Institute of Actuaries **Faculty of Actuaries**

EXAMINATIONS

September 1997

Subject C — Statistics

Paper Two

EXAMINERS' REPORT

PART ONE

1 C

2 B

3 C

4 D

5 C

6 B

7 The posterior distribution for θ has range $x \to \infty$ and density proportional to:

$$f(x \mid \theta) \,.\, g(\theta) = q^{-1} \,.\, qe^{-q} = e^{-q}$$

Hence, posterior density is ke^{-q} where

$$\int_x^{\infty} ke^{-q} dq = 1 \to [-ke^{-q}]_x^{\infty} = 1 \to k = e^{+x}$$

Bayes estimator w.r.t. absolute error loss function is the median of the posterior. Hence, estimator is y where

$$\int_x^{y} e^{-q} e^{x} dq = \tfrac{1}{2} \to e^{x}[-e^{-q}]_x^{y} = \tfrac{1}{2}$$

$$\to \quad e^{x}[e^{-x} - e^{-y}] = \tfrac{1}{2}$$

$$\to \quad \tfrac{1}{2}e^{-x} = e^{-y}$$

$$\to \quad y = \boldsymbol{x + \log 2}$$

8 (i) It is an ARMA(1,1) process.

(ii) $E[Y_t \, Z_t] \quad = [Z_t^2] = s^2$

$$E[Y_t\, Z_{t-1}] \;=\tfrac{1}{3}E[Y_{t-1}\, Z_{t-1}]+E[Z_t\, Z_{t-1}]+\tfrac{2}{3}E[Z_{t-1}^2]$$

$$=\tfrac{1}{3}s^{\,2}+0+\tfrac{2}{3}s^{\,2}\;=\;s^{\,2}$$

$$E[Y_t\, Z_{t-k}] \;=0 \text{ for } k\ge z$$

$$g_{\,0}\;=\;E[Y_t\,Y_t]\;=\;E[Y_t\{\tfrac{1}{3}Y_{t-1}+Z_t+\tfrac{2}{3}Z_{t-1}\}]$$

$$=\tfrac{1}{3}g_{\,1}+s^{\,2}+\tfrac{2}{3}s^{\,2}\rightarrow g_{\,0}-\tfrac{1}{3}g_{\,1}\;=\;\tfrac{5}{3}s^{\,2}$$

$$g_{\,1}\;=\;E[Y_t\,Y_{t-1}]\;=\;E[\{\tfrac{1}{3}Y_{t-1}+Z_t+\tfrac{2}{3}Z_{t-1}\}\,Y_{t-1}]$$

$$=\tfrac{1}{3}g_{\,0}+0+\tfrac{2}{3}s^{\,2}\rightarrow g_{\,0}-3g_{\,1}\;=-2s^{\,2}$$

$$\rightarrow\qquad \tfrac{8}{3}g_{\,1}\;=\;\tfrac{11}{3}s^{\,2}$$

$$\rightarrow\qquad \underline{\underline{g_{\,1}=\tfrac{11}{8}s^{\,2}}}$$

$$\rightarrow\qquad \underline{\underline{g_{\,0}=\tfrac{17}{8}s^{\,2}}}$$

$$k\ge 2\quad g_{\,k}\;=\;E[Y_t\,Y_{t-k}]\;=\;E[\{\tfrac{1}{3}Y_{t-1}+Z_t+\tfrac{2}{3}Z_{t-1}\}\,Y_{t-k}]$$

$$=\tfrac{1}{3}g_{\,k-1}\rightarrow g_{\,k}\;=\;\tfrac{11}{8}s^{\,2}\left(\tfrac{1}{3}\right)^{k-1}\quad k\ge 2$$

$$\rightarrow\qquad r_{\,0}=1;\;\; r_{\,1}\;=\;\frac{g_{\,1}}{g_{\,0}}\;=\;\tfrac{11}{17};\;\; r_{\,k}\;=\tfrac{11}{17}\left(\tfrac{1}{3}\right)^{k-1}$$

(iii) New process is ARIMA(1,1,1).

