Final Exam
December 14, 2002

Name (please print): X

Instructor (circle one):

Jing           Gerwyn           Justin           Mevin           Antonello           Joel           Gentry
Cao            Green            Hobart           Hooten          Loddo            Miller         White

150 points

Instructions:
*** DO NOT make any marks in the grading area (right margin.) ***

I. For full-work problems (#1 - #9), all work must be shown in order to receive credit. Your final answers should be clearly written in the blank provided or, if no blank is provided, circled.

II. Carry all computations to at least three decimal places unless otherwise indicated. Only round your final answer. Do not round during intermediate steps.

III. Multiple-choice questions (#10 - #29) are worth 2.5 points each. Point values are noted next to each full-work part. The entire exam is worth 150 points and counts for 30% of your course grade.

IV. For hypothesis tests you must: Clearly state $H_0$ and $H_A$; Show the computation of the test statistic; Indicate the rejection region by drawing a picture and labeling the critical value (or by computing a $p$-value); and Clearly indicate your conclusion.

V. You may make any assumptions which are necessary to work the problems using the methods which were developed during class. However, you should not make any assumptions beyond what is necessary.

VI. Final Exam papers are not returned to students, but you are welcome to look at your paper in your instructor's office. Please make an appointment with him or her during the first two weeks of the Winter 2003 semester. After that time, exam papers may be destroyed.
1. (10 points.) A newspaper article claims that 45% of Americans have brown eyes. If a sample of 80 Americans showed that 33 of them had brown eyes, is there reason to believe the newspaper's claim is incorrect? Test at the .10 level of significance.

1. \( H_0: \rho = .45 \) vs. \( H_a: \rho \neq .45 \)

Supporting work:
2. \( \alpha = .10 \)
3. \( \hat{p} = \frac{33}{80} = .4125 \)
4. \( z = \frac{.4125 - .45}{\sqrt{(1.45)(.55)}/80} = -0.67 \)

Should \( H_0 \) be rejected? (Circle one):
Yes [ ] No [X]

2. (7 points.) Heights of adult men in a certain ethnic group have a mean of 69 inches and a standard deviation of 3.5 inches. The distribution of these measurements is approximately normal. The tallest 5% of men in this ethnic group are _______ inches tall or taller.

\[ x = \mu + z\sigma = 69 + 1.64(3.5) = 74.74 \text{ (or } 74.775) \]
3. (10 points.) A group of 200 people were classified as to whether or not they received a flu shot last before last winter and whether or not they got the flu last winter. The results are summarized below.

<table>
<thead>
<tr>
<th></th>
<th>Got flu</th>
<th>No flu</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Got shot</td>
<td>(24)</td>
<td>(96)</td>
<td>120</td>
</tr>
<tr>
<td>No shot</td>
<td>(16)</td>
<td>(64)</td>
<td>80</td>
</tr>
<tr>
<td>TOTAL</td>
<td>40</td>
<td>160</td>
<td>200</td>
</tr>
</tbody>
</table>

At the .05 level of significance, does it appear that whether or not a person gets the flu is independent of whether or not he or she gets a flu shot?

(1) $H_0$: getting flu are independent vs. $H_A$: getting flu are not independent

Supporting work:

(2) $\alpha = .05$

(3) $\chi^2 = \frac{(24-20)^2}{24} + \frac{(96-96)^2}{96} + \frac{(16-16)^2}{16} + \frac{(64-64)^2}{64}$

$= 1.667 + 1.667 + 1 + 1.667 = 2.084$

$X^2_{(0.05,1)} = 3.84$

Should $H_0$ be rejected? (Circle one): Yes No

4. (7 points.) The diameters of Douglas fir trees have a mean of 4 inches and a standard deviation of 1.5 inches. The frequency curve for these measurements is approximately bell-shaped (normal). What percentage of Douglas fir trees have a diameter between 2.5 and 7 inches?

$z = \frac{7 - 4}{1.5} = 2.00$

$z = \frac{3.5 - 4}{1.5} = 1.00$

Answer: $.3413 + .4772 = .8185$
5. (10 points.) A frequent traveler believes that, on average, room prices at hotel chain #1 are more than those at hotel chain #2. Results from two independent random samples of room prices are shown below. 

