^2(10000)^2 = 110,250,000$ $OR \quad (10500)^2 = 110,250,000$ Adding \$1,000 does not affect σ or σ^2.	
2. $Y = 2X$ $\mu_Y = 2(1.7) = 3.4$ $\sigma_Y = 2(0.67) = 1.34$ $\sigma_Y^2 = 2^2(0.67)^2 = 1.34^2 = 1.7956$	3. $Y = 60X - 20$ $\mu_Y = 60(1.8) - 20 = 88$ $\sigma_Y = 60(1.08) = 64.8$ $\sigma_Y^2 = 60^2(1.08)^2 = 64.8^2 = 4199.04$

4.	The mean of *Y* is 30 and the standard deviation (width of the curve) remains at 10. Since the curve is symmetric, the mean is at the center (or peak) of the curve. 
5.	Both the mean and standard deviation are multiplied by 2, so (a) the graph of *Y* is centered on $\mu_Y = 4$, and (b) because $\sigma_Y = 4$, the distribution is twice as wide (and therefore half as tall) as *X*. 

7.3 Normal Density Curves

1.	2.
	

3. Use the normalpdf function: a) $\text{normalpdf}(0,0,1) = 0.399$ b) $\text{normalpdf}(1,0,1) = 0.242$ c) $\text{normalpdf}(2,0,1) = 0.054$ d) $\text{normalpdf}(3,0,1) = 0.004$ e) 0.004 (symmetric, so same as $x = 3$)	4. a) 0.4 b) 0.2 c) 0.1 d) 0.8 e) As σ increases by a factor of k, the center height decreases by a factor of $\frac{1}{k}$.
5. $(-1, 0.242)$ and $(1, 0.242)$	

7.4 <u>Empirical Rule</u>

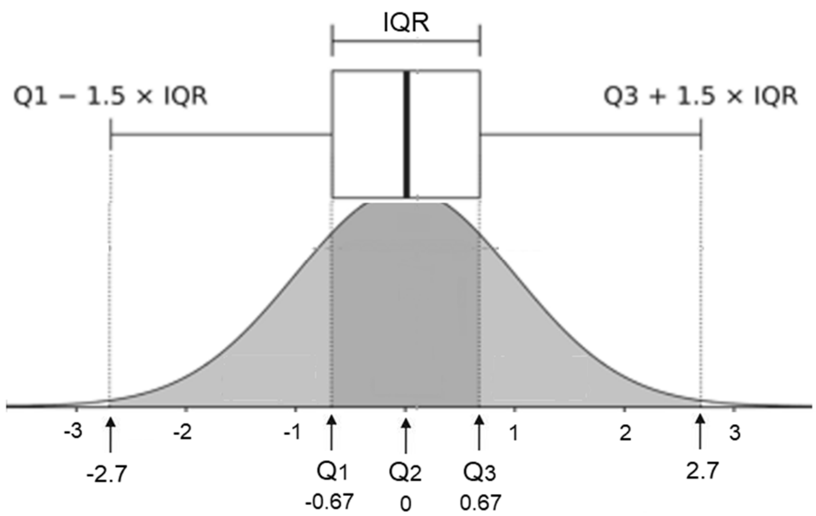
1. The interval from 115 to 125 is 1 standard deviation from the mean, which is about 68% of the data.	2. 95% of the data is within 2 standard deviations from the mean, so this is the interval between $66 - 2(4) = 58$ and $66 + 2(4) = 74$ inches.
3. $\frac{1}{2}(80 - 50) = 15$	4. $\frac{1}{4}(92 - 78) = 3.5$
5. $\frac{1}{4}(69 - 63) = 1.5$	6. $SD = \frac{1}{3}(81 - 57) = 8$ $57 + 8 = 65$ Mean = 65 (check: $81 - 2(8) = 65$ ✓)
7. From $56 - 2(5)$ to $56 + 2(5)$, or between 46 and 66.	8. Interval is within 1 SD of the mean, representing about 68% of the homes. $75 \times 68\% = 51$ homes.

7.5 Areas Under Normal Curves

1. normalcdf$(-1\text{E}99, -1.35, 0, 1) \approx 0.089$	2. normalcdf$(1.48, 1\text{E}99, 0, 1) \approx 0.069$
3. a) normalcdf$(0, 1.02, 0, 1) \approx 0.3461$ b) normalcdf$(1.02, 1\text{E}99, 0, 1) \approx 0.1539$ $OR\ 0.50 - 0.3461 = 0.1539$ c) normalcdf$(-1\text{E}99, 1.02, 0, 1) \approx 0.8461$ $OR\ 0.50 + 0.3461 = 0.8461$ $OR\ 1 - 0.1539 = 0.8461$	4. a) normalcdf$(-2.3, 1.8, 0, 1) \approx 0.9533$ b) normalcdf$(-1\text{E}99, -2.3, 0, 1) \approx 0.0107$ c) normalcdf$(1.8, 1\text{E}99, 0, 1) \approx 0.0359$ The three areas above add up to 1.
5. 1 – normalcdf$(-1.25, 1.25, 0, 1) \approx 0.21$ OR normalcdf$(-1\text{E}99, -1.25, 0, 1) +$ normalcdf$(1.25, 1\text{E}99, 0, 1) \approx 0.21$	6. The area below $-z$ is also 12%. So, the area between $-z$ and z is $A = 100\% - 2(12\%) = 76\%$.
7. normalcdf$(60, 73, 65, 5) \approx 0.787$	8. normalcdf$(620, 1\text{E}99, 500, 100) \approx 0.115$
9. normalcdf$(54.3, 63.5, 54.3, 4.6) \approx 48\%$	10. normalcdf$(74, 82, 80, 4) \approx 0.62$
11. normalcdf$(80, 1\text{E}99, 72, 9) \approx 19\%$	12. normalcdf$(12.5, 1\text{E}99, 11, 1.5) \approx 0.16$
13. normalcdf$(3, 1\text{E}99, 2.75, 0.42) \approx 0.28$	14. a) normalcdf$(90, 1\text{E}99, 75, 8) \approx 3.04\%$ b) normalcdf$(80, 90, 75, 8) \approx 23.56\%$ c) normalcdf$(-1\text{E}99, 60, 75, 8) \approx 3.04\%$
15. normalcdf $(42, 1\text{E}99, 35, 2.8) \approx 0.62\%$ $0.62\% \times 3000 \approx 19$	16. normalcdf$(550, 1\text{E}99, 510, 110) \approx .358$ $0.358 \times 1000 = 358$
17. $z =$ invNorm$(0.11, 0, 1) \approx -1.23$	18. $z =$ invNorm$(0.95, 0, 1) \approx 1.64$
19. $P_{90} =$ invNorm$(0.9, 65, 10) \approx 77.8$	
20. a) $P_{25} =$ invNorm$(0.25, 60, 12) \approx 52$ b) $P_{75} =$ invNorm$(0.75, 60, 12) \approx 68$ c) $IQR = Q_3 - Q_1 = P_{75} - P_{25}$ $= 68 - 52 = 16$	21. a) invNorm$(0.9, 75, 8) \approx 85$ b) invNorm$(0.65, 75, 8) \approx 78$
22. $P(z < 0.78) = P(z \geq 0.22)$ $z =$ invNorm$(0.22, 0, 1) \approx -0.77$	23. $z =$ invNorm$(0.25, 0, 1) \approx -0.6745$ $x = 500 + (-0.6745)(24) \approx 484$

24. a) $z = \text{invNorm}(0.85, 0, 1) \approx 1.04$ b) $x = 100 + (1.04)(15) \approx 116$ c) $x = \text{invNorm}(0.85, 100, 15) \approx 116$	25. a) $0.50 \pm \frac{1}{2}(0.80) \rightarrow 0.1, 0.9$ $\text{invNorm}(0.1, 0, 1) \approx -1.28$ $\text{invNorm}(0.9, 0, 1) \approx 1.28$ b) $x = 100 \pm (1.28)(15) \approx (80.8, 119.2)$ c) $x = \text{invNorm}(0.1, 100, 15) \approx 80.8$ $x = \text{invNorm}(0.9, 100, 15) \approx 119.2$
26. $x = \text{invNorm}(0.05, 71, 7.9) \approx 58$	27. $x = \text{invNorm}(0.92, 450, 13.6) \approx 469$ Yes, a score of 475 is high enough.
28. $z = \text{invNorm}\ (0.1, 0, 1) \approx -1.282$ $z = \frac{x - \mu}{\sigma}$ $-1.282 = \frac{40 - \mu}{2.5}$ $-1.282(2.5) - 40 = -\mu$ $\mu = 1.282(2.5) + 40 \approx 43.2$	29. $z = \text{invNorm}\ (0.667, 0, 1) \approx 0.432$ $z = \frac{x - \mu}{\sigma}$ $0.432 = \frac{110 - 100}{\sigma}$ $\sigma = \frac{10}{0.432} \approx 23$
30. a) $z = \text{invNorm}\ (0.05, 0, 1) \approx -1.645$ $-1.645 = \frac{1.2 - \mu}{25.4}$ $-1.645(25.4) - 1.2 = -\mu$ $\mu = 1.645(25.4) + 1.2 \approx 43.0$ b) $x = \text{invNorm}(0.95, 43, 25.4) \approx 84.8$ *OR* $43 - 1.2 = 41.8$ and $43 + 41.8 = 84.8$	31. a) left and right tails are 12.5% each $z = \text{invNorm}\ (0.125, 0, 1) \approx -1.15$ $-1.15 = \frac{50 - 100}{\sigma}$ $\sigma = \frac{50}{1.15} \approx 43.5$ b) for left and right tails of 25% each, $\text{invNorm}(0.25, 100, 43.5) \approx 71$ and $\text{invNorm}(0.75, 100, 43.5) \approx 129$

32. a) The area between Q_1 and Q_3 represents the middle 50% of the data, from 25% below to 25% above the mean. So, calculate invNorm$(0.75, 0, 1) = 0.67$.
The z-scores are –0.67 and 0.67 for Q_1 and Q_3, respectively.
b) For a standard normal, $IQR = Q_3 - Q_1 = 0.67 - (-0.67) = 1.34$.
$Q_3 + 1.5 \times IQR = 0.67 + (1.5)(1.34) \approx 2.7$.
$z_1 = -2.7$ and $z_2 = 2.7$.



7.6 Sampling Distributions

1. a) normalcdf $(-1\text{E}99, 33, 30, 3) \approx 0.84$
OR $z = \frac{x-\mu}{\sigma} = \frac{33-30}{3} = 1.00$, and normalcdf $(-1\text{E}99, 1, 0, 1) \approx 0.84$
[or the Standard Normal CP table yields 0.8413*]*
b) $\mu_{\bar{x}} = \mu = 30$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{36}} = 0.5$
normalcdf $(-1\text{E}99, 31, 30, 0.5) \approx 0.98$
OR $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{31-30}{0.5} = 2.00$, and normalcdf $(-1\text{E}99, 2, 0, 1) \approx 0.98$
[or the Standard Normal CP table yields 0.9772*]*
c) The probability that the sample mean of 36 students is less than 31 is greater than the probability that a single student is less than 33. The larger the sample size, the smaller the standard deviation of the sample mean, which means more of the data is closer to the mean.

2. a) normalcdf $(80, 1\text{E}99, 67, 7.9) \approx 0.05$ b) $\mu_{\bar{x}} = \mu = 67$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{7.9}{\sqrt{10}} \approx 2.5$, and normalcdf $(80, 1\text{E}99, 67, 2.5) \approx 0.00$ c) While there is a 5% probability of a single student having a score of at least 80, it would be *very unusual* for a random sample of 10 students to have a mean score of at least 80, considering the given population parameters.	
3. $\mu_{\bar{x}} = 400$ and $\sigma_{\bar{x}} = \frac{70}{\sqrt{100}} = 7 \to E$ normalcdf $(-1\text{E}99, 385, 400, E) \approx 0.016$ *OR* $z = \frac{385-400}{E} \approx -2.14 \to Z$ normalcdf $(-1\text{E}99, Z, 0, 1) \approx 0.016$ *[or Standard Normal CP table for* -2.14*]*	4. $\mu_{\bar{x}} = 28.3$; $\sigma_{\bar{x}} = \frac{2.3}{\sqrt{10}} = 0.727 \to E$ normalcdf $(-1\text{E}99, 27, 28.3, E) \approx 0.037$ *OR* $z = \frac{27-28.3}{E} \approx -1.79 \to Z$ normalcdf $(-1\text{E}99, Z, 0, 1) \approx 0.037$ *[or Standard Normal CP table for* -1.79*]*
5. $\mu_{\bar{x}} = 25000$; $\sigma_{\bar{x}} = \frac{1600}{\sqrt{64}} = 200 \to E$ normalcdf $(24600, 1\text{E}99, 25000, E) \approx 0.977$ *OR* $z = \frac{24600-25000}{E} = -2 \to Z$ normalcdf $(Z, 1\text{E}99, 0, 1) \approx 0.977$ *[On the Standard Normal CP table,* -2 *yields* 0.0228*, so calculate the area to the right as* $1 - 0.0228 \approx 0.977$.*]*	6. $\mu_{\bar{x}} = 302.8$; $\sigma_{\bar{x}} = \frac{56}{\sqrt{26}} \approx 10.98 \to E$ normalcdf $(-1e99, 324, 302.8, E) \approx 0.973$ *OR* $z = \frac{324-302.8}{E} = 1.93 \to Z$ normalcdf $(-1\text{E}99, Z, 0, 1) \approx 0.973$
7. $\mu_{\bar{x}} = 94.7$; $\sigma_{\bar{x}} = \frac{6.9}{\sqrt{32}} \approx 1.220 \to E$ normalcdf $(92, 1e99, 94.7, E) \approx 0.987$ *OR* $z = \frac{92-94.7}{E} = -2.21 \to Z$ normalcdf $(Z, 1\text{E}99, 0, 1) \approx 0.987$	8. $\mu_{\bar{x}} = 7.2$; $\sigma_{\bar{x}} = \frac{2.1}{\sqrt{12}} = 0.606 \to E$ normalcdf $(6, 8, 7.2, E) \approx 0.883$ *OR* $z_a = \frac{6-7.2}{E} \approx -1.98 \to A$ $z_b = \frac{8-7.2}{E} \approx 1.32 \to B$ normalcdf $(A, B, 0, 1) \approx 0.883$ *[On the Standard Normal CP table,* -1.98 *yields* 0.0233 *and* 1.32 *yields* 0.9066*, so the area between is* $0.9066 - 0.0233 \approx 0.883$.*]*

9. $\mu_{\bar{x}} = 202; \sigma_{\bar{x}} = \frac{14}{\sqrt{36}} = \frac{7}{3} \to E$ $202 \pm 4 = [198, 206]$ normalcdf $(198, 206, 202, E) \approx 0.914$ *OR* $z_a = \frac{198-202}{E} = -1.71 \to A$ $z_b = \frac{206-202}{E} = 1.71 \to B$ normalcdf $(A, B, 0, 1) \approx 0.914$ *[On the Standard Normal CP table,* -1.71 *yields* 0.0436 *and* 1.71 *yields* 0.9564*, so the area between is* $0.9564 - 0.0436 \approx 0.913$*. This difference is due to rounding z-scores.]*	10. $\mu_{\bar{x}} = 68; \sigma_{\bar{x}} = \frac{3}{\sqrt{9}} = 1 \to E$ $68 \pm 1 = [67, 69]$ normalcdf $(67, 69, 68, E) \approx 0.683$ *OR* $z_a = \frac{67-68}{E} = -1.00 \to A$ $z_b = \frac{69-68}{E} = 1.00 \to B$ normalcdf $(A, B, 0, 1) \approx 0.683$ *[On the Standard Normal CP table,* -1.00 *yields* 0.1587 *and* 1.00 *yields* 0.8413*, so the area between is* $0.8413 - 0.1587 \approx 0.683$*.]*
11. Even though the population is skewed, CLT applies since $n \geq 30$, so the distribution of $\bar{x}$ will be normal. $\mu_{\bar{x}} = 6; \sigma_{\bar{x}} = \frac{7}{\sqrt{49}} = 1$ normalcdf $(8, 1\text{E}99, 6, 1) \approx 0.023$	12. Even though the population is skewed, CLT applies since $n \geq 30$, so the distribution of $\bar{x}$ will be normal. $\mu_{\bar{x}} = 70; \sigma_{\bar{x}} = \frac{10}{\sqrt{64}} = 1.25$ normalcdf $(65, 1\text{E}99, 70, 1.25) \approx 1.000$
13. $\mu_{\bar{x}} = 43.5; \sigma_{\bar{x}} = \frac{7.2}{\sqrt{25}} = 1.44 \to E$ invNorm $(0.94, 43.5, E) \approx 45.7$ 94% should be less than 46 years old	14. $\mu_{\bar{x}} = 111.9; \sigma_{\bar{x}} = \frac{5.9}{\sqrt{30}} \approx 1.077 \to E$ normalcdf $(110, 1\text{E}99, 111.9, E) \approx 0.961 \to A$ $A(200) \approx 192$
15. CLT applies since $(117)(0.043) \geq 5$. $\mu_{\hat{p}} = p = 0.043$ and $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.043)(0.957)}{117}} \approx 0.0188 \to E$ normalcdf $(.09, 1\text{E}99, .043, E) \approx 0.006$ *OR* $z = \frac{.09-.043}{0.0188} = 2.5 \to Z$ normalcdf $(Z, 1\text{E}99, 0, 1) \approx 0.006$	16. CLT applies since $(100)(0.23) \geq 5$. $\mu_{\hat{p}} = p = 0.77$ and $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.77)(0.23)}{100}} \approx 0.0421 \to E$ normalcdf $(-1\text{E}99, 0.75, .77, E) \approx 0.317$ *OR* $z = \frac{.75-.77}{E} = -0.475 \to Z$ normalcdf $(-1\text{E}99, Z, 0, 1) \approx 0.317$

