Final Exam

Statistics 300:
Introduction to Probability and Statistics

Fall Semester 2011
Cosumnes College

Instructor: L.C. Larsen

Instructions

Time: 2 hours and 5 minutes on 12/9, 12/12, or 12/13.

Materials: Open book, notes, homework, etc.

Instruments: Calculator/Laptop of student's choice

No phones or consultants
Except to call the instructor : 346-6324.

Answers to confidence interval problems must include the expression (the formula) in symbolic form and the expression with all of the values inserted in the proper places. Then, the final answer can be calculated by any method or device.

Unless a p-value is given in the problem, each hypothesis test problem must include all four parts of the traditional approach to hypothesis tests, including the expression (the formula) for the test statistic in symbolic form and the expression with the values in the right places. The result can then be calculated by whatever method you like (TI-83, laptop computer, etc.).

If more space is needed for a problem, continue your work on the back of the page.

Statistics 300 **Fall 2011**
Exam #3 **Name:** _____ **Tue/Thu 7:00 - 9:05 pm**

(9 points; 10 minutes)

1. **Use the row percentages in the table to test the idea that the percentage of Phoenix Suns fans that live in California is the same as the percentage of Sacramento Kings fans that live in Arizona.**

   Use a 5% significance level for this test.

   The data represent truly random samples of Suns, Kings, and Sonics fans.

Favorite Basketball Team	Home State			Row Total
	AZ	CA	WA	
Phoenix Suns	68%	15%	12%	190
Sacramento Kings	21%	68%	7%	191
Seattle Sonics	11%	18%	81%	219

H_0: _____

H_1: _____

Statistics 300 **Fall 2011**
Exam #3 **Name:** _____ **Tue/Thu 7:00 - 9:05 pm**

(8 points; 8 minutes)

2. **Is there a linear relationship between daily average temperature and daily average wind speed? Use the data in the table for a random sample of five daily values to test the claim that mean temperature and mean wind speed are negatively correlated.** (Let $\alpha = 0.10$ for this test.)

Day	Mean Temp. °F	Mean Speed m/s
1	91.9	15.7
2	81.4	13.8
3	93.2	21.5
4	70.8	33.6
5	100	2.1

Claim: _____

_____

H_0: _____

H_1: _____

Statistics 300 **Fall 2011**
Exam #3 **Name:** _____ **Tue/Thu 7:00 - 9:05 pm**

(9 points; 10 minutes)

3. **Use the summary statistics for a random selection of Fridays and Saturdays to test the claim that the average number of cars on a Sacramento freeway is at least 1000 more on Fridays than it is on saturdays.** (Use a 0.025 significance level for this test.) **Differences in average traffic on Fridays are known to be larger than they are on Saturdays.**

H_0: _____

H_1: _____

Sample Statistic	Fridays	Saturdays
N =	10	16
Average =	38,378	36,811
Standard Deviation =	838	901

Statistics 300
Exam #3

Name: _____

Fall 2011
Tue/Thu 7:00 - 9:05 pm

(9 points; 10 minutes)

4. Use the survey results given in this problem to test the claim that the proportion of prison inmates who return to prison after being released is independent of the type of crime for which they were convicted. Use a Type I error rate of 0.05 for this test.

Type of Crime	Returned to Prison	
	Yes	No
Violent Felony	35	65
Non-violent Felony	26	74
Violent Misdemeanor	31	69
Non-violent Misdemeanor	32	68

H_0: _____

_____

_____

H_1: _____

_____

_____

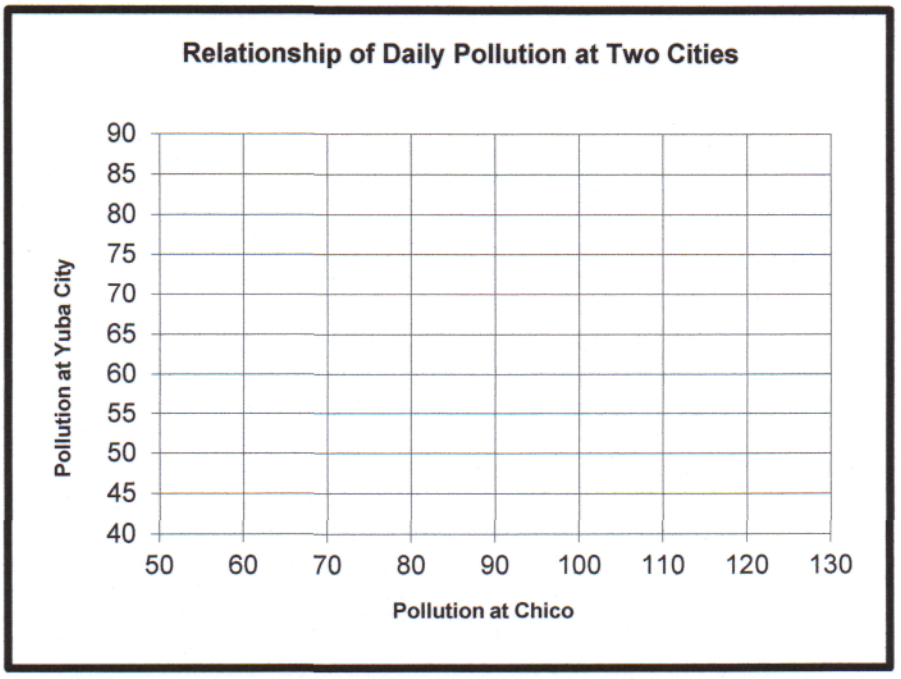
Statistics 300
Exam #3 **Name:** _____

Fall 2011
Tue/Thu 7:00 - 9:05 pm

(13 points; 13 minutes)

5. Daily air pollution measurements from communities that are near one another usually have a linear relationship to one another. Use the data for Chico and for Yuba City to answer the questions on this page and the next page.

Day	Yuba City	Chico
1	75	103
2	84	96
3	88	128
4	57	75
5	46	51
6	54	82



(a) Plot the points on the graph.

(b) Use your calculator to determine the equation of the line that best predicts pollution at Yuba City based on pollution at Chico.

equation = _____

(c) Plot your line on the graph.

(d) What is the predicted pollution at Yuba City when pollution at Chico is 100? _____

(e) Estimate the correlation of pollution at Chico and Yuba City on all days? _____

(f) What proportion of the variation in pollution levels at Yuba City for this set of six days is explained by the levels of pollution at Chico? _____

(g) For the "total" variation in the Yuba City pollution data:

The expression is: _____ **The value is:** _____

(h) For the "explained" variation in the Yuba City pollution data:

The expression is: _____ **The value is:** _____

(i) For the "unexplained" variation in the Yuba City pollution data:

The expression is: _____ **The value is:** _____

(2 points; 2 minutes)

6. *Continue using the Chico and Yuba City pollution data to answer the questions below.*

(a) For the "standard error of estimate" in relating Yuba City pollution to Chico pollution:

The expression is: _____ **The value is:** _____

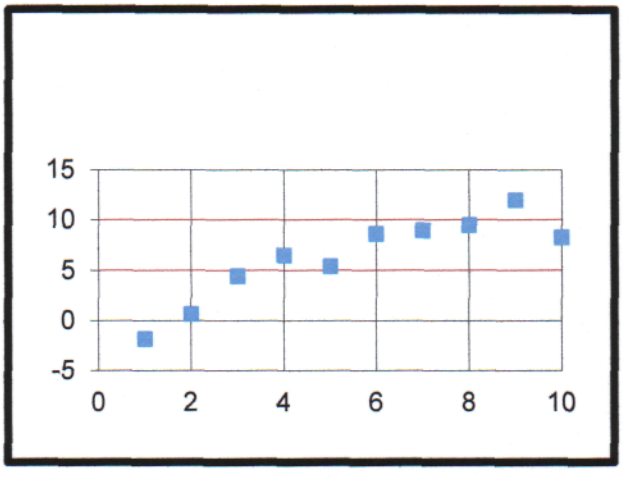
Statistics 300
Exam #3 **Name:** _____

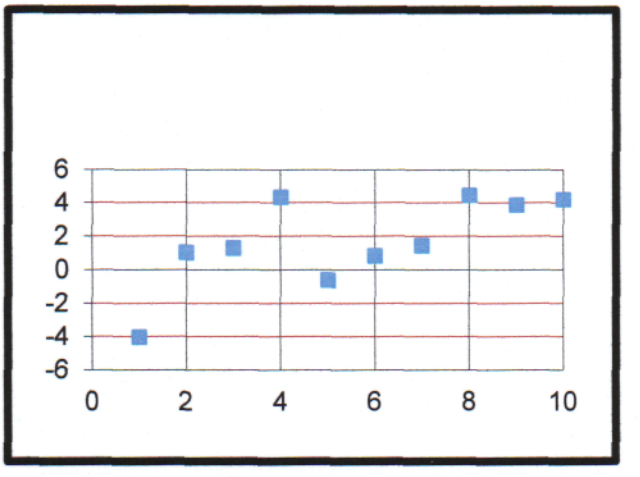
Fall 2011
Tue/Thu 7:00 - 9:05 pm

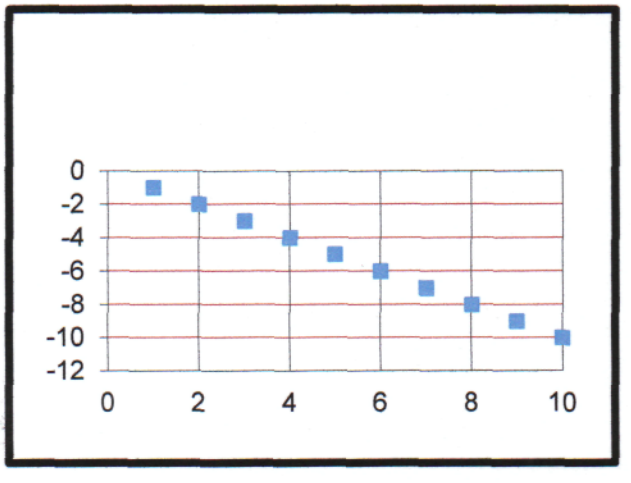
(6 points; 6 minutes)

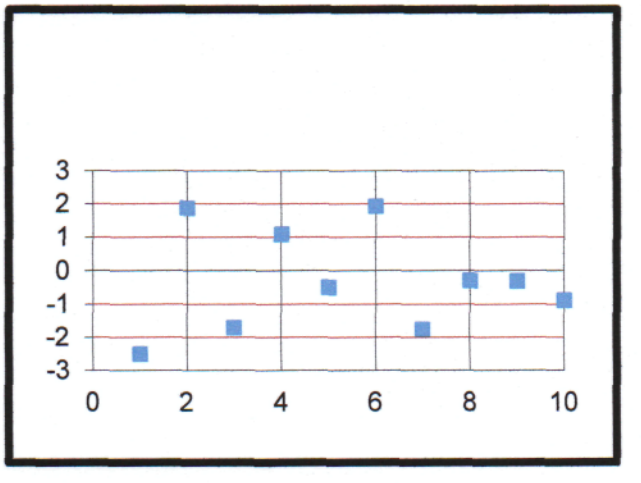
8. *Connect each picture with one of the candidate "r" values by writing the appropriate candidate "r" value in the space at the top of each graph.*

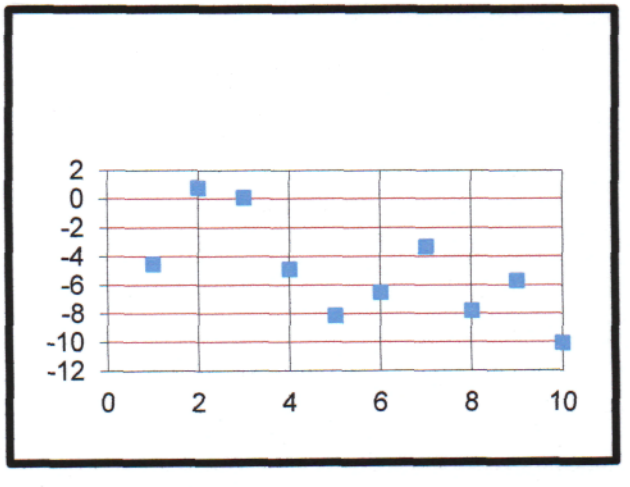
Candidate values of "r", the sample correlation coefficient.						
0.00	**-0.70**	**-0.90**	**-1.00**	**0.70**	**0.90**	**1.00**

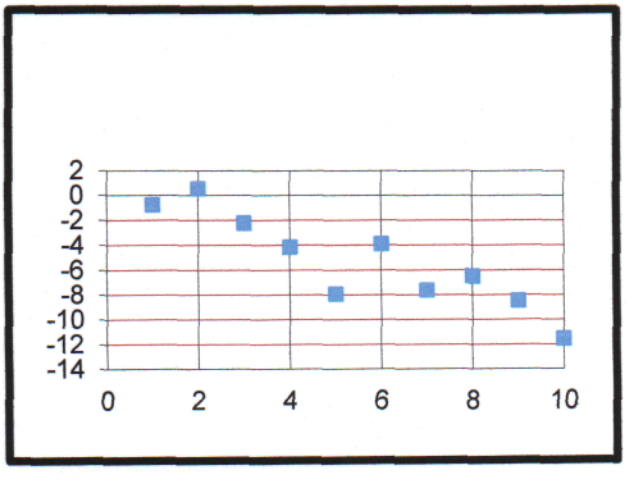












(9 points; 9 minutes)

9. Based on the statistics shown below, construct an 84% confidence interval for the difference between the percentage of 15 year old girls that have a personal cell phone and the percentage of 15 year old boys that have a personal cell phone. (For the test, let $\alpha = 0.05$.)

Sample Statistics

Personal Cell Phone	15 Year Old Girls	15 Year Old Boys
Yes	90	51
No	45	41

Based on your interval is it reasonable to claim that the percentage of 15 year old boys that have a personal cell phone is greater than the percentage of 15 year old girls that have a personal cell phone?

Yes **No** **Why?** _____

_____

_____

Statistics 300 **Fall 2011**
Exam #3 **Name:** _____ **Tue/Thu 7:00 - 9:05 pm**

(9 points; 7 minutes)

10. Use the information on this page to complete the Analysis of Variance table and test the claim that milk from nine different producers has the same average "shelf life" (number of days before milk goes bad). Use a 10% significance level for the test.

AOV Table

Source	SS	df	MS	F	p-value
Producer			9.818		0.0474
Error					
Total	671.41				

H_0: _____

H_1: _____

Shelf Lives (in "days") of Milk samples from Nine Producers

	A	B	C	D	E	F	G	H	I
	15.6	14.7	17.2	16.1	14.5	16.2	14.1	13.2	17.6
	11.1	11.7	13.5	15.5	16.0	12.7	11.7	15.7	14.0
	14.7	17.9	14.1	10.4	14.9	14.4	16.6	16.2	12.5
	16.2	16.1	14.2	11.8	13.9	13.4	14.4	18.4	12.7
	18.2	16.2	13.9	12.6	17.1	13.7	10.4	13.8	15.8
	14.8	13.1	17.1	12.0	14.1	16.4	11.9	18.5	12.6
	15.3	12.9	14.0	13.7	13.8	15.8	10.9	16.1	12.0
	14.5	12.0	14.9	14.4	15.2	8.7	14.3	9.0	13.7
	14.3	13.1	17.2	17.2	17.8	14.6	10.3	13.6	14.8
	16.3	12.3	15.4	11.4	16.3	17.7	14.1	16.6	12.6
	13.8	13.0	13.2	15.9		16.7	12.7	14.8	16.5
	18.8	13.9	18.2	16.6		14.3	12.9	13.5	17.6
	11.6	18.1	13.9	10.9			12.4	13.1	10.9
	12.6	14.7	18.7	14.2			7.3	14.5	16.1
	12.5	13.4	14.0	13.0			12.2	14.3	
	10.5	10.7					13.3	17.6	
		13.1					18.0		
n =	16	17	15	15	10	12	17	16	14
x =	14.4	13.9	15.3	13.7	15.4	14.6	12.8	14.9	14.2
s =	2.4	2.1	1.9	2.2	1.4	2.4	2.5	2.4	2.2

Statistics 300 **Fall 2011**
Exam #3 **Name:** _____ **Tue/Thu 7:00 - 9:05 pm**

(9 points; 10 minutes)

11. Two programs for encouraging school attendance were studied at five schools. At each shool, half of the students were randomly assigned to Method A and the other half were assigned to Method B. Use the data below to prepare a 98% confidence interval for the difference between the population means for the two methods.

1000's of Student-Days of Attendance		
School	**Method A**	**Method B**
1	70.4	69.4
2	74.9	78.9
3	64.3	68.3
4	80.8	83.8
5	76.3	78.3
mean =	73.3	75.7
st. dev. =	6.27	6.65

Statistics 300
Exam #3

Name: _____

Fall 2011
Tue/Thu 7:00 - 9:05 pm

(8 points; 8 minutes)

12. Five schools competed for best daily attendance. The competition lasted for 180 days. Use the results below to test the claim that all of the schools were equally likely to win on each of the 180 days during the competition. (Let alpha be 0.025 for this test.)

School	Count of Days school won
A	36
B	29
C	37
D	47
E	31
total =	180

H_0:

H_1: