STAT 1220          Spring 2007
Common Final Exam      May 3, 2007

Please print the following information:

Name: ___________________________  Instructor: ___________________________

Student ID #: ____________________  Section/Time: _______________________

THIS EXAM HAS TWO PARTS

PART I. Consists of 30 multiple choice questions worth a total of 60 points. Read all questions carefully. You may do calculations on the test paper. Mark the number of the opscan sheet corresponding to the test question number with a Number 2 pencil or a mechanical pencil with HB lead. Mark only one answer; otherwise the answer will be counted as incorrect. In case there is more than one answer, mark the best answer. Please make sure that your name appears on the opscan sheet in the spaces provided.

PART II. This part consists of 3 questions (40 points in total). You MUST show all work for each question in the space provided to receive full credit for that question. If you write your explanations in another part of the test, please indicate accordingly.

At the end of the examination, you MUST hand in this test booklet, your answer sheet and all scratch paper.

FOR DEPARTMENTAL USE ONLY:
PART II:

<table>
<thead>
<tr>
<th>Questions</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>14</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Part I          Part II          Total
Questions 1, 2 and 3. In a survey of daily internet game time usage by college school students, a sample of 7 students provided the following data:

3 1.5 2.2 1.5 0.5 0 2.5

1 Find the mean of the data.
(a) 1.4  (b) 1.6  (c) 1.8  (d) 2.0  (e) less than 1.4

2 Find the standard deviation of the sample data (rounded off to three decimal places).
(a) 0.994  (b) 1.600  (c) 1.074  (d) 4.140  (e) 1.122

3 Find the median of the data.
(a) 1.5  (b) 1.4  (c) 1.6  (d) 2.5  (e) 2.0

Questions 4, 5 and 6. A study at a university shows that the mean age of the students at their graduation is 24 with a standard deviation 3:

4 If a new graduate is 28 years old, what is his z-score?
(a) $z = 4$  (b) $z = -4$  (c) $z = 0.75$  (d) $z = 1.33$  (e) $z = -0.75$

5 If the distribution of the graduation ages at this university is approximately bell-shaped, then approximately what percentage of students at this university graduate at age younger than 21?
(a) At least 68%  (b) At most 68%  (c) Approximately 32%
(d) Approximately 16%  (e) Approximately 34%

6 If the distribution of the graduation ages at this university is normal, then which of the following is closest to the percentile rank of a student who graduated at the age of 31?
(a) 70  (b) 85  (c) 90  (d) 95  (e) 99

Question 7. Given that events $A$ and $B$ are independent and $P(A) = 0.8$, $P(B) = 0.2$, Find $P(A \cap B)$.
(a) 1.0  (b) 0.6  (c) 0.16  (d) 0.25  (e) 4.0
Question 8. A fair coin is tossed 5 times and the probability of getting at least two tails is $\frac{13}{16}$. Find the probability that at most one tail appears.

(a) $\frac{13}{16}$  (b) $\frac{3}{16}$  (c) $\frac{16}{3}$  (d) $\frac{16}{3}$  (e) $\frac{1}{16}$

Questions 9 and 10. To determine whether its service is satisfactory to its customers, a hotel surveyed 100 guests (44 females and 56 males) and the result is summarized in the table below. A guest is randomly selected from these 100 people.

<table>
<thead>
<tr>
<th></th>
<th>Satisfied</th>
<th>Not satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>42</td>
<td>2</td>
</tr>
<tr>
<td>Male</td>
<td>40</td>
<td>16</td>
</tr>
</tbody>
</table>

9  What is the probability that the selected guest is a female and is satisfied with the hotel?

(a) .84  (b) .82  (c) .58  (d) .56  (e) .42

10 If we know that this selected guest is a male, what is the probability that he is satisfied with the hotel?

(a) .82  (b) .84  (c) .49  (d) .71  (e) .4

Questions 11 and 12. The following table gives the possible net gains (in dollars) of a gambling game and the corresponding probabilities

<table>
<thead>
<tr>
<th>$x$</th>
<th>$-2$</th>
<th>$5$</th>
<th>$20$</th>
<th>$50$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(x)$</td>
<td>.90</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
</tbody>
</table>

11 If you play this game once, what is the probability that your net gain is $20$ or more?

(a) .04  (b) .2  (c) .1  (d) .9  (e) .03

12 Find the expected value of this game (that is, $\mu = E(x)$).

(a) $1.5$  (b) $18.25$  (c) $1.4$  (d) $-1.8$  (e) $-0.4$
Questions 13 and 14. It is known that the mean price of a college textbook was $85 in 2004 with the standard deviation of $16. Consider a random sample of 40 college textbooks marketed in 2004, and let $\bar{x}$ represent the mean of the sample.

13 Find the mean and standard deviation of $\bar{x}$, i.e., $\mu_{\bar{x}}, \sigma_{\bar{x}}$. (Round off to two decimal places.)

(a) 85, .4   (b) 13.44, 16   (c) 85, 2.53   (d) 85, 16   (e) 101, 2.53

14 What is the probability that the sample mean price $\bar{x}$ is below 69 (round off to two decimal places)?

(a) .00   (b) .03   (c) .16   (d) .32   (e) .50

Question 15. Suppose $x$ is a normally distributed random variable with $\mu = 6$ and $\sigma = 4$, find the probability $P(4 \leq x \leq 10)$:

(a) .1498   (b) .5000   (c) 1.0   (d) .4772   (e) .5328

Question 16. A random sample of 100 observations produced a mean $\bar{x} = 38.5$ and standard deviation $s = 3.2$. Find a 95% confidence interval for the population mean $\mu$ (round off to two decimal places).

(a) (32.23, 44.77)  
(b) (37.87, 39.13)  
(c) (37.68, 39.32)  
(d) (35.50, 41.70)  
(e) (38.44, 38.56)

Question 17. A random sample of 9 observations from a normally distributed population produced a mean $\bar{x} = 12.4$ and standard deviation $s = 2.8$. Find a 99% confidence interval for the population mean $\mu$ (round off to two decimal places).

(a) (11.36, 13.44)  
(b) (9.77, 15.03)  
(c) (9.70, 15.10)  
(d) (9.27, 15.53)  
(e) (9.37, 15.43)
Question 18. If no preliminary estimate is available, find the sample size needed to estimate a population proportion \( p \) with 95% confidence such that the estimate is accurate within 0.03 of \( p \).

(a) \( n = 800 \)  (b) \( n = 1010 \)  (c) \( n = 1068 \)  (d) \( n = 1547 \)  (e) \( n = 1987 \)

Question 19. A poll was taken of 588 residents in a county. The residents sampled were asked whether they think their local government did a good job overall. 490 responded “yes”. Let \( p \) denote the proportion of all residents in that county who think their local government did a good job. Construct a 95% confidence interval for \( p \). Round off to two decimal places.

(a) \((489.97, 490.03)\)
(b) \((0.80, 0.86)\)
(c) \((0.68, 0.92)\)
(d) \((0.10, 1.56)\)
(e) \((0.79, 0.87)\)

Questions 20, 21 and 22. A report from the office of the superintendent claims that the average reading test score of 4th grade students in the school district is 76. A group of parents suspect that the real mean is lower than this reported score so they draw a random sample consisting of 37 4th grade student reading exam scores. They find that the sample mean is 74 and the sample standard deviation is 4.5.

20 State the null hypothesis and the alternative hypothesis.

(a) \( H_0 : \mu = 76, \ H_a : \mu < 76 \)
(b) \( H_0 : \mu = 76, \ H_a : \mu > 76 \)
(c) \( H_0 : \mu = 74, \ H_a : \mu < 74 \)
(d) \( H_0 : \mu = 76, \ H_a : \mu = 74 \)
(e) \( H_0 : \mu = 74, \ H_a : \mu < 76 \)

21 The value of the test statistic is

(a) \(-2.70\)  (b) \(-0.44\)  (c) \(-16.44\)  (d) \(2.70\)  (e) \(0.44\)

22 For \( \alpha = 0.05 \), the rejection region is

(a) \( z < -1.645 \)  (b) \( z < -1.645 \)  (c) \( z > 1.645 \)  (d) \( z < -1.645 \) or \( z > 1.645 \)
(e) \( z < 1.645 \)
Questions 23, 24 and 25. An SAT prep school wishes to investigate whether using a specially designed computer course will lead to improved scores. The difference in the SAT scores (after − before) of five randomly selected students who used this course results in $\bar{x}_D = 98$ and $s_D = 144.59$. Assume that the SAT scores are approximately normally distributed.

23 Does the service help to improve the SAT scores? Choose the appropriate hypotheses to test the claim.

(a) $H_0 : \mu_D = 0, H_a : \mu_D \neq 0$
(b) $H_0 : \bar{x}_D = 0, H_a : \bar{x}_D > 0$
(c) $H_0 : \mu_D = 0, H_a : \mu_D > 0$
(d) $H_0 : \bar{x}_D = 0, H_a : \bar{x}_D < 0$
(e) $H_0 : \mu_D = 0, H_a : \mu_D < 0$

24 Calculate the value of the test statistic.

(a) 1.516  (b) 0.678  (c) 1.475  (d) 3.299  (e) $-1.516$

25 Find the rejection region and state your decision at $\alpha = .05$.

(a) Rejection Region: $t < -2.132$; Decision: Reject $H_0$
(b) Rejection Region: $t < -2.105$; Decision: Do not reject $H_0$
(c) Rejection Region: $t > 2.132$; Decision: Reject $H_0$
(d) Rejection Region: $t > 2.015$; Decision: Reject $H_0$
(e) Rejection Region: $t > 2.132$; Decision: Do not reject $H_0$

Question 26. Given that the $p$-value for a hypothesis test is $p = 0.044$, then at the significance level $\alpha = 0.05$, we

(a) reject $H_0$ because $\alpha$ is less than $p$
(b) do not reject $H_0$ because $p$ is less than $\alpha$
(c) do not reject $H_0$ because $\alpha$ is less than $p$
(d) reject $H_0$ because $p$ is less than $\alpha$
(e) cannot make a decision because there is not enough information

Question 27. If there is a strong negative linear relationship between $y$ and $x$ in a sample, then which of the following is a likely value for $r$, the coefficient of correlation?

(a) 0.88  (b) $-0.92$  (c) 0.22  (d) $-1.89$  (e) $-0.03$
Questions 28, 29 and 30. The prices of 6 stocks \((x, \text{ in dollars})\) and their corresponding annual earnings per share \((y, \text{ in dollars})\) are given below.

<table>
<thead>
<tr>
<th>price</th>
<th>32.98</th>
<th>49.72</th>
<th>70.29</th>
<th>97.34</th>
<th>43.58</th>
<th>52.44</th>
</tr>
</thead>
<tbody>
<tr>
<td>annual earning</td>
<td>4.50</td>
<td>5.28</td>
<td>10.00</td>
<td>16.00</td>
<td>5.50</td>
<td>7.00</td>
</tr>
</tbody>
</table>

The regression equation is \(\hat{y} = 0.19x - 2.73\), and the coefficient of determination is \(r^2 = 0.97\).

28 The interpretation of coefficient of determination, \(r^2\), is:

(a) We are 97% confident that the prices of the stocks are linearly related to their annual earnings.
(b) There is no correlation between the prices of the stocks and their annual earnings.
(c) We will observe a linear relation between the prices of the stocks and their annual earnings 97 percent of the time.
(d) There is a moderately strong negative correlation between the prices of the stocks and their annual earnings.
(e) 97% of the variation in the annual earnings of the stocks can be explained by the relationship between the prices of the stocks and their annual earnings.

29 Predict the annual earning of a stock if the price of the stock is \(x = \$50\).

(a) $0.19 \hspace{1cm} (b) $8.56 \hspace{1cm} (c) $10.25 \hspace{1cm} (d) $6.77 \hspace{1cm} (e) $4.35

30 If the predicted annual earning of a stock with price \(x_p = 61.21\) is \(\hat{y} = 8.90\) and the corresponding estimated standard error of prediction is \(s\sqrt{1 + \frac{1}{6} + \frac{(x_p - \bar{x})^2}{SS_{xx}}} = 1.41\), which of the following is a 95% prediction interval for \(y\) at \(x_p\)?

(a) 8.90 ± 3.45 \hspace{1cm} (b) 8.90 ± 3.63 \hspace{1cm} (c) 8.90 ± 3.91 \hspace{1cm} (d) 8.90 ± 3.01 \hspace{1cm} (e) 8.90 ± 2.84

End of Multiple Choice Section
1. GoLite, a food manufacturer, claims that each can of a new food it had just started to market contains at most 4 grams of fat. A small sample of 6 cans produced the following data (number of grams of fat in each can): 5.0  5.2  5.8  4.2  4.4  4.8

(a) (3 pts.) Compute the sample mean and sample standard deviation.

(b) (3 pts.) Find a 95% confidence interval for the mean amount of fat in a can of this food.

(c) (3 pts.) Set up an appropriate null and alternative hypothesis to test whether the company’s claim is valid.

(d) (3 pts.) Calculate the test statistic.

(e) (2 pts.) State your conclusion at the significance level $\alpha = 0.05$. 
2. Seven employees at a large corporation were randomly selected and their years of education \((x)\) and annual salaries (in \(y\) thousand dollars) were collected and yielded the following

\[
\sum x = 122, \quad \sum x^2 = 2210, \quad \sum y = 390, \quad \sum y^2 = 24826, \quad \sum xy = 7153.
\]

(a) (4 pts.) Compute \(SS_{xx}\), \(SS_{xy}\) and \(SS_{yy}\).

(b) (3 pts.) Calculate the correlation coefficient, \(r\), between the years of education of an employee and his/her annual salary.

(d) (4 pts.) Find the regression equation.

(e) (3 pts.) Use the equation in part (c) to predict \(y\) when \(x = 20\) years.
3. A shipping company claims that the average delivery time for its next day shipping service is 22 hours. However, a group of customers believe that the mean delivery time is longer than 22 hours. Suppose that a random sample of 78 next day deliveries by the company yields a mean delivery time of 23.5 hours, with a standard deviation of 4.8 hours.

(a) (3 pts.) State the null and the alternative hypotheses to the company’s claim.

(b) (3 pts.) In the context of the problem, explain Type I error.

(c) (3 pts.) Find the value of the test statistic.

(d) (3 pts.) Use the $p$-value or the rejection region to draw your decision at the significance level of $\alpha = 0.05$. 
The Keys to the multiple choice problems: