

NAME:

ID# (LAST 4 DIGITS):

**Math 114 Final Examination
December 13, 2006 8AM-noon**

Rules: Calculators, 2 8x11 pages personal notes permitted. Tables provided. No other assistance allowed. Exam is conducted under Tulane Honor Code—all work is to be your own.

Do all work on this exam paper. Show work, except in case of true-false or multiple choice. Unsupported answers receive no credit. If you used a calculator function, indicate clearly which function and indicate the specific numbers put into the function.

1. What does the phrase “this data is statistically significant” mean? (circle most appropriate answer)

- a) We have obtained statistical evidence FOR the null hypothesis.
- b) There is only a statistical chance our proposed conclusion is true
- c) This statistical conclusion is not only true but important
- d) The appropriate null hypothesis has been rejected
- e) None of a)-d) is particularly apt

2. Consider the following employment breakdown in a giant office building:

	Male	Female
Rank and File	200	80
Supervisor	30	10

We choose an employee at random.

- a) What is the probability we get a female supervisor? (Leave answer as a fraction—no arithmetic)
- b) What is the conditional probability we get a supervisor given that we get a female? (leave answer as a fraction)
- c) What is the conditional probability we get a female given that we get a supervisor? (leave answer as a fraction)

3. What do we mean when we say a random variable Z has the standard normal distribution? Be brief and specific.

4. Data is collected from a large sample of people on several variables including the following. For each variable indicate whether it is qualitative or quantitative by circling the appropriate word..

- | | | |
|-----------------------------------|-------------|--------------|
| a) ID Number | qualitative | quantitative |
| b) Hours per week on the Internet | qualitative | quantitative |
| c) Years of education | qualitative | quantitative |
| d) Gender | qualitative | quantitative |

5. We take a random sample from a normal population, get $\bar{x} = 65$, and seek a 95% confidence interval for the population mean μ . What should our confidence interval be

(a) if $n = 150$ and we calculate $s = 19$ from the data

(b) if $n = 14$ and we calculate $s = 19$ from the data.

c) if $n = 14$ and independently of the data we learn $\sigma = 19$ from a highly reliable source.

6. A discrete random variable X has this distribution:

X	1	2	3	4
Prob	.2	.3	.4	.1

Find the mean μ_X and standard deviation σ_X .

7. We want to obtain a 95% confidence interval for the mean μ of a large population, with sampling error (margin of error) no larger than 6. If we believe $\sigma \cong 30$, what (approximately) is the minimum necessary sample size?

8. A population of measurements has the normal distribution with mean $\mu = 95$ and standard deviation $\sigma = 17$.

a) What fraction of the population exceeds 100? Indicate clearly how you got your answer, and show work.

b) Find the 75th percentile of the population. Indicate clearly how you got your answer, and show work.

c) If we take a sample of 8 observations from this population, what is the probability the sample mean \bar{x} exceeds 100?

9. Consider the following partial regression printout from MS Excel for a regression model $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + e$, where e is the usual unobserved normal error term. Note our data consists of $n = 20$ cases.

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.860161
R Square	0.739876
Adjusted R Square	0.691103
Standard Error	4.399645
Observations	20

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	3	880.9161947	293.6387	15.16974	6.13518E-05
Residual	16	309.7100606	19.35688		
Total	19	1190.626255			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	1.380697	3.794267416	0.36389	0.720704
X Variable 1	1.521515	0.342693098	4.439878	0.000412
X Variable 2	1.903537	1.225610503	1.553133	0.139946
X Variable 3	-4.6139	0.985697338	-4.68085	0.00025

a) Of the variables x_1, x_2, x_3 which have predictive power and should be kept in our model, and which might as well be dropped from the model? Of course, justify your answer by reference to the printout.

b) Using the printout, give a 95% confidence interval for the true coefficient of x_1 .

10. The illiteracy rates in two populations are 35% and 25% respectively. We plan to take independent random samples of sizes 700 and 900 respectively from the two populations; let \hat{p}_1, \hat{p}_2 denote the two sample proportions we will obtain. Give the mean and standard deviation of $\hat{p}_1 - \hat{p}_2$.

11. Sixty-five percent of Atlanta voters favor the incumbent mayor for re-election. If we take a random sample of 8 Atlanta voters, what is the probability that 4 or fewer voters in the sample favor the incumbent mayor for re-election?

12. We shuffle a 52 card deck of playing cards and deal out 5 cards. What is the probability we get a “busted flush”? (Recall that a flush is a hand with all 5 cards in the same suit. A “busted flush” is a hand with 4 cards in the same suit but the other card in another suit—i.e., a hand which just misses being a flush.)

13. We want to obtain a 92% confidence interval for the proportion p of Louisiana voters who can name both US Senators from the state. We want our sampling error (margin of error) to be no larger than .035.

a) What is the minimum necessary sample size?

b) If we are very confident this proportion is between .10 and .30, what is now the minimum necessary sample size?

14. Suppose we take independent random samples from two populations and obtain $n_1 = 130, \bar{x}_1 = 345, s_1 = 23, n_2 = 190, \bar{x}_2 = 332, s_2 = 14$. Give a 95% confidence interval for the difference $\mu_1 - \mu_2$ of the population means.

15. Consider the following blood pressure data on 6 pairs of randomly chosen subjects who have been matched according to factors known to influence blood pressure. In each pair, one subject chosen at random has been subjected to Treatment A, the other to Treatment B. Do these data give evidence ($\alpha = .05$) that the mean blood pressure of the entire population of Treatment A subjects differs from the mean blood pressure of the population of Treatment B subjects? Set up and carry out an appropriate hypothesis test, answer whether these data present the sought for evidence, and give a p-value.

Pair Number	1	2	3	4	5	6
Treatment A	85	75	95	110	59	130
Treatment B	75	80	95	115	69	128

16. An urn contains 5 red balls and 3 green balls. We choose 4 balls at random from the urn. What is the probability we get exactly 2 red balls?

17. Suppose termites are present in 10% of houses in a large city. A new noninvasive termite detection machine will report the presence of termites in residential structures with probability .95 if termites are indeed present, but will report termites present with probability .05 if in fact no termites are present. We are interested in knowing the probability that termites are present, given that the new wonder machine says they are present. So answer these two questions:

a) If we choose a home at random from the city and use the new termite machine, what is the probability the machine will report termites present?

(b) If the machine reports termites are present in the home, what is the probability that there are termites present in the home?

18. Consider the following frequency table for a sample of quiz scores:

Score	Frequency	Relative Frequency
10	2	
9	6	
8	15	
7	30	
6	17	

a) Fill in the relative frequencies above.

b) Find the sample mean

c) Find the sample standard deviation s .

19. Suppose that we always use 90% confidence intervals. In 300 such intervals, how many on average will be correct?

20. We have seen in class that a fair pair of dice come up "7" $1/6$ of the time. Cheaters sometimes switch a loaded pair of dice into a game of craps, in order to improve their chance of rolling "7" at a particular time. If we roll a suspect pair of dice 400 times and get "7" 100 times, do we have evidence the dice are loaded? Set up an appropriate hypothesis and alternative, conduct the hypothesis test, answer the question and give a p-value.

21. Suppose we are conducting a hypothesis test with $H_0 : \mu = 40$, $H_a : \mu > 40$.

(a) For this specific test what would be a type I error? (be specific)

(b) For this specific test what would be a type II error? (be specific)

(c) If we accept H_0 what is an appropriate conclusion to report to the layman?

(d) If we reject H_0 what is an appropriate conclusion to report to the layman?

22. Consider these observations on the variables x, y for 7 subjects.

Subject	1	2	3	4	5	6	7
x	22	24	24	37	36	23	17
y	22	26	43	30	56	35	25

a) Draw a scatter plot for these data.

b) Suppose we want to summarize these data by a straight line formula of the form $y \cong \beta_0 + \beta_1 x$. Find the “best” β_0, β_1 (the least squares estimates.)

23. Consider the following sample observations from a large population:

43.3, 23.1, 27.6, 24.9, 18.5

a) Find the sample mean \bar{x} .

b) Find the sample median.

c) Find the sample standard deviation s .

24. A large collection of data has mean 150 and standard deviation 30. .

a) If we believe the data is mound-shaped, i.e., that the data has an approximately normal distribution, what can we say about the proportion falling between 90 and 210? Above 210? Justify your answers.

b) If we know nothing about the distribution of the data, what can we say about the proportion of data falling between 90 and 210? Justify your answers.

25.. Students in a sociology class want to describe and if possible quantify attitudes toward recovery in New Orleans. To do this they select 400 numbers from the telephone book, call each number a maximum of 3 times, and if possible briefly interview the person answering the phone, using 3 prepared questions. Based on these data, they report a numerical profile of percent of Orleanians optimistic, percent of Orleanians pessimistic etc. Assess the statistical methodology used by the students here. Confine yourself to the one or two most important points.

26. We obtain the following random sample from a normal population:

30.4, 28.4, 32.5, 33.6, 34.4, 34.5, 27.6, 30.5, 34.2, 28.8, 30.6, 27.6

At the significance level $\alpha = .05$, do these data provide evidence that $\mu > 30$? Set up the appropriate hypothesis and alternative, carry out the hypothesis test, answer the question, and give a p-value.