NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

B. COMM (HONOURS) DEGREE IN ACTUARIAL SCIENCE

ACTUARIAL STATISTICS II: CIN 2211

JULY/AUGUST 2004 SUPPLEMENTARY EXAMINATION

TIME ALLOWED: 3HOURS

INSTRUCTIONS TO CANDIDATES

(1) Candidates should answer all questions.
(2) In addition to this paper candidates should have a copy of statistical tables and an electronic calculator.

---

1. A random sample of size 50 is taken from a population with mean 100 and standard deviation 25. Find (approximately) the probability that the sample mean lies between 95 and 105. **[5marks]**

2. Let θ be the probability that a fuse of a particular type is defective. It is required to test $H_0 : \theta = 0{,}02$ v $H_1 : \theta = 0{,}01$. Suppose that 500 fuses are tested, and the critical region is taken to be "$x \leq 5$", where x is the number of defective fuses.

   Find the probabilities of the two types of decision error. **[6 marks]**

3. A random sample of 50 pairs of observations is taken from a bivariate normal distribution with correlation coefficient p. The sample correlation coefficient is 0,58. By using Fisher's transformation, calculate a 95% confidence interval for p. **[6 marks]**

4. (i) Claims arise on a certain type of policy as a Poisson process with rate λ per year, so that the number of claims which arise in one year follows a Poisson distribution with parameter λ. A random sample of n such policies is found to yield $X_1, X_2 \ldots, X_n$ claims during the last year.

   (a) Show that the maximum likelihood estimator of the claim rate λ is given by $\hat{\lambda} = \overline{X}$

   (b) If a random sample of 40 such policies yields a total of 68 claims, calculate the maximum likelihood estimate of the annual claim rate λ.

(c) Use a normal approximation to obtain a 95% confidence interval for annual claim rate λ.

[7 marks]

(ii) In a random sample of 200 current accounts held with bank A, 34% were found to be overdrawn. At bank B, a random sample of 200 current accounts contained only 23% which were overdrawn.

Calculate a 95% confidence interval for $P_A - P_B$, the difference in the underlying proportions of current accounts which are overdrawn for the two banks. **[5marks]**

[Total 12 marks]

5. The number of milligrams of tar in random samples of filter and non-filter cigarettes were recorded as follows:

Filter 0,9 1,1 0,1 0,7 0,3 0,9 0,8 1,0 0,4

$$\sum x = 6{,}2 \qquad \sum x^2 = 5{,}22$$

Non-filter: 1,5 0,9 1,6 0,5 1,4 1,9 1,0 1,2 1,3 1,6

$$\sum y = 15{,}0 \qquad \sum y^2 = 22{,}54$$

By making suitable assumptions which should be stated, find a 90% confidence interval for the underlying variances for the amounts of tar in the two types of cigarettes. Comment briefly on your result as regards the hypothesis of equal variances. **[5 marks]**

6. Claim amounts on a certain type of policy can be modelled by a gamma distribution with parameters α and λ, that is with density.

$$f(x) = \frac{\lambda^{\alpha} x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}, 0 < x < \infty$$

A random sample $X_1, X_2, \ldots, X_n$ of claim amounts is taken for investigation.

(i) Determine equations satisfied by the maximum likelihood estimators $\hat{\alpha}$ and $\hat{\lambda}$ and comment briefly on any difficulties that may be encountered solving these equations in practice. **[6 marks]**

(ii) Recent experience has shown that the value of the "shape" parameter α can be taken as $\alpha = 2$, so that the only "scale" parameter λ needs to be estimated.

(a) Use your results in part (i) to obtain maximum likelihood estimator $\hat{\lambda}$

(b)A sample of 25 claim amounts yields a mean of $62,40. Calculate the value of $\hat{\lambda}$ **[3marks]**

Total [9 marks]

7 In the surgical treatment of duodenal ulcers there are three different operations corresponding to the removal of various amounts of the stomach. The three operations are denoted A, B and C, with A being the least traumatic and C the most traumatic.

It is known that these operations have an undesirable side-effect for some patients. In cases where the side effect is present, it can be classified as being of "slight degree" or of "moderate degree".
The data in the following table relate to a group of 417 patients and specify the operation received and, where relevant, the degree of the side effects suffered.

	Existence/degree of side effects			
Operation	**None**	**Slight**	**Moderate**	**Total**
A	63	26	7	96
B	126	63	25	214
C	51	40	16	107
Total	**240**	**129**	**48**	**417**

(i) Perform a X^2 test on this table to investigate independence between level of operation and existence/degree of side-effects. **[6 marks]**

(ii) Analyse appropriate tables to test for evidence of:
(a) Whether presence/absence of the side-effect depends on the level of operation.
(b) Whether the degree of the side-effect when present depends on the level of operation. **[10 marks]**

(iii) Comment briefly on these three tests and any connections between them.
[2 marks]
[Total 18 marks]

8. For a particular insurance company, a sample of ten claims and payments on household contents policies is taken. The table gives the claim ($'00) and payment ($'00) for each of the claims considered.

Claim X	2,80	3,40	3,60	5,00	5,80	6,20	6,80	7,00	7,60	8,60
Payment Y	3,06	2,82	3,78	3,92	6,04	5,20	6,74	6,12	7,46	7,66

(a) Draw a scatter plot and comment briefly. **[3 marks]**
(b) Derive the equation of the line of best fit (y on x) using the method of least squares. **[5 marks]**

(c) Calculate a 95% C.I for the slope of the true regression line. **[4 marks]**

(d) Calculate (i) a 95 C.I for the average payment of claims of \$780, and (ii) a 95% prediction interval for the actual payment of a claim of \$780.

[3 and 4 marks]

(e) Calculate the correlation coefficient for the data and comment briefly.

[2 marks]

[Total 21 marks]

9. The probability θ of a claim arising under a certain type of policy is thought to be 0,1. To test $H_0 : \theta = 0{,}1$ v $H_1 : \theta = 0{,}4$, the following test procedures have been suggested.

Test 1: Observe X, the number of claims which arising in a group of ten independent policies. Accept H_0 if $x \geq 3$, accept H_1 if $x \geq 4$.

Test 2: Observe Y, the number of claims which arise in a group of one hundred independent policies. Accept H_0 if $Y \geq 16$, accept H_1 if $Y \geq 17$.

For each test find the probabilities of errors of the two types, and comment briefly on the critical regions used. **[9 marks]**

10. In a one-way analysis of variance the "between treatments" sum of squares is:

$$SS_B = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (\bar{y}_{i\bullet\bullet} \bar{y}_{\bullet\bullet})^2 = \sum_{i=1}^{K} \frac{y_{i\bullet}^2}{n_i} - \frac{y_{\bullet\bullet}^2}{n}$$

Define all symbols in these two expressions and show that they are equivalent.

[9 marks]

END OF EXAMINATION