<table>
<thead>
<tr>
<th>Chain</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>$100</td>
<td>$10</td>
<td>50</td>
</tr>
<tr>
<td>#2</td>
<td>$95</td>
<td>$10</td>
<td>50</td>
</tr>
</tbody>
</table>

Test at the .05 level.

$H_0: \mu_1 - \mu_2 = 0 \quad$ vs. \quad $H_a: \mu_1 - \mu_2 > 0$

Supporting work:

1. \[ \alpha = .05 \]

2. \[ Z = \frac{(100 - 95) - 0}{\sqrt{\frac{10^2}{50} + \frac{10^2}{50}}} = 2.50 \]

Should $H_0$ be rejected? (Circle one): Yes \[ \text{No} 

6. (7 points.) In a certain city, 40% of houses have two-car garages, 20% are painted white, and 9% have two-car garages and are painted white. If one house is selected at random, are the events “a house with a two-car garage is selected” and “a house that is painted white is selected” independent?

Circle one: Yes (the events are independent) \[ \text{No (the events are not independent)} \]

Justification (show a numeric comparison that supports your choice):

$P(T \cap W) = .09 \neq (.4)(.2) = .08$
7. Suppose that 40% of students have a final on the first day of finals week and that 27% of MU students have a final on the last day of finals week.

(a) (7 points.) If 10 students are selected at random, what is the probability that at least 6 of them will have a final on the first day of finals week?

\[ X \sim B(n, p) \]
\[ P(X \geq 6) = 1 - P(X \leq 5) = 1 - .834 = .166 \]

(b) (7 points.) If 7 students are selected at random, what is the probability that exactly 2 of them will have a final on the last day of finals week?

\[ P(2) = \binom{7}{2} (.27)^2 (.73)^5 
\]
\[ = .317 \]

(c) (7 points.) If 100 students are selected at random, what is the probability that more than 47 of them will have a final on the first day of finals week? (Use the normal distribution as an approximation and make the continuity correction.)

\[ \mu = np = 100 (.4) = 40 \]
\[ \sigma = \sqrt{npq} = \sqrt{100 (.4) (.6)} = 4.899 \]
\[ P(X > 47) = P\left(\frac{X - \mu}{\sigma} > \frac{47 - 40}{4.899}\right) \]
\[ \z = \frac{47.5 - 40}{4.899} = 1.53 \]

\[ P(X > 47) = .063 \]
8. Suppose that the number of cigarettes smoked per day by a particular smoker approximately follows a Poisson distribution with \( \lambda = 4 \).

\[ \mu = 4 \quad \sigma = 2 \]

(a) (7 points.) What is the probability that this smoker will smoke at most 6 cigarettes during a randomly selected day?

\[ X \sim \text{Poisson}(4) \]

\[ P(X \leq 6) = 0.889 \quad (\text{Table}) \]

(b) (7 points.) Use the Central Limit Theorem to approximate the probability that this smoker will smoke an average of at least 4.5 cigarettes during 36 randomly selected days.

\[
\frac{4.5 - 4}{\frac{2}{\sqrt{36}}} = 1.50
\]

9. The amount of time it takes a passenger on a particular airline to claim his or her luggage averages 12 minutes with a standard deviation of four minutes.

(a) (7 points.) What does Chebyshev's rule say about the proportion of passengers who take between 2 and 22 minutes to claim their luggage?

\[
Z = \frac{2 - 12}{4} = -2.5 \quad Z = \frac{22 - 12}{4} = 2.50
\]

\[
\text{at least } 1 - (\frac{2.50}{2})^2 = 0.84
\]

(b) (7 points.) If it may be assumed that the distribution of the amount of time it takes passengers to claim luggage is mound-shaped, what between what two amounts of time will approximately 95% of passengers take to claim their luggage.

\[ \mu \pm 2\sigma \Rightarrow 12 \pm 2(4) = 4 \text{ and } 20 \]
Multiple-choice section: PRINT the letter of the best choice in the blank provided.

The following information is for the next four questions. Suppose that, for houses built prior to 1945, the mean size is 2,000 square feet with a standard deviation of 200 square feet. Furthermore, 80% of these houses have only one bathroom. For samples of size \( n = 1,600 \), consider the sampling distributions of \( \hat{p} \) (the sample proportion of houses that have only one bathroom) and \( \bar{x} \) (the sample mean size of the houses).

