MATH 203 Final Examination April 27, 2009

Student Name:

Student Number:

McGill University
Faculty of Science
FINAL EXAMINATION

MATH 203
Principles of Statistics I
April 27th, 2009
9 a.m. - 12 Noon

Answer directly on the test (use front and back if necessary).

Calculators are allowed.

One 8.5" × 11" two-sided sheet of notes is allowed.

Language dictionaries are allowed.

There are 17 pages to this exam and 2 pages of tables.

The total number of marks for the exam is 100.

Examiner: Professor Russell Steele

Associate Examiner: Professor David Stephens

MATH 203 Final Examination April 27, 2009

Question 1: (10 points)

The Canadian government has decided that there was some suspicious, potentially illegal activity regarding a company's revenue stream. As part of their investigation, they wanted to establish that there was a significant increase in the mean transaction amount before and after a particular event. The government statistician took a random sample of transactions with 50 customers who were billed both before and after the date of interest (i.e. each customer was billed twice). The data are summarized in the table below:

	Mean	Std Dev	25%ile	Median	75%ile	Sample size
Before event	98.82	20.49	83.99	101.2	109.3	50
After event	108.3	17.94	96.25	107.8	120.1	50
Difference [After - Before]	9.47	19.81	5.188	9.01	24.72	50

Test for a significant increase in the mean transaction amount before and after the event at $\alpha = 0.05$.

MATH 203 Final Examination April 27, 2009

Question 2: (10 points)

Professor Steele gave out a surprise quiz during one of his classes. The quiz had 3 questions. Each question was worth **2 points**, however no partial credit was given. Assume that the probability distribution (i.e. probability mass function) for the number of correctly answered questions by a randomly selected student is:

Number of correctly answered questions	0	1	2	3
Probability	0.05	0.20	0.40	0.35

(a) What is the distribution (i.e. the probability mass function) for the *total score* on the quiz for a randomly selected student? [2 points]

(b) What is the expectation and variance of the total quiz score for a randomly selected student? [2 points]

3

(c) What is the distribution (i.e. the probability mass function) for the *average* (or *mean*) of **two** randomly selected student quiz scores? [3 points]

(d) What is the approximate distribution for the *average* (or *mean*) of **100** randomly selected student quiz scores? [3 points]

MATH 203 Final Examination April 27, 2009

Question 3: (6 points)

In a 1995 paper in Astrophysics and Space Science, Higgins and Henrikson discussed the distribution of gamma-ray bursts in halo neutron star-comet models. They based most of their conclusions on a simulation model where the velocities of comets ejected from globular cluster stellar systems were assumed to be Normally distributed with mean velocity (in km/s) equal to 200 and standard deviation 110.

(a) Find the proportion of comets ejected with velocities less than 150 km/s. [2 points]

(b) Find the proportion of comets ejected with velocities between 100 km/s and 240 km/s. [2 points]

(c) Find the proportion of comets ejected with velocity equal to *exactly (without rounding to whole numbers)* 200 km/s. [2 points]

MATH 203 Final Examination April 27, 2009

Question 4: (10 points)

In Alberta, a way for organizations to raise funds for charitable or religious purposes is to obtain a licence for a casino gaming event. A variety of card and roulette-based games are permitted, with much of the revenue going to the good cause. A gaming official claims that about 10% of these casino gaming licences are granted to support the arts.

(a) An auditor is testing the claim, at $\alpha = 0.05$, and selects a random sample of 190 licenses granted in the past year. She finds that 15 of licenses were used to support the arts. What conclusion should be reached? [6 points]

(b) If the auditor wanted the margin of error for the sample proportion to be $\pm 1\%$, how large would her sample size need to be? [4 points]

MATH 203 Final Examination April 27, 2009

Question 5: (10 points)

Animal behaviourists have discovered that the more domestic chickens peck at objects placed in their environment, the healthier the chickens seem to be. White string has been found to be a particularly attractive pecking stimulus. In one experiment, 72 chickens were exposed to a string stimulus. Instead of white string, blue-colored string was used. The number of pecks each chicken took at the blue string over a specified interval of time was recorded. Summary statistics for the 72 chickens were $\bar{x} = 1.13$ pecks and $s = 2.21$ pecks.

(a) Construct a 99% confidence interval to estimate the population mean number of pecks made by chickens pecking at blue string. Interpret the result. [7 points]

(b) Previous research has shown that $\mu = 7.5$ pecks if chickens are exposed to white string. Based on the results found in part (a), is there evidence that chickens are more apt to peck at white string than blue string? Explain. [3 points]

MATH 203 Final Examination April 27, 2009

Question 6: (12 points)

A bike rental chain owner owns rental locations in both Westmount and the Plateau. The owner was interested in trying to detect a difference between the two locations in the mean distance that the renters travelled on the bikes. The owner randomly selected a particular day and only rented bikes to customers that had odometers on the bikes so that he could measure the distance (in miles) travelled by each renter. The table and figures below contain the data for all of the renters at the two bike shops on that particular day.

	Mean	Std Dev	25%ile	Median	75%ile	Sample size
Westmount bike shop	13.15	10.56	8.478	11.47	14.97	13
Plateau bike shop	21.38	9.298	13.98	22.82	30.23	11

```
> stem(Westmount)

 The decimal point is 1 digit(s) to the right of the |

  0 | 14589
  1 | 011256
  2 | 8
  3 |
  4 | 0

> stem(Plateau)

The decimal point is 1 digit(s) to the right of the |

  0 | 6
  1 | 1179
  2 | 349
  3 | 113
```

MATH 203 Final Examination April 27, 2009

(a) Test to see if there was a statistically significant difference in mean bike distances between the two offices. (Use $\alpha = 0.05$). What assumption(s) need(s) to be made in order for the inference to be valid? Are they met in this particular problem? [9 points]

(b) Calculate the p-value (or an approximate p-value) related to the hypothesis test in part (a). [3 points]

MATH 203 Final Examination April 27, 2009

Question 7: (10 points)

The Canadian Food Inspection Agency has become worried about synthetic hormone levels in Canadian beef products. They would like to devise a control process that would allow them to detect when the mean synthetic hormone level of a particular supply of beef is above a particular level. Obviously, they cannot test every package of beef, so they randomly sample 200 packages from each supplier in order to decide whether the mean hormone level for that supplier is too high (anything over 200 ppm of the synthetic hormone would be considered unsafe). Assume that the distribution of the hormone levels from a supplier are normally distributed with unknown mean and with known standard deviation 10.

(a) Describe what a Type I error and what a Type II error would be in the context of their hormone testing. [4 points]

(b) What is the approximate distribution of the sample mean hormone level of the 200 randomly sampled packages from a particular supplier if the true mean hormone level is 203 ppm? [2 points]

CONTINUED ON NEXT PAGE

(c) If we assume that they want to limit themselves to a Type I error probability of 0.001, what would their approximate Type II error be if the true mean hormone level were actually 203 ppm? [4 points]

MATH 203 Final Examination April 27, 2009

Question 8: (10 points)

In a recent newspaper article on health costs, the author reported prevalence and incidence statistics for diabetes in Canada. The annual prevalence of a disease is the total number of cases of the disease in the population divided by the number of people in the population for a given year. The annual incidence of a disease is the total number of **new** cases in the population during the year. The newspaper article reported that the prevalence of diabetes in Canada is 350 cases per 10000 people in the population while the incidence of diabetes in Canada is 26 cases per 10000 people.

(a) Assuming the prevalence stated in the newspaper is the true prevalence, what is the probability of observing more than 40 people with diabetes in a random sample of 1000 Canadians?

(b) Assuming the incidence stated in the newspaper is the true prevalence, what is the probability of observing more than 5 people with diabetes in a random sample of 1000 Canadians?

Question 9:   (12 points)

A student who was interested in sociology and public health carried out an interesting survey of mothers to look for the predictive ability of some old wives' tales concerning pregnancy. In particular, she was interested in various cultural theories of attributes of the mother's pregnancy and the gender of the baby. There were two particular "tales" that she was interested in.

- Whether the mother had serious bouts of morning sickness (beyond just a general queasiness) which supposedly indicates a higher probability of having a girl

- What the mother thought she was having before finding out officially (either via ultrasound or at the birth)

She sent out a survey to a very large number of recent mothers (restricting her sample to only those who had non-multiple births, i.e. no twins, triplets, etc.) asking three questions:

1. Did you have very serious morning sickness during your pregnancy?

2. What did you think your baby's gender was before you had your 20-week ultrasound?

3. What was the eventual gender of your baby?

and found the following:

- 60% of the women who gave birth to a girl had very serious morning sickness during their pregnancies whereas only 25% of the women who gave birth to boys had very serious morning sickness

- 70% of the women who gave birth to girls thought they were going to have a girl and 80% of the women who gave birth to boys thought that they were going to have a boy

- 51% of the women surveyed gave birth to boys

- Whether or not one has morning sickness and whether one believes that they are going to have a girl (or a boy) are independent given the gender of the baby that the woman gives birth to

CONTINUED ON NEXT PAGE

(a) If a woman has serious morning sickness during her pregnancy, what does the study say that her probability of having a boy should be? [4 points]

(b) If a woman thinks that she is going to have a boy, what is the probability that she will have a boy? [4 points]

(c) If a woman has serious morning sickness during her pregnancy and thinks that she is going to have a boy, what is the probability that she eventually ends up having a boy? [4 points]

MATH 203 Final Examination April 27, 2009

Question 10: (10 points)

The table below contains the results for a study conducted on two treatments for kidney stones, Treatment A and Treatment B. The table contains the number of successes and failures for each treatment for two kinds of stones: small stones and large stones.

Small Stones			
	Successes	Failures	Total
Treatment A	162	12	174
Treatment B	468	72	540
Total	630	84	714
Large Stones			
	Successes	Failures	Total
Treatment A	384	142	526
Treatment B	110	50	160
Total	494	192	686

(a) Estimate the difference in the pooled (i.e. not stratified by type of stone) probability of success between the two treatments using a 90% confidence interval. Which of the two treatments do you believe is better to use based on the pooled probability of success with the two treatments? [3 points]

CONTINUED ON NEXT PAGE

(b) Estimate the difference in the probability of sucess between the two treatments using a 90% confidence interval **for each type of kidney stone [small and large] separately**. Which of the two treatments do you believe is better to use when looking at the two types of stone separately? [4 points]

(c) Did your conclusions differ in parts (a) and (b)? If so, state what phenomenon you've observed. [3 points]