STATISTICS 260
SAMPLE FINAL EXAMINATION 3
(from Spring Term 2003)

INSTRUCTIONS:

1. Use a PENCIL HB or softer.

2. Your NAME and STUDENT NUMBER must be recorded on BOTH your question paper and in CODED FORM on your green answer sheet.

3. Hand calculators are permitted; calculator memories must be cleared just before the examination begins. No other aids are permitted, except the Formula List and Statistical Tables which are provided.

4. Questions 1 through 28 are multiple-choice questions worth 2 marks each. Code your answers on the green answer sheet provided. For questions requiring numerical answers, the 10 choices are listed in numerically increasing order. Choose the value that is nearest your (unrounded) answer. In the special case that your (unrounded) answer is equidistant from the two nearest choices, choose the larger of these two choices. For verification purposes, show all calculations on your question paper. Unverified answers may be disallowed.

5. Questions 29, 30, 31 are full-answer questions worth 8 marks, 7 marks, 9 marks, respectively. For each of these questions write out your solution carefully and completely on the question paper. Marks will be deducted for incomplete or poorly presented solutions.

Maximum Score is 80 marks.

Note: To conserve paper, no working space is provided on this sample exam. Working space will be provided on the real exam.

Questions 1 through 3 refer to the following setup. The 20 houses in a certain city block were classified as follows:

<table>
<thead>
<tr>
<th>NUMBER OF STORIES</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>OWNER OCCUPIED</td>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>RENTED</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Suppose one of these 20 houses is selected at random.

1. What is the probability that the selected house is neither rented nor has fewer than two stories?
   (A) .05  (B) .10  (C) .15  (D) .20  (E) .25
   (F) .30  (G) .40  (H) .50  (I) .60  (J) .70
2. If the selected house is rented, what is the probability that it has fewer than three stories?

(A) .40  (B) .45  (C) .50  (D) .55  (E) .60
(F) .65  (G) .70  (H) .75  (I) .80  (J) .85

3. Which of the following three statements are true?

(i) The events “selected house is rented” and “selected house has an odd number of stories” are mutually exclusive.
(ii) The events “selected house is rented” and “selected house has an odd number of stories” are independent.
(iii) The events “selected house has an odd number of stories” and “selected house has an even number of stories” are independent.

(A) None  (B) Only (i)  (C) Only (ii)  (D) Only (iii)
(E) (i) & (ii)  (F) (i) & (iii)  (G) (ii) & (iii)  (H) All

Questions 4 through 6 refer to the following setup. Consider the system of components connected as in the following diagram.

The subsystem consisting of components 1 and 2 works if either 1 or 2 works. The subsystem consisting of components 3 and 4 works if either 3 or 4 works. The whole system works if both subsystems work. The probabilities that components 1 through 4 work are .1, .2, .3, .4, respectively. Assume the four components function independently.

4. What is the probability that the subsystem consisting of components 3 and 4 works?

(A) .05  (B) .15  (C) .25  (D) .35  (E) .45
(F) .55  (G) .65  (H) .75  (I) .85  (J) .95

5. What is the probability that the whole system works?

(A) .07  (B) .11  (C) .15  (D) .19  (E) .23
(F) .30  (G) .40  (H) .50  (I) .60  (J) .70

6. What is the expected number of components that work in the subsystem consisting of components 3 and 4?

(A) .50  (B) .65  (C) .80  (D) .95  (E) 1.10
(F) 1.25  (G) 1.40  (H) 1.55  (I) 1.70  (J) 1.95

7. A box contains 5 coins, of which two are honest and three are two-headed. One coin, selected at random from the box, is tossed and observed to come up heads. What is the probability that the selected coin is two-headed?

(A) .62  (B) .65  (C) .68  (D) .71  (E) .74
(F) .77  (G) .80  (H) .83  (I) .86  (J) .89
8. Let $A$ and $B$ be two events such that $P(A|B) = .33$, $P(B|A) = .55$, $P(A \cup B) = .66$. Find $P(A \cap B)$.

(A) .08  (B) .11  (C) .14  (D) .17  (E) .20
(F) .23  (G) .26  (H) .29  (I) .32  (J) .35

9. In a certain factory, accidents occur at random times at the average rate of 4 accidents every 6 working days. What is the probability that at least 2 accidents occur in the next 3 working days?

(A) .30  (B) .35  (C) .40  (D) .45  (E) .50
(F) .55  (G) .60  (H) .65  (I) .70  (J) .75

10. Suppose that 40% of adult males are overweight. In a random sample of 20 adult males what is the probability that between 8 and 10 (inclusive) are overweight?

