MA 180/418 Final Exam, Version A

Student Name (PRINT):.....

Student Signature:

The test consists of 20 questions.

Questions 1 through 10 are multiple-choice and worth 2.5 points each. Questions 11 through 20 are computational and worth 7.5 points each.

For questions 1 through 10, circle the correct answer $(\mathbf{a}, \mathbf{b}, \mathbf{c}, \text{ or } \mathbf{d})$ after each question. Each question is 2.5 points. No partial credit.

Q1 What is a **critical value**?

- (a) A critical value is the best point estimate of a population parameter.
- (b) A critical value is the degree of confidence when constructing a confidence interval.
- (c) A critical value is the value of one endpoint of a confidence interval.
- (d) A critical value is the number on the borderline separating sample statistics that are likely to occur from those that are unlikely to occur.

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

Q2 What is the meaning of the terms **population** and **sample**?

- (a) A population is the complete collection of all individuals of a large size, while a sample is the complete collection of all individuals of a small size.
- (b) A population is a set of parameter values, while a sample is a set of statistics.
- (c) A population is the complete collection of all individuals, while a sample is a sub-collection of members selected from a population.
- (d) A population is a set of categorical data, while a sample is a set of quantitative data.

Q3 How is the **t-distribution** different from the **z-distribution**?

- (a) The t-distribution has a mean $\mu = 1$, while the z-distribution has a mean $\mu = 0$.
- (b) The t-distribution has a standard deviation $\sigma = 0$, while the z-distribution has a standard deviation $\sigma = 1$.
- (c) The t-distribution is wider than the z-distribution.
- (d) The t-distribution is narrower than the z-distribution.

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

- Q4 Which of the following is true about the **chi-square distribution**?
 - (a) The shape of the chi-square distribution depends on two different degrees of freedom (numerator and denominator).
 - (b) The chi-square distribution has a mean of $\mu = 0$.
 - (c) Values of the chi-square distribution can be negative.
 - (d) The chi-square distribution is non-symmetric.

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

Q5 What is the difference between the null hypothesis and alternative hypothesis?

- (a) The null hypothesis is a statement that the value of a population parameter is equal to a claimed value, while the alternative hypothesis is a statement that the value of a sample statistic is different from the claimed value.
- (b) The null hypothesis is a statement that the value of a population parameter is equal to a claimed value, while the alternative hypothesis is a statement that the parameter has a value different from the claimed value.
- (c) The null hypothesis is a statement about a property of a population, while the alternative hypothesis is a statement about a property of a sample.
- (d) The null hypothesis is a statement that the value of a population mean is equal to a claimed value, while the alternative hypothesis is a statement that the value of a population proportion is equal to a claimed value.

Q6 What are the properties of the **F-distribution**?

- (a) It is symmetric and its values are positive.
- (b) It is not symmetric and the exact shape depends both on the numerator degrees of freedom and the denominator degrees of freedom.
- (c) It is symmetric and the exact shape depends on the smaller of the two degrees of freedom: one for the numerator and one for the denominator.
- (d) It is not symmetric and its values can be anything between -1 and 1.

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

- Q7 What is the meaning of the terms **regression line** and **regression equation**?
 - (a) The regression equation is an algebraic description of the linear relationship between two variables x and y, and the regression line is the graph of the regression equation.
 - (b) The regression equation is the formula to compute the margin of error when constructing a confidence interval, and the confidence interval is on the regression line.
 - (c) The regression equation is the formula to compute the linear correlation coefficient, and the regression line is known as the line of best fit.
 - (d) The regression equation is an algebraic description of the linear relationship between two variables x and y, and the regression line is a line separating the x and y variables.

Q8

What is a **prediction interval**?

- (a) A prediction interval is an interval estimate of a population parameter.
- (b) A prediction interval is an interval estimate of a predicted value of y, the response variable.
- (c) A prediction interval is an interval where x, the predictor variable, can take its value.
- (d) A prediction interval is the same as the critical region in a hypothesis test.

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

Q9 What is the meaning of the term **coefficient of determination**?

- (a) It is the ratio of the explained variation to the total variation.
- (b) It is the ratio of the explained variation to the unexplained variation.
- (c) It is the ratio of the unexplained variation to the explained variation.
- (d) It is the difference between the total variation and the explained variation.

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

Q10 Which of the following statements is true about the correlation and causality?

- (a) Correlation implies causality.
- (b) Correlation does not imply causality (i.e., there may be correlation without causality).
- (c) If the correlation is linear, then there is causality.
- (d) If the correlation coefficient is larger than the critical value, then causality is present.

For questions 11 through 20, write your answer in the space provided. Show your work. Each question is worth 7.5 points.

Q11 Scores on the SAT test have a mean of 1550 and a standard deviation of 330. Scores on the ACT test have a mean of 21.5 and a standard deviation of 4.9. Which is relatively better: a score of 1195 on the SAT test or a score of 16.3 on the ACT test? Why?

Q12 A cell phone company manager wants to determine the percentage of adults who live in a household with cell phones and no land-line phones. How many adults must he survey? Assume that he wants to be 99% confident that the sample percentage is within four percentage points of the true population percentage.

(a) Assume that nothing is known about the percentage of adults who live in a household with cell phones and no land-line phones.

(b) Assume that a recent survey suggests that about 9% of adults live in a household with cell phones and no land-line phones.

Q13 When Mendel conducted his experiments, he used peas with green pods and yellow pods. In one experiment, 25% (or 145) of the 580 offspring peas were expected to have yellow pods. Instead of getting 145 peas with yellow pods, he obtained 149. Assume that Mendel's 25% rate is correct.

(a) Find the mean number of peas with yellow pods, μ , and its standard deviation σ (round off the value of σ to three decimal places).

(b) Compute $\mu \pm 2\sigma$, give interpretation to these numbers.

(c) Use normal approximation to compute the probability of getting *at least* 149 peas with yellow pods (among 580 offspring peas).

[Bonus] Use normal approximation to compute the probability of getting *exactly* 149 peas with yellow pods (among 580 offspring peas).

Q14 A sample of 36 randomly selected M&M candies has a mean weight of 1.32g and a standard deviation of 0.05g. Construct a 90% confidence interval estimate for the mean of weights of all M&Ms.

Q15 Suppose that of 795 randomly selected adults, 580 said it is morally wrong to not report all income on tax returns. Use a 0.02 significance level to test the claim that 75% of all adults say that it is morally wrong to not report all income on tax returns:

(a) State the null and the alternative hypotheses.

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

(c) Compute the test statistic. Make an initial conclusion and a final conclusion.

(d) Find the P-value of the test. Make a conclusion by using the P-value.

Q16 What sample size is needed to estimate the mean white blood cell count for the population of adults? Assume that you want 90% confidence that the sample mean is within 0.14 of the population mean. The population standard deviation is 2.4.

Q17 A sample of 24 filtered cigarettes is obtained and the tar content in each cigarette is measured. The sample has a standard deviation 3.6mg. Use a 0.05 significance level to test a claim that the tar content in filtered cigarettes has a standard deviation more than 3.0mg.

(a) State the null and the alternative hypotheses.

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

(c) Compute the test statistic. Make an initial conclusion and a final conclusion.

[Bonus] For an extra credit, find an interval for the P-value by using a table.

Q18 Listed below are the measurements of the blood pressure of five subjects taken on their right arm and left arm. Use a 0.10 significance level to test for a difference between the measurements from the two arms.

Right arm:	105	112	101	95	107
Left arm:	114	125	108	99	110

(a) State the null and the alternative hypotheses.

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

(c) Compute the test statistic. Make an initial conclusion and a final conclusion.

[Bonus] For an extra credit, find the P-value of the test by calculator or find an interval for the P-value by using a table.

Q19 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:	105	112	101	95	107
Left arm:	114	125	108	99	110

(a) Find the linear correlation coefficient. Use 0.05 significance level to test the hypothesis that there is a linear correlation between x and y

(b) Find the equation of the regression line

(c) 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.)

Q20 Use the data from Question 19 (above) to do the following.

(a) Find the coefficient of determination r^2

(b) Find the standard error of estimate s_e

(c) Find the total variation

(d) Find the explained variation

(e) Find the unexplained variation

[Bonus] Find a 95% prediction interval for the estimate of the blood pressure in the left arm given that the blood pressure in the right arm is 100.