10. The mean of the sampling distribution of \( \bar{x} \) is
   (A) 2,000; \( \mu_{\bar{x}} = \mu = 2000 \)
   (B) 5; \( \mu_{\bar{x}} = \mu = 2000 \)
   (C) 80%; \( \mu_{\bar{x}} = \mu = 2000 \)
   (D) 1%; \( \mu_{\bar{x}} = \mu = 2000 \)
   (E) None of the above choices represents a suitable response.

   10. ANSWER: A

11. The standard deviation of the sampling distribution of \( \hat{p} \) is
   (A) more than .01; \( \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.8(1-.8)}{1600}} = .01 \)
   (B) .01; \( \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.8(1-.8)}{1600}} = .01 \)
   (C) .001; \( \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.8(1-.8)}{1600}} = .01 \)
   (D) less than .001; \( \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.8(1-.8)}{1600}} = .01 \)
   (E) None of the above choices represents a suitable response.

   11. ANSWER: B

12. The standard deviation of the sampling distribution of \( \bar{x} \) is
   (A) 2,000; \( \frac{\sigma}{\sqrt{n}} = \frac{200}{\sqrt{1600}} = 5 \)
   (B) 1,000; \( \frac{\sigma}{\sqrt{n}} = \frac{200}{\sqrt{1600}} = 5 \)
   (C) 5; \( \frac{\sigma}{\sqrt{n}} = \frac{200}{\sqrt{1600}} = 5 \)
   (D) 10; \( \frac{\sigma}{\sqrt{n}} = \frac{200}{\sqrt{1600}} = 5 \)
   (E) None of the above choices represents a suitable response.

   12. ANSWER: C

13. The mean of the sampling distribution of \( \hat{p} \) is
   (A) .001; \( \mu_{\hat{p}} = \mu = .001 \)
   (B) .010; \( \mu_{\hat{p}} = \mu = .001 \)
   (C) .100; \( \mu_{\hat{p}} = \mu = .001 \)
   (D) .800; \( \mu_{\hat{p}} = \mu = .001 \)
   (E) None of the above choices represents a suitable response.

   13. ANSWER: D
The following information is for the next four questions: A sample of 75 students was collected to estimate \( p \), the proportion of students who are enrolled in more than 12 credit hours for the current semester. The resulting confidence interval was

\[
\frac{25}{75} \pm 2.576 \sqrt{\frac{\frac{25}{75} \left( 1 - \frac{25}{75} \right)}{75}} \rightarrow 0.333 \pm 2.576 \left( 0.054 \right) \rightarrow [0.193, 0.474].
\]

Note that \( z_{0.05} = 2.576 \).

14. The confidence level of the interval is
   (A) 95\%; \quad \alpha = 0.05
   (B) 99\%; \quad \alpha = 0.01
   (C) 99.7\%; \quad 1 - \alpha = 0.99
   (D) 90\%; \quad \alpha = 0.10
   (E) None of the above choices represents a suitable response.

14. ANSWER: (B)

15. The sample proportion is (to three decimal places)
   (A) 0.667; \quad \hat{p} = \frac{25}{75} = 0.333
   (B) 0.750;
   (C) 0.250;
   (D) 0.333;
   (E) None of the above choices represents a suitable response.

15. ANSWER: (D)

16. The standard error of \( \hat{p} \) is (to three decimal places)
   (A) 0.667 \sqrt{\hat{p}(1-\hat{p})} = 0.054
   (B) 0.333
   (C) 0.054
   (D) 2.576(0.054) = 0.139
   (E) None of the above choices represents a suitable response.

16. ANSWER: (C)

17. Which one of the following claims would the interval tend to support?
   (A) "Most students are enrolled in more than 12 credit hours";
   (B) "Less than 25\% of students are enrolled in more than 12 credit hours";
   (C) "Exactly 60\% of students are enrolled in more than 12 credit hours";
   (D) "Most students are enrolled in 12 or fewer credit hours";
   (E) None of the above choices represents a suitable response.

17. ANSWER: (D)
18. When samples of size 25 are collected from a population that has a mean of 100 and a standard deviation of 20, the sampling distribution of $\bar{x}$ will have a standard deviation of

(A) 100;  \hspace{1cm} 18. \hspace{0.5cm} \text{ANSWER: } B
(B) 4;  \hspace{1cm} \sqrt{\frac{20}{\sqrt{25}}} = 4
(C) 16;
(D) 5;
(E) None of the above choices represents a suitable response.

