EXAMINATION

14 April 2005 (am)

Subject CT6 — Statistical Methods
Core Technical

Time allowed: Three hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 10 questions, beginning your answer to each question on a separate sheet.*

5. *Candidates should show calculations where this is appropriate.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

1 List the three main perils typically covered by employer's liability insurance. [3]

2 An insurer wishes to estimate the expected number of claims, λ, on a particular type of policy. Prior beliefs about λ are represented by a Gamma distribution with density function

$$f(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} \quad (\lambda > 0).$$

For an estimate, d, of λ the loss function is defined as

$$L(\lambda, d) = (\lambda - d)^2 + d^2.$$

Show that the expected loss is given by

$$E(L(\lambda, d)) = \frac{\alpha(\alpha+1)}{\beta^2} - \frac{2d\alpha}{\beta} + 2d^2$$

and hence determine the optimal estimate for λ under the Bayes rule. [5]

3 (i) Explain the disadvantages of using truly random, as opposed to pseudo-random, numbers. [3]

(ii) List four methods for the generation of random variates. [2]

[Total 5]

4 Y_t, $t = 1, 2, 3, \ldots$ is a time series defined by

$$Y_t - 0.8Y_{t-1} = Z_t + 0.2Z_{t-1}$$

where Z_t, $t = 0, 1, \ldots$ is a sequence of independent zero-mean variables with common variance σ^2.

Derive the autocorrelation ρ_k, $k = 0, 1, 2, \ldots$ [6]

5 An insurer believes that claim amounts, X, on its portfolio of pet insurance policies follow an exponential distribution with mean £200.

A reinsurance policy is arranged such that the reinsurer pays X_R, where

$$X_R = \begin{cases} 0 & \text{if } X \le £50 \\ X - 50 & \text{if } £50 < X \le M \\ M - 50 & \text{if } X > M \end{cases}$$

calculate M such that $E[X_R] = £100$. [8]

6 On 1 January 2001 an insurer in a far off land sells 100 policies, each with a five year term, to householders wishing to insure against damage caused by fireworks. The insurer charges annual premiums of £600 payable continuously over the life of the policy.

The insurer knows that the only likely date a claim will be made is on the day of St Ignitius' feast on 1 August each year, when it is traditional to have an enormous fireworks display. The annual probability of a claim on each policy is 40%. Claim amounts follow a Pareto distribution with parameters $\alpha = 10$ and $\lambda = 9{,}000$.

(i) Calculate the mean and standard deviation of the annual aggregate claims. [4]

(ii) Denote by $\psi(U, t)$ the probability of ruin before time t given initial surplus U.

(a) Explain why for this portfolio $\psi(U, t_1) = \psi(U, t_2)$ if $7/12 < t_1, t_2 < 19/12$. [1]

(b) Estimate $\psi(15{,}000, 1)$ assuming annual claims are approximately Normally distributed. [4]

[Total 9]

PLEASE TURN OVER

7 The no claims discount (NCD) system operated by an insurance company has three levels of discount: 0%, 25% and 50%.

If a policyholder makes a claim they remain at or move down to the 0% discount level for two years. Otherwise they move up a discount level in the following year or remain at the maximum 50% level.

The probability of an accident depends on the discount level:

Discount Level	*Probability of accident*
0%	0.25
25%	0.2
50%	0.1

The full premium payable at the 0% discount level is 750.

Losses are assumed to follow a lognormal distribution with mean 1,451 and standard deviation 604.4.

Policyholders will only claim if the loss is greater than the total additional premiums that would have to be paid over the next three years, assuming that no further accidents occur.

(i) Calculate the smallest loss for which a claim will be made for each of the four states in the NCD system. [2]

(ii) Determine the transition matrix for this NCD system. [6]

(iii) Calculate the proportion of policyholders at each discount level when the system reaches a stable state. [3]

(iv) Determine the average premium paid once the system reaches a stable state.[1]

(v) Describe the limitations of simple NCD systems such as this one. [2]
[Total 14]

8 (i) Write down the general form of a statistical model for a claims run-off triangle, defining all terms used. [5]

(ii) The table below shows the cumulative incurred claims on a portfolio of insurance policies.

Accident Year	*Delay Year*		
2000	2,748	3,819	3,991
2001	2,581	4,014	
2002	3,217		

The company decides to apply the Bornhuetter-Ferguson method to calculate the reserves, with the assumption that the Ultimate Loss Ratio is 85%.

Calculate the reserve for 2002, if the earned premium is 5,012 and the paid claims are 1,472. [9]

[Total 14]

9 $Y_1, Y_2, \ldots, Y_n$ are independent claims, which are assumed to be exponentially distributed, with

$$E[Y_i] = \mu_i .$$

(i) Show that the canonical link function is the inverse link function. [3]

(ii) It is decided that the canonical link function should not be used, but that the mean claim sizes should be modelled as follows:

$$\log\mu_i = \begin{cases} \alpha & i = 1, 2, \ldots, m \\ \beta & i = m+1, m+2, \ldots, n \end{cases}$$

(a) Show that the log-likelihood can be written as

$$-\left[m\alpha + (n-m)\beta + e^{-\alpha}\sum_{i=1}^{m} y_i + e^{-\beta}\sum_{i=m+1}^{n} y_i \right]$$

(b) Derive the maximum likelihood estimators of α and β.

(c) Show that the scaled deviance for this model is

$$2\left(\sum_{i=1}^{m} \log\left(\frac{\frac{1}{m}\sum_{j=1}^{m} y_j}{y_i} \right) + \sum_{i=m+1}^{n} \log\left(\frac{\frac{1}{n-m}\sum_{j=m+1}^{n} y_j}{y_i} \right) \right)$$ [12]

(iii) For a particular data set, $m = 20$, $n = 44$,

$$\frac{1}{20}\sum_{i=1}^{20} y_i = 14.2, \ \frac{1}{24}\sum_{i=21}^{44} y_i = 18.7.$$

Calculate the deviance residual for $y_1 = 7$. [3]

[Total 18]

10 (i) Explain what a conjugate prior distribution is. [2]

(ii) The random variables $X_1, X_2, \ldots, X_n$ are independent and have density function

$$f(x) = \lambda e^{-\lambda x} \quad (x > 0).$$

Show that the conjugate prior distribution for λ is a Gamma distribution. [3]

(iii) (a) The density function of λ is

$$f(\lambda) = \frac{s^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-s\lambda} \ (\lambda > 0).$$

Show that $E(1/\lambda) = s/(\alpha - 1)$.

(b) Hence if $X_1, X_2, \ldots, X_n$ is an independent random sample from an exponential distribution with parameter λ, show that the posterior mean of $1/\lambda$ can be expressed as a weighted average of the prior mean of $1/\lambda$ and the sample average.

[5]

(iv) An insurer is considering introducing a new policy to provide insurance against the failure of toasters within the first five years of purchase. Alan and Beatrice are underwriters working for the insurer. Based on his experience of similar products, Alan believes that toasters last three years on average. Beatrice believes that six years is the average lifetime. Both are adamant and are prepared to express their uncertainties about the average lifetime in terms of standard deviations of six months and one year respectively. They decide to resolve their differences by testing a sample of toasters large enough to ensure the difference in their posterior expectations for the average lifetime will be less than one year.

Calculate how many toasters they should test, assuming the exponential distribution is a good model for toaster lifetimes.

You may use the fact that if $\lambda \sim \Gamma(\alpha, s)$ then

$$\text{Var}(1/\lambda) = [E(1/\lambda)]^2 \times \frac{1}{\alpha - 2}.$$ [8]

[Total 18]

END OF PAPER