Stat 202 **Name (Print):** _____
Summer 2016
Midterm Practice Exam
7/14/16
Time Limit: 120 Minutes

This exam contains 6 pages (including this cover page) and 5 problems. Check to see if any pages are missing. Enter all requested information on the top of this page, and put your initials on the top of every page, in case the pages become separated.

You may *not* use your books, or notes, or cell phone. A calculator is OK as long as it has no internet. You may use the browser on the lab computer (but not your computer) to access StatCrunch and to access my web page http://www.seancarver.org/. Access to my web page is granted for the sole purpose of downloading the data that go with this exam. No other computer use is allowed.

You are required to show your work on each problem on this exam. The following rules apply:

- **Organize your work**, in a reasonably neat and coherent way, in the space provided.
- **Mysterious or unsupported answers will not receive full credit.** A correct answer, unsupported by calculations, explanation, or algebraic work will receive no credit; an incorrect answer supported by substantially correct calculations and explanations might still receive partial credit.
- If you need more space, use the back of the pages; clearly indicate when you have done this.
- Do not write in the table to the right.

Problem	Points	Score
1	20	
2	20	
3	30	
4	10	
5	20	
Total:	100	

1. (20 points) The cases of the motor trend cars data set (available on my website) are 32 different popular makes and models of cars. The first two columns of the data set (after the label) are "mpg" and "cyl." The first of these columns, "mpg," is the fuel efficiency of the car in miles per gallon. The second of these columns, "cyl," is the number of cylinders in its engine.

   (a) (10 points) Describe the distribution of the "mpg" variable.

   (b) (10 points) Describe the distribution of the "cyl" variable.

2. (20 points) This problem concerns the motor trend cars data set and the variable mpg. Specifically, we ask for suspected outliers where the usual 1.5 IQR Rule is used to identify suspected outliers.

   (a) (10 points) How many suspected outliers are there when all values of cyl are lumped together?

   (b) (10 points) How many suspected outliers are there for each value of cyl, considered separately?

3. (30 points) A four sided die has faces that are red, blue, yellow and purple. If you roll a red you get 10 dollars, if you roll a blue you lose 1 dollar, and if you roll a yellow also lose 1 dollar, and if you roll a purple, you lose 5 dollars. The die is weighted so that the purple side comes up twice as often as the other sides. The other sides have equal probability of appearing. Let X be the random variable whose value is the amount of money you win (i.e. X is negative if you lose money.

   (a) (10 points) Write a probability table for X.

   (b) (10 points) Find the mean of X.

   (c) (10 points) If you played the game repeatedly, would you expect to win money or lose money?

4. (10 points) In a particular game of chance, a six-sided fair die (with faces number 1-6) and a four-sided fair die (with faces numbered 1-4) are cast. You win points for each toss. The amount of points you win is the sum of the numbers on the faces minus 2. What is the mean of the amount of points you win?

5. (20 points) In this problem, the sample space is the integers between 1 and 10. The even outcomes are $\{2, 4, 6, 8, 10\}$. The odd outcomes are $\{1, 3, 5, 7, 9\}$. The prime outcomes are $\{1, 2, 3, 5, 7\}$.

   (a) (5 points) What is the complement of the prime outcomes?

   (b) (5 points) What is the intersection of the even outcomes and the prime outcomes?

   (c) (5 points) What is the union of the even outcomes and the odd outcomes?

   (d) (5 points) Are the prime outcomes and the even outcomes disjoint?

Stat 202
Summer 2016
Midterm Practice Exam
7/14/16
Time Limit: 120 Minutes

Name (Print): Solutions

This exam contains 6 pages (including this cover page) and 5 problems. Check to see if any pages are missing. Enter all requested information on the top of this page, and put your initials on the top of every page, in case the pages become separated.

You may *not* use your books, or notes, or cell phone. A calculator is OK as long as it has no internet. You may use the browser on the lab computer (but not your computer) to access StatCrunch and to access my web page http://www.seancarver.org/. Access to my web page is granted for the sole purpose of downloading the data that go with this exam. No other computer use is allowed.

You are required to show your work on each problem on this exam. The following rules apply:

- **Organize your work**, in a reasonably neat and coherent way, in the space provided.
- **Mysterious or unsupported answers will not receive full credit.** A correct answer, unsupported by calculations, explanation, or algebraic work will receive no credit; an incorrect answer supported by substantially correct calculations and explanations might still receive partial credit.
- If you need more space, use the back of the pages; clearly indicate when you have done this.
- Do not write in the table to the right.

Problem	Points	Score
1	20	
2	20	
3	30	
4	10	
5	20	
Total:	100	

1. (20 points) The cases of the motor trend cars data set (available on my website) are 32 different popular makes and models of cars. The first two columns of the data set (after the label) are "mpg" and "cyl." The first of these columns, "mpg," is the fuel efficiency of the car in miles per gallon. The second of these columns, "cyl," is the number of cylinders in its engine.

   (a) (10 points) Describe the distribution of the "mpg" variable.

You can generate a frequency table with stats → summary stats "pick n" for statistics "pick mpg" for columns

10 to 15	5
15 to 20	13
20 to 8	8
25 to 30	2
30 to 35	4

→ out of 32

"pick group by" mpg

hit ok for query about binning

distribution is what values it takes and how often it takes them

   (b) (10 points) Describe the distribution of the "cyl" variable.

cyl is a categorical variable

mpg is a quantitative variable

for categorical can also do bar plot / pie chart

for quantitative can also do histogram / stem plot

also box plot

pick for frequency table n for statistics

cyl for group by and cyl for columns

cyl	
4	11
6	7
8	14

2. (20 points) This problem concerns the motor trend cars data set and the variable mpg. Specifically, we ask for suspected outliers where the usual 1.5 IQR Rule is used to identify suspected outliers.

   (a) (10 points) How many suspected outliers are there when all values of cyl are lumped together?

no outliers

Use boxplot with fences

count outliers as dots/asterisks

↗

appears as dots

in Statcrunch

   (b) (10 points) How many suspected outliers are there for each value of cyl, considered separately?

only outlier is for 8 cyl

especially low

Use boxplot with fences and

Use group by cyl

3. (30 points) A four sided die has faces that are red, blue, yellow and purple. If you roll a red you get 10 dollars, if you roll a blue you lose 1 dollar, and if you roll a yellow also lose 1 dollar, and if you roll a purple, you lose 5 dollars. The die is weighted so that the purple side comes up twice as often as the other sides. The other sides have equal probability of appearing. Let X be the random variable whose value is the amount ot money you win (i.e. X is negative if you lose money.

(a) (10 points) Write a probability table for X.

Colors	red	blue	yellow	purple
Probability	p	p	p	$2p$

$p+p+p+2p=1$ Algebra: $p=.2$

value	10	-1	-5
probability	.2	.4	.4 ← purple

red ↑ blue or yellow

(b) (10 points) Find the mean of X.

$\mu_X = 10 \times .2 + (-1)(.4) + (-5)(.4)$

$= 2 - .4 - 2 = -.4$ $\boxed{\mu_X = -0.4}$

(c) (10 points) If you played the game repeatedly, would you expect to win money or lose money?

Lose money because mean is negative

4. (10 points) In a particular game of chance, a six-sided fair die (with faces number 1-6) and a four-sided fair die (with faces numbered 1-4) are cast. You win points for each toss. The amount of points you win is the sum of the numbers on the faces minus 2. What is the mean of the amount of points you win?

probability table 4 sided

Values	1	2	3	4
prob	.25	.25	.25	.25

mean of 4 sided $= 1*.25 + 2*.25 + 3*.25 + 4*.25$

$= \frac{1}{4}(1+2+3+4)$

$= \frac{5}{2}$

6 sided similar

$\text{mean} = \frac{1}{6}(1+2+3+4+5+6)$

$= 7/2$

points $= X + Y - 2$

mean points $\mu_{X+Y-2} = \mu_X + \mu_Y - 2$

$= 5/2 + 7/2 - 2 = 4$

5. (20 points) In this problem, the sample space is the integers between 1 and 10. The even outcomes are $\{2, 4, 6, 8, 10\}$. The odd outcomes are $\{1, 3, 5, 7, 9\}$. The prime outcomes are $\{1, 2, 3, 5, 7\}$.

   (a) (5 points) What is the complement of the prime outcomes?

   {4, 6, 8, 9, 10}

   (b) (5 points) What is the intersection of the even outcomes and the prime outcomes?

   {2}

   (c) (5 points) What is the union of the even outcomes and the odd outcomes?

   {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}

   (d) (5 points) Are the prime outcomes and the even outcomes disjoint?

   No have one element (2) in Common,