(A) .05  (B) .15  (C) .25  (D) .35  (E) .45
(F) .55  (G) .65  (H) .75  (I) .85  (J) .95

Questions 11 through 13 refer to the following setup. Let $X$ be a discrete random variable with the following probability distribution:

<table>
<thead>
<tr>
<th>$x$</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X = x)$</td>
<td>.04</td>
<td>.12</td>
<td>.36</td>
<td>.28</td>
<td>.20</td>
</tr>
</tbody>
</table>

11. Find the expected value of $X$.

(A) -.30  (B) -.20  (C) -.10  (D) 0  (E) .10
(F) .20  (G) .30  (H) .40  (I) .50  (J) .60

12. Find the variance of $X$.

(A) .80  (B) .90  (C) 1.00  (D) 1.10  (E) 1.20
(F) 1.30  (G) 1.40  (H) 1.50  (I) 1.60  (J) 1.70

13. Find $E(2X^2 + 6)$.

(A) 0  (B) 1  (C) 2  (D) 3  (E) 4
(F) 5  (G) 6  (H) 7  (I) 8  (J) 9

Questions 14 through 16 refer to the following setup. The yield $X$ (in litres) of a certain process is a continuous random variable with probability distribution given by the following probability density function:

$$f(x) = \begin{cases} \frac{x^3}{20} & \text{if } 1 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

14. Find the 10-th percentile of this distribution.

(A) 1.5  (B) 1.6  (C) 1.7  (D) 1.8  (E) 1.9
(F) 2.0  (G) 2.1  (H) 2.2  (I) 2.3  (J) 2.4
15. Find the standard deviation of $X$.
   (A) .1 (B) .2 (C) .3 (D) .4 (E) .5
   (F) .7 (G) .9 (H) 1.1 (I) 1.3 (J) 1.5

16. A random sample of 10 observations is to be drawn from the distribution of $X$. What is the expected number of these observations that will fall below 2 litres?
   (A) 1 (B) 2 (C) 3 (D) 4 (E) 5
   (F) 6 (G) 7 (H) 8 (I) 9 (J) 10

Questions 17 through 19 refer to the following setup. Two random variables $X$ and $Y$ have the following joint probability mass function:

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
<th>-2</th>
<th>0</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>.08</td>
<td>0</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>.20</td>
<td>.40</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.12</td>
<td>0</td>
<td>.18</td>
<td></td>
</tr>
</tbody>
</table>

17. Find the variance of $Y$.
   (A) 1.2 (B) 1.4 (C) 1.6 (D) 1.8 (E) 2.0
   (F) 2.2 (G) 2.4 (H) 2.6 (I) 2.8 (J) 3.0

18. Find the covariance of $X$ and $Y$.
   (A) -.5 (B) -.4 (C) -.3 (D) -.2 (E) -.1
   (F) 0   (G) .1  (H) .2  (I) .3  (J) .4

19. Find $P(X > Y \mid X + Y > -2)$.
   (A) .20 (B) .25 (C) .30 (D) .35 (E) .40
   (F) .45 (G) .50 (H) .55 (I) .60 (J) .65

20. Suppose the lifetime, $X$, of a certain type of component is a random variable having an exponential distribution with mean lifetime $\mu = 1/\lambda = 5$ years. What is the probability that a two-year-old component of this type will last an additional 4 years?
   (A) .20 (B) .25 (C) .30 (D) .35 (E) .40
   (F) .45 (G) .50 (H) .55 (I) .60 (J) .65

Questions 21 and 22 refer to the following setup. The amount of fill, $X$, for a certain size box of Brand K cereal is a normally distributed random variable with mean $\mu = 750$ grams and standard deviation $\sigma = 4$ grams.
21. What proportion of such cereal boxes will contain more than 755 grams of fill?
(A) 0.06  (B) 0.08  (C) 0.10  (D) 0.12  (E) 0.14  
(F) 0.16  (G) 0.18  (H) 0.20  (I) 0.22  (J) 0.24

22. Find \( w \) such that only 20% of these cereal boxes contain less than \( w \) grams of fill.
(A) 743.5  (B) 744.0  (C) 744.5  (D) 745.0  (E) 745.5  
(F) 746.0  (G) 746.5  (H) 747.0  (I) 747.5  (J) 748.0

Questions 23 and 24 refer to the following setup. Assume the time \( X \) between successive computer crashes at a certain facility is a random variable having an exponential distribution with unknown mean \( \mu = \frac{1}{\lambda} \) days. Suppose an observed random sample of 10 observations on \( X \) yielded the following data (in days): 0.3, 1.3, 1.1, 3.1, 4.0, 2.5, 0.9, 2.3, 0.8, 5.5

23. Based on these data, find the maximum likelihood estimate of the rate parameter \( \lambda \).
(A) 0.15  (B) 0.25  (C) 0.35  (D) 0.45  (E) 0.55  
(F) 0.65  (G) 0.75  (H) 0.85  (I) 0.95  (J) 1.05

24. Based of these data, find the maximum likelihood estimate of \( P(X > 2) \).
(A) 0.00  (B) 0.05  (C) 0.10  (D) 0.15  (E) 0.20  
(F) 0.25  (G) 0.30  (H) 0.35  (I) 0.40  (J) 0.50

25. Let \( X_1, X_2, X_3, X_4 \) be a random sample from a population distribution with unknown mean \( \mu \) and unknown standard deviation \( \sigma \). Consider the following three unbiased estimators for \( \mu \):
\[ R = \frac{1}{3} X_1 + \frac{1}{3} X_2 + \frac{1}{3} X_3 \, , \quad U = (1) X_1 + (2) X_2 + (3) X_3 + (4) X_4 \, , \quad \text{and} \, W = \frac{1}{2} X_1 + \frac{1}{2} X_2 + \frac{1}{2} X_3 + \frac{1}{2} X_4 \, . \]
Which of the following three statements are true?
(i) \( W \) is more efficient than \( U \) for estimating \( \mu \).
(ii) \( W \) is more efficient than \( R \) for estimating \( \mu \).
(iii) \( U \) is more efficient than \( R \) for estimating \( \mu \).
(A) None  (B) Only (i)  (C) Only (ii)  (D) Only (iii)  
(E) (i) & (ii)  (F) (i) & (iii)  (G) (ii) & (iii)  (H) All

Questions 26 and 27 refer to the following setup. A survey was conducted to estimate the proportion \( p \) of passengers on a certain airline who are dissatisfied with the food served. A random sample of 200 passengers using this airline was drawn, and, of these, 42 were dissatisfied with the food.

26. Compute the upper limit of a 90% confidence interval for \( p \).
(A) 0.21  (B) 0.22  (C) 0.23  (D) 0.24  (E) 0.25  
(F) 0.26  (G) 0.27  (H) 0.28  (I) 0.29  (J) 0.30

27. Use these data as a pilot study to determine the number of additional observations needed to estimate \( p \) within \( \pm 4 \) percentage points with 95% confidence (i.e. 95% CI length = .08).
(A) 0  (B) 50  (C) 100  (D) 150  (E) 200  
(F) 250  (G) 300  (H) 350  (I) 400  (J) 450
28. A random sample of 50 measurements of arsenic in copper yielded the following data (in percent):

Sample Mean = .180 percent  Sample Standard Deviation = .026 percent

Compute the lower limit of a 97% confidence interval for the true mean percent of arsenic in the sampled copper.

(A)  .170  (B)  .171  (C)  .172  (D)  .173  (E)  .174  (F)  .175  (G)  .176  (H)  .177  (I)  .178  (J)  .179

29. An experiment was performed to compare two overnight parcel delivery services. Each delivery service was given identical parcels at the same time for the same destination on each of 6 occasions. The delivery times in hours were as follows:

<table>
<thead>
<tr>
<th>Occasion</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service I</td>
<td>13</td>
<td>15</td>
<td>15</td>
<td>11</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Service II</td>
<td>10</td>
<td>10</td>
<td>15</td>
<td>8</td>
<td>12</td>
<td>11</td>
</tr>
</tbody>
</table>

Is there any evidence that the true mean delivery times for these delivery services differ? Let $\mu_1$ and $\mu_2$ denote the true mean delivery times for Delivery Service I and Delivery Service II, respectively; and let $\mu_D = \mu_1 - \mu_2$. Assume the relevant population distribution(s) is(are) normal.

(a) Specify the null and alternative hypotheses.

(b) Compute the observed value of the test statistic.

(c) Compute the P-value (or bracket the P-value) within Table accuracy, and indicate the probability distribution used for this computation.

(d) State your conclusion, and report the estimated value of the parameter being tested and the estimated standard error.

(e) Determine whether or not the null hypothesis should be rejected at significance level $\alpha = 0.05$.

30. At a certain gasoline station, the amount of gasoline sold to a randomly selected customer has mean 20 litres and standard deviation 5 litres.

(a) On a day when 300 customers independently purchase gasoline at this station, what is the probability that the total amount of gasoline sold exceeds 6100 litres?

(b) What is the probability that on a day with 300 customers, the total amount of gasoline sold is less than the total amount sold on a day with 290 customers?
31. A certain brand of microwave oven was priced at random samples of stores in Toronto and in Vancouver with the following results:

<table>
<thead>
<tr>
<th>City</th>
<th>Sample Size</th>
<th>Mean Price</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toronto</td>
<td>5</td>
<td>470</td>
<td>80</td>
</tr>
<tr>
<td>Vancouver</td>
<td>7</td>
<td>560</td>
<td>50</td>
</tr>
</tbody>
</table>

Let \( \mu_1 \) denote the true mean price of these microwave ovens in Toronto, and let \( \mu_2 \) denote the true mean price in Vancouver. Assume the relevant population distribution(s) is(are) normal.

