Final Exam

Statistics 300:
Introduction to Probability and Statistics

Summer 2016 Day Class
Cosumnes College

Instructor: L.C. Larsen

Instructions

Time: 2 hours and 10 minutes
Date: 28-Jul-16

Materials: Open book, notes, homework, quiz sets, example problems, study guides, etc.

Instruments: Calculator/tablet/Laptop of student's choice

No phone calls or consultants except for questiods addressed to the instructor.

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 a p-value is given in the problem, the conclusion must be based on a proper comparison of the p-value to the significance level.

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

Statistics 300 **Summer 2016**
Exam #3 **Name:** ______________________ **Mon/Tue/Wed/Thu 8:00 - 10:50 am**

(9 points; 9 minutes)

1. **Use the data below to make a 98% confidence interval for the difference between the proportion of people in New York that say they "mainly use public transit" and the proportion of people in Chicago that say they "mainly use public transit".**

	I mainly use public transit	
City	**Yes**	**No**
New York	**730**	**270**
Los Angeles	**230**	**770**
Chicago	**675**	**325**

Based on your confidence interval, is it reasonable to claim that "The proportion of New York residents that mainly use public transit is the same as the proportion of Chicago residents that mainly use public transit" ?

Yes **No** **Why?** ______________________________

____

____

(8 points; 8 minutes)

2. **Use the data shown in the table to test the claim that the time (hours) needed for paint to dry is negatively correlated with temperature.**
   (Let α = 0.025 for this test.)

Temperature ^{o}C	Hours Needed to Dry Paint
5	41
10	35
15	28
20	31
25	27

Claim: ____________________

____

H_0: ____________________

H_1: ____________________

Statistics 300 **Summer 2016**
Exam #3 **Name:** ______________________ **Mon/Tue/Wed/Thu 8:00 - 10:50 am**

(10 points; 10 minutes)

3. **Cancer patients were treated with two anti-cancer drugs: Drug A and Drug B. Use the results shown in the table to test the hypothesis that the two drugs are equally effective in curing cancer, so the percent of patients cured is the same.**
(Use a 10% significance level for this test.)

H_0: ______________________

H_1: ______________________

Response of Cancer Patients

	Drug A	Drug B
Cured	291	136
Not Cured	119	219
Total	410	226

(9 points; 9 minutes)

4. **Pollutants in the water from the City water treatment plant are measured in samples taken during the Spring (May) and Summer (August). Use the statistics given here to make a 95% confidence interval for the difference between the the population mean for pollutant measurments in May and the population mean for pollutant measurementsin August. The variation in May measurements is similar to the variation in August measurements.**

Amount of Pollutants		
Sample Statistics	Month	
	May	August
N =	18	14
Average =	78	84
Std. Deviation =	11	10

Based on your interval is it reasonable to say that the average amount of pollutant in the water in May is 74 and the average amount of pollutant in the water in August is 88?

Yes **Why?** ______________________________

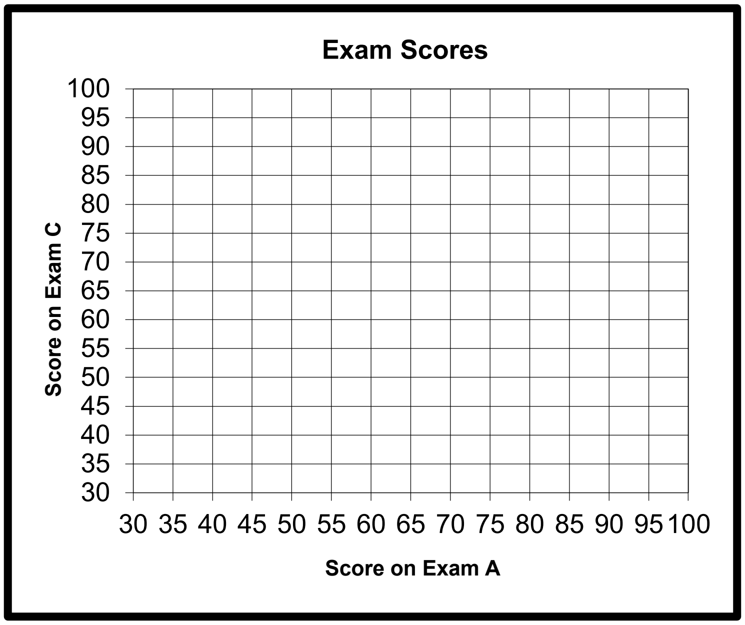
No ______________________________

Statistics 300 **Summer 2016**
Exam #3 **Name:** ______________________ **Mon/Tue/Wed/Thu 8:00 - 10:50 am**

(13 points; 13 minutes)

5. Plot the appropriate test scores as points (score A as X, and score C as Y) on the graph axes. Each row in the dataset is for a different person. Then answer parts b, c, d, e, f, and g.

Person	Exam Scores		
	A	B	C
1	90	84	100
2	78	70	78
3	73	82	62
4	59	52	48
5	38	35	40
6	50	42	65
7	98	92	98



(a) Plot the points on the graph.

(b) Use your calculator to determine the equation of the line that best predicts the Test C score based on a known score for Test A.

Equation of your line : ______________________

(c) Plot your line on the graph.

(d) What is the linear correlation between the scores for Test A and Test C? ____________

(e) Provide the symbolic expressions for Total, Explained, and Unexplained variation in "Y".

____________ = ____________ + ____________
Total Variation **Explained Variation** **Unexplained Variation**

(f) Provide the values for Total, Explained, and Unexplained variation in "Y" for the graphed data.

____________ = ____________ + ____________
Total Variation **Explained Variation** **Unexplained Variation**

(g) Provide the symbolic expression and the value for the Standard Error of Estimate.

____________ ____________
Symbolic Expression **Value**

(9 points; 10 minutes)

6. **Random samples of 600 people from three age groups were asked what size of car they drove -- Large, Medium, or Small. Use the data below to test the claim that all people in the three age-groups drive Large, Medium, and Small cars in the same proportions.**
(Use a 0.025 significance level for this test.)

	Size of Car			Row
Age Group	**Large**	**Medium**	**Small**	**Total**
< 30	**142**	**206**	**252**	**600**
30 to 40	**150**	**220**	**230**	**600**
> 40	**218**	**180**	**202**	**600**
Col. Total	**510**	**606**	**684**	**1800**

Claim: ______________________

H_0: ______________________

____

H_1: ______________________

____

(9 points; 10 minutes)

7. **A City studied the effects of visible police presence on average driving speed. Use the data below to test the claim that the average driving speed is more than five miles per hour faster when police are not visibly present compared to when police are visibly present. Variability in driving speeds increases when police are not visibly present.** (Treat the data as "simple random samples", and let α = 5%.)

Hourly Values for Driving Speed with Police "Visible" and "Not Visible"

Sample Statistic	Visible	Not Visible
n =	31	21
mean =	66.4 mph	72.8 mph
st. dev. =	3.4 mph	5.2 mph

Claim: ______________________

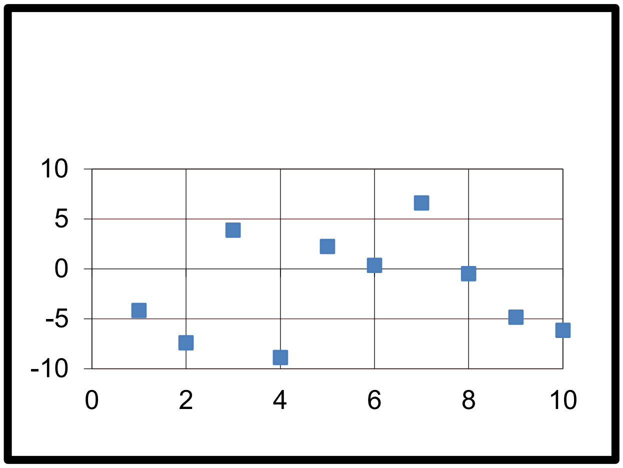
H_0: ______________________

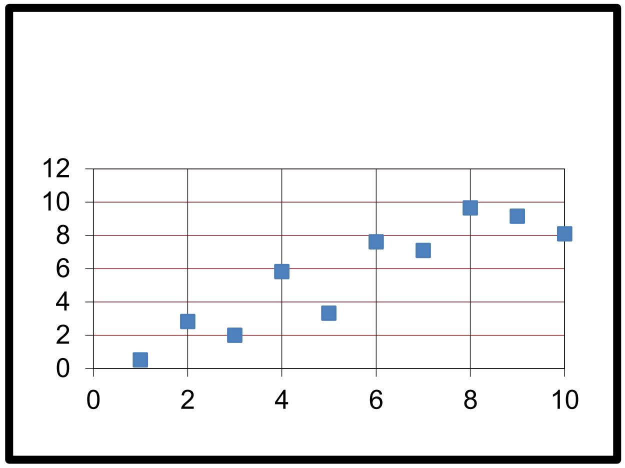
H_1: ______________________

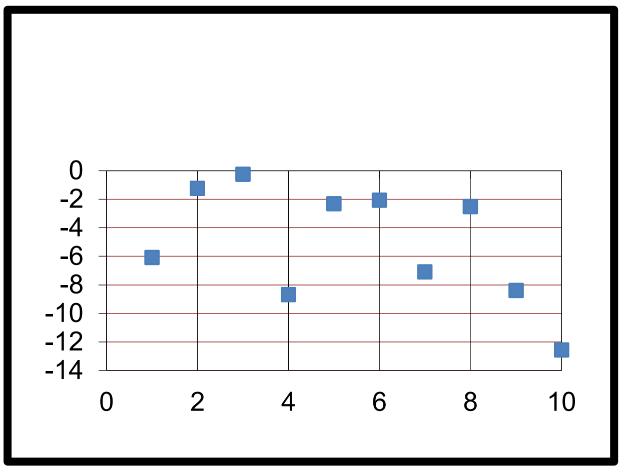
(6 points; 6 minutes)

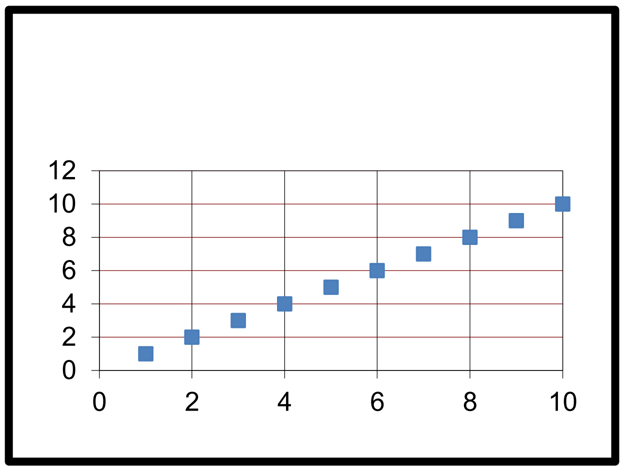
8. Connect each picture with one of the candidate correlation (r) values by writing the appropriate candidate "r" value in the space at the top of each graph.

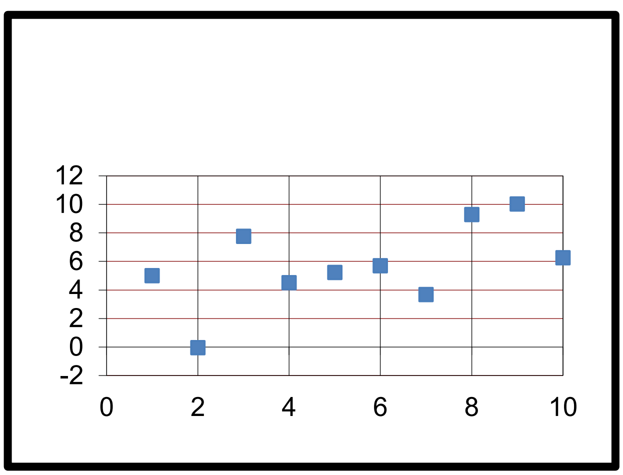
Candidate values of "r", the sample correlation coefficient.							
0.00	**-0.50**	**-0.90**	**-1.00**	**0.50**	**0.90**	**1.00**	**2.00**

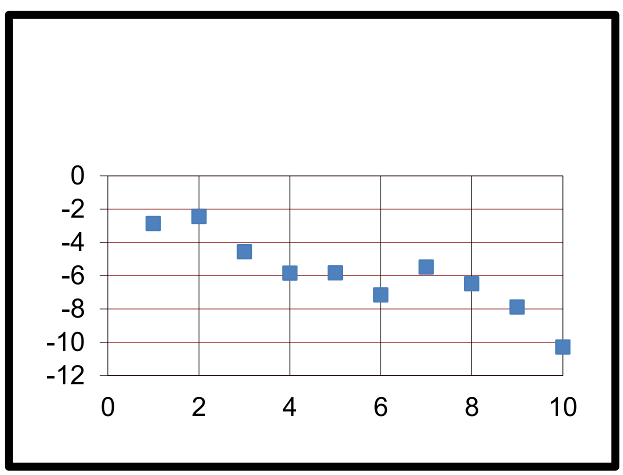












(9 points; 7 minutes)

10. The Analysis of Variance table below is based on the 304 data values on the next page. Complete the AOV table and test the claim that types of glass made by ten different makers of windows lose the same average amount of energy. Use a 2% significance level for the test based on the p-value approach to hypothesis testing.

AOV Table

Source	Sum of Squares	df	Mean Square	F	p-value
Producers	544				0.099718
Error					
Total	11288				

H_0: ______________________

H_1: ______________________

Energy losses through windows made from 10 different types of glass

	Type_1	Type_2	Type_3	Type_4	Type_5	Type_6	Type_7	Type_8	Type_9	Type_10
	96	92	107	99	102	95	103	102	93	98
	88	97	102	101	92	109	95	98	106	107
	107	101	110	99	109	94	93	105	110	106
	100	106	103	108	93	90	94	94	102	90
	96	106	108	104	107	110	95	107	99	95
	90	97	97	102	110	105	98	106	99	91
	91	97	101	98	94	98	90	100	104	98
	93	93	95	102	103	94	105	98	106	93
	91	104	99	105	100	99	98	92	108	101
	97	106	105	92	97	108	97	101	102	109
	98	93	99	90	90	92	105	90	104	104
	90	98	105	99	90	95	110	100	91	99
	95	99	90	99	107	98	100	104	90	93
	96	93	110	109	102	109	108	105	99	92
	99	96	102	91	105	97	96	93	97	93
	106	101	92	96	100	92	90	105	99	99
	90	104	93	100	98	93	97	95	90	95
	89	92	110	90	103	92	106	103	110	99
	99	98	92	98	109	108	101	102	96	103
	89	98	91	96	94	93	108	95	93	92
	103	106	95	108	103	96	110	103	90	90
	100	110	108	105	93	102	96	106	97	93
	103	109	90	91	108	101	95	93	108	92
	94	95	99	93		109	98	91	110	105
	101	94	95	96		100	109	102	106	96
	102	105	102	96		93	96	91	108	
	94	101	110	108		104	104	103	110	
	93	107	93	108		95	95	108	94	
	95	109	95	110		97	96	99	102	
	94	104	108	105		104	99	105	92	
	105	103	96				110	101		
	101	99	98				107	91		
		90					105			
		104					108			
		100					101			
ı =	32	35	32	30	23	30	35	32	30	25
ı =	96.4	100.2	100.0	99.9	100.4	99.1	100.5	99.6	100.5	97.3
/ =	5.4	5.6	6.6	6.1	6.5	6.3	6.0	5.5	6.8	5.7

Statistics 300 | **Fall 2014**
Exam #3 **Name: ______________________** **Tue/Thu 6:00 - 8:05 pm**

(8 points; 8 minutes)

11. **A local newspaper surveyed 900 likely voters and asked each person how likely they were to vote for the candidate currently representing their district. Use the data below to test the claim that the proportions today are the same as the proportions from another taken before the previous election for district representative.** Let $\alpha = 0.05$.

Results of current election survey		Last Survey
Very Likely	284	32%
Somewhat Likely	308	27%
Equally Likely and Unlikely	100	12%
Somewhat Unlikely	120	18%
Very Unlikely	88	11%
Total	900	100%

H_0: ______________________

____

H_1: ______________________

____