

STAT 1222

Common Final Exam

Fall 2010
Dec 10, 2010

Please print the following information:

Name: _____

Instructor: _____

Student ID #: _____

Section/Time: _____

THIS EXAM HAS TWO PARTS

PART I. Consists of 32 multiple choice questions worth a total of 64 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 2 questions (36 points in total). You MUST show all the 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
Maximum	16	20
Score		

Part I

Part II

Total

Use the following information to answer questions 1, 2, 3 and 4.

The recommended prices (in dollars) for a sample of stocks that an analyst predicts will have 10% annual returns is:

4 3 1 6 9 2 3

1. Find the mean of the data set.
 - (a) 3.6
 - (b) 4
 - (c) 4.8
 - (d) 6
 - (e) 3

2. Find the sample variance of the data set is closest to
 - (a) 6.29
 - (b) 53.78
 - (c) 7.33
 - (d) 2.51
 - (e) 9

3. Find the third quartile of the data set.
 - (a) 3
 - (b) 2.5
 - (c) 1.5
 - (d) 6
 - (e) 3.5

4. Find the mode of the data set.
 - (a) 3
 - (b) 1
 - (c) 9
 - (d) 2
 - (e) 3.5

5. In a recent study of incomes in Allegheny county in Pennsylvania, it was found that the distribution of family incomes is skewed to the right (i.e., it has a long right tail). What can we say about the relationship between mean and median.
- (a) mean and median are same.
 - (b) mean is greater than median.
 - (c) mean is less than median.
 - (d) the median must be thrice the mean.

Use the following information to answer questions

The table below shows the results of a study on 175 people in which researchers examined the relationship between the presence of a certain mutated gene and colon cancer.

	Gene present	Gene absent	Total
Patient has cancer	70	20	90
Patient does not have cancer	25	60	85
Total	95	80	175

6. The probability that a randomly selected person is cancer free but has the mutated gene is closest to
- (a) .49
 - (b) .71
 - (c) .50
 - (d) .29
 - (e) .14
7. The probability that a randomly selected person either has the mutated gene or has cancer is closest to
- (a) .66
 - (b) 1.06
 - (c) .54
 - (d) .51
 - (e) .32

8. The mean sale per customer for 72 customers at a gas station is \$ 36.00 with a standard deviation of \$ 4.00. Using Chebychev's rule, determine at least how many of these sixty customers spent between \$ 24.00 and \$ 48.00.
- (a) 36
 - (b) 68
 - (c) 72
 - (d) 64
 - (e) 54

Use the following information to answer questions 9 and 10

The mean rate for satellite television from a sample of households was \$49.00 with a standard deviation \$2.50. Assume that the distribution of satellite television rates is bell-shaped.

9. Approximately 68% of households pay between
- (a) \$46.50 and \$49.00.
 - (b) \$49.00 and \$54.00.
 - (c) \$41.50 and \$46.50.
 - (d) \$44.00 and \$ 54.00.
 - (e) \$46.50 and \$51.50.
10. Approximately what percentage of households pay more than \$ 54.00.
- (a) 2.5%
 - (b) 5%
 - (c) 95%
 - (d) 68%
 - (e) 45%

Use the following for questions 11, 12 and 13.

The probability distribution $P(x)$ of the number of dogs x per household in a certain county in US is given in the following table.

x	$P(x)$
0	.45
1	.25
2	.15
3	.10
4	.05

11. What is the probability that a randomly chosen household has at least one and at most three dogs?
- (a) .45
 - (b) .40
 - (c) .50
 - (d) .05
 - (e) .95
12. Find the expected number of dogs per household.
- (a) 0
 - (b) 1.05
 - (c) 2
 - (d) 2.5
 - (e) 3.75
13. The standard deviation of x is closest to
- (a) 20.449
 - (b) 3.243
 - (c) 35.812
 - (d) 0
 - (e) 5.984

Use the following information to answer questions 14, 15 and 16.

The weights of adult male beagles are normally distributed, with a mean of 25 pounds and a standard deviation of 4 pounds.

14. The probability that a randomly selected beagle weighs between 15 pounds and 33 pounds is closest to
- (a) 0.5000
 - (b) 0.0062
 - (c) 0.9570
 - (d) 0.9710
 - (e) 1.2334
15. If 75 beagles are randomly selected, then approximately how many are expected to weigh more than 30 pounds?
- (a) 8
 - (b) 15
 - (c) 30
 - (d) 12
 - (e) 25
16. To be in the bottom 10% of the weights, a beagle should weigh at most
- (a) 17.39 pounds.
 - (b) 25.27 pounds.
 - (c) 15.25 pounds.
 - (d) 22.35 pounds.
 - (e) 19.88 pounds.

Use the following for questions 17 and 18.

In a random sample of 15 computers, the mean repair cost was \$ 100 and the standard deviation was \$ 42.50. Assume that the repair cost of computers follow a normal distribution.

17. Construct a 95% confidence interval for μ , the mean repair cost of computers.

- (a) (70.25, 105.49)
- (b) (78.49, 121.51)
- (c) (76.46, 123.54)
- (d) (57.50, 142.50)
- (e) (15, 185)

18. All other information remaining unchanged, which of the following would produce a narrower interval than the 95% confidence interval constructed?

- (a) The sample size is 10 instead of 15.
- (b) The sample size is 28 instead of 15.
- (c) Compute a 98% confidence interval rather than a 95% confidence interval.
- (d) The sample standard deviation is computed to be 30 instead of 42.50

Use the following to answer question 19.

The mean height of men in the US (ages 20-29) is 69.5 inches and the standard deviation is 3.0 inches. A random sample of 49 men between ages 20 - 29 is drawn from this population.

19. Find the probability that the sample mean height \bar{x} is more than 70.5 inches.

- (a) .0099
- (b) .3707
- (c) .0294
- (d) .9901
- (e) .0250.

Use the following information to answer questions 20, 21, 22.

A restaurant claims that its hamburgers have less than 10 grams of fat. In a random sample of 19 hamburgers, a nutritional agency finds that the mean fat content to be 8.5 grams with a standard deviation of 2 grams. Assume that the fat content of bananas follow a normal distribution.

20. Set up the null and alternate hypothesis to test the restaurant's claim.

- (a) $H_0 : \mu = 10$ versus $H_a : \mu \neq 10$
- (b) $H_0 : \mu \geq 10$ versus $H_a : \mu < 10$
- (c) $H_0 : \mu < 10$ versus $H_a : \mu \geq 10$
- (d) $H_0 : \mu \leq 10$ versus $H_a : \mu > 10$
- (e) $H_0 : \mu \neq 10$ versus $H_a : \mu = 10$

21. The value of the standardized test statistic is

- (a) 3.269
- (b) -1.75
- (c) -0.75
- (d) 0.75
- (e) -3.269

22. Find the rejection region at 5% significance level and state your decision.

- (a) Decision rule: Reject H_0 if $z < 1.734$. Decision: Reject H_0 .
- (b) Decision rule: Reject H_0 if $t > 1.734$. Decision: Do not reject H_0 .
- (c) Decision rule: Reject H_0 if $t < -1.734$ or $t > 1.734$. Decision: Reject H_0 .
- (d) Decision rule: Reject H_0 if $t < -1.734$. Decision: Reject H_0 .
- (e) Decision rule: Reject H_0 if $t < -2.101$. Decision: Do not reject H_0 .

23. A machine is supposed to cut plastic into sheets that are 600 inches long. The company wants to estimate the mean length the machine is cutting the plastic, accurate to within 0.04 inch. Determine the sample size required to construct a 92% confidence interval. Assume that the population is normal with a standard deviation of 0.25 inch.

- (a) 120
- (b) 32
- (c) 100
- (d) 945
- (e) 115

Use the following for questions 24, 25 and 26.

A research center on nutrition and public health claims that no more than 40% of US adults eat breakfast every day. In a random sample of 250 US adults, it was found that 120 eat breakfast every day.

24. Set up the null and alternate hypothesis to test the research center's claim.

- (a) $H_0: p \geq .4$ versus $H_a: p < .4$
- (b) $H_0: p \leq .4$ versus $H_a: p > .4$
- (c) $H_0: p = .4$ versus $H_a: p \neq .4$
- (d) $H_0: p < .4$ versus $H_a: p \geq .4$
- (e) $H_0: p \leq .48$ versus $H_a: p > .48$

25. The value of the standardized test statistic is closest to

- (a) 3
- (b) -1.96
- (c) 1.96
- (d) -2.58
- (e) 2.58

26. Find the rejection region and state your decision at $\alpha = .01$.

- (a) Rejection Region: $z > 2.33$; Decision: Reject H_0 .
- (b) Rejection Region: $z > -2.33$; Decision: Fail to reject H_0 .
- (c) Rejection Region: $z < -1.96$ or $z > 1.81$; Decision: Reject H_0 .
- (d) Rejection Region: $z > -2.33$; Decision: Fail to reject H_0 .
- (e) Rejection Region: $z > 1.96$; Decision: Reject H_0 .

27. In a hypothesis testing problem, a researcher wants to test the hypothesis

$$H_0 : \mu = 20 \quad \text{versus} \quad H_a : \mu \neq 20.$$

A sample of size 100 yielded a sample mean $\bar{x} = 19.25$ and sample standard deviation $s = 2.5$. The P-value for the above test is

- (a) .0294
- (b) .0147
- (c) .0013
- (d) .0026
- (e) .5000

Use the following for questions 28, 29 and 30.

