Carleton UNIVERSITY

FINAL
EXAMINATION
December 4, 2008

DURATION: 3 HOURS **No. of Students: approx 350**

Department Name & Course Number: Mathematics & Statistics, STAT2606
Course Instructor(s): Dr. Horn, Dr. Park, Dr. Sinha

AUTHORIZED MEMORANDA
Non-Programmable, Non-Graphing Calculators Allowed

Students MUST count the number of pages in this examination question paper before beginning to write, and report any discrepancy to a proctor. This question paper has six two-sided pages.

Students will be provided with a formula sheet and a sheet of tables by the professor.

This examination question paper may not be taken from the examination room.

This examination question paper may not be released to the library.

In addition to this question paper, students require a formula sheet and statistical tables provided by the instructors.

Name: _____

Student Number: _____

Section (circle one): **A (Sinha)** **B (Horn)** **C (Park)**

There are **21** multiple choice questions worth a total of **21 marks**.
There are **8** written questions worth a total of **54 marks**.
The exam is out of **75 marks**.
You have **3 hours** to complete the exam.

DO NOT OPEN THE EXAM UNTIL YOU HAVE BEEN TOLD TO DO SO

Part A – Multiple Choice Questions

INSTRUCTIONS: Correct answers are worth **1 mark**. Incorrect and blank answers are worth **0 marks**. Please read each question carefully and **clearly circle only one answer**.

1. In order to determine the quality of a shipment of 20 parts, a sample of 3 items is randomly selected without replacement from the shipment. Four of the 20 items in the shipment are actually defective. Let Y be a random variable representing the total number of defective items in the sample. Then P(Y = 1) is

(A) 0.48 (B) 0.60 (C) 0.08 (D) 0.42 (E) 0.92

2. The value of Z representing the first quartile of the standard normal distribution is

(A) 0.67 (B) -0.67 (C) 1.28 (D) -1.28

3. In testing $H_0 : \mu = 50$ versus $H_1 : \mu \neq 50$, a random sample of 9 observations from a normally distributed population yielded a test statistic of $t_0 = 2.03$. The p-value for this test is

(A) greater than 0.10.
(B) between 0.05 and 0.10.
(C) between 0.025 and 0.05.
(D) between 0.01 and 0.025.
(E) less than 0.01.

4. Suppose a random variable follows a binomial distribution with 10 trials and probability of success as 0.8. Which of the following statements is true?

(A) The shape of the distribution is skewed to the right and normal approximation to binomial is not appropriate.

(B) The shape of the distribution is symmetrical and normal approximation to binomial is not appropriate.

(C) The shape of the distribution is skewed to the left and normal approximation to binomial is appropriate.

(D) The shape of the distribution is skewed to the left and normal approximation to binomial is not appropriate.

5. A normal random variable *X* has an unknown μ and variance is 4. If the probability that *X* exceeds 7.5 is 0.8023, find μ.

(A) 10.9 (B) 5.8 (C) 9.2 (D) 8.6

6. A researcher wants to conduct a hypothesis test at the $\alpha = 0.01$ significance level. She designs her study so that the probability of a Type II error is 0.05 (for some fixed value of the parameter of interest). If the alternative hypothesis is correct, the probability that the researcher will reject the null hypothesis is

(A) 0.05 (B) 0.01 (C) 0.99 (D) 0.95

7. Which of the following statements about the Central Limit Theorem (CLT) is correct?

(A) The CLT states that the sample mean $\overline{X}$ is always equal to the population mean μ.

(B) The CLT states that the sample mean $\overline{X}$ is equal to the population mean μ provided that the sample size is large enough ($n \geq 30$).

(C) The CLT states that the sampling distribution of the sample mean $\overline{X}$ is exactly normal for large sample sizes ($n \geq 30$).

(D) The CLT states that the sampling distribution of the population mean μ is approximately normal when $n \geq 30$.

(E) The CLT states that the sampling distribution of the sample mean $\overline{X}$ is approximately normal for large sample sizes ($n \geq 30$).

8. The p-value for a two-sided test of the null hypothesis $H_0 : \mu = 10$ is calculated as 0.06. Does the 90% confidence interval constructed include the value 10?

(A) Yes (B) No (C) Not enough information to tell.

9. Suppose you have obtained a confidence interval for μ, but wish to obtain a greater degree of precision. Which of the following would result in a narrower confidence interval?

(1) Increasing the sample size while keeping the confidence level fixed.

(2) Decreasing the sample size while keeping the confidence level fixed.

(3) Increasing the confidence level while keeping the sample size fixed.

(4) Decreasing the confidence level while keeping the sample size fixed.

(A) 1 or 4 (B) 1 or 3 (C) 2 or 4 (D) 2 or 3 (E) None of these.

10. You have measured the systolic blood pressure of a random sample of 25 employees of a company. A 95% confidence interval for the mean systolic blood pressure for the employees is computed to be (122, 138). Which of the following statements gives a valid interpretation of this interval?

(A) 95% of the sample of employees has a systolic blood pressure between 122 and 138.

(B) 95 % of the employees in the company have a systolic blood pressure between 122 and 138.

(C) If the sampling procedure were repeated 100 times, then approximately 95 of the sample means would be between 122 and 138.

(D) If the sampling procedure were repeated 100 times, then approximately 95 of the resulting 100 confidence intervals would contain the true mean systolic blood pressure for all employees of the company.

(E) We are 95% confident the sample mean is between 122 and 138.

11. The management of a local restaurant wants to estimate the average amount their customers spend at the restaurant to within \$0.50, with a 95% confidence. What is the minimum sample size required, if the true standard deviation is assumed to be \$3.50?

(A) 188 (B) 189 (C) 325 (D) 326 (E) 133

12. What measure of central tendency would be most appropriate for a measurement of salaries when there are a few people in the sample who make over one million dollars, but most of the employees sampled made under \$50,000?

(A) Mean (B) Median (C) Mode (D) Standard Deviation (E) IQR

13. The probability that a customer will buy a product given that he or she has seen an advertisement for the product is 0.15. The probability that a consumer will see an ad for this particular product is 0.20. What is the probability that a consumer will both see the ad and buy the product?

(A) 0.75 (B) 0.05 (C) 0.35 (D) 0.03 (E) 0.075

14. A very bad basketball player only has a 20% chance of successfully making a free throw shot. He will keep shooting until he makes a shot. What is the probability he will make his first shot within four attempts?

(A) 0.5904 (B) 0.1024 (C) 0.0546 (D) 0.1280 (E)

15. If $P(A \cup B) = 0.8$, $P(A) = 0.3$, and $P(B) = 0.6$, what is $P(A \cap B^C)$?

(A) 0.1 (B) 0.2 (C) 0.3 (D) 0.4 (E) 0.6

16. When is the t-distribution used in the construction of confidence intervals?

(A) When the population standard deviation is not known and the sample is small.

(B) When the population standard deviation is known.

(C) Either the standard normal or t-distribution may be used at anytime.

(D) Always.

(E) When the sample is not random or when the population distribution is strongly skewed.

17. You are told that the p-value for a hypothesis test was 0.03. This means:

(A) There is only a probability of 0.03 that the null hypothesis is true.

(B) If H_1 is true, we have observed outcomes that occur 3% of the time.

(C) If H_0 is true, test statistic values that are at least as extreme as the value found from this sample occur 3% of the time.

(D) There is a probability of 0.03 that the alternative hypothesis is true.

18. Weekly sales of regular ground coffee at a supermarket have in the recent past varied according to a normal distribution with mean of 354 units per week. The store reduces the price of coffee by 5%. Is there good evidence that average sales are now higher? If a test is carried out, which of the following hypotheses should be tested to answer this question?

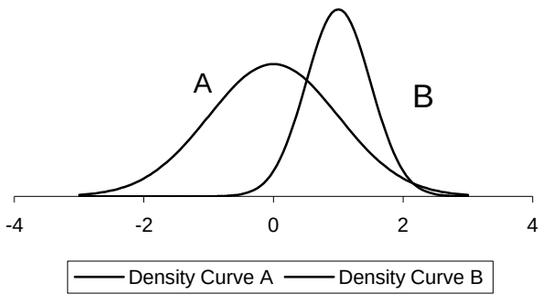
(A) $H_0 : \mu \geq 0.05$ versus $H_1 : \mu < 0.05$

(B) $H_0 : \mu \leq 0.05$ versus $H_1 : \mu > 0.05$

(C) $H_0 : \mu \geq 354$ versus $H_1 : \mu < 354$

(D) $H_0 : \mu \leq 354$ versus $H_1 : \mu > 354$

19. Two normal density curves are given in the graph below.



If μ_A and σ_A are the mean and standard deviation from density curve A and μ_B and σ_B are the mean and standard deviation from density curve B, which of the following statements is true?

(A) μ_A is less than μ_B and σ_A is equal to σ_B.
(B) μ_A is less than μ_B and σ_A is less than σ_B.
(C) μ_A is greater than μ_B and σ_A is greater than σ_B.
(D) μ_A is less than μ_B and σ_A is greater than σ_B.
(E) μ_A is greater than μ_B and σ_A is less than σ_B.

20. Your professor tells bad jokes in class according to a Poisson process with an average of 2 bad jokes every 30 minutes. What is the probability that your professor will tell exactly 5 bad jokes during a 90 minute class?

(A) 0.5420 (B) 0.1606 (C) 0.0030 (D) 0.4124 (E) 0.0361

21. The sample mean is an unbiased estimator for the population mean. This means:

(A) The sample mean always equals the population mean.

(B) The average sample mean, over all possible samples, equals the population mean.

(C) The sample mean is always very close to the population mean.

(D) The sample mean will only vary a little from the population mean.

(E) The sample mean has a normal distribution.

Part B – Written Questions

INSTRUCTIONS: Each question outlines its own marking scheme. Please read each question carefully. **For full marks, be sure to show all of your steps.**

1. Suppose the weights of packages of Oreo cookies have a normal distribution with a mean of 252 grams and a standard deviation of 9 grams.

(A) (**3 marks**) What is the probability that a randomly selected package will weigh between 245 grams and 265 grams?

(B) (**2 marks**) When selecting a random sample of 4 packages of Oreo cookies, what is the sampling distribution of the sample mean $\overline{X}$ and why?

(C) (**3 marks**) If you select 4 packages of Oreo cookies at random, what is the probability the average weight of these 4 packages will be between 250 grams and 255 grams?

(D) (**3 marks**) What weight should be advertised on the packages so that only 5% of packages are below that weight?

2. An agricultural researcher plants 25 randomly selected plots with a new variety of corn. From these 25 plots, the average yield was 150 bushels per acre. The population standard deviation is known to be 20 bushels per acre.

(A) (**4 marks**) Give a 99% confidence interval for the population mean yield μ bushels per acre. Interpret your answer in terms of the question. What do you need to assume in order for your confidence interval to be valid?

(B) (**2 marks**) Suppose the researcher wanted to estimate the true average yield, to within 5 bushels per acre, with 99% confidence. How many plots of land should he use in order to achieve the desired margin of error?

3. An experiment was conducted to test the effect of a new drug on a viral infection. The infection was induced in 100 mice, and the mice were randomly split into two groups of 50. The first group, the control group, received no treatment for the infection. The second group, the treatment group, received the drug. After a 30-day period, the number of survivors in the control group and treatment group were 18 and 30, respectively.

(A) (**3 marks**) Give a 95% confidence interval for the difference between the cure rates. Interpret the result in terms of the question.

(B) (**4 marks**) Is there sufficient evidence to indicate that the drug is effective in treating the viral infection? Conduct a hypothesis test at the 5% level of significance.

4. (**5 marks**) A research manager at Coca-Cola claims that the proportion, p, of cola drinkers that prefer Coca-Cola over Pepsi is greater than 50%. In a consumer taste test, 100 randomly selected people were given blind samples of Coca-Cola and Pepsi. 58 of these subjects preferred Coca-Cola. Is there sufficient evidence at the 5% level of significance to validate Coca-Cola's claim? Conduct an appropriate hypothesis test.

5. (**5 marks**) Tony's Tire Company claims that his tires last more than 1,500 miles longer than his competitor Willy's Wheels. Two independent samples of 10 tires from each company are subjected to wear and the average life of Tony's tires is found to be 16,700 miles, with a standard deviation of 1,700 miles. Willy's tires lasted a mean of 15,100 miles, with a standard deviation of 1,350 miles. Test Tony's claim at the 10% level of significance. Assume that Tony's tires correspond to "Population 1" and Willy's tires correspond to "Population 2". You may assume that $\sigma_1^2 = \sigma_2^2$.

6. The following is the probability distribution for the number of cameras X sold by a very unpopular camera store in a given day. Unfortunately, some tuna fish got spilled on the probability for $x = 2$ and it is now illegible (and smells bad).

x	0	1	2	3	4
$P(X=x)$	0.40	0.30	???	0.10	0.05

(A) (**4 marks**) Compute μ_X and σ_X^2.

(B) (**2 marks**) The store is open 225 days per year. Consider these 225 days to be a random sample of days, and $\overline{X}$ to be the average number of cameras sold. What can you say about the sampling distribution of $\overline{X}$ and why?

(C) (**3 marks**) Approximate the probability that the stores sells at total of at least 260 cameras in one year.

7. 24% of the patients who come to a local clinic have high blood pressure (B), 12% are overweight (W) and 6% suffer from both conditions.

(A) (**2 marks**) What proportion of the patients have at least one of these conditions?

(B) (**2 mark**) Are the events B and W independent? Why or why not?

(C) (**2 marks**) Are the events B and W mutually exclusive? Why or why not?

8. (**5 marks**) In a test of the effectiveness of a device intended to increase gas mileage, four cars were operated with and without the device. The mileages (in mpg), as well as some sample statistics, are given in the table below.

CAR	WITHOUT DEVICE	WITH DEVICE
1	18.5	22.5
2	30.7	31.7
3	27.8	28.8
4	31.5	33.5

The researcher wishes to show that the device actually increases the average mileage. Conduct an appropriate hypothesis test at $\alpha = 0.10$.