**MA 180/418 Final Exam, Version A Spring 2011**

Student Name (PRINT):. . . . . .

Student Signature: . . . . . .

The test consists of 20 questions.
Questions 1 through 10 are multiple-choice and worth 5 points each.
Questions 11 through 20 are computational and worth 15 points each.

For questions 1 through 10, circle the correct answer (**a**, **b**, **c**, or **d**) after each question. Each question is 5 points. No partial credit.

Q1 In the following list, one random variable is discrete and three others are continuous. Which one is **discrete**?

(a) The weight of a randomly selected person from a population.

(b) The age of a randomly selected person from a population.

(c) The duration (in minutes) of a randomly selected speech at university meetings.

(d) The length (in words) of a randomly selected speech at university meetings.

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

Q2 What is the meaning of the **continuity correction**?

(a) If the random variable is discrete, then after the correction it becomes continuous.

(b) Each whole number $x$ in the binomial distribution is represented by the interval from $x - 5$ to $x + 5$.

(c) Each whole number $x$ in the binomial distribution is represented by the interval from $x - 0.5$ to $x + 0.5$.

(d) It allows us to apply the Central Limit Theorem for small samples ($n \leq 30$).

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

Q3 Of the following, which is the correct interpretation regarding a **95% confidence interval** for a population proportion?

(a) The sample proportion $\hat{p}$ will be correct 95% of the time.

(b) The sample proportion $\hat{p}$ will be the best point estimate 95% of the time.

(c) If we repeat the procedure many times, approximately 95% of the intervals so constructed will contain the true population proportion.

(d) The confidence interval contains 95% of the sample data.

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

Q4 What is a **Type-I error** in hypothesis testing?

(a) A Type-I error occurs when the test statistic is outside the critical region.

(b) A Type-I error occurs when the test statistic falls in the critical region.

(c) A Type-I error occurs when the null hypothesis is accepted when, in fact, the null hypothesis is false.

(d) A Type-I error occurs when the null hypothesis is rejected when, in fact, the null hypothesis is true.

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

Q5 Which statement about the **margin of error** is correct?

(a) It is the probability of Type-I error.

(b) It is the probability of Type-II error.

(c) It is the length of a confidence interval.

(d) It is half-length of a confidence interval.

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

Q6 Which statement about the **null hypothesis** is correct?

(a) The null hypothesis is rejected when the P-value is high ($> \alpha$)

(b) The null hypothesis is accepted when the P-value is high ($> \alpha$)

(c) The null hypothesis always agrees with the original claim.

(d) The null hypothesis is accepted when the test statistic falls in the critical region.

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

Q7 How do you compute the **number of degrees of freedom** when testing hypothesis about means in two independent samples?

(a) It is the product of $(n_1 - 1)$ and $(n_2 - 1)$.

(b) It is the smaller of $(n_1 - 1)$ and $(n_2 - 1)$.

(c) It is the larger of $(n_1 - 1)$ and $(n_2 - 1)$.

(d) It is the quotient of $(n_1 - 1)$ and $(n_2 - 1)$.

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

Q8 What is the best description of **linear correlation coefficient**?

(a) The linear correlation coefficient is the first coefficient (y-intercept) of the regression equation.

(b) The linear correlation coefficient is the second coefficient (slope) of the regression equation.

(c) The linear correlation coefficient measures the strength of the linear correlation between two variables.

(d) The linear correlation coefficient is the square of the coefficient of determination.

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

Q9 Why do we use the word "squares" in the **least squares property** of a regression line?

(a) Because this property involves the squares of residuals.

(b) Because this property is based on the square of the linear correlation coefficient, i.e., on $r^2$.

(c) Because the regression line cuts through given squares.

(d) Because the regression line has to pass between given squares.

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

Q10 How do we compute **expected frequencies** if they are not all equal?

(a) We use the rule $E = n/k$.

(b) We use the rule $E = np$.

(c) We use the rule $E = n/p$.

(d) We use the rule $E =$(upper confidence limit - lower confidence limit)/2.

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

For questions 11 through 20, write your answer in the space provided. Show your work. Each question is worth 15 points.

Q11 In the accompanying table, the random variable $x$ represents the number of computers per household.

(a) Find the mean $\mu$ and standard deviation $\sigma$ of the given probability distribution. (Round off to **two** decimal places.)

| $x$ | $P(x)$ |
|---|---|
| 0 | 0.04 |
| 1 | 0.09 |
| 2 | 0.25 |
| 3 | 0.20 |
| 4 | 0.18 |
| 5 | 0.16 |
| 6 | 0.04 |
| 7 | 0.03 |
| 8 | 0.01 |
| 9 | 0+ |

(b) Use the Range Rule of Thumb to determine the minimum usual value and the maximum usual value.

(c) List all the values of $x$ that are unusual by the above rule.

Q12 Use the random variable from question Q11 to answer the following questions:

(a) Find the probability that a randomly selected household has 6 computers.

(b) Find the probability that a randomly selected household has at least 6 computers.

(c) Find the probability that a randomly selected household has more than 6 computers.

(d) Which question, (a) or (b) or (c), is relevant for determining wether 6 is an unusually high number of computers?

(e) Determine wether 6 is an unusually high number of computers. Why or why not?

Q13 Use the random variable from question Q11 to answer the following questions:

(a) Determine all unusually high values of $x$ by using probabilities.

(a) Determine all unusually low values of $x$ by using probabilities.

[Bonus] Do the unusual values of $x$ that you find here in question Q13 have to be the same as you found in question Q11?

Q14 By using a gender selection method designed to increase the probability of conceiving a girl, 497 babies have been born of which 343 were girls.

(a) What is the best point estimate for the population proportion of girls born to parents using this method?

(b) Construct a 98% confidence interval estimate of the population proportion of girls born to parents using this method. (Round off the confidence limits to **three** decimal places.) If you use technology, name the respective software or the calculator function.

(c) Does this gender selection method appear to be effective? Why or why not?

Q15 Use the data from question Q14 and the significance level 0.5% to test the claim that the described gender selection method is effective.

(a) State the null and the alternative hypotheses.

(b) Compute the test statistic (round off to **two** decimal places).

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

(d) Make an initial conclusion and a final conclusion.

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

Q16 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 98% confident that the sample percentage is within two 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 26% of adults live in a household with cell phones and no land-line phones.

Q17 The measurements of the lead in the air shown below were recorded on 9 consecutive days:

3.15 1.98 0.99 2.11 1.12 1.94 0.96 1.24 1.57

Assume that the population has normal distribution.

(a) Compute the sample mean, sample standard deviation, and the quartiles (round off to **two** decimal places):

$\bar{x} =$ $s =$ $Q_1 =$ $Q_3 =$

(b) Construct a 90% confidence interval for the population variance. (Round off the confidence limits to **two** decimal places.) If you use technology, name the respective software or the calculator function.

(c) What is the requirement on the population in this test? How strict is it?

Q18 Reserachers collected a random sample of the times that 101 college students required to earn their bachelor's degrees. The sample has a mean of 4.7 years and a standard deviation of 2.3 years. Use a 0.10 significance level to test a claim that the mean time for all college students is less than 5 years.

(a) State the null and the alternative hypotheses.

(b) Compute the test statistic (round off to **two** decimal places).

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

(d) Make an initial conclusion and a final conclusion.

(e) What are the requirements on the population/sample in this test? How strict are they?

[Bonus] Find the P-value of the test (round off to **three** decimal places). Or use a table to determine an interval for the P-value.

Q19 Claim: weights of babies born to mothers given placebos vary more than weights of babies born to mothers given zinc supplements. Test it at a 0.05 significance level. Sample results are summarized below:

| | | | |
|---|---|---|---|
| Placebo group: | n=61 | $\bar{x} = 3066$ g | $s = 801$ g |
| Treatment group: | n=41 | $\bar{x} = 3013$ g | $s = 704$ g |

(a) State the null and the alternative hypotheses.

(b) Compute the test statistic (round off to **four** decimal places).

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

(d) Make an initial conclusion and a final conclusion.

[Bonus] Find the P-value of the test. Make a conclusion by using the P-value.

Q20 Records of randomly selected births were categorized according to the day of the week that they occurred. Use a 0.05 significance level to test the claim that the births occur with the same frequency on every day of the week:

| Day | Sun | Mon | Tue | Wed | Thu | Fri | Sat |
|---|---|---|---|---|---|---|---|
| Number of births | 88 | 125 | 111 | 127 | 119 | 105 | 97 |

(a) State the null and the alternative hypotheses.

(b) Compute the expected frequencies.

(b) Compute the test statistic (round off to **four** decimal places).

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

(d) Make an initial conclusion and a final conclusion.

[Bonus] Find the P-value of the test (round off to **three** decimal places). Or use a table to determine an interval for the P-value.