An economist wants to test if there is any difference in income between Escambia county and Miami-Dade county in Florida. To do this, she collects random samples of residents from both these counties. The information collected from these samples is tabulated below. Assume that both the populations are normally distributed with **equal** population variances.

Escambia County (Population 1)	Miami-Dade County (Population 2)
$n_1 = 11$	$n_2 = 15$
$\bar{x}_1 = \$36,700$ calories	$\bar{x}_2 = \$34,700$ calories
$s_1 = \$7800$ calories	$s_2 = \$7375$ calories

28. Set up the null and alternative hypotheses to test the economist's claim.
- (a) $H_0 : \mu_1 - \mu_2 \leq 0$ versus $H_a : \mu_1 - \mu_2 > 0$
 - (b) $H_0 : \mu_1 < \mu_2$ versus $H_a : \mu_1 \geq \mu_2$
 - (c) $H_0 : \mu_1 \neq \mu_2$ versus $H_a : \mu_1 = \mu_2$
 - (d) $H_0 : \mu_1 - \mu_2 \geq 0$ versus $H_a : \mu_1 - \mu_2 < 0$
 - (e) $H_0 : \mu_1 - \mu_2 = 0$ versus $H_a : \mu_1 - \mu_2 \neq 0$
29. Find the value of the standardized test statistic.
- (a) 0.64
 - (b) -0.64
 - (c) -2.713
 - (d) 2.713
 - (e) -0.67
30. At 5% significance level, find the rejection region and state your decision.
- (a) Reject H_0 if $t < -1.711$. Decision: Reject H_0 .
 - (b) Reject H_0 if $t < -2.064$. Decision: Reject H_0 .
 - (c) Reject H_0 if $t < -2.064$ or $t > 2.064$. Decision: Do not reject H_0 .
 - (d) Reject H_0 if $t < -1.711$ or $t > 1.711$. Decision: Do not reject H_0 .
 - (e) Reject H_0 if $t > 2.064$. Decision: Do not reject H_0 .

31. In a survey of 1037 US adults ages 65 and over, 643 were concerned about getting the flu. A 99% confidence interval for the proportion p of US adults ages 65 and over who are concerned about getting the flu is closest to
- (a) (.3815, .4227).
 - (b) (.5315, .7935).
 - (c) (.5773, .6025).
 - (d) (.5813, .6589).
 - (e) (.25, .75).

Use the following for question 30.

Data is collected on 11 randomly chosen trees to study the relationship between tree height x (in feet) and its trunk diameter y (in inches). The data is shown below.

x	70	72	75	76	71	73	85	78	77	80	82
y	8.3	10.5	11.0	11.4	9.2	10.9	14.9	14.0	16.3	18.0	15.8

The equation of the best fit line relating x to y is:

$$\hat{y} = 0.560x - 29.922$$

32. The predicted trunk diameter for a tree which has height 74 feet is
- (a) 11.518 inches
 - (b) 25.33 inches
 - (c) cannot be predicted
 - (d) 0.056 inches
 - (e) 29.922 inches

End of Multiple Choice Section

1. A manufacturing plant wants to determine whether its manager performance rating (0-100) changed from last month to this month. The table shows the performance ratings of the same six managers for last month and this month.

Manager	1	2	3	4	5	6
Last month's rating	85	96	70	76	81	78
This month's rating	88	85	89	86	92	89

- (a) (2 pts.) State the correct hypotheses to test the claim that there was no difference in the ratings between last month and this month.

$$H_0 : \quad \quad \quad H_a :$$

- (b) (4 pts.) In the context of the problem, explain Type I error and Type II error.

Type I Error:

Type II Error:

- (c) (5 pts.) Find the value of the standardized test statistic.

(d) (3 pts.) Find the rejection region at $\alpha = .05$

(e) (2 pts.) State your conclusion in the context of the problem.

3. Are older drivers safer than younger drivers? A Wall Street Journal survey studied the connection between x (the age in years of a licensed driver) and y (the percentage of fatal accidents for drivers of that age which are caused by speeding). They collected the following data:

x	17	27	37	47	57	67	77
y	36	25	20	12	10	7	5

$$\sum x = 329, \sum y = 115, \sum x^2 = 18263, \sum y^2 = 2639, \sum xy = 4015, s_e = 3.455$$

- (a) (6 pts.) Find the equation for the regression line (best fit line) relating x to y .

(b) (4 pts.) Construct a 95% prediction interval for the percentage of fatal accidents caused by speeding for a 30 year-old driver.

(c). (3 pts.) Calculate the correlation coefficient, r , between the amount of the bill and the amount of the tip.

(d). (7 pts.) Test for the significance of the correlation coefficient at $\alpha = .05$. Also, fill in the relevant information.

H_0 : H_a :

Standardized Test Statistic:

Rejection Region:

Conclusion: