NATIONAL UNIVERSITY OF SINGAPORE

EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE

(Semester I: 2002–2003)

ST2334 PROBABILITY AND STATISTICS

November 2002 — Time Allowed : 2 Hours

INSTRUCTIONS TO CANDIDATES

1. This examination paper contains **FIVE (5)** questions and comprises **TEN (10)** printed pages.
2. Answer **ALL** the questions. The number in [ ] indicates the number of marks allocated for that part. The total number of marks for this paper is 60.
3. Write your answers in the spaces provided.
4. Candidates may use calculators. However, they should lay out systematically the various steps in the calculations.
5. Candidates may bring in <u>one</u> handwritten A4-size (210 × 297 mm) help sheet.
6. Statistical tables are provided.

Matriculation No.: | | | | - | | | | | | - | | |

Question No.	1	2	3	4	5	**Total**
Score						

QUESTION 1

The number of bacteria colonies of a certain type in a sample of polluted water has a Poisson distribution with a mean of 3 per cubic-centimeter.

(a) If five 2-cubic-centimeter samples are independently selected from this water, what is the probability that at least two samples will contain four or more bacteria colonies? [5 marks]

(b) Suppose a k-cubic-centimeter sample is to be selected from this water. How large should k be in order to have a probability of at least 0.90 of seeing one or more bacteria colonies? [3 marks]

(a) A wall is to be built with 100 cement blocks. The length of the cement blocks has mean 20cm and standard deviation 0.2cm.

(i) Use Chebyshev's inequality to obtain a lower bound on the probability that the length of the wall is within 5cm of 2000cm. [4 marks]

(ii) Use Central Limit Theorem to evaluate the probability of the event mentioned in part (i), and compare with the probabilistic statement given in part (i). [4 marks]

(b) A farmer wants to estimate the area of a square field. When he measures the length of the field, his measurement of the observed length is normally distributed with an expected value equal to the true length μ and a standard deviation σ. Concerned about his measurement error, he makes two independent measurements X_1 and X_2. He isn't sure whether he should average these values and then square,

$$Y = \left(\frac{X_1 + X_2}{2}\right)^2$$

or square first and then average,

$$W = \frac{X_1^2 + X_2^2}{2}$$

For which procedure is the expected value of the estimate of the area closer to the true area μ^2? Justify your answer. [4 marks]

Let X and Y be the amount of calories, in hundreds of milligrams, found in a tuna sandwich and a cheese burger of a fast-food chain MacDonny, respectively. The joint probability density function of X and Y is given by

$$f(x, y) = \begin{cases} \frac{1}{20}(x+y) & \text{if} \quad 1 < x < 3,\ 2 < y < 4, \\ 0 & \text{otherwise.} \end{cases}$$

If you buy a tuna sandwich and a cheese burger from MacDonny, find the probabilities that

(a) the amount of calories in the burger is higher than the amount of calories in the sandwich;

[5 marks]

(b) the total amount of calories in both the sandwich and the burger exceeds 500 milligrams.
[5 marks]

Let $X_1, X_2, \ldots, X_{10}$ be a random sample of size 10 taken from a normal population with mean μ and standard deviation 5.

(a) Peter, a statistician, constructs a confidence interval, $(\bar{X} - 1, \bar{X} + 2)$, for the unknown μ. Determine the confidence level of this interval. [3 marks]

(b) Peter also wants to test the hypotheses

$$H_0 : \mu = 4 \qquad \text{versus} \qquad H_1 : \mu \neq 4$$

He decides to reject H_0 if $\bar{X} < 2$ or $\bar{X} > 7$.

(i) Show that the significance level of Peter's test is 0.1325. [3 marks]

(b) (cont'd)

(ii) If the true μ is 5.5, evaluate the power of Peter's test. [3 marks]

(iii) John, another statistician, claims that tests of the form

$$\text{Reject } H_0 \text{ if } \bar{X} < 4 - k \text{ or } \bar{X} > 4 + k$$

for some $k > 0$, have better power than the test designed by Peter. Determine the value k such that the significance levels of John's test and Peter's test are the same. For this value of k, is John's claim true at $\mu = 5.5$? Justify your answer. [6 marks]
(Note: $z_{0.05} = 1.6449$, $z_{0.06} = 1.5548$, $z_{0.07} = 1.4758$, $z_{0.08} = 1.4051$)

An experiment was carried out to test whether weight gain for pigs fed ration A is higher than those fed ration B. Eight pairs of pigs were used. The rations were assigned at random to the two animals within each pair. The gains (in pounds) after 45 days, assuming normally distributed, are given below:

Pairs	1	2	3	4	5	6	7	8	$\sum x$	$\sum x^2$
Ration A	65	37	40	47	49	65	53	59	415	22319
Ration B	58	39	31	45	47	55	59	51	385	19187
Difference A - B	7	-2	9	2	2	10	-6	8	30	342

(a) If the pigs within each pair were littermates, test the hypothesis that ration A is better, in terms of weight gain, than ration B at a 10% significance level. [5 marks]

(b) Suppose you are told now that the pigs within each pair are not littermates and they are in fact independently selected. Test the same hypothesis as in (a) at a 10% significance level. (Hint: You need to test the hypothesis of equal variances first. Then use your result to perform an appropriate test for the hypothesis in (a)) [10 marks]

END OF PAPER