(a) Do these data provide evidence that the true mean price is higher in Vancouver than in Toronto?

(i) Specify the null and alternative hypotheses.

(ii) Compute the P-value (or bracket the P-value) within Table accuracy, and indicate the probability distribution used for this computation.

(iii) State your conclusion, and report the estimated value of the parameter being tested and the estimated standard error.

(b) Construct a 90% confidence interval for \( \mu_1 - \mu_2 \).

---

ANSWERS FOR SAMPLE FINAL EXAMINATION 3

1. .3 (F)  
2. .75 (H)  
3. (A)  
4. .58 (F)  
5. .1624 (C)  
6. .70 (B)  
7. .75 (E)  
8. .1715 (D)  
9. .594 (G)  
10. .456 (E)  
11. .48 (I)  
12. 1.1296 (D)  
13. 8.88 (J)  
14. 1.732 (C)  
15. .4585 (E)  
16. 1.875 (B)  
17. 2.56 (H)  
18. .08 (G)  
19. .3478 (D)  
20. .449 (F)  
21. .1056 (C)  
22. 746.64 (G)  
23. .4587 (D)  
24. .3996 (I)  
25. (H)  
26. .2574 (F)  
27. 199 (E)  
28. .172 (C)  
29. (a) \( H_0: \mu_d = 0 \) vs \( H_a: \mu_d \neq 0 \)  
   (b) \( T_{obs} = \frac{\bar{d} - 0}{S_d / \sqrt{n}} = \frac{2.333 - 0}{2.338 / \sqrt{6}} = 2.44 \)  
   (c) \( P-value = P(T_{(5)} \geq 2.44 \text{ or } \leq -2.44) = 2P(T_{(5)} \geq 2.44) \therefore .054 < p \nu < .062 \)  
   (d) There is moderate evidence (.054 < p \nu < .062 ) that \( \mu_d \neq 0 \). The estimated value of \( \mu_d \) is 2.33 hours with estimated standard error = .95 hours.  
   (e) Since \( p \nu > .05 \), retain \( H_0 \) at level \( \alpha = .05 \).
30. (a) Let $X_i$ = amount of gasoline sold to $i$-th customer on a 300-customer day

\[ T_0 = X_1 + X_2 + \cdots + X_{300}, \quad E(T_0) = (300)(20) = 6000, \quad V(T_0) = (300)(25) = 7500 \]

$T_0 \sim N\left(\mu_{T_0} = 6000, \quad \sigma_{T_0} = \sqrt{7500}\right)$ by the Central Limit Theorem

\[ P(T_0 > 6100) \approx P\left(Z > \frac{6100 - 6000}{\sqrt{7500}}\right) = P(Z > 1.15) = .1251 \]

(b) Let $Y_i$ = amount of gasoline sold to $i$-th customer on a 290-customer day

\[ W_0 = Y_1 + Y_2 + \cdots + Y_{290}, \quad E(W_0) = (290)(20) = 5800, \quad V(W_0) = (290)(25) = 7250 \]

$T_0 - W_0 \sim N\left(\mu_{T_0-w_0} = 6000 - 5800 = 200, \quad \sigma_{T_0-w_0} = \sqrt{7500 + 7250} = 121.45\right)$

\[ P(T_0 - W_0 < 0) \approx P\left(Z < \frac{0 - 200}{121.45}\right) = P(Z < -1.65) = .0495 \]

31. (a) (i) $H_0: \mu_1 - \mu_2 = 0$ vs $H_a: \mu_1 - \mu_2 < 0$

(ii) Since $\frac{s_1}{s_2} = 1.6 > 1.4$, use unpooled-$t$ procedures

degrees of freedom = integer part of \[ \frac{\left(\frac{80^2}{5} + \frac{50^2}{7}\right)^2}{\left(\frac{(80^2/5)^2}{4} + \frac{(50^2/7)^2}{6}\right)} = 6 \]

$T_{obs} = \frac{470 - 560 - 0}{\sqrt{\frac{80^2}{5} + \frac{50^2}{7}}} = \frac{-90 - 0}{40.46} = -2.22$

$P$-value $= P\left(T_{(6)} \leq -2.22\right) = P\left(T_{(6)} \geq 2.22\right)$ by symmetry

$\therefore .031 < \alpha < .035$

(iii) There is strong evidence (.031 < $p_\alpha > .035$) that $\mu_1 - \mu_2 < 0$. The estimated value of $\mu_1 - \mu_2$ is -90 dollars with estimated standard error $= 40.46$.

(b) \[ \bar{x} - \bar{y} \pm (t_{0.05,6}) \sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}} \]

\[ 470 - 560 \pm (1.943)(40.46) \]

or (-168.61, -11.39)