

Stat 31-951 Final Exam

Tuesday May 7th, 2002

Name: _____

Instructions:

1. Answer all questions and **show all work** on the exam in the space provided (you may use the backs of pages if necessary). You will **not** receive credit if you do not justify your answers.
2. All multiple choice questions have only one correct answer.
3. Please draw a **circle** or **box** around your final answers.
4. Unless otherwise specified, give exact answers.
5. Remember to sign the Honor Pledge.

I have neither given nor received any unauthorized help on this exam and I have conducted myself within the guidelines of the University Honor Code.

Pledge: _____

1. (10 points) *Aladdin and the flying carpets*

Aladdin asked his Genie to bring him the fastest flying carpet in the kingdom. The genie knows that the speed of flying carpets has a normal distribution with mean $\mu = 150$ leagues per hourglass, and standard deviation $\sigma = 7$ leagues per hourglass.

(a) (10 points) The flying carpet named lightning flies at 164 leagues per hourglass. How likely is it to find a faster flying carpet in the kingdom?

(b) Lightning is from the SilkyFly brand of flying carpets. To test if carpets from this brand are faster than the general flying carpets population, the genie takes a SRS of size 4 of SilkyFly carpets. The sample average speed of this SRS is 160 leagues per hourglass. Assuming that the standard deviation of the SilkyFly carpets is the same as the general population, do you have enough evidence that SilkyFly flying carpets are faster than the general flying carpets population?

2. (15 points) *Ali Baba and the forty thieves*

Ali Baba discovered the cave with the treasure of the forty thieves. He thought that if he stole small portions of the treasure, the thieves would not notice a difference. The captain of the forty thieves weighs all his treasure every week. Knowing that there are always some measurement errors, he likes to weigh his gold 13 times to get the most accurate measurement. This week's 13 weights (in pounds) are given in the following stemplot

21		4			
21		6	8		
22		3	4	4	4
22		5	6	7	
23		0	4		
23					
24		0			

- (a) (15 points) Describe the distribution of the weights.
- (b) Find the Median and the IQR the data.
- (c) If the data were in cells A1 to A13 in your Excel spreadsheet. What Excel command could you write to find the sample mean of the data?
- (d) The average weight of the treasure this week is 225 pounds and the sample standard deviation is 7.036 pounds. Is 225 pounds a good estimate of this week's true weight? and is 7.036 a good estimate of the spread in the measurements? Why or why not?

Continued...

- (e) The forty thieves didn't use any of their treasures this month. For the past month, the weight of their treasure was $\mu = 229$ pounds. The captain of the forty thieves believes some of his treasure was stolen. Do you have enough evidence to support his suspicion?

Continued...

3. (24 points) *Sheherazade the story teller*

Sheherazade, the vizir's daughter, is the best story teller in the kingdom. She speaks at the perfect pace to keep her audience in suspense. This talent, which saved her life after 1001 nights of story telling, was the result of some training. She perfected it by telling stories at different speech paces and see how long it took the audience to fall asleep. In the following data, X are the different speech paces (in words per hourglass) and Y is the time (in hourglass) the audience was listening before falling asleep. The average and standard deviations of the two variables are given as well.

	X	Y
	105	15
	110	75
	115	110
	120	125
	125	105
	130	80
Average	$\bar{x} = 117.5$	$\bar{y} = 47.5$
Standard deviation	$s_x = 9.35$	$s_y = 39.11$

The two following parts are independent of each other but use the same data above.

Part I

The following represents the output of the Excel linear regression on the above data

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-203.71429	199.3056484	-1.02	0.3645	-757.07662	349.64805	-757.07662	349.64805
X Variable 1	2.4571429	1.691756725	1.452	0.22	-2.2399365	7.1542223	-2.2399365	7.1542223

(a) (24 points) Write the linear regression equation.

(b) Is the slope of the linear regression significant? Why or why not?

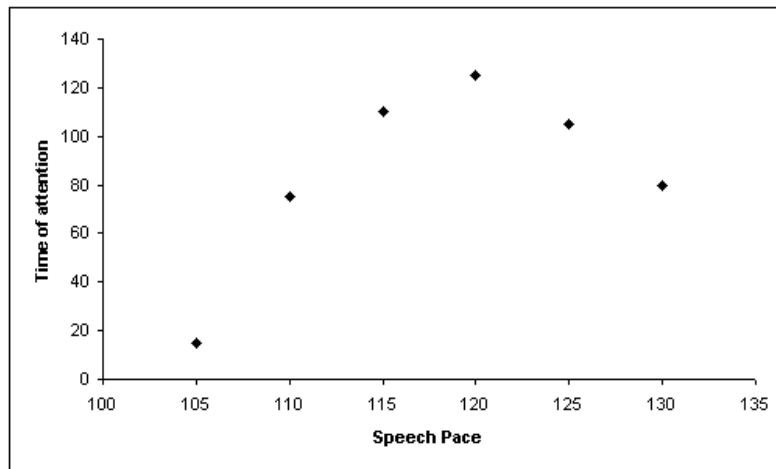
Continued...