19. When samples of size 10 are collected from a normal population, the sampling distribution of $\bar{x}$ will be

(A) normal;
(B) skewed to the right;
(C) skewed to the left;
(D) symmetric, but not normal;
(E) The form of the sampling distribution cannot be determined from the information provided.

19. \hspace{0.5cm} \text{ANSWER: } A

20. Suppose that 20% of cars have at least one under-inflated tire. If two cars are independently selected, at random, what is the probability both will have at least one under-inflated tire?

(A) .36; \hspace{1cm} \binom{20}{2} \cdot 20 \cdot 20 = .04
(B) .40;
(C) .20;
(D) .04;
(E) None of the above choices represents a suitable response.

20. \hspace{0.5cm} \text{ANSWER: } D

21. If a hypothesis test rejects $H_0$, then

(A) it is not possible that the test committed a type I error;
(B) it is not possible that the test committed a type II error;
(C) $H_0$ must be true;
(D) $H_0$ must be false;
(E) both (A) and (B).

21. \hspace{0.5cm} \text{ANSWER: } B
The following information is for the next two questions: A marketing research firm asked consumers which one of four potential slogans they felt would be most effective in marketing a new automobile. The firm will contact 100 consumers and use the resulting data to test \( H_0: p_1 = p_2 = p_3 = p_4 = \frac{1}{4} \) vs. \( H_A: \) At least one inequality.

22. If \( H_0 \) is true, \( E(n_i) = \) ?
(A) 100;
(B) 96;
(C) 25;
(D) 3.84;
(E) None of the above choices represents a suitable response.

23. If this test were conducted at \( \alpha = .05 \), then the critical value would be
(A) 9.48773 \( (X^2_{.050} \text{ with df } = 4); \)
(B) 11.1433 \( (X^2_{.025} \text{ with df } = 4); \)
(C) 7.81473 \( (X^2_{.050} \text{ with df } = 3); \)
(D) 9.34840 \( (X^2_{.025} \text{ with df } = 3); \)
(E) .351846 \( (X^2_{.050} \text{ with df } = 3). \)

The following information is for the next two questions: A sample of five graduating seniors showed that they had the following numbers of job offers: 0, 0, 1, 4, 10. Note that \( \sum x = 15 \) and \( \sum x^2 = 117. \)

24. The sample mean is \( \bar{x} = \frac{15}{5} = 3 \)
(A) 3;
(B) 23.4;
(C) 3.75;
(D) 29.25;
(E) None of the above choices represents a suitable response.

25. The sample variance is
(A) \( \sqrt{14.4}; \)
(B) 14.4;
(C) \( \sqrt{18}; \)
(D) 18;
(E) None of the above choices represents a suitable response.
The following information is for the next two questions: A criminal trial is, essentially, a test of $H_0$: The defendant is innocent vs. $H_a$: The defendant is guilty. The jury may decide to convict the defendant (i.e., reject $H_0$) or to acquit the defendant (i.e., not reject $H_0$). Suppose that, overall, 18% of guilty defendants are acquitted and 3% of innocent defendants are convicted.

26. The probability that the jury will commit a type I error is
(A) .05;
(B) .03;
(C) .97;
(D) .82;
(E) None of the above choices represents a suitable response.

27. The probability that the jury will commit a type II error is
(A) .18;
(B) .82;
(C) .03;
(D) .97;
(E) None of the above choices represents a suitable response.

The following information is for the next two questions: It is of interest to estimate $p_1 - p_2$ to within .1 of the true value with 95% certainty. Assume that equal sample sizes are desired.

28. If no other information is available, then each sample should be at least
(A) 19; $n_1 = n_2 = \frac{(1.96)^2((.5)(.5) + (.5)(.5))}{.01}$
(B) 20;
(C) 193;
(D) 215;
(E) None of the above choices represents a suitable response.

29. If we suspect $p_1$ is approximately .20 and $p_2$ is approximately .30, then each sample should be at least
(A) 73; $n_1 = n_2 = \frac{(1.96)^2((.2)(.8) + (.3)(.7))}{.1}$
(B) 193;
(C) 143;
(D) 215;
(E) None of the above choices represents a suitable response.