STAT 1220   FALL 2010

Common Final Exam   December 10, 2010

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, chose the one that is most complete or most accurate. make
sure that your name and ID number are written and correctly bubbled on the OPSCAN sheet.

PART II.

Part II consists of 3 free response questions, with values as indicated. You must show all work in
the space provided or elsewhere on the exam paper in a place that you clearly indicate. Work on
loose sheets will not be graded.

FOR DEPARTMENT USE ONLY:

Part II.

Question	1	2	3
Score			

Part I	Part II	TOTAL

Part I

1. In order to determine the average price per gallon of regular gasoline, the U.S. department of energy conducted a survey of 950 gas stations in the country and found the average price to be \$ 2.77 per gallon in the Thanksgiving weekend last year. The *population* of interest is

    (a) All the 950 gas stations from which price data were collected.

    (b) All gas stations in the country.

    (c) All gas stations excluding the 950 gas stations from which price data were collected.

    (d) All gas stations where the price per gallon is close to \$ 2.77 per gallon.

    (e) None of the above.

For questions 2, 3 and 4:   The following sample lists the number of minutes used by some Stat-1220 students in their last internet session:

    49, 17, 47, 29, 47, 30, 27, 34, 44, 37, 38, 44, 34

2. The sample mean is

    (a) 40.51   (b) 37.12   (c) 36.69   (d) 38.32   (e) 39.23

3. The sample median is

    (a) 34   (b) 37   (c) 33   (d) 36   (e) 35

4. The sample standard deviation is

    (a) 8.35   (b) 8.46   (c) 8.05   (d) 9.46   (e) 9.12

5. Heights of women follow an approximately bell shaped distribution with a mean of 64 inches and standard deviation of 2.75 inches. The percentage of women whose height is between 61.25 inches and 69.5 inches is about

    (a) 95%   (b) 81.5%   (c) 68%   (d) 99.7%   (e) 50%

6. Two fair dice are rolled. The probability that the sum of of scores on the two dice is 8 or more is

(a) 10/36   (b) 26/36   (c) 15/36   (d) 20/36   (e) 6/36

For questions 7, 8, 9 and 10: Suppose all 120 employees (70 males and 50 females) of a company were asked whether they are in favor or oppose paying high salaries to CEOs of Wall Street companies and had the following responses:

Types	Favor	Oppose	Total
Male	25	45	70
Female	14	36	50
Total	39	81	120

An employee of this company is picked at random.

7. The probability that the person is a male and opposed to paying high salaries to CEOs is about

(a) 0.55   (b) 0.64   (c) 0.58   (d) 0.38   (e) 0.47

8. The probability that the person is a male, given that the person is in favor of paying high salaries to CEOs is about

(a) 0.55   (b) 0.72   (c) 0.64   (d) 0.43   (e) 0.32

9. The probability that the person is a female or opposed to paying high salaries to CEOs is about

(a) 0.35   (b) 0.46   (c) 0.68   (d) 0.55   (e) 0.79

10. Which of the following statements about the events $M : Males$ and $R : In favor$ is true?

(a) M are R are independent events because $P(M \text{ or } R) = P(M) + P(R)$
(b) M and R are dependent events because $P(M \text{ or } R) \neq P(M) + P(R)$
(c) M and R are independent events because $P(M \text{ and } R) = P(M)P(R)$
(d) M and R are dependent events because $P(M \text{ and } R) \neq P(M)P(R)$
(e) M and R are independent events because $P(M \mid R) = P(M)$

For questions 11, 12 and 13:   The probability distribution of a discrete random variable x is as shown in the following table

x	1	2	3	4	5
$P(x)$	0.22	0.22	0.37	0.05	0.14

11. $P(x \geq 3)$ is

   (a) 0.22   (b) 0.19   (c) 0.56   (d) 0.44   (e) 0.34

12. The mean of the random variable x is

   (a) 2.67   (b) 3.25   (c) 1.52   (d) 4.12   (e) 3.96

13. The standard deviation of the random variable x is about

   (a) 0.59   (b) 2.27   (c) 1.27   (d) 3.12   (e) 1.96

14. Suppose that a poll of 18 voters is taken in a large city. The random variable x denotes the number of voters who favor a certain candidate for mayor. Suppose that 61% of all the city's voters favor the candidate. The mean and standard deviation of x are respectively

   (a) (10.98, 4.28)   (b) (10.98, 2.07)   (c) (11.12, 1.53)   (d) (9.31, 2.32)
(e) (10.22, 3.12)

15. Let z denote the standard normal random variable. The $P(z < -2.47 \ \ or \ \ z > 2.47)$ is

   (a) 0.0136   (b) 0.0347   (c) 0.0244   (d) 0.4622   (e) 0.0068

For questions 16 and 17: Weights of newborn babies in a certain state have normal distribution with mean 5.33 lb and standard deviation 0.65 lb. A newborn weighing less than 4.85 lb is considered to be *at-risk*, that is, has a higher mortality rate.

16. A baby just born in this state is picked at random. The probability that the baby is *at-risk* is about

(a) 0.43   (b) 0.33   (c) 0.23   (d) 0.13   (e) 0.53

17. The hospital wants to take pictures of the heaviest 10% of the newborn babies. The minimum weight (in lbs) required for a picture to be taken is about

(a) 5.89   (b) 6.16   (c) 9.12   (d) 8.12   (e) 7.27

18. We select a random sample of 39 observations from a population with mean 81 and standard deviation 5.5. The probability that the sample mean is more than 83 is

(a) 0.0501   (b) 0.0714   (c) 0.0627   (d) 0.0116   (e) 0.2315

19. A Gallup Organization poll of $n = 856$ randomly selected American adults in October 2010 found that 60% of those surveyed felt that their weight was about right. A 95% confidence interval for the percentage of American adults who think their weight is about right is

(a) (0.741, 0.852)   (b) (0.342, 0.451)   (c) (0.443, 0.564)   (d) (0.567, 0.633)
(e) (0.572, 0.528)

20. Suppose an insurance company wants to determine the average speed of cars passing through an intersection. They randomly selected 85 cars and found their average speed to be 42 miles per hour with standard deviation of 4.2 miles per hour. A 90% confidence interval for the average speed of all the cars passing through the intersection is

(a) (38.12, 45.12)   (b) (39.52, 44.72)   (c) (42.53, 43.67)   (d) (41.25, 42.75)
(e) (40.09, 43.91)

21. You randomly select 16 restaurants and measure the temperature of the coffee sold at each. The sample mean temperature is $162\,^\circ$F with a standard deviation of $10\,^\circ$F. Assuming that the temperature is approximately normally distributed a 95% confidence interval for the mean temperature is

   (a) (144.7, 156.5)   (b) (156.7, 167.3)   (c) (168.9, 172.5)   (d) (171.5, 181.3)   (e) (172.6, 182.9)

22. You are running a political campaign and wish to estimate, with 95% confidence, the proportion of registered voters who will vote for your candidate. Your estimate must be accurate within 3% of the true population. the minimum sample size required is

   (a) 757   (b) 1725   (c) 1812   (d) 1066   (e) 1068

For questions 23, 24, and 25: A market research analyst claims that 32% of the people who visit the mall actually make a purchase. You think that less than 32% buy something and decide to test this claim of the analyst. You stand buy the exit door of the mall starting at noon and ask 82 people as they are leaving whether they bought anything. You find that actually 20 people made a purchase.

23. State the null and alternative hypotheses to test that less than 32% of mall visitors make a purchase.

   (a) $H_0 : p = 0.32$ versus $H_1 : p < 0.32$
   (b) $H_0 : p = 32$ versus $H_1 : p \geq 32$
   (c) $H_0 : p = 0.32$ versus $H_1 : p > 0.32$
   (d) $H_0 : p = 0.32$ versus $H_1 : p \neq 0.32$
   (e) $H_0 : p \geq 32$ versus $H_1 : p \leq 32$

24. Find the value of the test statistics z is

   (a) 1.48   (b) -1.48 (c) 1.60   (d) -1.60 (e) 1.74

25. The $p-$value of the test is about

   (a) 0.4521   (b) 0.3446   (c) 0.3520   (d) 0.0694   (e) 0.3211

26. A *Type I error* in a test of hypothesis occurs

(a) if we do not reject H_0 when H_0 is really true

(b) if we do not reject H_0 when H_1 is really true

(c) if we reject H_0 when H_1 is really true

(d) if we reject H_0 when H_0 is really true

(e) if we do not reject H_0 when H_0 is really false

27. Suppose the weights of fish caught from pier 1 and pier 2 are normally distributed with equal population standard deviations. The natural assumption is that the mean weights of fish caught at the two piers are equal. The owner of pier 1 thinks that the average at his pier is greater. To test at the 5% level of significance that the average weights of the fish in pier 1 is more than pier 2 the the null and the alternative hypotheses are

(a) $H_0 : \mu_1 - \mu_2 > 0$ versus $H_1 : \mu_1 - \mu_2 \leq 0$

(b) $H_0 : \mu_1 - \mu_2 = 0$ versus $H_1 : \mu_1 - \mu_2 \geq 0$

(c) $H_0 : \mu_1 - \mu_2 = 0$ versus $H_1 : \mu_1 - \mu_2 > 0$

(d) $H_0 : \mu_1 - \mu_2 \geq 0$ versus $H_1 : \mu_1 - \mu_2 < 0$

(e) $H_0 : \mu_1 - \mu_2 = 0$ versus $H_1 : \mu_1 - \mu_2 \neq 0$

28. In a dieting program 8 participants were randomly selected and interviewed. Their weights as they entered the program and their weights (in pounds) after 10 weeks are as follows:

Before	214	211	195	153	165	168	175	180
After	211	208	190	154	165	169	170	179

The population of paired difference of the weights has normal distribution with unknown standard deviation. To test the claim that the program enables the participants to lose weight the formula for the test statistics is

(a) $z = \dfrac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$ (b) $t = \dfrac{b-B}{s_e/\sqrt{SS_{xx}}}$ (c) $t = \dfrac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ (d) $t = \dfrac{\bar{d}-\mu_d}{s_d/\sqrt{n}}$ (e) $t = \dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}$

29. The following is the regression equation for income and food expenditure (in thousands of dollars) for a sample of six households: $\hat{y} = 1.36 + 0.26x$

Given that $s_e = 1.74$ and $SS_{xx} = 1608.83$, a 95% confidence interval for the slope B of the regression line is closest to

(a) $(0.06, \ 0.46)$
(b) $(0.14, \ 0.37)$
(c) $(0.23, \ 0.29)$
(d) $(-0.15, \ 0.65)$
(e) $(3.02, \ 7.91)$

30. The equation of a regression line is $\hat{y} = 13.2 - 2.5x$. If x increases by one unit then

(a) y decreases by about 2.5 units
(b) y increases by about 2.5 units
(c) y decreases by about 13.2 units
(d) y increases by about 13.2 units
(e) the response of y cannot be predicted.

End of Multiple Choice Section

Part II

1. For the past few years the average score on a common final exam in Psychology has been 80. In a sample of 92 students, the mean score on this year's exam was 82 with standard deviation 8.7. Test at 5% level of significance whether the mean score on this year's exam is higher than that of previous years.

a) (3 pts) State the null and the alternative hypotheses.

H_0 :

H_1 :

b) (4 pts) Write down the formula of the test statistics and find its value.

c) (3 pts) Determine the rejection region and make a decision.

d) (2 pts) State you conclusion in the context of the problem.

2. A business school wants to compare a new method of teaching reading to slow learners to the current standard method. They decide to base this comparison on the results of a reading test given at the end of a learning period of six months. Of a random sample of 11 slow learners, 5 are taught by the new method and 6 are taught by the standard method. All 11 children are taught by qualified instructors under similar conditions for a six month period. Assume that the populations are normally distributed and the population variances are equal.

New Method Score: 81   80 79 81 76

Standard Method Score: 69 68 71 68 73 72

Test at 1% level of significance that the new method is worse than the old method, i.e., average score obtained by the new method is less than the average score obtained by the standard method.

a) (3 pts) Compute the mean and standard deviation of *each* data set.

b) (3 pts) State the null and the alternative hypotheses.

H_0 :

H_1 :

c) (3 pts) Write down the formula of the test statistics and find its value.

d) (2 pts) Determine the rejection region and make a decision.

e) (2 pts) Make a conclusion in the context of the problem.

3. Data collected from a sample of five children relating annual per capita sugar consumption (x) (in kilograms) and the number of cavities (y) yielded

$\bar{x} = 5,\ \bar{y} = 2.6,\ SS_{xx} = 16,\ SS_{xy} = 10,\ SS_{yy} = 9.2.$

Values of x ranged from 2 to 7.

    a) (3 pts) Calculate the sample correlation coefficient.

    b) (4 pts) Find the equation of the regression line.

    c) (2 pts) Predict the average number of cavitiesin children who consume 5 kilogram of sugar per year.

    d) (6 pts) Is there sufficient evidence in the data to conclude, at 10% significance level, that the slope of the regression line is nonzero? Interpret the result in the context of the problem.