- (c) What do you think about the estimate of the intercept and its margin of error?
- (d) Find the correlation of the two variables.
- (e) How much of the variation in the time to sleep is explained by the linear regression of speech pace on time to sleep?
- (f) According to the linear regression model, how long would it take the audience to sleep if Sheherazade spoke at a pace of 117.5 words per hourglass? Are you confident on this estimate? Why or Why not?

Continued...

Part II

The plot of the data is



- (a) (24 points) Describe the relationship between the two variables.
- (b) Is the correlation a good measure of the relationship between the two variables? Why or Why not?

Continued...

4. (6 points) *Where are the Genies ?*

The probability to find a Genie in a lamp is of 0.4, and the probability to find a Genie in Cheb El Hillal town is 0.54.

(a) (6 points) Is the probability to find a Genie in a lamp and in Cheb El Hillal town $(0.4) \cdot (0.54)$? Why or why not?

(b) Is the probability to find a Genie in a lamp or in Cheb El Hillal town $(0.4) + (0.54)$? Why or why not?

Continued...

5. (10 points) *The largest Bird on earth*

In his fifth voyage, Sindbad the sailor encountered the Rukh, the largest bird on earth. The number of eggs the Rukh can lay varies from year to year. Let X be the exact number of eggs the Rukh lays in a year. The probability table of X is given below

X	0	1	2	3	4	5
p	0.08	0.10	0.42	0.25	0.10	0.05

(a) (10 points) What is the probability that the Rukh lays at least 2 eggs in a year?

(b) Find the expected number μ of eggs a Ruck could lay.

(c) Find the standard deviation of X .

Continued...

6. (10 points) *The forty thieves strike back*

Ali Baba hid the treasure he stole from the forty thieves in 13 out of the 200 houses in the village. The forty thieves start looking for the treasure. Each thief chooses randomly one house to search thoroughly. The thieves make their choice independently of each other, so it is possible that two thieves might pick the same house.

- (a) (10 points) What is the probability that at least one of the forty thieves finds a house with some treasure?

- (b) Let X be the number of thieves who find a house with some treasure. What is the expected value of X ?

- (c) What is the standard deviation of X ?

7. (5 points) *Why use large samples?*

In the stat31 class, we have learned how to construct several types of Confidence Intervals based on large sample. The availability of a large sample allows us to ...

- (a) (5 points) ...treat \bar{x} as an (approximately) Normal r.v., regardless of the shape of X's original distribution.
- (b) ...use the Normal distribution as an approximation to the Binomial distribution.
- (c) ...use s as a reasonably good estimate of σ .
- (d) All of the above.
- (e) None of the above.

Continued...

8. (20 points) *Sampling*

A UNC student would like to know how familiar are the Arabian Nights stories in North Carolina. She finds that 20 out of the 50 people in her sample were familiar with some of the Arabian Nights stories.

Part I : Analysis of the data

- (a) (20 points) She uses the Stat31 formula to compute a 98% Confidence Interval of the proportion of people in North Carolina who were familiar with the arabian night stories. What is the confidence interval that she finds?

- (b) Using Stat31 hypothesis testing techniques, she have a strong evidence that the population proportion is significantly different from 60%? Why or Why not?

Part II *Checking the assumptions*

- (a) (20 points) Did she have enough data to use the Normal approximation to the Binomial distribution (use the sample statistic as a guess of the population parameter)? Why or why not?
- (b) She tells you that she selected the people in her sample by taking a SRS of UNC students from the phone book, and 20% didn't return her call or refused to answer the question. Are you confident on her results in Part I based on this sampling procedure? Why or why not (be specific)?
- (c) Based on your answers to the two previous questions and using her results as a preliminary study (i.e. If needed you can use her sample statistic as a guess for the population parameter). First, explain briefly how you would select your sample to conduct the survey. Finally, find how many individuals should be in your sample to get a 98% confidence interval of the target parameter she is interested in with a 0.05 margin of error.