

MA 180/418 Midterm Test 3, Version A Fall 2010

Student Name (PRINT):

Student Signature:

The test consists of 10 questions.

Questions 1 through 5 are multiple-choice and worth 5 points each.

Questions 6 through 10 are computational and worth 15 points each.

For questions 1 through 5, circle the correct answer (**a**, **b**, **c**, or **d**) after each question. Each question is 5 points. No partial credit.

Q1 How do you compute the **number of degrees of freedom** when testing hypotheses about means in two independent samples, whose sizes are n_1 and n_2 ?

- (a) It is $n_1 + n_2 - 1$.
- (b) It is the larger of $n_1 - 1$ and $n_2 - 1$.
- (c) It is the smaller of $n_1 - 1$ and $n_2 - 1$.
- (d) It is the product of $n_1 - 1$ and $n_2 - 1$.

Correct answer (circle one): (a) (b) (c) (d)

Q2 Which of the following statements is true about the **correlation** and **causality**?

- (a) If the correlation is linear, then causality exists.
- (b) If the correlation coefficient is larger than the critical value, then causality is present.
- (c) There cannot be correlation without causality.
- (d) There may be correlation without causality.

Correct answer (circle one): (a) (b) (c) (d)

Q3 What is the meaning of **least-squares property** of a regression line ?

- (a) The sum of the squares of the residuals is the smallest sum possible.
- (b) The sum of the residuals of the squares is the smallest sum possible.
- (c) The square of the correlation coefficient is the smallest square possible.
- (d) The square of the regression equation is the smallest square possible.

Correct answer (circle one): (a) (b) (c) (d)

Q4 What is the meaning of the term **linear correlation coefficient**?

- (a) Linear correlation coefficient, r , measures the amount of variation in y that is explained by the regression line.
- (b) Linear correlation coefficient, r , measures the strength of the linear correlation between the x and y variables.
- (c) Linear correlation coefficient is b_1 , the slope of the regression line.
- (d) Linear correlation coefficient is b_0 , the y -intercept of the regression line.

Correct answer (circle one): (a) (b) (c) (d)

Q5 When do you use the **F-distribution** and when do you use the **chi-square distribution**?

- (a) Use F-distribution to test hypotheses about means in two populations and chi-square distribution to test hypotheses about mean in one population.
- (b) Use F-distribution to test hypotheses about mean in one population and chi-square distribution to test hypotheses about means in two populations.
- (c) Use F-distribution to test hypotheses about variances in two populations and chi-square distribution to test hypotheses about variance in one population.
- (d) Use F-distribution to test hypotheses about variance in one population and chi-square distribution to test hypotheses about variances in two populations.

Correct answer (circle one): (a) (b) (c) (d)

For questions 6 through 10, write your answer in the space provided. Show your work. Each question is worth 15 points.

Q6 In a 2003 survey of 670 college students, 164 said they used illegal drugs. In a more recent survey of 1045 college students, 204 said they used illegal drugs. Use a 0.01 significance level to test the claim that the proportion of college students using illegal drugs has decreased since 2003.

- (a) State the null and the alternative hypotheses.

- (b) Compute the pooled sample proportion. (Round off to **four** decimal places.)

- (c) Compute the test statistic. (Round off to **two** decimal places.)

- (d) Draw a diagram. Find and mark the critical value(s).

- (e) Make an initial conclusion and a final conclusion.

- (f) Find the P-value. (Round off to **four** decimal places.)

- (g) Make an initial conclusion by using the P-value. Clearly state the rule you use.

Q7 A statistician collected a random sample of the cents portion from 71 checks and from 110 credit card charges. The cents portions of the checks have a mean of 46 cents and a standard deviation of 22 cents. The cents portions of the credit card charges have a mean of 25 cents and a standard deviation of 29 cents. Construct a 98% confidence interval estimate of the difference between the mean of the cents portions from checks and the mean of the cents portions from credit card charges. Do not assume that σ_1 and σ_2 are known or equal.

(a) Write down an appropriate formula for the interval (copy it from the formula card even if you use a calculator).

(b) Write down an appropriate formula for the margin of error E , including the rule for the degrees of freedom (copy those from the formula card even if you use a calculator).

(c) Find the interval, i.e., the left confidence limit and the right confidence limit. (Round off to **two** decimal places.) If you use a calculator, name the respective function.

Q8 Listed below are the measurements of the blood pressure of five subjects taken on their right arm and left arm. Construct a 90% confidence interval for a difference between the measurements from the two arms.

Right arm:	108	114	100	98	109
Left arm:	119	126	107	99	114

(a) Write down an appropriate formula for the interval (copy it from the formula card even if you use a calculator).

(b) Write down an appropriate formula for the margin of error E , including the rule for the degrees of freedom (copy those from the formula card even if you use a calculator).

(c) Find the interval, i.e., the left confidence limit and the right confidence limit. (Round off to **two** decimal places.) If you use a calculator, name the respective function.

Q9 Listed below are the measurements of the blood pressure of five subjects taken on their right arm and left arm. Use the pressure in the right arm as the x variable and the pressure in the left arm as the y variable.

Right arm:	108	114	100	98	109
Left arm:	119	126	107	99	114

(a) Find the linear correlation coefficient. (Round off to **three** decimal places.)

(b) Use 0.05 significance level to test the hypothesis that there is a linear correlation between x and y . Write down the critical value. Make the conclusion.

(c) Find the equation of the regression line. (Round off the regression coefficients to **two** decimal places.)

Q10 Use the data from Question 9 (above) to do the following.

(a) Find the best predicted blood pressure in the left arm given that the blood pressure in the right arm is 100. Show which formula you use and explain why. (Round off to **one** decimal place.)

(b) Find the coefficient of determination r^2 . (Round off to **four** decimal places.)

(c) Find the standard error of estimate s_e . (Round off to **four** decimal places.)

(Bonus) Find the total variation

(Bonus) Find the explained variation

(Bonus) Find the unexplained variation