STAT 1222 Fall 2018
Common Final Exam December 7, 2018

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:

Questions	1	2	3
Maximum	12	12	16
Score			

Part I Part II Total

Part I

1. What is the proportion of adults who are still reading newspapers regularly in USA? In a recent survey, 3000 adults in the United States were asked if they read newspapers regularly. Three hundred of the adults said yes. Which of the following is true?

   (a) The sample is the 300 adults who are reading newspapers regularly.

   (b) The population in this question consists of 3000 adults.

   (c) The sample size is 3000.

   (d) The true proportion of adults who are still reading newspapers regularly in USA is $300/3000 = 0.10$.

Use the following information to answer the questions 2 to 3.
Consider the sample data: 116, 115, 112, 127, 112, 120.

2. What is the sample **mean** of the data?

   (a) 6 (b) 117 (c) 115.5 (d) 115 (e) 116

3. What is the **sample standard deviation** of the data (approximately)?

   (a) 32.8 (b) 27.33 (c) 5.23 (d) 8.37 (e) 5.73

Problems 4 and 5 pertain to the following situation.

The distribution of the lengths of a commercially caught fish in a lake is bell-shaped with mean 12 inch and standard deviation 3 inch.

4. The local authority rules that any fish measuring less than 6 inch must be released. The proportion of fish that must be released is about

   (a) .05 (b) .32 (c) .16 (d) .025 (e) .68

5. Which of the following intervals (length measured in inch) will include approximately 95% of all this commercially caught fish in the lake?

   (a) (12, 15)

   (b) (6, 15)

   (c) (9, 18)

   (d) (6, 18)

   (e) (9, 15)

6. Find $P(-1.25 \leq Z \leq 0.80)$ where Z denotes the standard normal random variable.

   (a) 2.05 (b) 0.1056 (c) 0.8937 (d) 0.7881 (e) 0.6825

7. Linda took a test with her class. Later she learned that the class averaged 80 on the test with a standard deviation of 4, and the z-score of her exam is -2.50. What is her actual score on this test?

   (a) 90

   (b) 70

   (c) 80

   (d) 60

   (e) There is no way to tell Linda's actual score.

For questions 8 and 9: Assume that children's IQs (Age 6-12) follow a normal distribution with mean 100 and standard deviation of 12.

8. Find the probability that a randomly selected child has IQ above 109.

   (a) 0.7734 (b) 0.2266 (c) 0.5428 (d) 0.2734 (e) 0.8543

9. If a local school district wants to set up a gifted program aiming for those children whose IQs are in top 10%, then the minimum IQ score required by this program needs to be (round off to nearest integer)

   (a) 112 (b) 120 (c) 123 (d) 124 (e) 115

The following table denotes the probability distribution for a discrete random variable X. **Use this information to answer questions 10-11.**

X	0	1	2	9
$P(X)$	0.1	0.2	0.2	0.5

10. Find the mean of X.

    (a) 5.1 (b) 3.0 (c) 1.5 (d) 1 (e) 2

11. The standard deviation of X is closest to

    (a) 15.49 (b) 19.9 (c) 4.46 (d) 4.08 (e) 3.94

12. A newspaper wants to estimate the proportion of residents in a large city who approve the job of their current mayor by using the confidence interval method. The ideal confidence interval should have a margin of error (E) of 0.03 with 95% confidence level. The newspaper needs to conduct a survey to achieve this. Find the minimum sample size for this survey.

    (a) 4267 (b) 1068 (c) 752 (d) 1842 (e) 456

13. Type I error in hypothesis test is

(a) Fail to reject H_0 when H_0 is actually false.

(b) Reject H_a when H_a is actually false.

(c) Reject H_0 when H_0 is actually true.

(d) Reject H_a when H_a is actually true.

Use the following information to answer questions 14-15.
The following two-way contingency table gives the breakdown of voters in a particular locale according to gender and political party preference. A person is selected at random from this population.

	Vote for Democratic Party	Vote for Republic Party
Male	240	280
Female	290	190

14. What is the probability that the person will vote for the Republic party?

(a) 0.28 (b) 0.19 (c) 0.24 (d) 0.53 (e) 0.47

15. Find the probability that this person is a female **or** will vote for the Republic party.

(a) 0.95 (b) 0.48 (c) 0.47 (d) 0.76 (e) 0.53

16. In a clinical trial, 50 patients who received a new medication are randomly selected. It was found that 10 of them suffered serious side effects from this new medication. Let p denote the population proportion of patients suffered serious side effects from this new medication. The 95% confidence interval for p is closest to

(a) $10 \pm .093$ (b) $.1 \pm .111$ (c) $.2 \pm .146$ (d) $.2 \pm .111$ (e) $.1 \pm .093$

17. Given that the population standard deviation is $\sigma = 2$, determine the minimum sample size needed in order to estimate the population mean so that the margin of error is $E = .2$ at 80% level of confidence.

(a) 164 (b) 271 (c) 100 (d) 385 (e) 664

18. If a hypothesis is rejected at significance level $\alpha = .001$, is it possible that the hypothesis is **not** rejected if the same test was done at significance level $\alpha = .01$ (with everything else staying the same)?

(a) Yes (b) No (c) No way to tell

19. In a random sample of 20 iPads brought in for repairs, the average repair cost was \$80 with the standard deviation of \$15. Assume that the repair costs are normally distributed. Construct a 95% confidence interval for μ, where μ represents the average repair cost for all iPads.

(a) $80 \pm (1.96)(\frac{15}{\sqrt{20}})$

(b) $80 \pm (2.086)(\frac{15}{\sqrt{20}})$

(c) $20 \pm (1.96)(\frac{15}{\sqrt{80}})$

(d) $20 \pm (1.645)(\frac{15}{\sqrt{80}})$

(e) $80 \pm (2.093)(\frac{15}{\sqrt{20}})$

Use the following information for questions 20 to 21. The equation of the regression line between two variables x (independent variable) and y (dependent variable) is given by $\hat{y} = -3x + 2$; and the correlation coefficient is $r = -.95$. The possible x-values range from 1 to 10.

20. Based on the given r, which of the following conclusions may be made?

(a) x and y are very weakly correlated.

(b) x and y are strongly correlated, and y tends to increase as x is decreased.

(c) x and y are strongly correlated, and y tends to increase as x is increased.

(d) x and y are moderately correlated, and y tends to increase as x is decreased.

21. Predict y if $x = 3.5$.

(a) The predicted y value is -8.5.

(b) The predicted y value is 11.

(c) The predicted y value is 9.

(d) The predicted y value is -10.5.

(e) The y value cannot be predicted.

22. One wants to estimate the population mean μ by using the confidence interval method. A random sample of size $n = 100$ yielded a sample mean 60 with sample standard deviation 10. Which of the following confidence levels will result in the widest confidence interval?

(a) 99% (b) 95% (c) 90% (d) 80% (e) There is no way to tell.

23. In the test of hypothesis $H_0 : \mu = 10$ vs $H_a : \mu \neq 10$, a sample of size 100 yields the standardized test statistic $z = 1.36$. Find the P-value for this test.

(a) 0.9131 (b) 1.8262 (c) 0.0869 (d) 0.1360 (e) 0.1738

For questions 24 and 25: Scores on the common final exam in a statistics course are normally distributed with mean 72 and standard deviation 6. A random sample of 64 students is selected. Let $\bar{x}$ represent the mean score of the sample.

24. Find the mean and standard deviation of $\bar{x}$, i.e., $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$.

    (a) $\mu_{\bar{x}} = 64, \quad \sigma_{\bar{x}} = 6$

    (b) $\mu_{\bar{x}} = 64, \quad \sigma_{\bar{x}} = 0.75$

    (c) $\mu_{\bar{x}} = 72, \quad \sigma_{\bar{x}} = 0.75$

    (d) $\mu_{\bar{x}} = 72, \quad \sigma_{\bar{x}} = 6$

    (e) $\mu_{\bar{x}} = 72, \quad \sigma_{\bar{x}} = 0.125$

25. The probability that the sample mean score $\bar{x}$ is more than 74 is closest to

    (a) 0.0038 (b) 0.3707 (c) 0.6293 (d) 0.9962 (e) 0.4013

The following is used for questions 26 and 27.
In a random sample of 200 adults, 136 say they are in favor of outlawing cigarettes in certain areas. Let p be the proportion of all adults who are in favor of outlawing cigarettes. One is interested for the following hypotheses: $H_0 : p = 0.6$ versus $H_a : p > 0.6$.

26. The standardized test statistics is about

    (a) -2.31 (b) 2.31 (c) 2.43 (d) -2.43 (e) 2.00

27. Assume the significant level for this test is 5%. Find the rejection region for the test.

    (a) $(-\infty, 1.645)$

    (b) $(-\infty, -1.645)$

    (c) $(1.645, \infty)$

    (d) $(1.96, \infty)$

    (e) $(-\infty, 1.96)$

(Continue to the next page.)

28. Two training procedures are compared by measuring the time that it takes trainees to assemble a device. Two groups of trainees are taught, each using a different method. Is there a difference in the two methods? Assume both populations follow normal

Method 1	Method 2
$n_1 = 10$	$n_2 = 12$
$\bar{x}_1 = 35$	$\bar{x}_2 = 31$
$s_1 = 4.9$	$s_2 = 4.5$

distributions with equal variances. Which of the following formula should be used to calculate the standardized test statistic?

(a) $Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

(b) $Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2 + s_2^2}{n_1 + n_2}}}$

(c) $T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

(d) $T = \frac{\bar{d} - D_0}{s_D/\sqrt{n}}$

(e) $Z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$

Use the following information for questions 29 to 30. In order to test hypotheses $H_0 : \mu = 80$ vs. $H_a : \mu < 80$, a random sample of size 25 is taken from a normally distributed population of unknown variance. The sample yields $\bar{x} = 78$ and $s = 8.7$.

29. The standardized test statistics for the above test is about

(a) 1.15 (b) -1.15 (c) -0.05 (d) 1.92 (e) 0.05

30. At the significance level of 0.05, find the rejection region for the test.

(a) $(-\infty, 1.711)$

(b) $(1.711, \infty)$

(c) $(-\infty, -1.645)$

(d) $(-\infty, -1.711)$

(e) $(1.645, \infty)$

End of Multiple Choice Section

Part II

1. Six randomly selected students took an IQ test A, and the next day they took a very similar IQ test B. Their scores are shown in the table below.

Student	1	2	3	4	5	6
Test A	121	93	71	119	104	100
Test B	121	91	72	122	108	100

Assume that the scores in Tests A and B are normally distributed.

(a) (3 pts) Let $d =$ Test A score - Test B score and denote by μ_d the mean of the differences. Find the sample mean and sample standard deviation for d.

(b) (3 pts) On average, do people score better on the second test than in the first test they take? Set up appropriate hypotheses.

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

(d) (3 pts) Find the rejection region at $\alpha = .05$ and state your conclusion in the context of the problem.

2. The American Automobile Association claims that the average daily cost for meals and lodging for vacationing in Texas is larger than the same average costs for vacationing in North Carolina. The table below shows the results of a survey of vacationers in each state. The two samples are independent.

Texas	North Carolina
$n_1 = 35$	$n_2 = 50$
$\bar{x}_1 =$ \$218	$\bar{x}_2 =$ \$208
$s_1 =$ \$28	$s_2 =$\$15

(a) (3 pts) Set up the null and alternative hypotheses to test whether the new treatment **increases** the mean survival period.

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

(c) (3 pts) Find the rejection region at $\alpha = .05$ and state your conclusion in the context of the problem.

(d) (3 pts) Find the P-value for this test.

3. Suppose a experiment is conducted to study the relationship between the percentage of a certain drug in the bloodstream (X) and the length of the time it takes to react to a stimulus (Y). The results are below.

X	2	1	4	3	5
Y	1	1	2	2	4

$\sum x = 15, \quad \sum x^2 = 55, \quad \sum y = 10, \quad \sum y^2 = 26, \quad \sum xy = 37.$

(a) (3 pts) Draw a scatter plot for the data. What can we learn from this plot?

(b) (3 pts) Find the coefficient of correlation between x and y and interpret its meaning in the context of the problem.

(c) (4 pts) Find the equation of the regression line between y and x.

(d) (3 pts) Given a future observation with $x = 2.5$, predict its y value.

(e) (3 pts) Can we predict y value for $x = 0$? Why or why not?