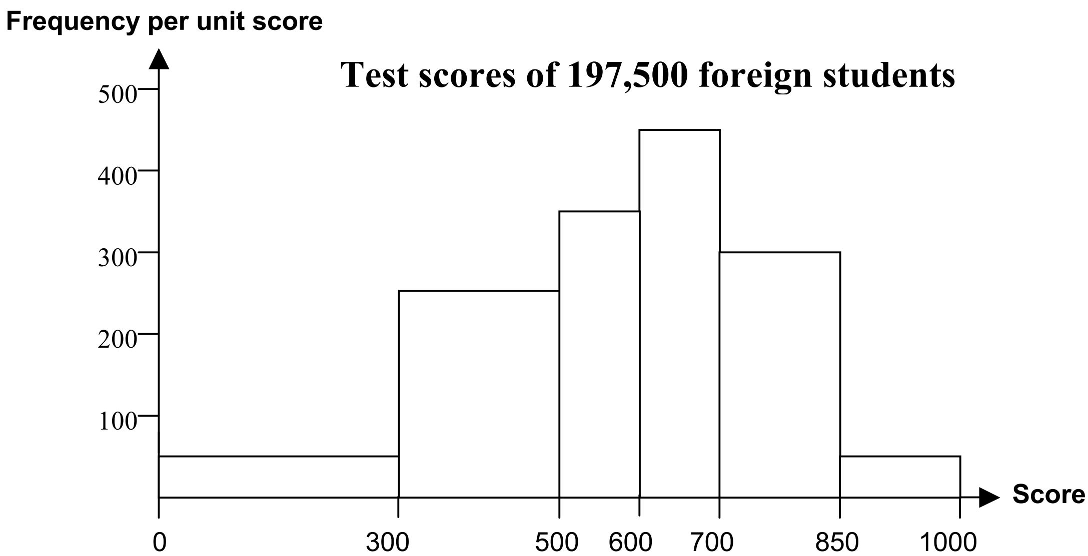
MATH 244 Applied Statistics

Final Examination May 24, 2001

Time allowed : 3 hours
Answer all six questions.

1. (*10%*) A disease is known to affect one out of every 100 people in a population and there is a clinical test being used to test for the presence of this disease. When a person has the disease, the test comes back positive 99% of the time. The test also produces some false positives: 3% of the uninfected people will be tested positive. John has just been tested positive. What is the chance that he gets the disease?

2. (*20%*) Scores on a language test for 197,500 foreign students in year 2000 are summarized by the following histogram. The standard deviation of these scores is 207.



(a) Sketch a boxplot for these scores. (*8 marks*)

(b) Calculate the mean of these scores. (*4 marks*)

(c) If a random sample with size 50 is drawn from these students, what is the probability that their average score will be between 500 and 600? (*4 marks*)

(d) If the top 5% among those students who had taken the test are to be given special recognition, what score does a student need in order to receive special recognition? (*4 marks*)

P. T. O.

3. (*15%*) Electrical power failures in a factory have been modelled as a Poisson experiment with a rate of 1.5 per month. Suppose each power failure will result in a loss of USD400 .

   (a) What is the expected number of power failures in a year? (*2 marks*)

   (b) What is the probability that the number of failures in a year will differ by more than a standard deviation from the expected number? (*5 marks*)

   (c) Find the probability that at least eight months during the year will have monthly loss less than USD1000. (*5 marks*)

   (d) Let T be the waiting time (in months) until there is a accumulative loss of USD10000 due to power failures. Write down the distribution of T . (*3 marks*)

4. (*15%*) In Hong Kong, the H5N1 virus had been detected in 13 out of 30 chicken specimens randomly selected from the markets. The following statement was found in a newspaper: "Based on this result, more than 40% of the chickens in Hong Kong should have been infected by the H5N1 virus."

   (a) Find a 95% confidence interval for p, the proportion of chickens in Hong Kong that were infected by the H5N1 virus. (*5 marks*)

   (b) Critize the statement. (*5 marks*)

   (c) How many specimens should we test in order to obtain a 95% confidence interval for p with length less than 0.05? (*5 marks*)

5. (*25%*) An experiment was conducted to compare the mean power level readings (in watts) on a type of military electronic tube by two purportedly identical pieces of test equipments. Sixteen electronic tubes, randomly selected from production, were randomly divided into two groups. The power output of each electronic tube in one group was measured by the first piece of test equipment and the power output of each electronic tube in the other group was measured by the second piece of test equipment. The data are summarized below.

Power Level Readings

Tester 1	2563	2665	2460	2650	2610	2657	2590	2427
Tester 2	2556	2579	2426	2619	2617	2591	2570	2446

Assume that the populations are normal. Let μ_1, μ_2, σ_1^2, σ_2^2 be the population means and population variances of the power outputs measured by tester 1 and tester 2 respectively.

(a) Test $H_0: \sigma_1^2 = \sigma_2^2$ against $H_1: \sigma_1^2 \neq \sigma_2^2$ at 10% significance level. What can be concluded from this test? (*5 marks*)

(b) Test $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$ at 5% significance level. What can be concluded from this test? (*5 marks*)

(c) Construct a 95% confidence interval for μ_1 and a 95% confidence interval for μ_2. (*6 marks*)

(d) Do we have 95% confidence that **both** of the two intervals in (c) contain the corresponding parameters μ_1 and μ_2? Explain briefly. (*4 marks*)

(e) Construct a 90% confidence interval for $\mu_1 - \mu_2$. (*5 marks*)

P. T. O.

6. (*15%*) The following table shows the result in an experiment to investigate the effect of vaccination of laboratory animals against a particular disease. Let p_V be the probability that a vaccinated animal will get the disease and p_N be the probability that a non-vaccinated animal will get the disease.

	Got disease	Did not get disease
Vaccinated	9	42
Not vaccinated	17	28

(a) Suppose we want to establish the assertion that the vaccination is effective in protecting the laboratory animals. What hypotheses should we test? Write down the hypotheses in terms of p_V and p_N. (*5 marks*)

(b) Test your hypotheses in (a) at 5% significance level. What can be concluded? (*5 marks*)

(c) Construct a 95% confidence interval for $p_V - p_N$. What can be concluded from this confidence interval? (*5 marks)*

< E N D >