STAT 1220  
Common Final Exam  

Fall 2021  
December 10, 2021  

Please print the following information  

Name: ___________________________  
Instructor: ________________________  

Student ID #: _____________________  
Section/Time: _____________________  

THIS EXAM HAS TWO PARTS  

PART I.  
Part I consists of 30 multiple choice questions. Each correct answer is scored 2 points; each incorrect (or blank) answer is scored 0, so there is no penalty for guessing. You may do calculations on the test paper, but your answers must be marked on the OPSCAN sheet with a soft lead pencil (HB or No. 2 lead). Any question with more than one choice marked will be counted as incorrect. If more than one choice seems correct, choose the one that is most complete or most accurate. Make sure that your name appears on the OPSCAN sheet in the spaces provided.  

PART II.  
Part II 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>12</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part I</th>
<th>Part II</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1
The following is used for questions 1-3.

Given the following sample data set \( \{6, 2, 1, 8, 5, 9, 9, 9, 5, 2\} \)

1. The value of the median is
   \( (a)\ 4.0 \quad (b)\ 4.5 \quad (c)\ 5.0 \quad (d)\ 5.5 \quad (e)\ 6.0 \)

2. The value of the mean is
   \( (a)\ 5.00 \quad (b)\ 5.50 \quad (c)\ 5.60 \quad (d)\ 5.75 \quad (e)\ 6.00 \)

3. The value of the mode is
   \( (a)\ 2.0 \quad (b)\ 5.0 \quad (c)\ 8.0 \quad (d)\ 9.0 \quad (e)\ 9.5 \)

4. A given sample of 12 data points sums to 166. When each of those data points are squared and the squared values are added together, the sum is 2554. Give this, what is the sample standard deviation?
   \( (a)\ 4.6338 \quad (b)\ 4.8399 \quad (c)\ 15.1962 \quad (d)\ 21.4722 \quad (e)\ 23.4242 \)

5. A fair coin is tossed twice. If \( A \) is the event to see at least one head and \( B \) is the event to see exactly one tail, find \( P(A \cup B) \).
   \( (a)\ 0.25 \quad (b)\ 0.50 \quad (c)\ 0.75 \quad (d)\ 1.00 \quad (e)\ 1.25 \)
Use the following to answer questions 6-10.

The breakdown of 200 cyclists (F: Female, M: Male) and the routes they prefer (L: Lake Path, H: Hilly Path, W: Wooded Path) is shown in the following table:

<table>
<thead>
<tr>
<th>Gender</th>
<th>Route</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lake Path</td>
<td>Hilly Path</td>
<td>Wooded Path</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>45</td>
<td>38</td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>26</td>
<td>52</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A cyclist is selected at random. Adjoin the row and column totals to the table and use the expanded table to find the probability of each of the following events.

6. The cyclist is female.
   (a) 0.355  (b) 0.409  (c) 0.450  (d) 0.500  (e) 0.550

7. The cyclist prefers a lake path.
   (a) 0.195  (b) 0.225  (c) 0.355  (d) 0.450  (e) 0.789

8. The cyclist is female who prefers a lake path.
   (a) 0.130  (b) 0.225  (c) 0.409  (d) 0.500  (e) 0.634

9. The cyclist is either a female or prefers a lake path.
   (a) 0.145  (b) 0.225  (c) 0.455  (d) 0.680  (e) 0.905

10. The cyclist prefers a lake path given that s/he is female.
    (a) 0.225  (b) 0.409  (c) 0.634  (d) 0.645  (e) 0.905
Use the following to answer questions 11-12.

Let $X$ be a discrete random variable with the following probability mass function, where $c$ is an unknown constant.

\[
\begin{array}{c|c|c|c}
  x & 1 & 2 & 3 \\
p(x) & c & 4c & 9c \\
\end{array}
\]

Find the unknown $c$, fill out the probabilities, and calculate the following.

11. The expected value of $X$ is:
   (a) 0.333  (b) 0.857  (c) 2.571  (d) 3.000  (e) 6.000

12. The probability that $X$ is greater than 1.5 is:
   (a) 0.000  (b) 0.286  (c) 0.643  (d) 0.929  (e) 1.000

13. According to a recent survey, 80% of students are planning to attend at least one football game at their university. If a random sample of 10 students is selected, what is the probability that at least 9 of those students are planning to attend at least one game?
   (a) 0.1074  (b) 0.1342  (c) 0.3490  (d) 0.3758  (e) 0.5000

Use the following to answer questions 14-15.

A population data set with a bell-shaped distribution has mean $\mu = 40$ and standard deviation $\sigma = 11$. Find the approximate percentage of observations in the data set that lie:

14. below 29
   (a) 13.5%  (b) 16.0%  (c) 32.0%  (d) 68.0%  (e) 84.0%

15. between 29 and 62
   (a) 50.0%  (b) 68.0%  (c) 81.5%  (d) 95.0%  (e) 99.7%
Use the following to answer questions 16-17.

The amount of time people spend going through airport security for planes at a busy airport is normally distributed with mean of 21 minutes and a standard deviation of 4.2 minutes.

16. Find the probability that the wait time for security of a randomly selected person is more than 22 minutes.
   (a) 0.159  (b) 0.313  (c) 0.406  (d) 0.594  (e) 0.687

17. Find the probability that the average wait time for security of 40 randomly chosen people is between 20 and 23 minutes.
   (a) 0.277  (b) 0.500  (c) 0.683  (d) 0.933  (e) 1.000

Use the following to answer questions 18-19.

Among a sample of 63 students selected at random from a university, the mean number of siblings is 1.2 with a standard deviation of 1.3.

18. A 95% confidence interval for the mean number of siblings for all students at this university is:
   (a) (-1.348, 3.748)
   (b) (0.783, 1.617)
   (c) (0.850, 1.550)
   (d) (0.879, 1.521)
   (e) (0.931, 1.469)

19. The minimum sample size required to estimate the mean number of siblings for all students with a 95% confidence and within 0.25 accuracy is
   (a) 21  (b) 74  (c) 104  (d) 293  (e) 416
20. A survey of 900 residents showed that 40% of them are in favor of higher taxes for improving schools. Which one of the following best approximates a 90% confidence interval for the population proportion \( p \)?

(a) 0.40 ± 0.021  
(b) 0.40 ± 0.027  
(c) 0.40 ± 0.032  
(d) 0.40 ± 0.068  
(e) 0.40 ± 0.960

21. In a random sample of 14 American adults, the mean waste recycled per person per day was 4.3 pounds and the standard deviation was 0.3 pounds. Assume that the waste is normally distributed. A 90% confidence interval for the population mean waste is given by:

(a) 4.3 ± 1.28 * (0.3/14)  
(b) 4.3 ± 1.35 * (0.3/14)  
(c) 4.3 ± 1.64 * (0.3/14)  
(d) 4.3 ± 1.77 * (0.3/14)  
(e) 4.3 ± 1.96 * (0.3/14)

Use the following to answer questions 22-25.

It is suspected that a machine, used for filling plastic bottles with an average net volume of 16.0 oz does not perform according to specifications. An engineer will collect 45 measurements and will reset the machine if there is evidence that the mean fill volume is different from 16 oz. The resulting data yield \( \bar{X} = 16.0167 \) and \( S = 0.0551 \). Perform the test using the information collected at the 5% level of significance.

22. Set up the null and alternative hypotheses to test the engineer’s claim.

(a) \( H_0: \mu = 16 \) versus \( H_a: \mu ≠ 16 \)  
(b) \( H_0: \mu = 16 \) versus \( H_a: \mu > 16 \)  
(c) \( H_0: \mu = 16 \) versus \( H_a: \mu < 16 \)  
(d) \( H_0: \mu ≤ 16 \) versus \( H_a: \mu > 16 \)  
(e) \( H_0: \mu ≥ 16 \) versus \( H_a: \mu < 16 \)

23. The value of the standardized test statistic is

(a) -2.033  
(b) -0.303  
(c) 0.303  
(d) 2.033  
(e) 2.248
24. The p-value of this test is
(a) 0.021  (b) 0.042  (c) 0.762  (d) 0.979  (e) cannot be determined

25. The decision of the test is
(a) Rejection region: \((-\infty, -1.645]\) \(\cup\) \([1.645, \infty)\); Decision: Reject the null hypothesis
(b) Rejection region: \((-\infty, -1.645]\); Decision: Reject the null hypothesis
(c) Rejection region: \((-\infty, -1.96]\) \(\cup\) \([1.96, \infty)\); Decision: Reject the null hypothesis
(d) Rejection region: \((-\infty, -1.96]\) \(\cup\) \([1.96, \infty)\); Decision: Fail to reject the null hypothesis
(e) Rejection region: \((-\infty, -1.96]\); Decision: Fail to reject the null hypothesis

Use the following to answer questions 26-27.

To test the hypothesis \(H_0: p = 0.7\) Vs \(H_1: p > 0.7\), where \(p\) represents the proportion of customers who were satisfied with the cell phone services being provided by their company. A survey of 400 customers found that 292 of them were satisfied with the company’s cell phone services provided. The test will be performed at the 5% level of significance.

26. The value of the test statistic is closest to which of the following?
(a) -1.35  (b) -1.31  (c) 0.03  (d) 1.31  (e) 1.35

27. The decision of the test is:
(a) Reject \(H_0\) because the p-value is greater than 0.05.
(b) Reject \(H_0\) because the p-value is less than 0.05.
(c) Do not reject \(H_0\) because the p-value is greater than 0.05.
(d) Do not reject \(H_0\) because the p-value is less than 0.05.
(e) No decision can be made from the data because the sample size is not large enough.
Use the following to answer questions 28-30.

It's Friday night and you want to watch a movie. You want to decide between *Interstellar* and *Star Wars: The Last Jedi* by performing a hypothesis test to compare the average ratings between the two movies. The results are given by:

*Interstellar*: $n_1 = 45, \bar{X}_1 = 4.35, S_1 = 0.76$

*Star Wars: The Last Jedi*: $n_2 = 44, \bar{X}_2 = 4.5, S_2 = 0.62$

Test whether there is a difference on the average ratings for the two movies at the 5% level of significance.

28. Set up the null and alternative hypotheses for the test.

   (a) $H_0: \mu_1 - \mu_2 = 0$ vs $H_1: \mu_1 - \mu_2 > 0$

   (b) $H_0: \mu_1 - \mu_2 = 0$ vs $H_1: \mu_1 - \mu_2 < 0$

   (c) $H_0: \mu_1 - \mu_2 = 0$ vs $H_1: \mu_1 - \mu_2 \neq 0$

   (d) $H_0: \bar{X}_1 - \bar{X}_2 = 0$ vs $H_1: \bar{X}_1 - \bar{X}_2 > 0$

   (e) $H_0: \bar{X}_1 - \bar{X}_2 = 0$ vs $H_1: \bar{X}_1 - \bar{X}_2 < 0$

29. The standardized test statistics is

   (a) -4.842

   (b) -1.021

   (c) -0.852

   (d) -0.153

   (e) Cannot be determined

30. The conclusion about the test is

   (a) **Rejection region**: $(-\infty, -1.96] \cup [1.96, \infty)$; **Decision**: Fail to reject the null hypothesis

   (b) **Rejection region**: $(-\infty, -1.96] \cup [1.96, \infty)$; **Decision**: Reject the null hypothesis

   (c) **Rejection region**: $(-\infty, -1.645] \cup [1.645, \infty)$; **Decision**: Fail to reject the null hypothesis

   (d) **Rejection region**: $(-\infty, -1.645] \cup [1.645, \infty)$; **Decision**: Reject the null hypothesis

   (e) **Rejection region**: $(-\infty, -1.28] \cup [1.28, \infty)$; **Decision**: Fail to reject the null hypothesis

End of Multiple Choice Section
Part II

1. Fourth-four sixth graders were randomly selected from a school district. Then, they were divided into 22 matched pairs, each pair having equal IQ’s. One member of each pair was randomly selected to receive special training. Then, all of the students were given an IQ test. Test results are summarized below:

\[ \sum_{i=1}^{22} D_i = 22, \quad \sum_{i=1}^{22} (D_i - \bar{D})^2 = 270, \]

where \( D_i \) is the difference between the score on the IQ test for trained and no trained groups. Do these results provide evidence that the special training helped students’ performance? Use a 0.05 significance level. Assume that the scores are normally distributed.

(Remember that, \( S_D^2 = \frac{\sum_{i=1}^{n} (D_i - \bar{D})^2}{n-1} \))

(a) [3 pts.] Set up the null and alternative hypothesis to test whether the special training helped students’ performance.

(b) [3 pts.] Find the value of the standardized test statistic.

(c) [3 pts.] Find the rejection region and state the decision.

(d) [3 pts.] State your conclusion in the context of the problem.
2. Price $x$ (in 2005 dollars per pound) and consumption $y$ (in pounds per capita) of beef were sampled for ten randomly selected years. Summary information is:

$$
\sum_{i=1}^{10} x_i = 36.19, \quad \sum_{i=1}^{10} x_i^2 = 134.17, \quad 2.9 \leq x \leq 6.2
$$

$$
\sum_{i=1}^{10} y_i = 774.74, \quad \sum_{i=1}^{10} y_i^2 = 60739.23
$$

$$
\sum_{i=1}^{10} x_i y_i = 2832.21
$$

(a) [4 pts.] Find the correlation coefficient $r$.

(b) [3 pts.] Find the equation of the regression line between price and consumption.

(c) [2 pts.] Predict consumption ($y$) when $x=3.1$. 

(d) [4 pts.] Using the value $s_e = 7.619$, construct a 95% confidence interval for the slope of the population regression line.

3. We want to investigate whether there is a difference between the proportion of male and female students that have more than 19 credits this semester at UNCC. We randomly select 400 male students and 320 female students, from which 42 and 30 admitted in having more than 19 credits, respectively.

(a) [3 pts.] Find a point estimate for the difference in the proportion of all male students who have more than 19 credits and the proportion of all female students who have more than 19 credits.

(b) [3 pts.] State the correct hypotheses to test whether the proportion of male students who have more than 19 credits is greater than the proportion of female students who have more than 19 credits.

(c) [3 pts.] Find the value of the standardized test statistic.
(d) [3 pts.] Find the rejection region at $\alpha = 0.05$ and state your decision.

(e) [3 pts.] State your conclusion in the context of the problem.