STAT 1222 FALL 2011
Common Final Exam December 09, 2011

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 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	3
Maximum	15	10	15
Score			

Part I Part II Total

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

$$5 \quad 2 \quad 0 \quad 1 \quad -2 \quad 6$$

1. Find the mean of the data.

(a) 2.4
(b) 2
(c) 4.8
(d) 3
(e) 0

2. The sample standard deviation of the data set is closest to

(a) -1
(b) 4
(c) 1.852
(d) 2.170
(e) 3.033

3. Find the median of the data.

(a) 1
(b) 1.5
(c) 2
(d) 2.5
(e) 6

4. Find the first quartile of the data.

(a) 6
(b) 5
(c) 2
(d) 1
(e) 0

Use the following information to answer questions 5, 6 and 7

The table below shows the results of a study on 102 women in which researchers examined the association between the occurrence of a mutation of the BRCA gene and breast cancer.

	Mutated Gene present	**Mutated Gene absent**	**Total**
Has cancer	33	19	52
Does not have cancer	39	11	50
Total	72	30	102

5. The probability that a randomly selected woman does not have cancer is closest to

(a) .49
(b) .71
(c) .29
(d) .34
(e) .67

6. The probability that a randomly selected woman has cancer and the mutated gene is present is closest to

(a) .49
(b) .71
(c) .27
(d) .32
(e) .46

7. The probability that a randomly selected woman has cancer **or** has the mutated gene, is closest to

(a) 1.216
(b) .892
(c) .36
(d) .27
(e) .75

The weights of bags of chips produced by Great Home Snacks company has a bell-shaped distribution with a mean of 20 ounces and a standard deviation of .07 ounces. **Use this information to answer questions 8, 9 and 10.**

8. Approximately 68% of chips bags weigh between

(a) 15 ounces and 25 ounces.

(b) 19.79 ounces and 20.21 ounces.

(c) 18.15 ounces and 22.5 ounces.

(d) 19.86 ounces and 20.14 ounces.

(e) 19.93 ounces and 20.07 ounces.

9. Approximately what percentage of chips bags weigh between 19.86 and 20.07 ounces.

(a) 95% **(b)** 81.5% **(c)** 99.7% **(d)** 67.2% **(e)** 90.5%

10. If 200 chips bags are selected at random, about how many do you expect to weigh more than 20.14 ounces.

(a) 95 **(b)** 10 **(c)** 32 **(d)** 25 **(e)** 5

The probability distribution of the number of pets x per household in a certain locality is given in the following table. **Use this information to answer questions 11 and 12.**

x	0	1	2	3	4
$P(x)$	?	.28	.40	.10	.07

11. What is the probability that a randomly chosen household in this locality has no pets.

(a) .5 **(b)** .25 **(c)** .15 **(d)** .95 **(e)** .35

12. Find the expected number of pets in this locality.

(a) 0 **(b)** 0.74 **(c)** 1.25 **(d)** 1.66 **(e)** 2.27.

Use the following information to answer questions 13, 14, 15 and 16.
The weights of adult Fox Terriers in US (a dog breed) are normally distributed, with a mean of 15 pounds and a standard deviation of 3 pounds.

13. The probability that a randomly chosen Fox Terrier weighs between 9 pounds and 18 pounds is closest to

(a) 0.7961
(b) 0.2514
(c) 0.8185
(d) 0.0475
(e) .9725

14. To be in the top 13% of the weights, a Fox Terrier should weigh at least

(a) 18.39 pounds.
(b) 25.27 pounds.
(c) 17.52 pounds.
(d) 20.64 pounds.
(e) 15 pounds.

15. A random sample of 100 Fox Terriers is drawn from this population. Identify the mean $\mu_{\bar{x}}$, and standard error $\sigma_{\bar{x}}$, of the sample mean weight $\bar{x}$.

(a) $\mu_{\bar{x}} = 15, \sigma_{\bar{x}} = .3$
(b) $\mu_{\bar{x}} = 1.5, \sigma_{\bar{x}} = 3$
(c) $\mu_{\bar{x}} = 15, \sigma_{\bar{x}} = 3$
(d) $\mu_{\bar{x}} = 15, \sigma_{\bar{x}} = 30$
(e) $\mu_{\bar{x}} = 1.5, \quad \sigma_{\bar{x}} = .3$

16. Find the probability that the sample mean weight $\bar{x}$ exceeds 16 pounds.

(a) .9996
(b) .5000
(c) .6293
(d) .3707
(e) .0004

Use the following information to answer questions 17 and 18.
A car dealer wants to get information about the number of years car owners keep their cars. A random sample of 25 car owners resulted in $\bar{x} = 7.01$ years, and $s = 3.74$ years. Assume that the sample is drawn from a normally distributed population.

17. Construct a 95% confidence interval for the mean number of years of car ownership.

(a) (5.25, 8.19)
(b) (5.55, 8.46)
(c) (5.41, 8.29)
(d) (5.47, 8.55)
(e) (5, 8)

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

(a) The sample size is 29 instead of 25.
(b) The sample size is 10 instead of 25.
(c) Compute a 80% confidence interval rather than a 95% confidence interval.
(d) Compute a 90% confidence interval rather than a 95% confidence interval.
(e) The sample standard deviation is computed to be 1.52 instead of 3.74.

19. An economist wants to estimate the mean income during the first year of employment for a college graduate who has had a statistics course. Find the minimum sample size needed to estimate the mean μ with 88% confidence. The estimate must be accurate to within \$500 of μ. Assume $\sigma = \$6250$.

(a) $n = 375$
(b) $n = 257$
(c) $n = 601$
(d) $n = 378$
(e) $n = 1037$.

20. In a survey of 2563 adults from France, 1666 said that they believed that the activities of humans are contributing to an increase in global temperatures. A 95% confidence interval for the proportion p of adults in France, who think humans are contributing to an increase in global temperatures, is closest to

(a) $(.3815, .4227)$ **(b)** $(.5315, .7935)$ **(c)** $(.6315, .6685)$ **(d)** $(.1727, .7819)$
(e) $(.2500, .7500)$

Use the following for questions 21, 22 and 23.

A Western blot assay is a blood test for the presence of HIV. It has been found that this test some times gives false positive results for HIV. A medical researcher claims that the rate of false positives is less than 2%. A recent study of 300 randomly selected US blood donors who do not have HIV found that 3 received false positive test result.

21. Set up the null and alternate hypothesis to test the researcher's claim.

(a) $H_0 : p = .02$ vs $H_a : p \neq .02$.

(b) $H_0 : p < .02$ vs $H_a : p \geq .02$.

(c) $H_0 : p \geq .02$ vs $H_a : p < .02$.

(d) $H_0 : p \leq .02$ vs $H_a : p > .02$.

(e) $H_0 : p \neq .02$ vs $H_a : p = .02$.

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

(a) 2.21

(b) -2.21

(c) 1.055

(d) -1.24

(e) 1.24

23. Find the rejection region and state your decision at $\alpha = .05$.

(a) Rejection Region: $z > 1.645$; Decision: Reject H_0.

(b) Rejection Region: $z < -1.645$; Decision: Fail to reject H_0.

(c) Rejection Region: $z < -1.96$ or $z > 1.96$; Decision: Reject H_0.

(d) Rejection Region: $z < 1.555$; Decision: Reject H_0.

(e) Rejection Region: $z > -1.555$; Decision: Fail to reject H_0.

Use the following for questions 24, 25 and 26.

A nutritionist in the FDA wants to compare the caloric content of medium french fries sold by Wendonald (Population 1) and McKing (Population 2) fast-food chains, to see if there is any difference between them. To test this, random samples from each chain is taken and the caloric contents of french fries are measured. Their finding is summarized in the table below.

Wendonald	McKing
$n_1 = 40$	$n_2 = 50$
$\bar{x}_1 = 380$ calories	$\bar{x}_2 = 360$ calories
$s_1 = 30$ calories	$s_2 = 38$ calories

24. Set up the null and alternative hypotheses to test whether there is any difference between the caloric content of french fries sold by the two chains.

(a) $H_0 : \mu_1 \geq \mu_2$ versus $H_a : \mu_1 < \mu_2$

(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) $H_0 : \mu_1 \leq \mu_2$ versus $H_a : \mu_1 > \mu_2$

(e) $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 \neq \mu_2$

25. Find the value of the standardized test statistic.

(a) -1.848 **(b)** 1.848 **(c)** -2.79 **(d)** 2.79 **(e)** -1.23

26. Find the P-value for the above test and state your decision at 1% significance level.

(a) P=.0052; Decision: Reject H_0.

(b) P=.0026; Decision: Reject H_0.

(c) P=.0052; Decision: Do not reject H_0.

(d) P=.5000; Decision: Reject H_0.

(e) P=.0322; Decision: Do not reject H_0.

Use the following information to answer questions 27–30.

The data below are the number of absences and the final grades of 9 randomly chosen students from a statistics class.

Number of absences, x	0	3	6	4	9	2	15	8	5
Final grade, y	98	86	80	82	71	92	55	76	82

$$\sum x = 52,\ \sum y = 722,\ \sum x^2 = 460,\ \sum y^2 = 59154,\ \sum xy = 3732.$$

27. Find the sample correlation coefficient r.

(a) .5000 **(b)** $-.6255$ **(c)** .6255 **(d)** .9908 **(e)** $-.9908$

28. We would like to test whether there is significant linear relationship between number of absences and Final Grade, i.e., $H_0 : \rho = 0$ versus $H_a : \rho \neq 0$. The value of the test statistic for this test is closest to

(a) -19.37 **(b)** 19.37 **(c)** $-.9908$ **(d)** 2.5719 **(e)** 0

29. Find the rejection region at $\alpha = .02$ significance level and state your conclusion.

(a) Rejection region: $t < -2.998$ or $t > 2.998$; Decision: Reject H_0.

(b) Rejection region: $t < -2.998$; Decision: Do not reject H_0.

(c) Rejection region: $t < -3.499$ or $t > 3.499$; Decision: Reject H_0.

(d) Rejection region: $t < -3.499$ or $t > 2.998$; Decision: Do not reject H_0.

(e) Rejection region: $t < -2.998$ or $t > 2.998$; Decision: Do not reject H_0.

30. The equation of the best fit line relating x, the number of absences, to y, the Final grade is $\hat{y} = -2.755x + 96.139$. Student A missed 7 classes while student B missed 25 classes. The predicted Final Grades for these students are as follows:

(a) student A: 76.854; student B: 27.264.

(b) student A: 89.33; student B: 50.235

(c) student A: 76.854; student B: predicting using the regression line is not meaningful.

(d) student A: predicting using the regression line is not meaningful; student B: predicting using the regression line is not meaningful.

(e) student A: predicting using the regression line is not meaningful; student B: 27.264.

End of Multiple Choice Section

1. Consider the following table that lists SAT scores before and after a sample of five students took a preparatory course.

Student	1	2	3	4	5
SAT score before course	700	840	830	860	690
SAT score after course	720	840	820	900	700

(a) (2 pts.) State the correct hypotheses to test the effectiveness of the course in improving the SAT scores.

H_0 : H_a :

(b) (2 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) (3 pts.) State your conclusion in the context of the problem.

2. The data below are the ages (in years) of 10 men and their systolic blood pressures.

Age, x	16	25	39	45	49	64	70	29	57	22
Systolic blood pressure, y	109	122	143	132	199	185	199	130	175	118

$\sum x = 416,\ \sum y = 1512,\ \sum x^2 = 20398,\ \sum y^2 = 239514,\ \sum xy = 68173.$

(a) (7 pts.) Find the equation of the regression line.

Regression line: $\hat{y} = mx + b$ where **m=** _____ **b =** _____

(b) (3 pts) Find the coefficient of determination. Explain what it represents?

3. It is believed that mean bumper repair costs for low speed crashes is less for small cars than it is for midsize cars. In order to test this claim, the following data was collected on a sample of small cars and a sample of midsize cars.

Small Car	Midsize Car
$n_1 = 12$	$n_2 = 16$
$\bar{x}_1 = \$473$	$\bar{x}_2 = \$741$
$s_1 = \$190$	$s_2 = \$205$

Assume that the repair costs are normally distributed **with equal population variances**.

(a) (**2 points**) Set up the null and alternative hypotheses to test the claim.

H_0 : H_a :

(b) (**5 points**) Find the value of the appropriate standardized test statistic and say whether it is a t or z.

(c) (**5 points**) Find the rejection region at 1% significance level.

(d) (**3 points**) State your conclusion in the context of the problem.