7.7 Normal Approximations

1. a) area to the right of 4.5
   b) area to the left of 8.5
   c) area between 4.5 and 8.5

2. a) area to the right of 5.5
   b) area to the left of 7.5
   c) area between 5.5 and 7.5

3. $n = 618$ and $p = 0.1$
   $\mu = np = 618(0.1) = 61.8$
   $\sigma = \sqrt{npq} = \sqrt{618(0.1)(0.9)} = 7.5$
   $P(x < 50) \rightarrow$ left of 49.5
   $\text{normalcdf}(-1\text{E}99, 49.5, 61.8, 7.5) \approx 0.051$
   $OR\ \ z = \frac{49.5-61.8}{7.5} = -1.64$
   $\text{normalcdf}(-1\text{E}99, -1.64, 0, 1) \approx 0.051$

4. $n = 600$ and $p = 0.366$
   $\mu = 600(0.366) = 219.6$
   $\sigma = \sqrt{600(0.366)(0.634)} = 11.8$
   $P(x \geq 200) \rightarrow$ right of 199.5
   $\text{normalcdf}(199.5, 1\text{E}99, 219.6, 11.8) \approx 0.96$
   $OR\ \ z = \frac{199.5-219.6}{11.8} = -1.70$
   $\text{normalcdf}(-1.7, 1\text{E}99, 0, 1) \approx 0.96$

5. $n = 160$ and $p = 0.1$
   $\mu = 160(0.1) = 16$
   $\sigma = \sqrt{160(0.1)(0.9)} \approx 3.795$
   $P(x = 18) \rightarrow$ between 17.5 and 18.5
   $\text{normalcdf}\ (17.5, 18.5, 16, 3.795) \approx 0.09$

6. $n = 120$ and $p = 0.8$
   $\mu = 120(0.8) = 96$
   $\sigma = \sqrt{120(0.8)(0.2)} \approx 4.382$
   $P(86 \leq x \leq 98) \rightarrow$ between 85.5 and 98.5
   $\text{normalcdf}\ (85.5, 98.5, 96, 4.382) \approx 0.71$

7. $n = 180$ and $p = \frac{1}{6}$
   $\mu = np = 180\left(\frac{1}{6}\right) = 30$
   $\sigma = \sqrt{npq} = \sqrt{180\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)} = 5$
   a) $\text{normalcdf}\ (19.5, 40.5, 30, 5) \approx 0.964$
   b) $\text{normalcdf}\ (34.5, 1\text{E}99, 30, 5) \approx 0.184$
   c) $\text{normalcdf}\ (34.5, 35.5, 30, 5) \approx 0.048$
   d) ${}_{180}C_{35} \cdot \left(\frac{1}{6}\right)^{35} \left(\frac{5}{6}\right)^{145} =$
   $\text{binompdf}\ (180, 1/6, 35) \approx 0.046$
   The actual probability is about 0.002 less.

8. $n = 100$ and $p = 0.8$
   $\mu = np = 100(0.8) = 80$
   $\sigma = \sqrt{npq} = \sqrt{100(0.8)(0.2)} = 4$
   a) $\text{normalcdf}\ (90.5, 1\text{E}99, 80, 4) \approx 0.004$
   b) $\text{normalcdf}\ (-1\text{E}99, 74.5, 80, 4) \approx 0.085$
   c) $\text{normalcdf}\ (75.5, 89.5, 80, 4) \approx 0.861$
   d) $0.004 + 0.085 + 0.861 = 0.95$
   No, it does not equal 1 because it does not include $P(x = 75)$ or $P(x = 90)$.
   $\text{normalcdf}\ (74.5, 75.5, 80, 4) \approx 0.046$ and
   $\text{normalcdf}\ (89.5, 90.5, 80, 4) \approx 0.004$,
   which add up to the other 0.05.

9. $\sigma = \sqrt{42.5} \rightarrow S$
   a) $\text{normalcdf}(39.5, 40.5, 42.5, S) \approx 0.0568$ $\quad$ $\text{poissonpdf}(42.5, 40) = 0.0584$
   b) $\text{normalcdf}(-1\text{E}99, 39.5, 42.5, S) \approx 0.3226$ $\quad$ $\text{poissoncdf}(42.5, 39) = 0.3300$
   c) $\text{normalcdf}(50.5, 1\text{E}99, 42.5, S) \approx 0.1099$ $\quad$ $1 - \text{poissoncdf}(42.5, 50) = 0.1119$

Chapter 8 Confidence Intervals

8.1 Critical Value and Margin of Error

1. $ME = CV \cdot SE = (1.28)(5) = 6.4$	2. $75 \pm 2.5 \rightarrow (72.5, 77.5)$
3. $50 \pm 8 \rightarrow (42, 58)$	4. $15 \pm 1.2 \rightarrow (13.8, 16.2)$
5. (2) higher confidence levels lead to wider intervals	6. The point estimate is in the middle of the interval, so $\frac{22.5 + 31.5}{2} = \frac{54}{2} = 27$. $ME = 31.5 - 27 = 4.5$
7. $z^* = 1.96$	8. $z^* = 1.65$
9. ZInterval$(1, 0, 1, 0.98) \rightarrow 2.33$ *OR* $98\% + \left(\frac{100-98}{2}\right)\% = 99\%$ invNorm $(0.99, 0, 1) \approx 2.33$	10. ZInterval$(1, 0, 1, 0.75) \rightarrow 1.15$ *OR* $75\% + \left(\frac{100-75}{2}\right)\% = 87.5\%$ invNorm $(0.875, 0, 1) \approx 1.15$

8.2 CI for Proportions

1. a) $\hat{p} = \frac{245}{350} = 0.7$ b) $\hat{q} = 1 - \hat{p} = 0.3$ c) $SE = \sqrt{\frac{(0.7)(0.3)}{350}} \approx 0.024$	2. a) $\hat{p} = \frac{290}{500} = 0.58$ b) $\hat{q} = 1 - \hat{p} = 0.42$ c) $SE = \sqrt{\frac{(0.58)(0.42)}{500}} \approx 0.022$
3. a) $\hat{p} = \frac{32}{80} = 0.4$ b) $SE = \sqrt{\frac{(0.4)(0.6)}{80}} \approx 0.055$	4. a) $\hat{p} = \frac{102}{121} \approx 0.84$ b) $SE = \sqrt{\frac{(0.84)(0.16)}{121}} \approx 0.033$

5. a) $z^* = 1.96$

b) $\hat{p} = \frac{9}{64} \approx 0.141$

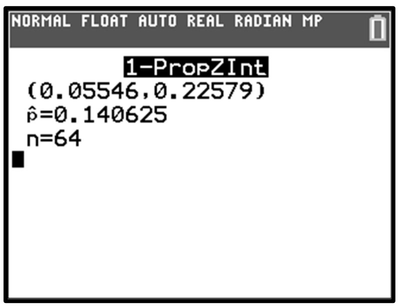
c) $SE = \sqrt{\frac{(0.141)(0.859)}{64}} \approx 0.044$

d) $ME = (1.96)(0.044) \approx 0.085$

e) $0.141 - 0.085 < p < 0.141 + 0.085$

$0.055 < p < 0.226$

f)



6. a) $z^* = 1.65$

b) $\hat{p} = \frac{38}{169} \approx 0.225$

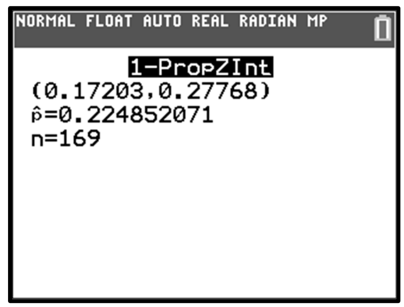
c) $SE = \sqrt{\frac{(0.225)(0.775)}{169}} \approx 0.032$

d) $ME = (1.65)(0.032) \approx 0.053$

e) $0.225 - 0.053 < p < 0.225 + 0.053$

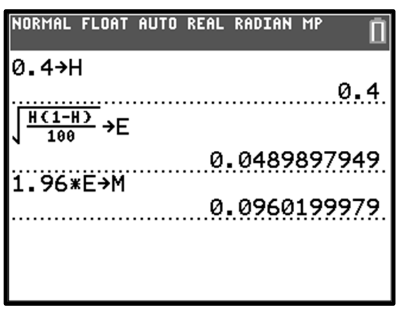
$0.172 < p < 0.278$

f)



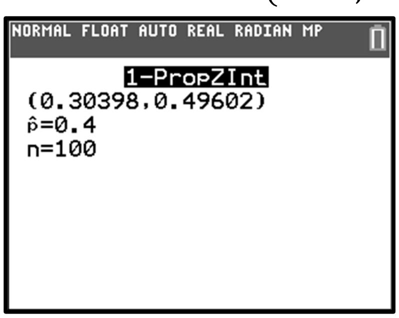
7. $SE = \sqrt{\frac{(0.4)(0.6)}{100}} \approx 0.049$

$ME = (1.96)(0.049) \approx 0.096$



$0.4 - 0.096 < p < 0.4 + 0.096$

Population proportion should fall within the interval $(0.304, 0.496)$.



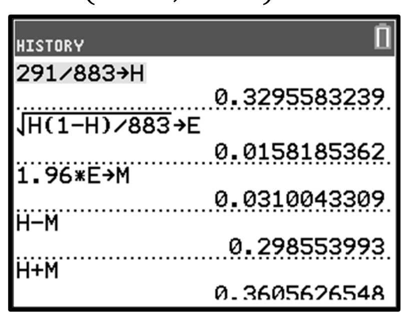
8. $\hat{p} = \frac{291}{883} \approx 0.330$

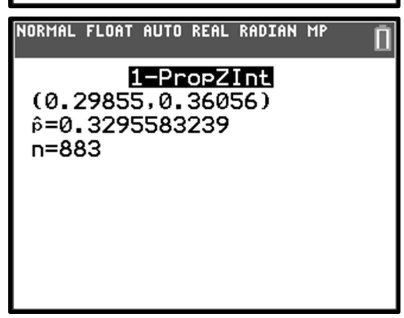
$SE = \sqrt{\frac{(0.33)(0.67)}{883}} \approx 0.0158$

$ME = (1.96)(0.0158) \approx 0.031$

$0.330 - 0.031 < p < 0.330 + 0.031$

$CI = (0.299, 0.361)$





9. $\hat{p} = \frac{12}{19} = 0.632$

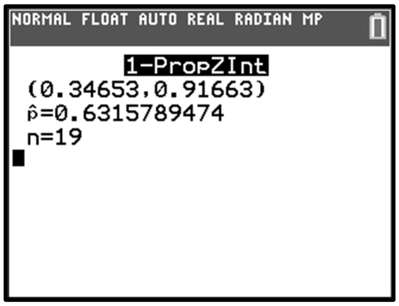
For 99% C-level, $z^* = 2.58$

$$SE = \sqrt{\frac{(0.632)(0.368)}{19}} \approx 0.111$$

$ME = (2.58)(0.111) \approx 0.285$

$0.632 - 0.285 < p < 0.632 + 0.285$

$(0.35, 0.92)$



10. $\hat{p} = \frac{17}{26} = 0.654$

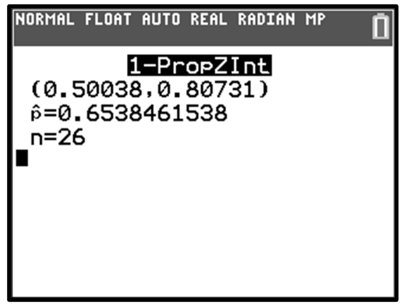
For 90% C-level, $z^* = 1.65$

$$SE = \sqrt{\frac{(0.654)(0.346)}{26}} \approx 0.093$$

$ME = (1.65)(0.093) \approx 0.154$

$0.654 - 0.154 < p < 0.654 + 0.154$

$(0.50, 0.81)$



11. $\hat{p} = \frac{372}{547} = 0.68$

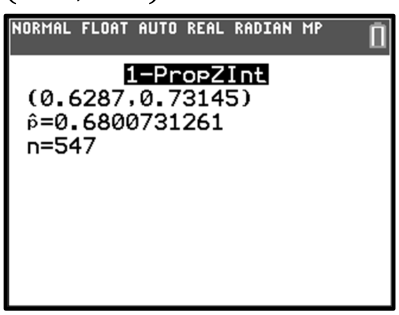
For 99% C-level, $z^* = 2.58$

$$SE = \sqrt{\frac{(0.68)(0.32)}{547}} \approx 0.020$$

$ME = (2.58)(0.020) \approx 0.05$

$0.68 - 0.05 < p < 0.68 + 0.05$

$(0.63, 0.73)$



12. $\hat{p} = \frac{180}{300} = 0.6$

$$SE = \sqrt{\frac{(0.6)(0.4)}{300}} \approx 0.028$$

$ME = (1.96)(0.028) \approx 0.055$

$0.6 - 0.055 < p < 0.6 + 0.055$

$(0.545, 0.655)$

$54.5\% \times 1{,}800 = 981$ and

$65.7\% \times 1{,}800 = 1{,}179$, so between 981 and 1,179 students are estimated to own a pet.

13. $n \geq \left(\frac{z^*}{ME}\right)^2 \hat{p}\hat{q}$

$n \geq \left(\frac{1.65}{0.03}\right)^2 (0.58)(0.42)$

$n \geq 736.9$

The minimum sample size is 737.

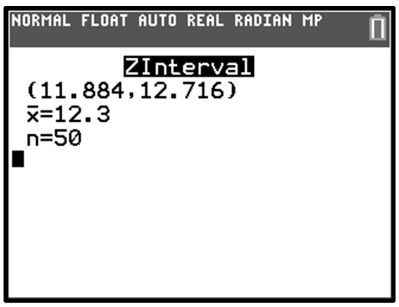
14. $n \geq \left(\frac{z^*}{ME}\right)^2 \hat{p}\hat{q}$

$n \geq \left(\frac{1.96}{0.01}\right)^2 (0.5)(0.5)$

$n \geq 9604$

The minimum sample size is 9604.

8.3 CI for Means (z-Scores)

1. $SE = \frac{s}{\sqrt{n}} = \frac{12}{\sqrt{200}} \approx 0.848$ $ME = (1.96)(0.848) \approx 1.66$	2. $CV = 1.96$ $SE = \frac{s}{\sqrt{n}} = \frac{0.125}{\sqrt{50}} \approx 0.018$ $ME = (1.96)(0.018) \approx 0.035$
3. $SE = \frac{s}{\sqrt{n}} = \frac{1.5}{\sqrt{50}} \approx 0.212$ $ME = (1.96)(0.212) \approx 0.42$ $12.3 \pm 0.42 \rightarrow (11.88, 12.72)$  	4. $SE = \frac{s}{\sqrt{n}} = \frac{2.2}{\sqrt{100}} \approx 0.22$ $ME = (1.65)(0.22) \approx 0.36$ $10.5 \pm 0.36 \rightarrow (10.14, 10.86)$
5. $SE = \frac{s}{\sqrt{n}} = \frac{10.2}{\sqrt{100}} = 1.02$ $ME = (1.96)(1.02) \approx 2.00$ $44.25 - 2.00 < \mu < 44.25 + 2.00$, or $(42.25, 46.25)$	6. $CV = 1.96$ $SE = \frac{s}{\sqrt{n}} = \frac{2.5}{\sqrt{49}} \approx 0.357$ $ME = (1.96)(0.357) \approx 0.7$ $12 - 0.7 < \mu < 12 + 0.7$, or $(11.3, 12.7)$
7. $CV = 1.65$ $SE = \frac{0.5}{\sqrt{100}} = 0.05$ $ME = (1.65)(0.05) \approx 0.08$ $3.55 - 0.08 < \mu < 3.55 + 0.08$, or $(3.47, 3.63)$	8. $CV = 2.58$ $SE = \frac{2.4}{\sqrt{49}} \approx 0.343$ $ME = (2.58)(0.343) \approx 0.88$ $70.3 - 0.88 < \mu < 70.3 + 0.88$, or $(69.42, 71.18)$
9. $CV = 1.65$ $SE = \frac{51.7}{\sqrt{49}} \approx 7.386$ $ME = (1.65)(7.386) \approx 12.19$ $307.5 - 12.19 < \mu < 307.5 + 12.19$, or $(295.31, 319.69)$	10. $CV = 2.58$ $SE = \frac{6.9}{\sqrt{79}} \approx 0.776$ $ME = (2.58)(0.776) \approx 2.00$ $111.1 - 2.0 < \mu < 111.1 + 2.0$, or $(109.1, 113.1)$

11. $n \geq \left(z^* \cdot \frac{\sigma}{ME}\right)^2$ $n \geq \left(1.96 \cdot \frac{18.1}{5}\right)^2$, so $n \geq 50.3$ We need to sample at least 51 bags.	12. $n \geq \left(z^* \cdot \frac{\sigma}{ME}\right)^2$ $n \geq \left(1.96 \cdot \frac{6}{2}\right)^2$, so $n \geq 34.6$ We need to sample at least 35.
13. $n \geq \left(2.58 \cdot \frac{7}{3}\right)^2$, so $n \geq 36.2$ We need to sample at least 37.	14. $n \geq \left(1.65 \cdot \frac{2.5}{1}\right)^2$, so $n \geq 17.01$ We need to sample at least 18.

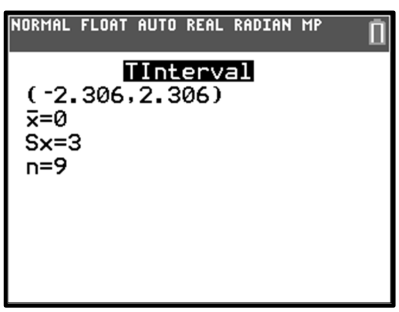
8.4 t-Distributions

1.

population standard deviation σ is known	sample size $n > 30$	z-score or t-score?
Yes	Yes	z-score
Yes	No	z-score
No	Yes	z-score
No	No	t-score

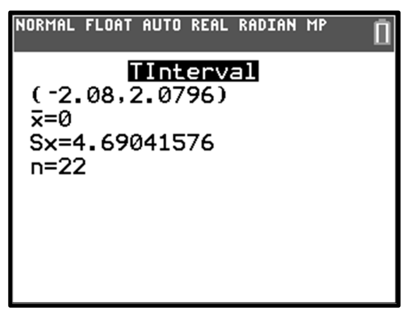
2. $df = 9 - 1 = 8$

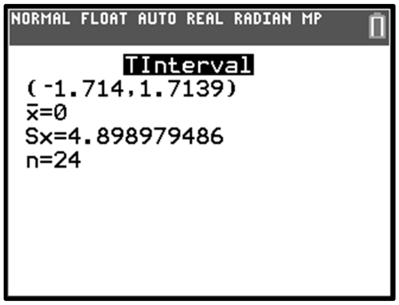
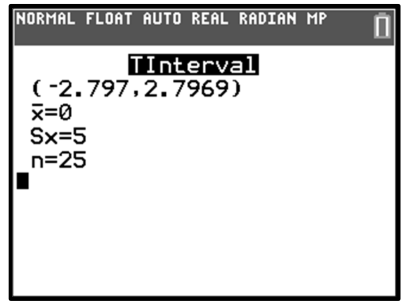
$t^* = 2.306$



3. $df = 22 - 1 = 21$

$t^* = 2.080$



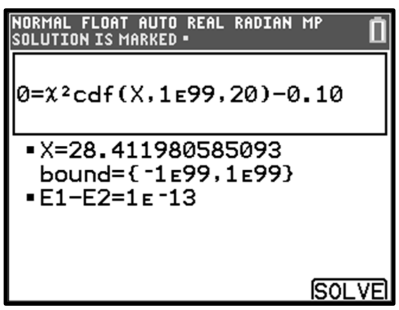
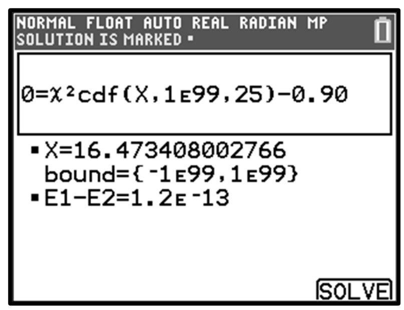
4. $df = 24 - 1 = 23$ $t^* = 1.714$  	5. $df = 25 - 1 = 24$ $t^* = 2.797$ 
6. $\text{tpdf}(1, 10) \approx 0.230$	7. $\text{tpdf}(0, 20) \approx 0.394$
8. $\text{tcdf}(-1, 1, 16) \approx 0.668$	9. $\text{tcdf}(-1\text{E}99, -1, 24) \approx 0.164$
10. $\text{invT}(0.4, 20) \approx -0.26$	11. $100\% - 5\% = 95\%$ $\text{invT}(0.95, 25) \approx 1.71$

8.5 CI for Means (t-Scores)

1. $df = 15 - 1 = 14$ $CV = 2.145$ $SE = \frac{s}{\sqrt{n}} = \frac{0.125}{\sqrt{15}} \approx 0.032$ $ME = (2.145)(0.032) \approx 0.07$	2. $df = 17 - 1 = 16$ $CV = 1.746$ $SE = \frac{s}{\sqrt{n}} = \frac{5.6}{\sqrt{17}} \approx 1.358$ $ME = (1.746)(1.358) \approx 2.37$
3. $df = 20 - 1 = 19$ $CV = 2.093$ $SE = \frac{s}{\sqrt{n}} = \frac{8.4}{\sqrt{20}} \approx 1.878$ $ME = (2.093)(1.878) \approx 3.93$ $44.7 - 3.93 < \mu < 44.7 + 3.93$, or $(40.77, 48.63)$	4. $df = 14 - 1 = 13$ $CV = 1.771$ $SE = \frac{s}{\sqrt{n}} = \frac{8.9}{\sqrt{14}} \approx 2.379$ $ME = (1.771)(2.379) \approx 4.21$ $68.5 - 4.21 < \mu < 68.5 + 4.21$, or $(64.29, 72.71)$
5. $df = 24 - 1 = 23$ $CV = 2.807$ $SE = \frac{s}{\sqrt{n}} = \frac{8.18}{\sqrt{24}} \approx 1.670$ $ME = (2.807)(1.670) \approx 4.69$ 14.75 ± 4.69, or $(10.06, 19.44)$	6. $df = 10 - 1 = 9$ $CV = 1.833$ $SE = \frac{s}{\sqrt{n}} = \frac{0.5}{\sqrt{10}} \approx 0.158$ $ME = (1.833)(0.158) \approx 0.29$ 3.55 ± 0.29, or $(3.26, 3.84)$

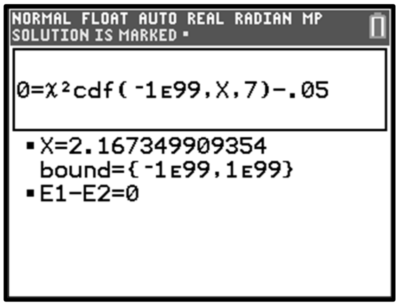
7. $df = 5 - 1 = 4$ $CV = 1.778$ $SE = \frac{s}{\sqrt{n}} = \frac{4.5}{\sqrt{5}} \approx 2.012$ $ME = (1.778)(2.012) \approx 3.58$ 22.2 ± 3.58, or $(18.62, 25.78)$	8. $df = 25 - 1 = 24$ $CV = 2.797$ $SE = \frac{s}{\sqrt{n}} = \frac{16}{\sqrt{25}} = 3.2$ $ME = (2.979)(3.2) \approx 9.5$ 988 ± 9.5, or $(978.5, 997.5)$ No, the desired mean life span of 1,000 cycles is not within the interval.

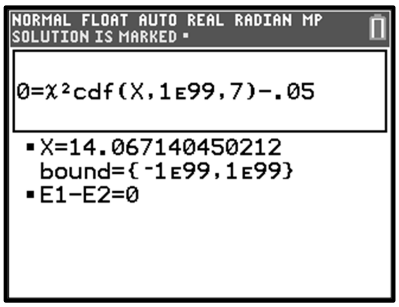
8.6 Chi-Square Distributions

1. $\chi^2\text{pdf}(5, 5) \approx 0.122$	2. $\chi^2\text{pdf}(10, 15) \approx 0.063$
3. $\chi^2 c\text{df}(10, 20, 16) \approx 0.646$	4. $\chi^2 c\text{df}(20, 1\text{E}99, 24) \approx 0.697$
5. 28.412 	6. $100\% - 10\% = 90\%$ to the right 16.473 

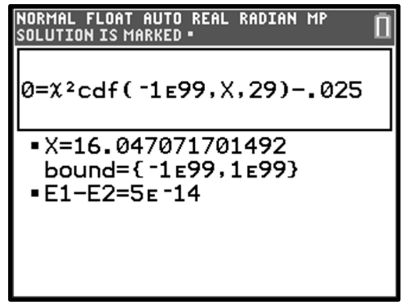
8.7 CI for Variance

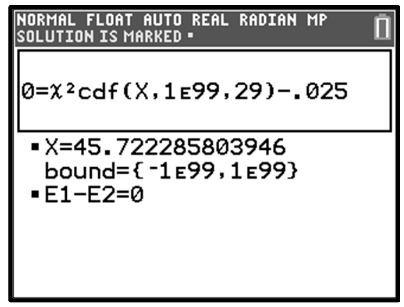
1. $\chi_L^2 = 2.167$ and $\chi_R^2 = 14.067$





2. $\chi_L^2 = 16.047$ and $\chi_R^2 = 45.722$





3. $s^2 = 35^2 = 1225$

$\chi_L^2 = 8.672$ and $\chi_R^2 = 27.587$

$$\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$

$$\frac{(17)(1225)}{27.587} < \sigma^2 < \frac{(17)(1225)}{8.672}$$

$755 < \sigma^2 < 2401$

4. $s^2 = 0.8^2 = 0.64$

$\chi_L^2 = 1.735$ and $\chi_R^2 = 23.589$

$$\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$

$$\frac{(9)(0.64)}{23.589} < \sigma^2 < \frac{(9)(0.64)}{1.735}$$

$0.244 < \sigma^2 < 3.320$

$\sqrt{0.244} < \sigma < \sqrt{3.320}$

$0.5 < \sigma < 1.8$

5. a) $s^2 = 1.43^2 = 2.04$

$\chi_L^2 = 5.009$ and $\chi_R^2 = 24.736$

$$\frac{(13)(2.04)}{24.736} < \sigma^2 < \frac{(13)(2.04)}{5.009}$$

$1.07 < \sigma^2 < 5.29$

b) $\sqrt{1.07} < \sigma < \sqrt{5.29}$

$1.03 < \sigma < 2.30$

6. a) $s^2 = 23.6^2 = 556.96$

$\chi_L^2 = 5.892$ and $\chi_R^2 = 22.362$

$$\frac{(13)(556.96)}{22.362} < \sigma^2 < \frac{(13)(556.96)}{5.892}$$

$324 < \sigma^2 < 1229$

b) $\sqrt{324} < \sigma < \sqrt{1229}$

$18 < \sigma < 35$

Chapter 9 Hypothesis Testing

9.1 HT for Proportions

1. a) No; $np = (15)(0.28) = 4.2 < 5$ b) Yes; $np = (14)(0.4) = 5.6 \geq 5$ and $nq = (14)(0.6) = 8.4 \geq 5$ c) No; $nq = (10)(0.35) = 3.5 < 5$	2. a) $H_0: p = 0.28, H_a: p \neq 0.28$ two-tailed b) $H_0: p \leq 0.4, H_a: p > 0.4$ right-tailed c) $H_0: p \geq 0.65, H_a: p < 0.65$ left-tailed
3. $H_0: p \geq 0.08$ and $H_a: p < 0.08$ Claim is H_a.	4. $H_0: p = 0.65$ and $H_a: p \neq 0.65$ Claim is H_0.
5. The claim is $H_0: p = 0.015$. a) There *is* sufficient evidence to *reject* the claim that the percentage of hourly paid workers earning at or below the minimum wage was 1.5% in 2020. b) There *is not* enough evidence to *reject* the claim that the percentage of hourly paid workers earning at or below the minimum wage was 1.5% in 2020.	6. The claim is $H_a: p > 0.9$. a) There *is* sufficient evidence to *support* the claim that more than 90% of Texas students graduate high school. b) There *is not* enough evidence to *support* the claim that more than 90% of Texas students graduate high school.
7. a) $\hat{p} = \frac{84}{200} = 0.42$ b) $SE = \sqrt{\frac{(0.4)(0.6)}{200}} = 0.0346$ c) $z = \frac{0.42-0.4}{0.0346} = 0.58$	8. a) $\hat{p} = \frac{63}{90} = 0.7$ b) $SE = \sqrt{\frac{(0.75)(0.25)}{90}} = 0.0456$ c) $z = \frac{0.7-0.75}{0.0456} = -1.10$

9. H_0: $p \leq 0.32$ and H_a: $p > 0.32$

Claim is H_0; $\alpha = 0.05$

$\hat{p} = \frac{350}{1000} = 0.35 \rightarrow H$

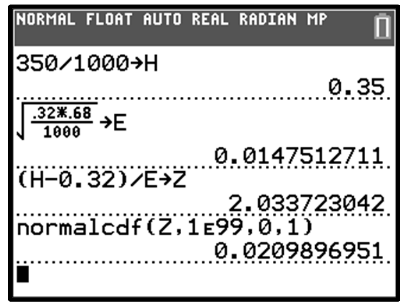
$SE = \sqrt{\frac{(0.32)(0.68)}{1000}} \approx 0.0148 \rightarrow E$

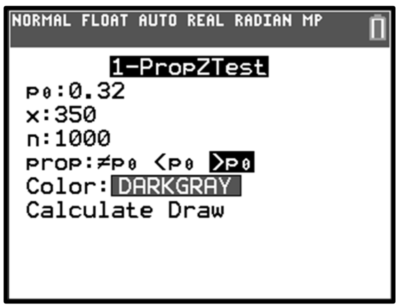
$z = \frac{H - 0.32}{E} \approx 2.03 \rightarrow Z$

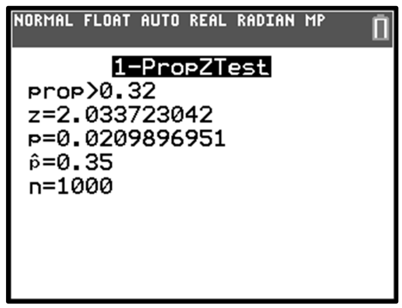
P-value = normalcdf $(Z, 1\text{E}99, 0, 1) \approx 0.02$

$0.02 \leq 0.05$, so reject H_0

There *is* sufficient evidence to *reject* the claim that at most 32% of Americans watched the Super Bowl. *[check using 1-PropZTest below]*







10. H_0: $p \leq 0.75$ and H_a: $p > 0.75$

Claim is H_a; $\alpha = 0.10$

$\hat{p} = \frac{123}{150} = 0.82 \rightarrow H$

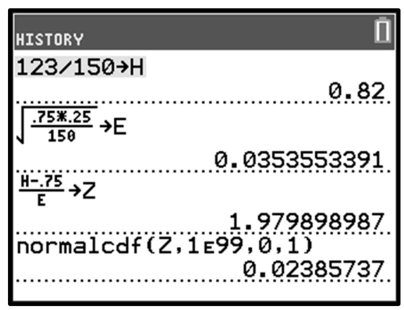
$SE = \sqrt{\frac{(0.75)(0.25)}{150}} \approx 0.0354 \rightarrow E$

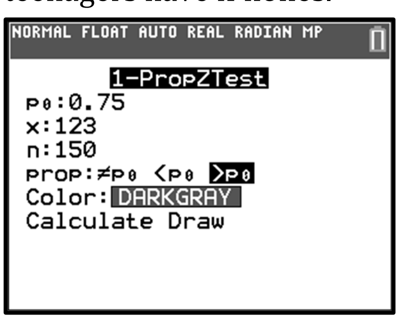
$z = \frac{H - 0.75}{E} \approx 1.98 \rightarrow Z$

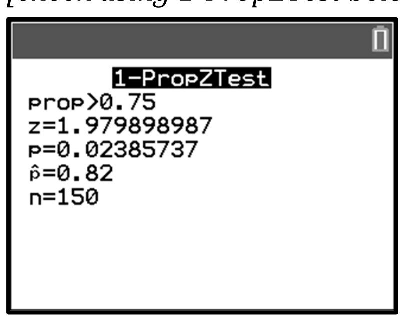
P-value = normalcdf $(Z, 1\text{E}99, 0, 1) \approx 0.02$

$0.02 \leq 0.10$, so reject H_0

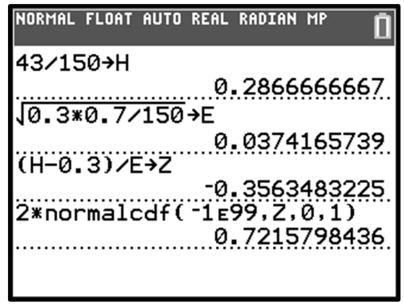
There *is* sufficient evidence to *support* the claim that more than 75% of U.S. teenagers have iPhones. *[check using 1-PropZTest below]*

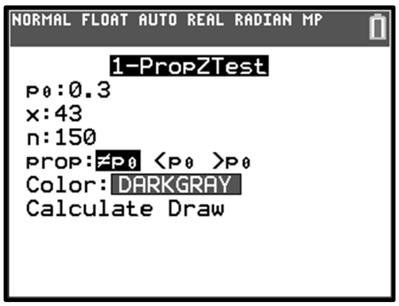


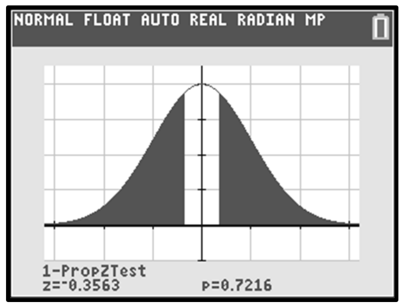




11. H_0: $p = 0.30$ and H_a: $p \neq 0.30$
Claim is H_0; $\alpha = 0.05$
$\hat{p} = \frac{43}{150} \approx 0.2867 \to H$
$SE = \sqrt{\frac{(0.30)(0.70)}{150}} \approx 0.0374 \to E$
$z = \frac{H - 0.30}{E} \approx -0.36 \to Z$
P-value = $2 \times$ normalcdf $(-1\text{E}99, Z, 0, 1) \approx 0.72$
$0.72 > 0.05$, so do *not* reject H_0
There *is not* enough evidence to *reject* the claim that 30% of households subscribe to two or more streaming services. *[check using 1-PropZTest below]*







9.2 HT for Means (z-Tests)

1. a) $H_0: \mu \geq 50, H_a: \mu < 50$ left-tailed
b) $H_0: \mu = 750, H_a: \mu \neq 750$ two-tailed
c) $H_0: \mu \leq 12, H_a: \mu > 12$ right-tailed

3. $H_0: \mu = 12$ and $H_a: \mu \neq 12$ Claim is H_0.	4. $H_0: \mu \geq 15$ and $H_a: \mu < 15$ Claim is H_a.
5. $H_0: \mu \geq 36$ and $H_a: \mu < 12$ Claim is H_0.	6. $H_0: \mu \leq 20$ and $H_a: \mu > 20$ Claim is H_0.

7. a) $SE = \frac{s}{\sqrt{n}} = \frac{0.8}{\sqrt{60}} \approx 0.103 \to E$ $z = \frac{\bar{x} - \mu_0}{E} = \frac{99.2 - 98.7}{E} \approx 4.84$
b) $SE = \frac{s}{\sqrt{n}} = \frac{9}{\sqrt{120}} \approx 0.822 \to E$ $z = \frac{\bar{x} - \mu_0}{E} = \frac{323.6 - 325}{E} \approx -1.70$

8.

a) normalcdf(-1E$99, -1.47, 0, 1$) ≈ 0.071.
$0.071 \leq 0.10$, so reject H_0.

b) normalcdf($2.22, 1$E$99, 0, 1$) ≈ 0.013.
$0.013 > 0.01$, so fail to reject H_0.

c) $2 \times$ normalcdf($1.8, 1$E$99, 0, 1$) ≈ 0.072.
$0.072 > 0.05$, so fail to reject H_0.

9.

a) invNorm($0.10, 0, 1$) ≈ -1.28, so the rejection region is to the left of -1.28. -1.47 is in this region, so reject H_0.

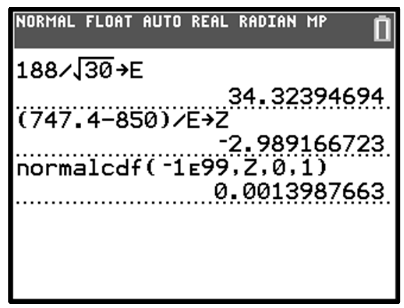
b) invNorm($0.99, 0, 1$) ≈ 2.33, so the rejection region is to the right of 2.33. 2.22 is *not* in this region, so fail to reject H_0.

c) ZInterval($1, 0, 1, 0.95$) $\rightarrow (-1.96, 1.96)$, so the rejection region is to the left of -1.96 or to the right of 1.96. 1.80 is *not* in this region, so fail to reject H_0.

10. $H_0: \mu \geq 850$ and $H_a: \mu < 850$. Claim is H_a. $\alpha = 0.05$.

$SE = \frac{188}{\sqrt{30}} \approx 34.324$

$z = \frac{747.4-850}{E} \approx -2.99$



METHOD 1

P-value = normalcdf (-1E$99, -2.99, 0, 1$) ≈ 0.001

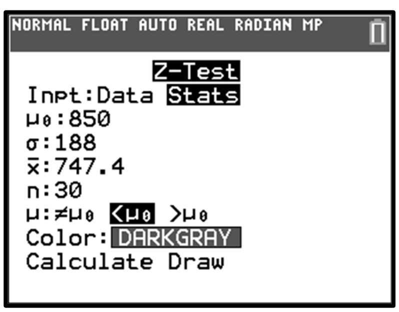
$0.001 \leq 0.05$, so reject H_0

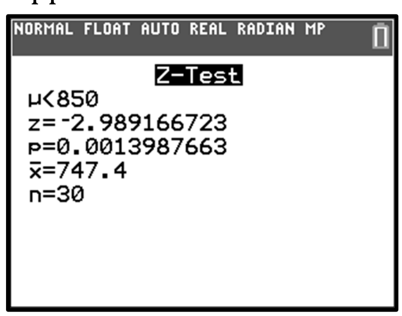
METHOD 2

CV = invNorm ($0.05, 0, 1$) ≈ -1.65

-2.99 is in the rejection region to the left of -1.65, so reject H_0

There is sufficient evidence to support the claim

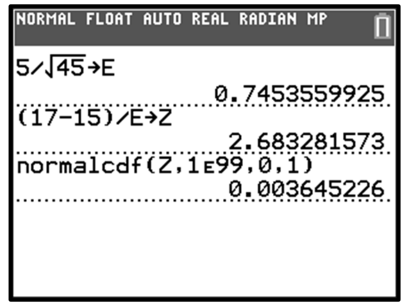




11. $H_0: \mu \leq 15$ and $H_a: \mu > 15$. Claim is H_0. $\alpha = 0.05$.

$SE = \frac{5}{\sqrt{45}} \approx 0.745$

$z = \frac{17-15}{E} \approx 2.68$



METHOD 1

P-value = normalcdf $(2.68, 1E99, 0,1) \approx 0.004$

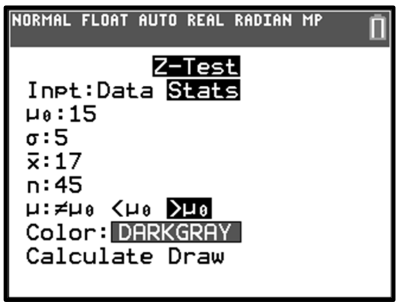
$0.004 \leq 0.05$, so reject H_0

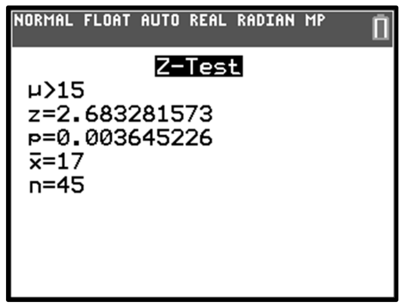
METHOD 2

$CV = \text{invNorm}\ (0.95, 0, 1) \approx 1.65$

2.68 is in rejection region to the right of 1.65 so reject H_0

There is sufficient evidence to reject the claim

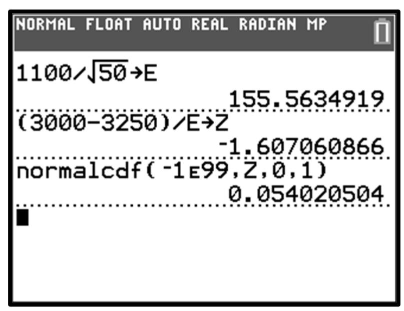




12. $H_0: \mu \geq 3250$ and $H_a: \mu < 3250$. Claim is H_a. $\alpha = 0.05$.

$SE = \frac{1100}{\sqrt{50}} \approx 155.56$

$z = \frac{3000-3250}{E} \approx -1.61$



METHOD 1

P-value = normalcdf $(-1E99, -1.61, 0,1) \approx 0.054$

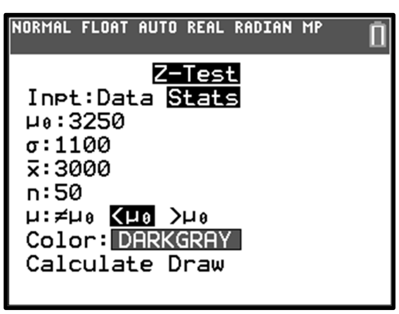
$0.054 > 0.05$, so we fail to reject H_0

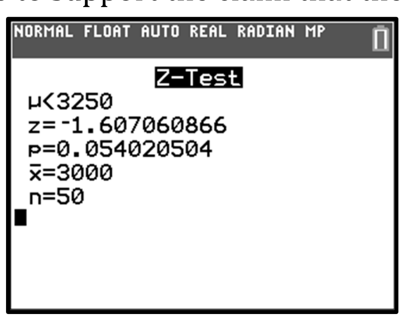
METHOD 2

$CV = \text{invNorm}\ (0.05, 0, 1) \approx -1.65$

-1.61 is *not* in rejection region, so we fail to reject H_0

There is *not* sufficient evidence to support the claim that the debt is less than \$3,250.





9.3 HT for Means (t-Tests)

1. a) $SE = \frac{s}{\sqrt{n}} = \frac{0.4}{\sqrt{9}} \approx 0.133 \rightarrow E$ $\qquad t = \frac{\bar{x}-\mu_0}{E} = \frac{67.7-67.1}{E} = 2.25$

   b) $SE = \frac{s}{\sqrt{n}} = \frac{5}{\sqrt{12}} \approx 1.443 \rightarrow E$ $\qquad t = \frac{\bar{x}-\mu_0}{E} = \frac{122.75-125}{E} \approx -1.56$

2. a) tcdf$(-1\text{E}99, -1.47, 19) \approx 0.079$. $\quad 0.079 > 0.01$, so fail to reject H_0.

   b) tcdf$(2.22, 1\text{E}99, 24) \approx 0.018$. $\quad 0.018 \leq 0.05$, so reject H_0.

   c) $2 \times$ tcdf$(1.8, 1\text{E}99, 15) \approx 0.092$. $\quad 0.092 \leq 0.10$, so reject H_0.

3. a) invT$(0.01, 19) \approx -2.54$, so the rejection region is to the left of -2.54.
   -1.47 is *not* in this region, so fail to reject H_0.

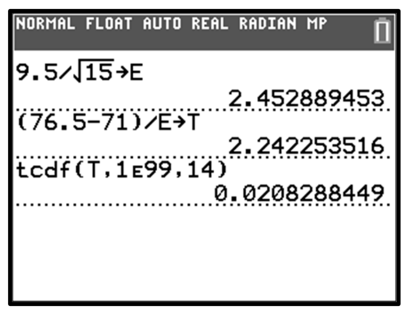
   b) invT$(0.95, 0, 1) \approx 1.71$, so the rejection region is to the right of 1.71.
   2.22 is in this region, so reject H_0.

   c) TInterval$(0, 4, 16, 0.9) \rightarrow (-1.75, 1.75)$, so the rejection region is to the left of -1.75 or to the right of 1.75. 1.80 is in this region, so reject H_0.

4. H_0: $\mu \leq 71$ and H_a: $\mu > 71$. Claim is H_a. $\alpha = 0.01$.

   $SE = \frac{9.5}{\sqrt{15}} \approx 2.453$

   $t = \frac{76.5-71}{E} \approx 2.24$, $\qquad df = 15 - 1 = 14$

   METHOD 1

   P-value = tcdf $(2.242, 1\text{E}99, 14) \approx 0.021$

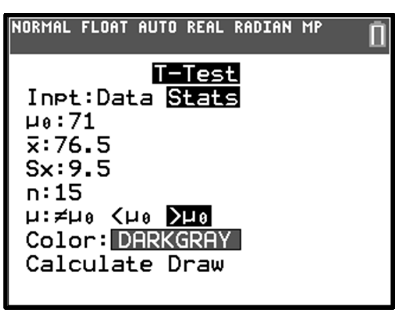
   $0.021 > 0.01$, so we do *not* reject H_0

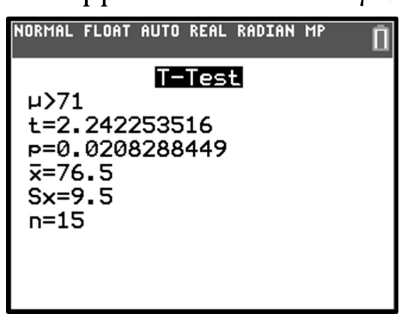
   METHOD 2

   $1 - \alpha = 0.99$, $\qquad CV =$ invT $(0.99, 13) \approx 2.624$

   2.24 is *not* in the rejection region, so we do *not* reject H_0

   There is *not* sufficient evidence to support the claim that $\mu > 71$.

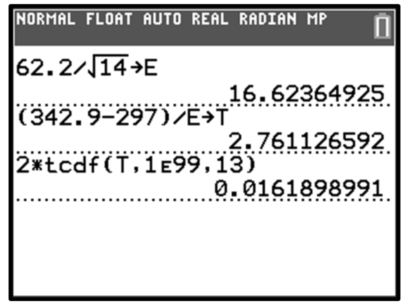
   

5. $H_0: \mu = 297$ and $H_a: \mu \neq 297$. Claim is H_0. $\alpha = 0.05$.

$SE = \frac{62.2}{\sqrt{14}} \approx 16.624$

$t = \frac{342.9 - 297}{E} \approx 2.761, \quad df = 14 - 1 = 13$



METHOD 1

P-value = $2 \times$ tcdf $(2.761, 1E99, 13) \approx 0.016$

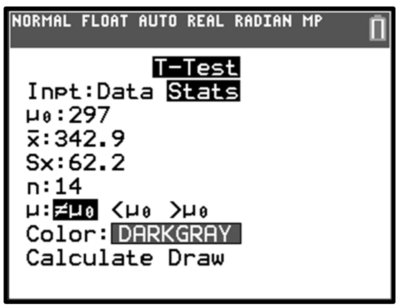
$0.016 \leq 0.05$, so reject H_0

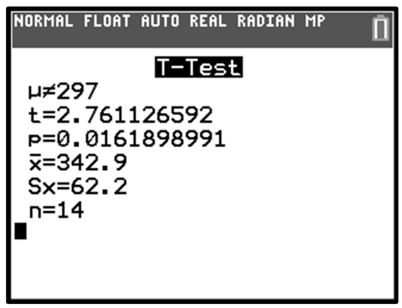
METHOD 2

$1 - \frac{\alpha}{2} = 0.975, \qquad CV = \text{invT}\ (0.975, 13) \approx 2.16$

2.761 is in the rejection region to the right of 2.16, so reject H_0

There is sufficient evidence to reject the claim that the population mean is 297





9.4 HT for Variance

1. a) $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{50 \cdot 29^2}{25^2} = 67.28$

   b) $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{23 \cdot 35^2}{40^2} \approx 17.609$

   c) $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{14 \cdot 0.75^2}{0.8^2} \approx 12.305$

2. a) χ^2cdf $(-1e99, 37, 27) \approx 0.90$

   $0.90 > 0.20$, so fail to reject H_0

   b) χ^2cdf $(28, 1e99, 17) \approx 0.045$

   $0.045 \leq 0.05$, so reject H_0

c) $2 \times \chi^2$cdf $(-1E99, 22.75, 34) \approx 0.142$

   $0.142 > 0.10$, so fail to reject H_0

3. a) $\chi_L^2 = 20.70$; region is $\chi^2 < 20.70$

   $\chi^2 > 20.70$, so fail to reject H_0

   b) $\chi_R^2 = 27.59$; region is $\chi^2 > 27.59$

   $\chi^2 > 27.59$, so reject H_0

c) $\chi_L^2 = 21.66$ and $\chi_R^2 = 48.60$; region is the area outside $(21.66, 48.60)$.

χ^2 is inside the interval, so fail to reject H_0

4. $H_0: \sigma^2 \leq 0.0004$, $H_a: \sigma^2 > 0.0004$ (claim), $\alpha = 0.05$.

$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(29)(0.0005)}{0.0004} \approx 36.25$, $df = n - 1 = 29$

METHOD 1

$P = \chi^2\text{cdf}(36.25, 1e99, 29) = 0.166$

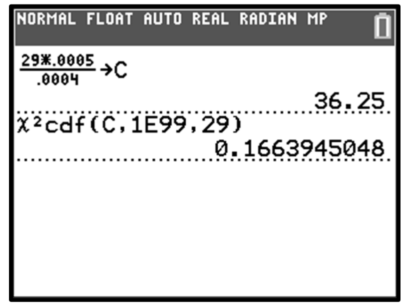
$P > 0.05$, so we fail to reject H_0

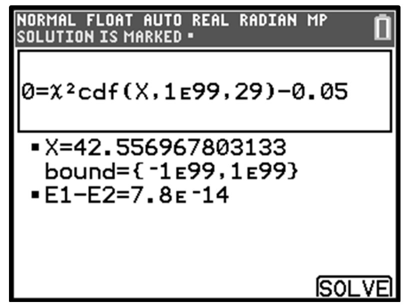
METHOD 2

$\chi^2_R = 42.56$, so rejection region is $\chi^2 > 42.56$.

$\chi^2 = 36.25$ does *not* lie in the rejection region, so we fail to reject H_0.

There *is not* enough evidence to *support* the claim that the variation is too high.





5. $H_0: \sigma \geq 7.2$, $H_a: \sigma < 7.2$ (claim), $\alpha = 0.05$.

$\sigma^2 = 7.2^2 = 51.84$, $s^2 = 3.5^2 = 12.25$.

$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(24)(12.25)}{51.84} \approx 5.67$, $df = n - 1 = 24$

METHOD 1

$P = \chi^2\text{cdf}(-1e99, 5.67, 24) = 0.00004$

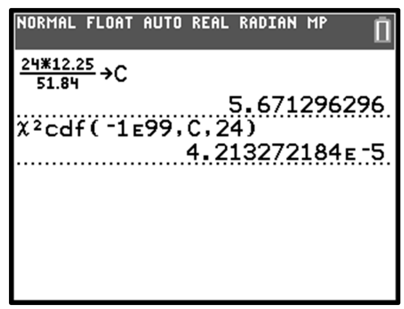
$P \leq 0.05$, so reject H_0

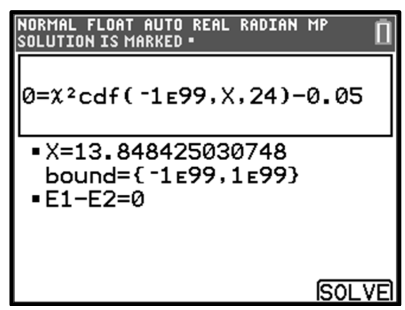
METHOD 2

$\chi^2_L = 13.85$, so rejection region is $\chi^2 < 13.85$.

$\chi^2 = 5.67$ lies in the rejection region, so reject H_0.

There *is* sufficient evidence to *support* the claim that the variation in wait times has been reduced.





9.5 Types of Errors

1. (2)	2. (1)
3. a) A Type I error occurs if the actual proportion of patients who experience side effects is at most 3%, but the null hypothesis, $H_0: p \leq 0.03$, is rejected. b) A Type II error occurs if the actual proportion of patients who experience side effects is greater than 3%, but the test fails to reject the null hypothesis, $H_0: p \leq 0.03$.	4. a) A Type I error occurs if the actual proportion of residents who approve of the mayor is 65%, but the null hypothesis, $H_0: p = 0.65$, is rejected. b) A Type II error occurs if the actual proportion of residents who approve of the mayor is not 65%, but the test fails to reject the null hypothesis, $H_0: p = 0.65$.
5. a) A Type I error occurs if the actual mean is 12 mg, but the null hypothesis, $H_0: \mu = 12$, is rejected. b) A Type II error occurs if the actual mean is not 12 mg, but the test fails to reject the null hypothesis, $H_0: \mu = 12$.	6. a) A Type I error occurs if the actual mean is at least 15 minutes, but the null hypothesis, $H_0: \mu \geq 15$, is rejected. b) A Type II error occurs if the actual mean is less than 15 minutes, but the test fails to reject the null hypothesis, $H_0: \mu \geq 15$.

Chapter 10 Two Samples

10.1 Combine Random Variables

1. $\mu_{X+Y} = \mu_X + \mu_Y$ and $\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$

	μ	σ^2
X	12	2
Y	33	4
Z	45	6

2. $\sigma^2_Z = \sigma^2_X + \sigma^2_Y = 16 + 9 = 25$

   $\sigma_Z = \sqrt{25} = 5$

3. Let $X =$ weight of chocolate per bar and let $Y =$ weight of caramel per bar

   Let $B = X + Y =$ weight of each bar

   $\mu_B = \mu_X + \mu_Y = 32 + 18 = 50\text{g}$

   $\sigma_B = \sqrt{\sigma^2_X + \sigma^2_Y} = \sqrt{0.75^2 + 0.25^2} \approx 0.79\text{g}$

4. $\mu_Z = \mu_X + \mu_Y = 10 + 10 = 20$

   $\sigma^2_Z = \sigma^2_X + \sigma^2_Y = 1.5^2 + 2^2 = 6.25$

   So, $\sigma_Z = \sqrt{\sigma^2_X + \sigma^2_Y} = \sqrt{6.25} = 2.5$

   $Z \sim N(20, 2.5)$

   $P(Z > 25) = \text{normalcdf}(25, 1\text{E}99, 20, 2.5) \approx 0.023$

5. $P(W > M) = P(M - W < 0)$

   Let $D = M - W$

   $\mu_D = \mu_M - \mu_W = 70 - 67 = 3$

   $\sigma^2_D = \sigma^2_M + \sigma^2_W = 3.2^2 + 2.4^2 = 16$

   So, $\sigma_D = \sqrt{16} = 4$

   $D \sim N(3, 4)$

   $P(D < 0) = \text{normalcdf}(-1\text{E}99, 0, 3, 4) \approx 0.227$

10.2 Compare Proportions

1. $\hat{p}_1 = \frac{23}{50} = 0.46$ and $\hat{p}_2 = \frac{11}{50} \approx 0.22$

   $\hat{q}_1 = 1 - \hat{p}_1 = 0.54$ and $\hat{q}_2 = 1 - \hat{p}_2 = 0.78$

   $(\hat{p}_1 - \hat{p}_2) = 0.46 - 0.22 = 0.24$

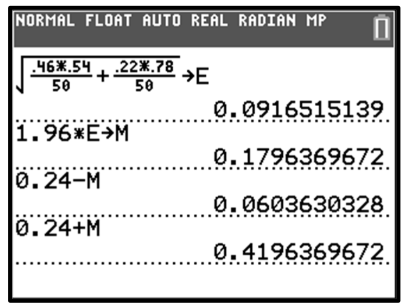
   For $CL = 95\%, z^* = 1.96$

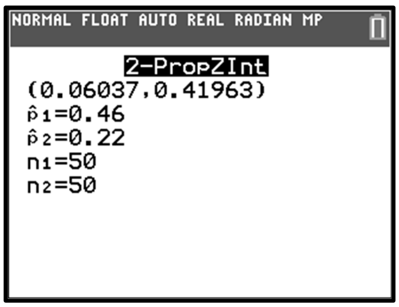
   $$SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \sqrt{\frac{(0.46)(0.54)}{50} + \frac{(0.22)(0.78)}{50}} \approx 0.092$$

   $ME = (CV)(SE) = (1.96)(0.092) \approx 0.180$

   $0.24 - 0.18 < (p_1 - p_2) < 0.24 + 0.18$

   $0.06 < (p_1 - p_2) < 0.42$





2. $\hat{p}_1 = \frac{53}{100} = 0.53$ and $\hat{p}_2 = \frac{37}{100} \approx 0.336$

   $\hat{q}_1 = 1 - \hat{p}_1 = 0.47$ and $\hat{q}_2 = 1 - \hat{p}_2 = 0.664$

   $(\hat{p}_1 - \hat{p}_2) = 0.53 - 0.336 = 0.194$

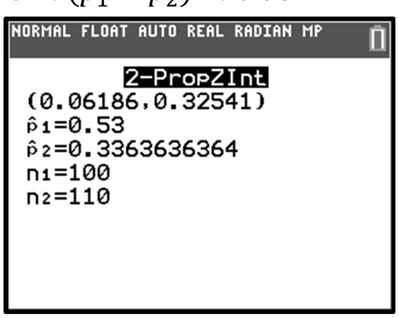
   For $CL = 95\%, z^* = 1.96$

   $$SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \sqrt{\frac{(0.53)(0.47)}{100} + \frac{(0.336)(0.664)}{110}} \approx 0.067$$

   $ME = (CV)(SE) = (1.96)(0.067) \approx 0.132$

   $0.194 - 0.132 < (p_1 - p_2) < 0.194 + 0.132$

   $0.06 < (p_1 - p_2) < 0.33$



3. H_0: $p_1 \le p_2$ and H_a: $p_1 > p_2$

$\hat{p}_1 = \frac{12}{40} = 0.3$ and $\hat{p}_2 = \frac{9}{60} = 0.15$

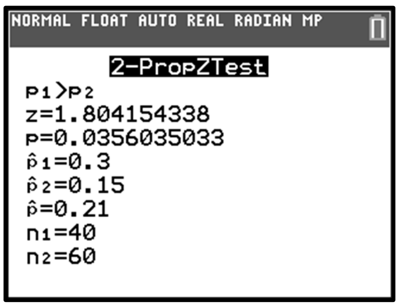
pooled proportion is $\bar{p} = \frac{12+9}{40+60} = 0.21$ and $\bar{q} = 1 - 0.21 = 0.79$

$SE(\bar{p}) = \sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{(0.21)(0.79)\left(\frac{1}{40} + \frac{1}{60}\right)} = 0.0831$

$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{SE} = \frac{(0.3 - 0.15) - 0}{0.0831} = 1.804$

P-value is normalcdf $(1.804, 1\text{E}99, 0, 1) \approx 0.036$

$0.036 > 0.01$, so fail to reject H_0



4. H_0: $p_1 = p_2$ and H_a: $p_1 \ne p_2$

$\hat{p}_1 = \frac{42}{150} = 0.28$ and $\hat{p}_2 = \frac{75}{200} = 0.375$

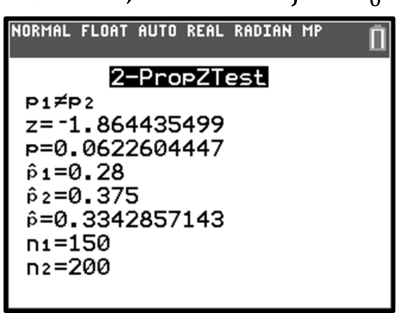
pooled proportion is $\bar{p} = \frac{42+75}{150+200} \approx 0.334$ and $\bar{q} = 1 - 0.334 = 0.666$

$SE(\bar{p}) = \sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{(0.334)(0.666)\left(\frac{1}{150} + \frac{1}{200}\right)} = 0.05094$

$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{SE} = \frac{(0.28 - 0.375) - 0}{0.05094} = -1.86$

P-value is $2 \times$ normalcdf $(-1e99, -1.86, 0, 1) \approx 0.06$

$0.06 > 0.05$, so fail to reject H_0



5. H_0: $p_1 = p_2$ and H_a: $p_1 \neq p_2$

$\hat{p}_1 = \frac{882}{1000} = 0.882$ and $\hat{p}_2 = \frac{892}{1200} = 0.743$

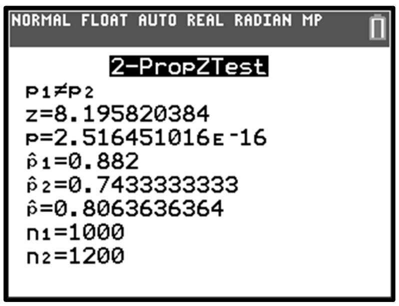
pooled proportion is $\bar{p} = \frac{882+892}{1000+1200} = 0.8064$ and $\bar{q} = 1 - 0.8064 = 0.1936$

$SE(\bar{p}) = \sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{(0.8064)(0.1936)\left(\frac{1}{1000} + \frac{1}{1200}\right)} = 0.01692$

$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{SE} = \frac{(0.882 - 0.743) - 0}{0.01692} = 8.2$

P-value is $2 \times \text{normalcdf}\,(8.2, 1\text{E}99, 0, 1) = 0.000$

$0.000 < 0.05$, so there is sufficient evidence that the true proportion of Colorado internet users differs from North Carolina internet users



6. a) H_0: $p_1 - p_2 = 0$ and H_a: $p_1 - p_2 \neq 0$
Using the 2-PropZTest function, $p \approx 0.315$.
$0.315 > 0.05$, so there is *not* enough evidence to reject H_0.
There is *not* sufficient evidence of a difference in the true free throw rates.

b) Using the 2-PropZInt function, $CI \approx (-0.0473, 0.1474)$.
We are 95% confident that the true difference is between -0.0473 and 0.1474.

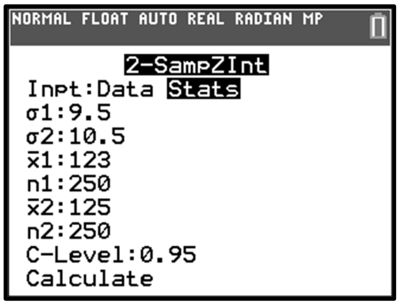
c) Yes, they are consistent. A difference of zero is within the confidence interval, which explains why we can't conclude that their true free throw rates differ.

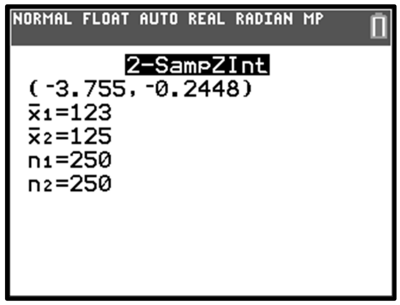
d) 88% of 400 is 352, 83% of 400 is 332. Using the 2-PropZTest function, $p \approx 0.045$.
$0.045 < 0.05$, so there is sufficient evidence of a difference in their rates.
Using the 2-PropZInt function, $CI \approx (0.0013, 0.0987)$, so we are 95% confident that the true difference in their free throw rates is between 0.0013 and 0.0987.
This is also consistent: a difference of zero is *not* in the confidence interval, so the evidence supports that their true free throw rates are different.

10.3 Compare Means (z-Scores and z-Tests)

1. $CI = (\bar{x}_1 - \bar{x}_2) \pm z^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

$$= (123 - 125) \pm 1.96 \cdot \sqrt{\frac{9.5^2}{250} + \frac{10.5^2}{250}} \approx -2 \pm 1.76 \rightarrow (-3.76, -0.24)$$





2. a) $(68.9 - 63.4) \pm 1.96 \cdot \sqrt{\frac{2.7^2}{1545} + \frac{2.5^2}{1781}} \approx 5.5 \pm 0.18 \rightarrow (5.32, 5.68)$

b) $(194.0 - 157.7) \pm 1.96 \cdot \sqrt{\frac{33.8^2}{1612} + \frac{34.6^2}{1894}} \approx 36.3 \pm 2.27 \rightarrow (34.03, 38.57)$

c) $(28.8 - 27.6) \pm 1.96 \cdot \sqrt{\frac{4.6^2}{1545} + \frac{5.9^2}{1781}} \approx 1.2 \pm 0.36 \rightarrow (0.84, 1.56)$

d) $(192.4 - 207.1) \pm 1.96 \cdot \sqrt{\frac{35.2^2}{1544} + \frac{36.7^2}{1766}} \approx -14.7 \pm 2.45 \rightarrow (-17.15, -12.25)$

3. $H_0: \mu_1 = \mu_2$ (claim), $H_a: \mu_1 \neq \mu_2$

$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{3^2}{30} + \frac{1.5^2}{30}} \approx 0.6124$

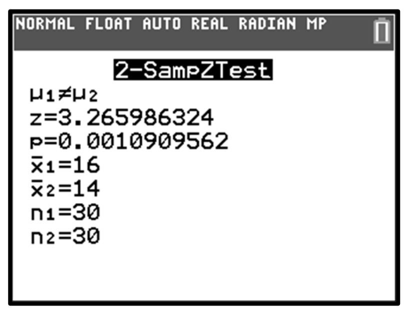
$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE} = \frac{(16-14) - 0}{SE} = \frac{2}{0.6124} \approx 3.27$

P-value $= 2 \times$ normalcdf $(3.27, 1\text{E}99, 0, 1) \approx 0.001$

$0.001 < 0.01$, so reject H_0.

OR $CV = \pm 2.576$ *[*ZInterval *for* $CL = 0.99$*]*

$3.27 > 2.576$ is in the rejection region, so reject H_0.



4. $H_0: \mu_1 \geq \mu_2, H_a: \mu_1 < \mu_2$ (claim)

$SE = \sqrt{\frac{75^2}{40} + \frac{100^2}{90}} \approx 15.866$

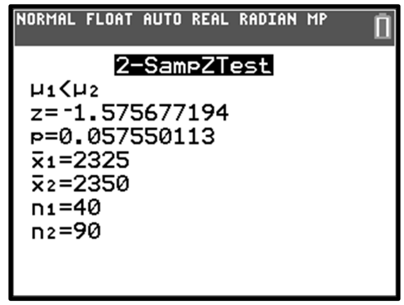
$z = \frac{(2325-2350) - 0}{SE} = \frac{-25}{15.866} \approx -1.5757$

P-value = normalcdf $(-1\text{E}99, -1.5757, 0, 1) \approx 0.058$

$0.058 > 0.05$, so fail to reject H_0.

OR $CV = -1.645$ *[*invNorm$(0.05, 0, 1)$*]*

$-1.5757 > -1.645$ is *not* in the rejection region, so fail to reject H_0.



5. $H_0: \mu_1 \leq \mu_2, H_a: \mu_1 > \mu_2$ (claim)

$SE = \sqrt{\frac{4.2^2}{200} + \frac{6.1^2}{200}} \approx 0.5237$

$z = \frac{(14.4-13.7) - 0}{SE} = \frac{0.7}{0.5237} \approx 1.34$

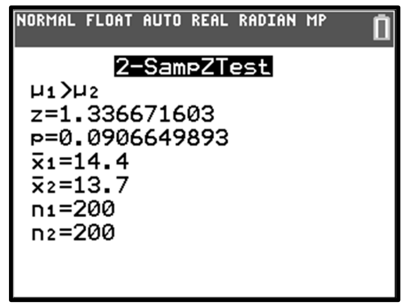
P-value = normalcdf $(1.34, 1\text{E}99, 0, 1) \approx 0.09$

$0.09 < 0.1$, so reject H_0.

OR $CV = 1.28$ *[*invNorm$(0.9, 0, 1)$*]*

$1.34 > 1.28$ is in the rejection region, so reject H_0.

There *is* sufficient evidence to *support* the claim that men have a higher mean concentration of the mineral.



10.4 Compare Variances

1. $H_0: \sigma_1^2 = \sigma_2^2$ and $H_a: \sigma_1^2 \neq \sigma_2^2$

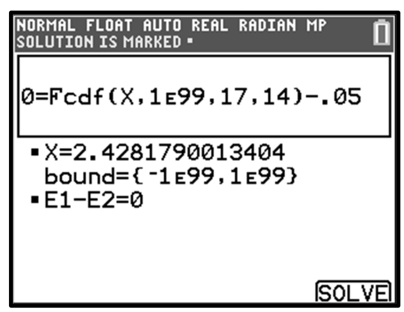
$df_1 = n_1 - 1 = 17$ and $df_2 = n_2 - 1 = 14$

$CV = 2.43$; rejection region is where $F > 2.43$

$F = \frac{s_1^2}{s_2^2} = \frac{300}{150} = 2$

$F < CV$, so fail to reject H_0. There *is not* enough evidence to *reject* the claim that $\sigma_1^2 = \sigma_2^2$.

OR Using 2-SampFTest, $p = 0.196 > 0.10$, so fail to reject H_0.



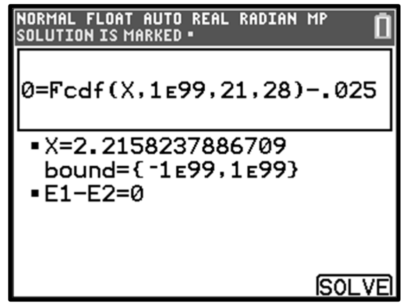
2. $H_0: \sigma_1^2 = \sigma_2^2$ and $H_a: \sigma_1^2 \neq \sigma_2^2$

$df_1 = n_1 - 1 = 21$ and $df_2 = n_2 - 1 = 28$

$CV = 2.22$; rejection region is where $F > 2.22$

$F = \frac{s_1^2}{s_2^2} = \frac{445}{190} = 2.34$

$F > CV$, so reject H_0. There *is* sufficient evidence to *reject* the claim that $\sigma_1^2 = \sigma_2^2$.



OR Using 2-SampFTest, $p = 0.036 < 0.05$, so reject H_0.

3. $H_0: \sigma_1^2 = \sigma_2^2$ and $H_a: \sigma_1^2 \neq \sigma_2^2$

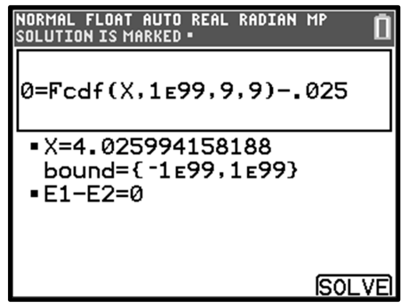
$df_1 = n_1 - 1 = 9$ and $df_2 = n_2 - 1 = 9$

$CV = 4.03$; rejection region is where $F > 4.03$

$s_1^2 = 0.75^2 = 0.5625$ and $s_2^2 = 0.683^2 \approx 0.4665$

$F = \frac{s_1^2}{s_2^2} = \frac{0.5625}{0.4665} \approx 1.21$

$F < CV$, so fail to reject H_0. There *is not* enough evidence to *reject* the claim that $\sigma_1^2 = \sigma_2^2$.



OR Using 2-SampFTest, $p = 0.785 > 0.05$, so fail to reject H_0.

4. $H_0: \sigma_1^2 = \sigma_2^2$ and $H_a: \sigma_1^2 \neq \sigma_2^2$

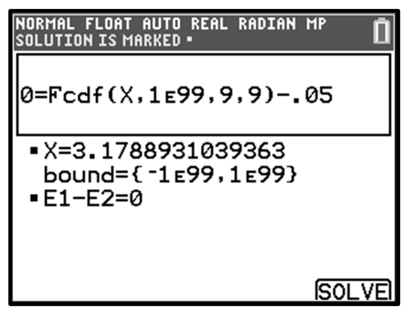
Let s_1^2 represent the larger variance, 89.9, and let s_2^2 represent the smaller variance, 52.3.

$df_1 = n_1 - 1 = 9$ and $df_2 = n_2 - 1 = 9$

$CV = 3.18$; rejection region is where $F > 3.18$

$F = \frac{s_1^2}{s_2^2} = \frac{89.9}{52.3} \approx 1.72$

$F < CV$, so fail to reject H_0. There *is not* enough evidence to *reject* the claim that $\sigma_1^2 = \sigma_2^2$.



OR Using 2-SampFTest, $p = 0.432 > 0.1$, so fail to reject H_0.

5. $H_0: \sigma_1^2 = \sigma_2^2$ and $H_a: \sigma_1^2 \neq \sigma_2^2$

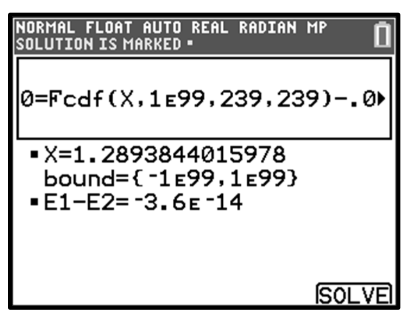
$df_1 = n_1 - 1 = 239$ and $df_2 = n_2 - 1 = 239$

$CV = 1.29$; rejection region is where $F > 1.29$

$s_1^2 = 65.55^2 \approx 4297$ and $s_2^2 = 61.85^2 \approx 3825$

$F = \frac{s_1^2}{s_2^2} = \frac{4297}{3825} \approx 1.12$

$F < CV$, so fail to reject H_0. There *is not* enough evidence to *reject* the claim that $\sigma_1^2 = \sigma_2^2$.



OR Using 2-SampFTest, $p = 0.370 > 0.05$, so fail to reject H_0.

10.5 Compare Means (t-Scores and t-Tests)

1. a) $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{(24)(11.3)^2 + (31)(9.1)^2}{(24) + (31)} = 102.394$

   b) $SE = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = \sqrt{\frac{102.394}{25} + \frac{102.394}{32}} \approx 2.701$

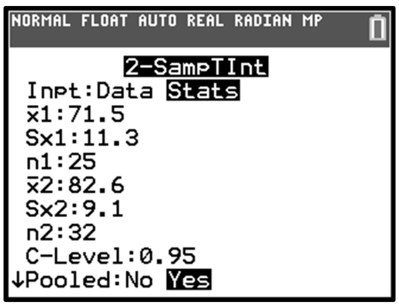
   c) $df = n_1 + n_2 - 2 = 25 + 32 - 2 = 55$

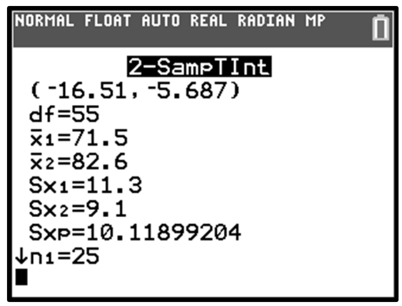
   d) $t^* = \text{TInterval}\ (\bar{x} = 0, Sx = \sqrt{56}, n = 56, CL = 0.95) \approx 2.004$

   e) $CI = (71.5 - 82.6) \pm (2.004)(2.701) \approx -11.1 \pm 5.4$

   Estimated difference of population means is between −16.5 and −5.7.

   f) The calculator's 2-SampTInt function calculates the same interval.





2. $s_p^2 = \frac{(14)(45)^2 + (14)(30)^2}{(14)+(14)} = 1462.5$ *OR* $s_p^2 = \frac{(45)^2 + (30)^2}{2} = 1462.5$

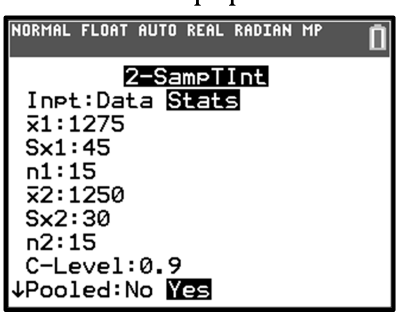
   $SE = \sqrt{\frac{1462.5}{15} + \frac{1462.5}{15}} \approx 13.964$ *OR* $SE = \sqrt{1462.5} \cdot \sqrt{\frac{2}{15}} \approx 13.964$

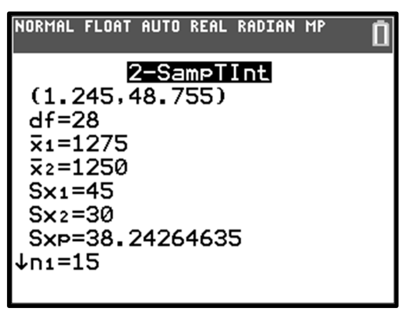
   $df = 15 + 15 - 2 = 28$

   $t^* = \text{TInterval}\ (\bar{x} = 0, Sx = \sqrt{29}, n = 29, CL = 0.90) \approx 1.701$

   $CI = (1275 - 1250) \pm (1.701)(13.964) \approx 25 \pm 23.75$

   Estimated difference of population means is between 1.25 and 48.75.





3. METHOD 1 (Using the *smaller* of $n_1 - 1$ or $n_2 - 1$ for df)

$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{0.60^2}{9} + \frac{0.12^2}{5}} \approx 0.207$

$df = 5 - 1 = 4$

$t^* = \text{TInterval}\left(\bar{x} = 0, Sx = \sqrt{5}, n = 5, CL = 0.95\right) = 2.776$

$ME = (CV)(SE) = (2.776)(0.207) = 0.57$

$CI = (27 - 24) \pm ME = 3 \pm 0.57 = (2.43, 3.57)$

METHOD 2 (Using the long formula for df)

$A = \frac{s_1^2}{n_1} = \frac{0.6^2}{9} = 0.04$ and $B = \frac{s_2^2}{n_2} = \frac{0.12^2}{5} = 0.00288$

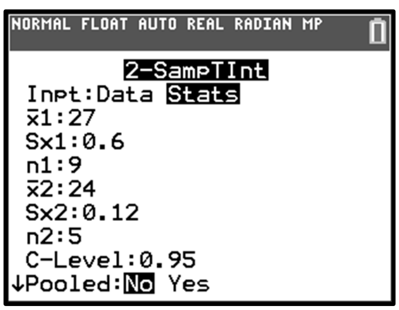
$SE = \sqrt{A + B} \approx 0.207$

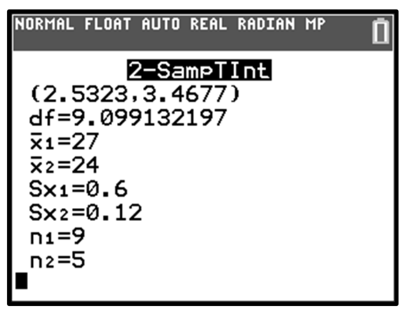
$$df = \frac{(A+B)^2}{\frac{A^2}{8} + \frac{B^2}{4}} \approx 9$$

$t^* = \text{TInterval}\ (\bar{x} = 0, Sx = \sqrt{10}, n = 10, CL = 0.95) \approx 2.26$

$ME = (CV)(SE) = (2.26)(0.207) = 0.47$

$CI = (27 - 24) \pm ME \approx 3 \pm 0.47 = (2.53, 3.47)$





4. $H_0: \mu_1 = \mu_2$ and $H_a: \mu_1 \neq \mu_2$. The claim is H_0. $\alpha = 0.1$.

$df = 15 + 15 - 2 = 28$

$$\hat{\sigma} = \sqrt{\frac{s_1^2 + s_2^2}{2}} = \sqrt{\frac{45^2 + 30^2}{2}} = 38.243$$

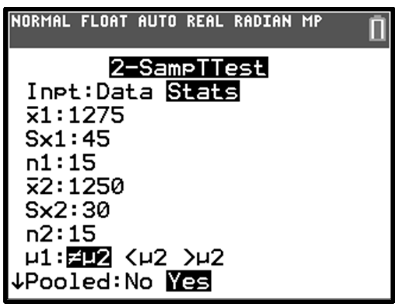
$$SE = 38.243\sqrt{\frac{2}{15}} = 13.964$$

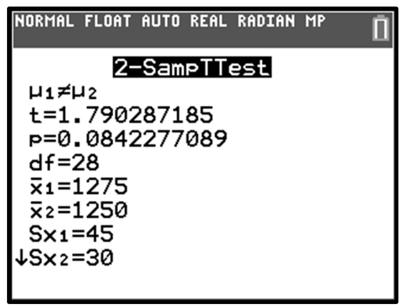
$$t = \frac{(1275 - 1250) - 0}{13.964} = 1.79$$

TInterval($\bar{x} = 0, s_x = \sqrt{29}, n = 29, CL = 0.90$) gives us $(-1.70, 1.70)$.

$t = 1.79$ is outside this interval, so it lies in the rejection region and we reject H_0.

There is sufficient evidence to reject the claim that the population means are equal.





5. $H_0: \mu_1 \leq \mu_2$ and $H_a: \mu_1 > \mu_2$. The claim is H_a. $\alpha = 0.05$.

METHOD 1 (Using the *smaller* of $n_1 - 1$ or $n_2 - 1$ for df)

$df = 5 - 1 = 4$

$$SE = \sqrt{\frac{0.6^2}{9} + \frac{0.12^2}{5}} \approx 0.207$$

$$t = \frac{(27-24)-0}{0.207} = 14.49$$

TInterval $(0, \sqrt{5}, 5, 0.95)$ gives us $(-2.776, 2.776)$.

$t = 14.49$ is outside this interval, so we reject H_0.

There is enough evidence to reject the claim that the population means are equal.

METHOD 2 (Using the long formula for df)

$A = \frac{s_1^2}{n_1} = \frac{0.6^2}{9} = 0.04$ and $B = \frac{s_2^2}{n_2} = \frac{0.12^2}{5} = 0.00288$

$SE = \sqrt{A + B} \approx 0.207$

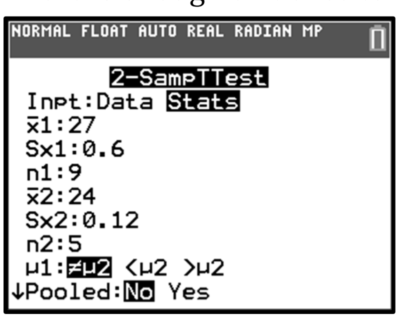
$$df = \frac{(A+B)^2}{\frac{A^2}{8} + \frac{B^2}{4}} \approx 9$$

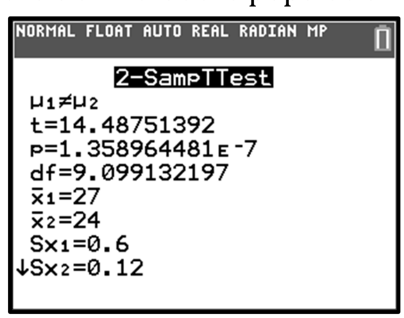
$$t = \frac{(27-24)-0}{0.207} = 14.49$$

TInterval $(0, \sqrt{10}, 10, 0.95)$ gives us $(-2.262, 2.262)$.

$t = 14.49$ is outside this interval, so we reject H_0.

There is enough evidence to reject the claim that the population means are equal.





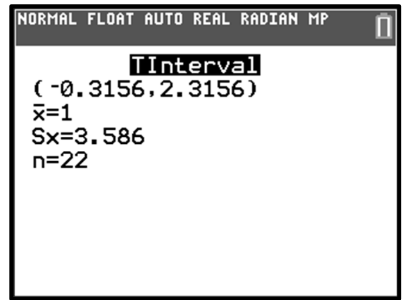
10.6 Dependent Samples

1. TInterval$(0, \sqrt{22}, 22, 0.9)$ or the t-Distribution CV table gives us $t^* = 1.721$.

   $SE = \frac{s_d}{\sqrt{n}} = \frac{3.586}{\sqrt{22}} = 0.765$

   $ME = (CV)(SE) = (1.721)(0.765) = 1.3$

   $1 \pm 1.3 = (-0.3, 2.3)$



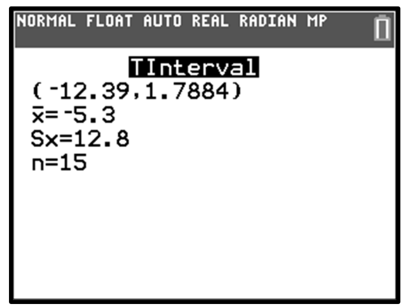
2. a) TInterval$(0, \sqrt{15}, 15, 0.95)$ or the t-Distribution CV table gives us $t^* = 2.145$.

   $SE = \frac{s_d}{\sqrt{n}} = \frac{12.8}{\sqrt{15}} = 3.305$

   $ME = (CV)(SE) = (2.145)(3.305) = 7.1$

   $-5.3 \pm 7.1 = (-12.4, 1.8)$

   b) No. The *CI* includes zero, so a null hypothesis of $\mu_d = 0$ would not be rejected.



3. $H_0: \mu_d \leq 0$ (claim), $H_a = \mu_d > 0$

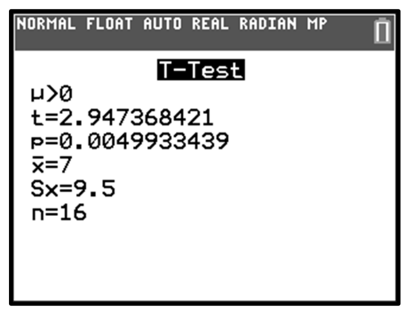
   $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{7 - 0}{\frac{9.5}{\sqrt{16}}} \approx 2.95$

   P-value = tcdf $(2.95, 1\text{E}99, 15) \approx 0.005$

   $0.005 < 0.10$, so reject H_0.

OR $CV = 1.34$ *[*invT$(0.10, 15)$*]*

   $2.95 > 1.34$ is in the rejection region, so reject H_0.



4. $H_0: \mu_d \geq 0$ (claim), $H_a = \mu_d < 0$

   $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{-2.5 - 0}{\frac{3}{\sqrt{14}}} \approx -3.12$

   P-value = tcdf $(-1\text{E}99, -3.12, 13) \approx 0.004$

   $0.004 < 0.01$, so reject H_0.

OR $CV = -2.65$ *[*invT$(0.01, 13)$*]*

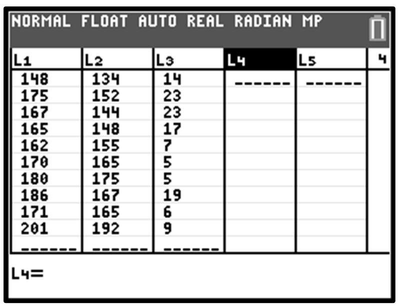
   $-3.12 < -2.65$ is in the rejection region, so reject H_0.

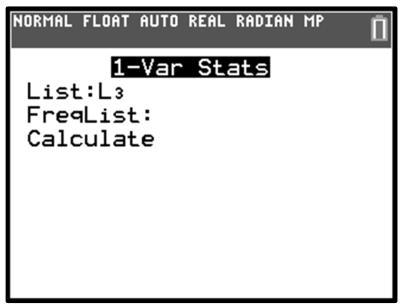
5.

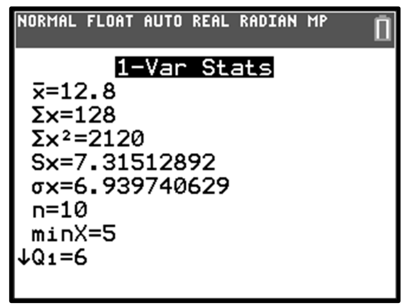
Patient	1	2	3	4	5	6	7	8	9	10	Σ
Before	148	175	167	165	162	170	180	186	171	201	
After	134	152	144	148	155	165	175	167	165	192	
$\boldsymbol{d}$	14	23	23	17	7	5	5	19	6	9	**128**
$\boldsymbol{d-\bar{d}}$	1.2	10.2	10.2	4.2	−5.8	−7.8	−7.8	6.2	−6.8	−3.8	
$\boldsymbol{(d-\bar{d})^2}$	1.44	104.04	104.04	17.64	33.64	60.84	60.84	38.44	46.24	14.44	**481.6**

$$\bar{d} = \frac{\sum d}{n} = \frac{128}{10} = 12.8 \qquad s_d = \sqrt{\frac{\sum(d-\bar{d})^2}{n-1}} = \sqrt{\frac{481.6}{9}} \approx 7.315$$

OR







Claim is that the mean difference is greater than 0, so

H_0: $\mu_d \leq 0$ (claim), $H_a = \mu_d > 0$

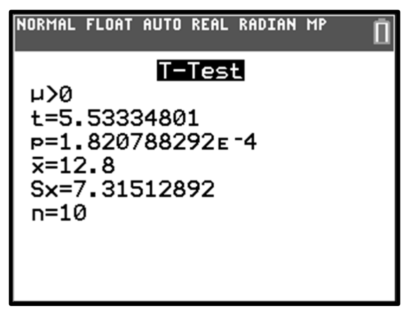
$$t = \frac{\bar{d}-\mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{12.8-0}{\frac{7.315}{\sqrt{10}}} \approx 5.53$$

P-value = tcdf (5.53, 1E99, 9) ≈ 0.00018

$0.00018 < 0.05$, so reject H_0.

OR $CV = 1.83$ [invT(0.05, 9)]

$5.53 > 1.83$ is in the rejection region, so reject H_0.



There is sufficient evidence to support the claim that the drug reduces blood pressure.

Chapter 11 Regression

11.1 Correlation Coefficient

1. (1) 0.89	2. (4) 0.90
3. (2) There is a positive slope.	4. (3) There is a negative slope.
5. (4)	6. (2) -0.24 It is a weak correlation.
7. a. 0.90 b. –0.40 c. 0.99 d. –0.85 e. 0.50 f. 0	
8. (1) III only I is false because correlation does not imply causality. II is false because this is a survey, not an experiment	
9. $r = b\left(\frac{s_x}{s_y}\right) = 1.75\left(\frac{19.5}{38.0}\right) \approx 0.90$	10. $b = -4.3$ $r = b\left(\frac{s_x}{s_y}\right) = -4.3\left(\frac{2.26}{12.48}\right) \approx -0.78$
11. $r = 1$; perfect positive correlation	
12. $r \approx 0.371$; weak positive correlation	
13. $r \approx -0.860$; strong negative correlation	
14. $r \approx 0.986$; very strong positive correlation	
15. $r \approx -0.999$; nearly perfect negative correlation	

11.2 HT for Correlation Coefficient

1. $H_0: \rho = 0$ and $H_a: \rho \neq 0$

   $df = 13 - 2 = 11$

   $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.923}{\sqrt{\frac{1-0.923^2}{11}}} \approx 7.955$

   $P = 2 \times \text{tcdf}\,(7.955, 1\text{E}99, 11) \approx 0.000007$

   $P < 0.01$ so reject H_0

   There is sufficient evidence to support that a significant linear correlation exists.

2. $H_0: \rho = 0$ and $H_a: \rho \neq 0$

   $df = 8 - 2 = 6$

   $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.623}{\sqrt{\frac{1-0.623^2}{6}}} \approx 1.951$

   $P = 2 \times \text{tcdf}\,(1.951, 1\text{E}99, 6) \approx 0.099$

   $P > 0.01$ so fail to reject H_0

   There is *not* enough evidence to conclude that a significant linear correlation exists.

3. $H_0: \rho = 0$ and $H_a: \rho \neq 0$

   $df = 50 - 2 = 48$

   $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.405}{\sqrt{\frac{1-0.405^2}{48}}} \approx 3.069$

   $P = 2 \times \text{tcdf}\,(3.069, 1\text{E}99, 48) \approx 0.0035$

   $P < 0.05$ so reject H_0

   There is sufficient evidence to support that a significant linear correlation exists.

4. $H_0: \rho = 0$ and $H_a: \rho \neq 0$

   $df = 10 - 2 = 8$

   $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{-0.15}{\sqrt{\frac{1-(-0.15)^2}{8}}} \approx -0.429$

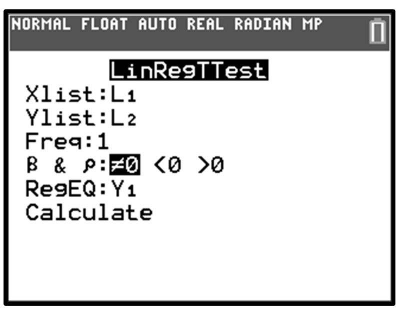
   $P = 2 \times \text{tcdf}\,(-1\text{E}99, -0.429, 8) \approx 0.340$

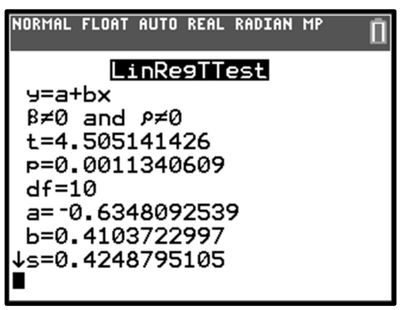
   $P > 0.05$ so fail to reject H_0

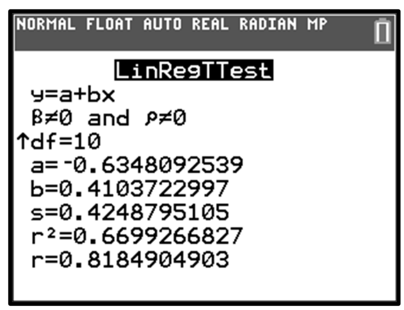
   There is *not* enough evidence to conclude that a significant linear correlation exists.

3. $r \approx 0.818$ and $P \approx 0.001 < 0.01$ so reject H_0

   There is sufficient evidence to conclude that a significant linear correlation exists.







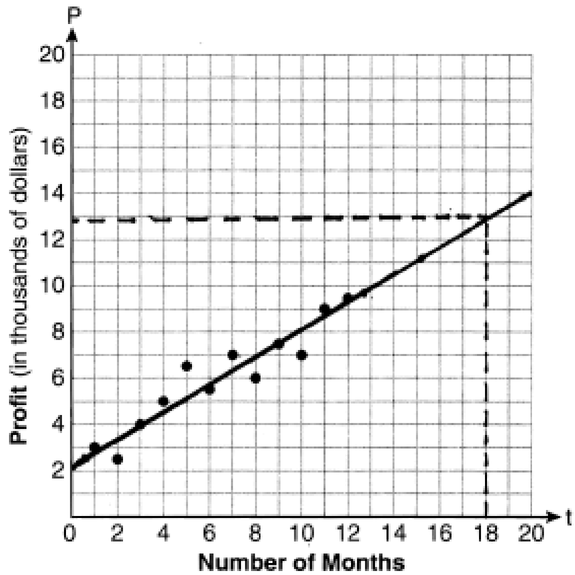
11.3 Linear Regression

1. Line A. Most of the points are closer to Line A than to Line B.

2. a) 80 wpm b) 9 wpm

3. $\hat{y} = 5.14 + 2x$

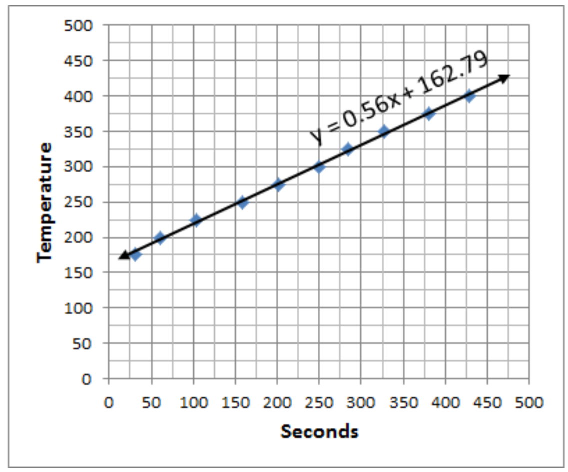
4. $\hat{y} = 0.57x + 2.32$



No, the line crosses near (18,13)

$\hat{y} = 0.57(18) + 2.32 = 12.58$

5.



$\hat{y} = 0.56x + 162.79$

6. a) $\hat{y} = -0.112x + 23.448$

b) $\hat{y} = -0.112(255) + 23.448 \approx -5°C$

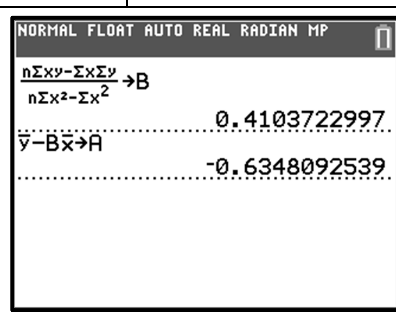
7. a) $\hat{y} = -35.5x + 457.5$

b) $\hat{y} = -35.5(10) + 457.5 \approx 103$

8. $b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} \approx 0.410$

$a = \bar{y} - b\bar{x} \approx -0.635$

$\hat{y} = -0.635 + 0.410x$



9. $b = r\left(\frac{s_y}{s_x}\right) = 0.755\left(\frac{11.35}{13.66}\right) = 0.627$

$a = \bar{y} - b\bar{x} = 78.4 - 0.627(74.7) = 31.5$

$\hat{y} = 31.5 + 0.627x$

11.4 Residuals

1. Actual value – Predicted value = 12,550 – 14,050 = –1,500

2.

Study Time in Hours (x)	Test Score (y)	Predicted Test Score	Residual
0.5	63	62.8	0.2
1	67	67.2	-0.2
1.5	72	71.6	0.4
2	76	76.0	0
2.5	80	80.4	-0.4
3	85	84.8	0.2
3.5	89	89.2	-0.2

3.

a) $y = 0.75(22) - 0.25 = 16.25.$ (Scores cannot be fractional, so 16 is a valid answer.)

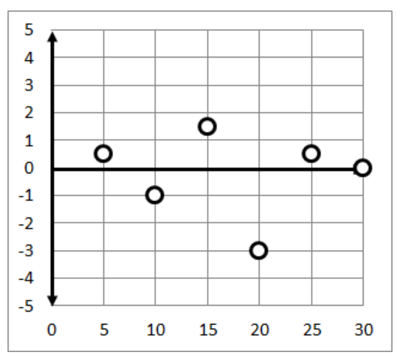
b) $y = 0.75(34) - 0.25 = 25.25$ Residual = 32 – 25.25 = 6.75

c) $y = 0.75(28) - 0.25 = 20.75$ $p - 20.75 = -0.75$

$p = 20$ They scored 20 points.

4.

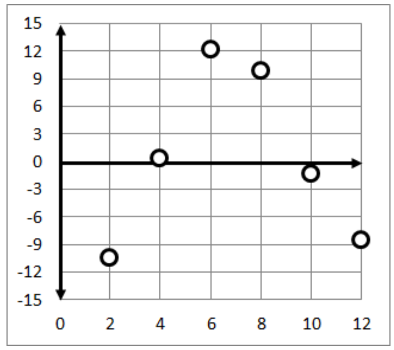
x	y	Predicted Value	Residual
5	3	2.5	0.5
10	4	5.0	-1
15	9	7.5	1.5
20	7	10.0	-3
25	13	12.5	0.5
30	15	15.0	0



Yes. There is no clear pattern in the residual plot.

5.

x	y	Predicted Value	Residual
2	5	15.5	-10.5
4	15	14.7	0.3
6	26	13.9	12.1
8	23	13.1	9.9
10	11	12.3	-1.3
12	3	11.5	-8.5



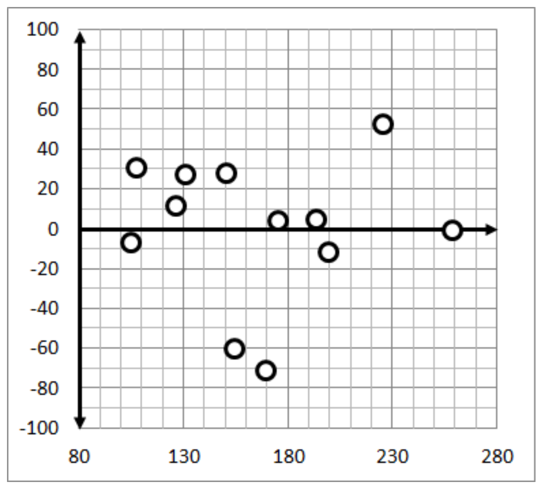
No. There appears to be a parabola-like pattern in the residual plot.

6. a) $y = 0.117x + 83.267$

b) and c)

Distance (miles)	Airfare ($)	Predicted Price ($)	Residual
576	178	150.7	27.3
370	138	126.6	11.4
612	94	154.9	-60.9
1,216	278	225.5	52.5
409	158	131.1	26.9
1,502	258	259.0	-1.0
946	198	193.9	4.1
998	188	200.0	-12.0
189	98	105.4	-7.4
787	179	175.3	3.7
210	138	107.8	30.2
737	98	169.5	-71.5

d)



11.5 Variation

1. For $x_i = 20$, $\hat{y}_i = 52 + 4(20) = 132$
   explained deviation:
   $\hat{y}_i - \bar{y} = 132 - 100 = 32$
   unexplained deviation:
   $y_i - \hat{y}_i = 150 - 132 = 18$
   total deviation: $32 + 18 = 50$
   $OR \quad y_i - \bar{y} = 150 - 100 = 50$

2. a) The point $(\bar{x}, \bar{y})$ is on the regression line, so $\bar{y} = 18 + 2.5(312) = 798.$
   b) For $x_i = 300$, $\hat{y}_i = 18 + 2.5(300) = 768$
   explained deviation:
   $\hat{y}_i - \bar{y} = 768 - 798 = -30$
   unexplained deviation:
   $y_i - \hat{y}_i = 725 - 768 = -43$
   total deviation: $(-30) + (-43) = -73$
   $OR \quad y_i - \bar{y} = 725 - 798 = -73$

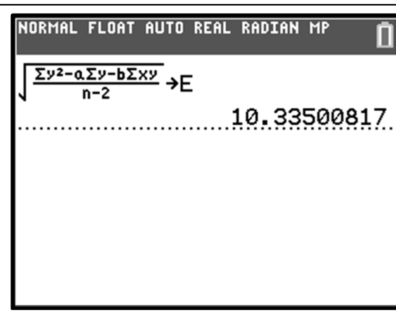
3. total variation $= 2.64829 + 0.08004 = 2.72833$
   $r^2 = \frac{explained\ variation}{total\ variation} = \frac{2.64829}{2.72833} = 0.97066$
   Since the slope is positive, r is positive, so $r = \sqrt{0.97066} = 0.985$

4. a) $b = r\left(\frac{s_y}{s_x}\right) = 0.9\left(\frac{1.2}{3.6}\right) = 0.30$
   Every month, a child's weight is expected to increase by 0.30 kg.
   b) $r^2 = (0.9)^2 = 0.81 = 81\%$

5. (4) $r^2 = (0.718)^2 \approx 0.516 = 51.6\%$

11.6 Prediction Intervals

1. $SE = \sqrt{\frac{\sum y^2 - a\sum y - b\sum xy}{n-2}} \approx 10.335$



2. $df = 10 - 2 = 8$

$t^* = 1.397$

$\hat{y} = 49.78 + 6.27x$

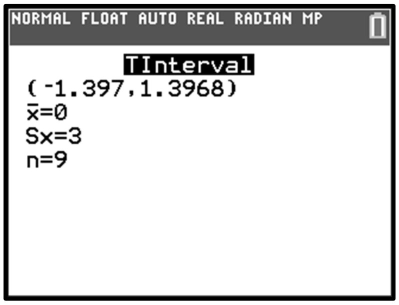
$s = 10.335$

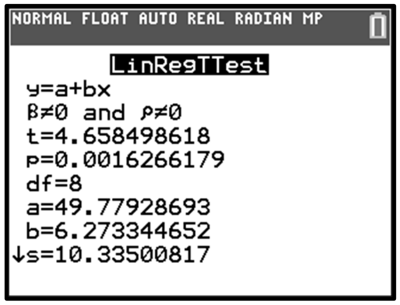
$\hat{y} = 49.78 + 6.27(4) \approx 74.9$

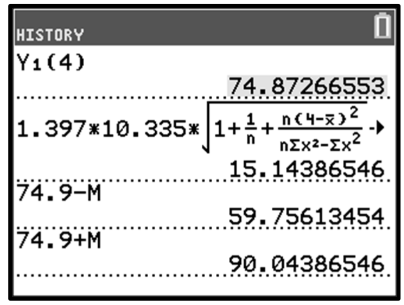
$$ME = t^* s \sqrt{1 + \frac{1}{n} + \frac{n(4-\bar{x})^2}{n\sum x^2 - (\sum x)^2}} \approx 15.1$$

$74.9 - 15.1 < \hat{y} < 74.9 + 15.1$

$(59.8, 90.0)$







3. $n = 6$, so $df = 6 - 2 = 4$

$t^* = 2.776$

$\hat{y} = 0.0298 + 0.9894x$

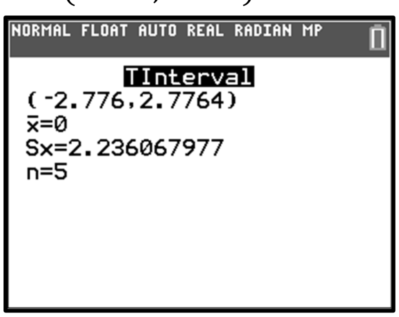
$s = 0.199$

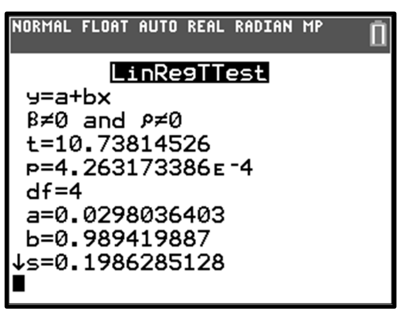
$\hat{y} = 0.0298 + 0.9894(3.15) \approx 3.1465$

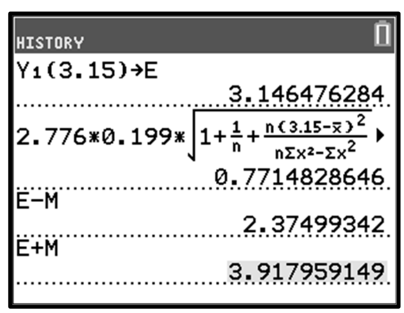
$$ME = t^* s \sqrt{1 + \frac{1}{n} + \frac{n(3.15-\bar{x})^2}{n\sum x^2 - (\sum x)^2}} \approx 0.7715$$

$3.1465 - 0.7715 < \hat{y} < 3.1465 + 0.7715$

$(2.375, 3.918)$







Chapter 12 Chi-Square Tests

12.1 Goodness-of-Fit

1. H_0: The distribution of games played fits the expected proportions. (claim)
   H_1: The distribution of games played differs from the expected proportions.

games played	4	5	6	7
observed (O_i)	9	10	13	18
expected (E_i)	6.25	12.5	15.625	15.625
$\frac{(O_i - E_i)^2}{E_i}$	1.21	0.5	0.441	0.361

   $df = 4 - 1 = 3$
   $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 2.512$ *[the sum of the last row above]*
   P-value = χ^2cdf $(2.512, 1\text{E}99, 3) \approx 0.47$
   P-value > 0.05, so fail to reject H_0.
   There *is not* enough evidence to reject the claim that the distribution fits.

2. H_0: the distribution fits a normal distribution with $\mu = 7$ and $\sigma = 2.415$

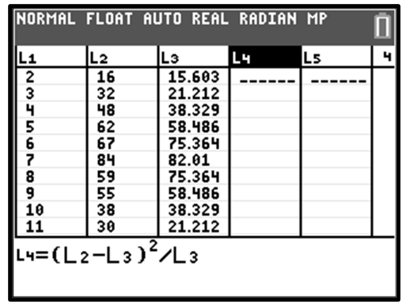
H_a: the distribution does *not* fit a normal distribution with $\mu = 7$ and $\sigma = 2.415$

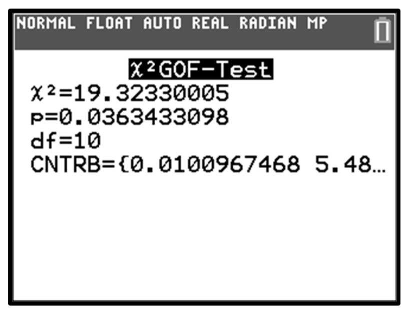
Calculate the expected frequencies using $E_i = n \times \text{normalcdf}\,(lb, ub, \mu, \sigma)$, where the (lb, ub) for each are $(-1\text{E}99, 2.5), (2.5, 3.5), (3.5, 4.5), \ldots (11.5, 1\text{E}99)$.

(If done correctly, $\sum E_i = n = 500$.)

Then calculate $\frac{(O_i - E_i)^2}{E_i}$ for each using the formula as shown in the screenshot below.

roll (x)	**observed** (O_i)	**expected** (E_i)	$\frac{(O_i - E_i)^2}{E_i}$
2	16	15.60	0.010
3	32	21.21	5.486
4	48	38.33	2.440
5	62	58.49	0.211
6	67	75.36	0.928
7	84	82.01	0.048
8	59	75.36	3.553
9	55	58.49	0.208
10	38	38.33	0.003
11	30	21.21	3.641
12	9	15.60	2.794





$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \approx 19.32$ and $df = 11 - 1 = 10$

P-value = $\chi^2\text{cdf}\,(19.32, 1e99, 10) \approx 0.036$

$0.036 > 0.01$, so there *is not* enough evidence to reject claim (H_0) of a normal distribution with $\mu = 7$ and $\sigma = 2.415$.

3. H_0: LeBron James' free throw successes, when awarded two shots, follows a binomial distribution with $p = 0.75$.

H_a: The distribution does *not* follow a binomial distribution with $p = 0.75$.

Using $P = {}_nC_x \cdot p^x q^{n-x}$,

$P(X = 0) = {}_2C_0(0.75^0)(0.25^2) = 0.0625$ $E = (0.0625)(200) = 12.5$

$P(X = 1) = {}_2C_1(0.75^1)(0.25^1) = 0.375$ $E = (0.375)(200) = 75$

$P(X = 2) = {}_2C_2(0.75^2)(0.25^0) = 0.5625$ $E = (0.5625)(200) = 112.5$

successes	0	1	2
observed (O_i)	10	60	130
expected (E_i)	12.5	75	112.5
$\frac{(O_i - E_i)^2}{E_i}$	0.5	3.0	2.722

$df = 3 - 1 = 2$

$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 6.22$

P-value = χ^2cdf $(6.22, 1\text{E}99, 2) \approx 0.045$

P-value < 0.05, so reject H_0.

There *is* sufficient evidence to reject the claim that the distribution fits.

12.2 Independence

1. H_0: Children's gender and attendance at the event are independent.

   H_a: Children's gender and attendance at the event are dependent.

EXPECTED	did not attend	attended
boys	40.97	76.03
girls	42.03	77.97

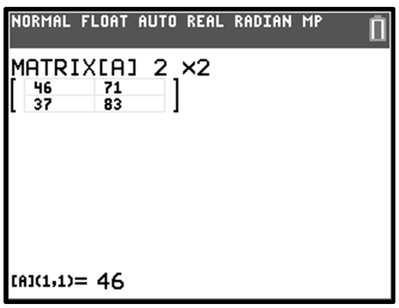
All expected frequencies are at least 5, so the χ^2 test may be used.

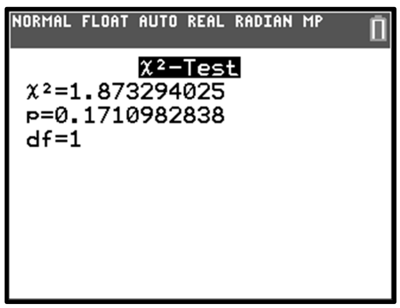
$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \approx 1.873$

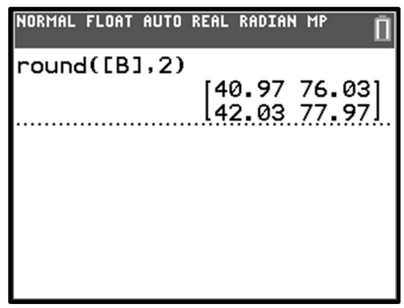
$df = (r-1)(c-1) = (1)(1) = 1$

P-value = $\chi^2\text{cdf}\,(1.873, 1\text{E}99, 1) \approx 0.17$

P-value > 0.05, so fail to reject H_0. There *is not* enough evidence to reject independence.







2. H_0: Students' residence and their opinion on the proposal are independent.

   H_a: Students' residence and their opinion on the proposal are dependent.

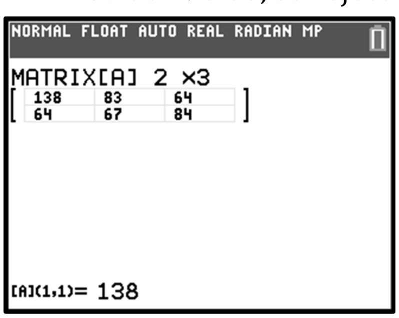
EXPECTED	approve	undecided	disapprove
on-campus residents	115.14	85.5	84.36
off-campus residents	86.86	64.5	63.64

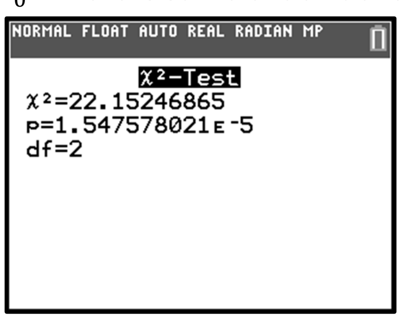
$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \approx 22.152$

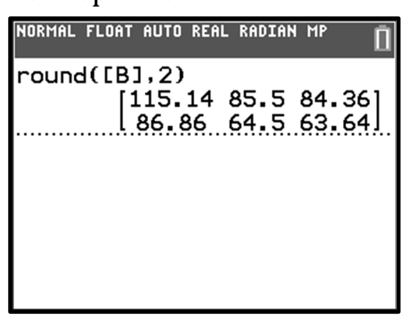
$df = (r-1)(c-1) = (1)(2) = 2$

P-value = $\chi^2\text{cdf}\,(22.152, 1\text{E}99, 2) \approx 0.000015$

P-value < 0.05, so reject H_0. There *is* sufficient evidence of dependence.







3. H_0: Level of education is independent of neighborhood.
H_a: Level of education is dependent on neighborhood.

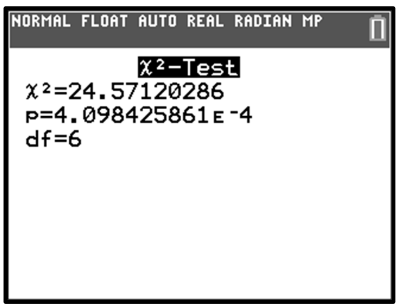
EXPECTED	A	B	C	D
never attended college	80.54	80.54	107.38	80.54
some college	34.85	34.85	46.46	34.85
college graduate	34.62	34.62	46.15	34.62

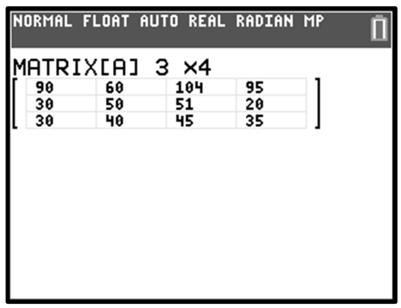
$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \approx 24.571$

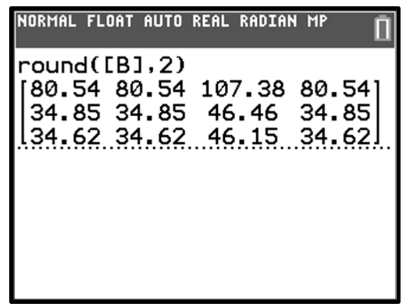
$df = (3-1)(4-1) = 6$

P-value = χ^2cdf $(24.571, 1\text{E}99, 6) \approx 0.0004$

P-value < 0.1, so reject H_0. There *is* sufficient evidence of dependence.

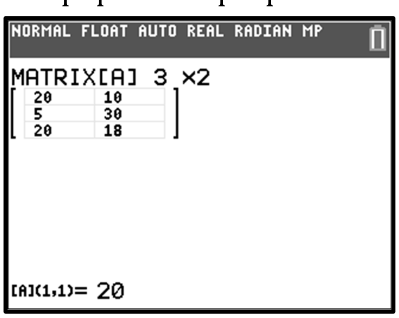


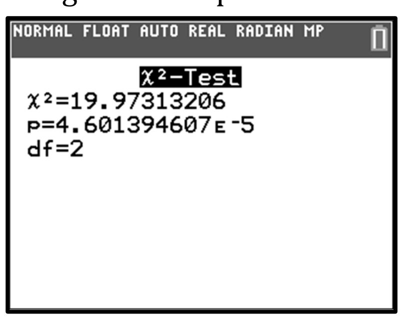


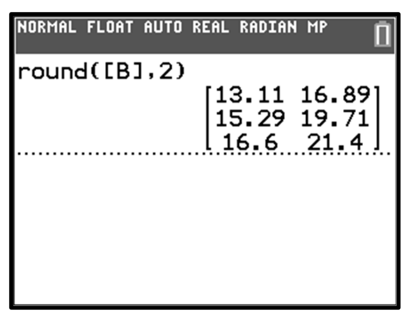


12.3 Homogeneity

1. H_0: The population proportions of those in favor of instituting an estate tax are the same among the three political affiliations.
H_a: At least one of the populations has a different proportion than the others.
$\chi^2 \approx 19.973$, $df = 3 - 1 = 2$, *P-value* = χ^2cdf $(19.973, 1\text{E}99, 2) \approx 0.000046$
P-value < 0.05, so reject H_0. There *is* sufficient evidence to *reject* the claim that the population proportions among the three political affiliations are the same.







2. H_0: The distributions of technology use are the same for high school and college students.
   H_a: The distributions of technology use are *not* the same for high school and college students.
   $\chi^2 \approx 5.423$, $df = (2-1)(3-1) = 2$, *P-value* = $\chi^2\text{cdf}\,(5.423, 1\text{E}99, 2) \approx 0.066$
   P-value > 0.05, so fail to reject H_0.
   There *is not* enough evidence to reject the claim that the distributions of technology use are the same for high school and college students.