$$W_t-\tfrac{4}{3}W_{t-1}+\tfrac{1}{3}W_{t-2}\;=\;Z_t+\tfrac{2}{3}Z_{t-1}$$

9 The range of the posterior distribution of N is r, $r+1$, $r+2$, ...

The probability of r males, given n children is

$${}^{n}C_r\;p^r(1-p)^{n-r}$$

Hence, the posterior distribution of N is proportional to

$${}^{n}C_r\;p^r(1-p)^{n-r}\,.\,\frac{2^n e^{-2}}{n!}\qquad n=r,\,r+1,\,\ldots$$

$$\propto \quad \frac{n!}{(n-r)!}\frac{(2q)^n}{n!} \qquad q = 1-p$$

$$\propto \quad (2q)^n / (n-r)!$$

Hence, the posterior distribution is

$$P[N=n\,|\,r] = K(2q)^n/(n-r)! \qquad n = r, r+1, \ldots$$

where $K(2q)^r\left[1+2q+\frac{(2q)^2}{2!}+\frac{(2q)^3}{3!}+\ldots\right]=1$

$$\rightarrow \quad K = e^{-2q}/(2q)^r$$

$$\rightarrow \quad P[N-r=x\,|\,r] = P[N=x+r\,|\,r]$$

$$= \frac{e^{-2q}}{(2q)^r}\frac{(2q)^{x+r}}{x!} = (2q)^x\, e^{-2q}/x! \qquad x = 0, 1, 2, \ldots$$

which is Poisson with parameter $2q (= 2(1-p))$

$E[N\,|\,r] = E[N-r\,|\,r] + r = 2q + r = 2(1-p) + r$

10 (i) Let number of deaths in each group be $D_i\ (i = 1, 2, 3)$

Then $N = D_1 + 2D_2 + 3D_3$

Hence $M_N(t) = E(e^{D_1 t})\, E(e^{2D_2 t})\, E(e^{3D_3 t})$

$= M_{D_1}(t)\, M_{D_2}(2t)\, M_{D_3}(3t)$

$= (0.01e^t + 0.99)^{10}(0.01e^{2t} + 0.99)^{15}\,(0.01e^{3t} + 0.99)^5$

(ii) $E(N) = E(D_1) + 2E(D_2) + 3E(D_3)$

$= 10 \times 0.01 + 2 \times 15 \times 0.01 + 3 \times 5 \times 0.01 = 0.55$

$\text{Var}(N) = \text{Var}(D_1) + 4\text{Var}(D_2) + 9\text{Var}(D_3)$

$= 10 \times 0.01 \times 0.99 + 4 \times 15 \times 0.01 \times 0.99 + 9 \times 5 \times 0.01 \times 0.99^6$

$= 1.1385$

11 $(1 - 0.3B)\,(1 - 0.4B)\,(1 - 0.5B)$

$= 1 - 1.2B + 0.47B^2 - 0.06B^3$

$\therefore \quad Y_t = 1.2Y_{t-1} - 0.47Y_{t-2} + 0.06Y_{t-3} + Z_t$

multiply by Y_{t-1} and take expectations:

$g_1 = 1.2g_0 - 0.47g_1 + 0.06g_2$

$\Rightarrow \quad r_1 = 1.2 - 0.47r_1 + 0.06r_2$

$\therefore \quad 1.47r_1 - 0.06r_2 = 1.2 \qquad (1)$

multiply by Y_{t-2} and take expectations:

$g_2 = 1.2g_1 - 0.47g_0 + 0.06g_1$

$\Rightarrow \quad r_2 = 1.2r_1 - 0.47 + 0.06r_1$

\ $\quad 1.26r_1 - r_2 = 0.47 \qquad (2)$

Solve (1), (2) for r_1, r_2:

$1.3944r_1 = 1.1718$

$\therefore \quad r_1 = 0.8404$

$r_2 = 1.26r_1 - 0.47 = 0.5889$

12 (i) $y_1(U_0, 1) = P[U_1 < 0] = P[(1 + 1.2) \times 1.05 - S_1 < 0]$

$= P[S_1 > 2.31]$

$= P[Z > (2.31 - 1)/1] = \underline{0.0951}$

(ii) $P[U_2 < 0] = P[(U_1 + 1.2) \times 1.05 - S_2 < 0]$

$= P[[(1 + 1.2) \times 1.05 - S_1 + 1.2] \times 1.05 - S_2 < 0]$

$= P[3.6855 - (1.05S_1 + S_2) < 0]$

$1.05S_1 + S_2 \sim N(2.05, 2.1025 = 1.45^2)$

$$= P\left[Z > \frac{3.6855 - 2.05}{1.45}\right] = \underline{\underline{0.12924}}$$

(iii) $y_1(U_0, 2)$ is the probability of all sample paths for which **either** $U_1 < 0$ or $U_2 < 0$ or both.

$P[U_2 < 0]$ is the probability of a subset of these paths, i.e. only those for which $U_2 < 0$.

Hence $P[U_2 < 0] \le y_1(U_0, 2)$.

Since there will be positive probability that $U_1 < 0$ and $U_2 > 0$, we can write:

$$P[U_2 < 0] < y_1(U_0, 2)$$

(iv) $y_1(U_0, 2) - y_1(U_0, 1)$ is the probability of all sample paths for which $U_2 < 0$ but $U_1 > 0$. From time 1 to time 2 the process is probabilistically identical to the process from time 0 to time 1.

However, given that $U_1 > 0$, the expected surplus at time 1 is greater than 1. Hence $y_1(U_0, 2) - y_1(U_0, 1) < y_1(U_0, 1)$.

(v) Consider $y_1(U_0, n+1)$ and condition on what happens at time 1.

Either ruin occurs: probability $y_1(U_0, 1)$

or ruin does not occur

$\rightarrow$ $S_1 = x$ $(x \in (-\infty, 2.31))$
probability $f(x)dx$

and ruin occurs subsequently from an "initial" surplus of $2.31 - x$ in n years: probability $y_1(2.31 - x, n)$

Hence, ruin occurs $y_1(U_0, 1)$

or does not occur at time 1 but does occur in next n years

$$\int_{x-\infty}^{(U_0+1.2)\times 1.05} f(x) y_1((U_0 + 1.2) \times 1.05 - x, n) dx$$

$\uparrow$
Summing over all possible values of x.

$$\rightarrow \quad y_1(U_0, n+1) = y_1(U_0, 1) + \int_{-\infty}^{(U_0+1.2)\times 1.05} f(x) y_1((U_0 + 1.2) \times 1.05 - x, n) dx$$

13 (i) Discount level:

	Premiums *No Claim*	*Claim*
0%	320	400
	240	320

Difference 160. $P[\text{Claim}] = e^{-\frac{160}{500}} = \underline{0.726}$

20%	240	320
	160	240

Difference 160. $P[\text{Claim}] = \underline{0.726}$

40%	160	240
	160	160

Difference 80. $P[\text{Claim}] = e^{-\frac{80}{500}} = \underline{0.852}$

60%	160	160
	160	160

Difference 0. $P[\text{Claim}] = \underline{1}$

(ii)

0%	400	400	
	320	320	Prob = 1
20%	320	400	
	240	320	Prob = 0.726
40%	240	400	
	160	320	Prob = 0.527
60%	160	400	
	160	320	Prob = 0.449

(iii) $P[0\% \to 20\%] = P[\text{No claims}]$

$= 0.7 + 0.2 \times 0.274 + 0.1 \times 0.274 \times 0.274 = \underline{0.7623}$

$P[0\% \to 0\%] = 1 - P[0\% \to 20\%] = \underline{0.2377}$

$P[20\% \to 40\%] = P[\text{No claims}] = 0.7 + 0.2 \times 0.274 + 0.1 \times 0.274^2$

$= \underline{0.7623}$

$P[20\% \to 0\%] = P[2 \text{ claims}] = 0.1 \times 0.726 \times 0.726$

$= \underline{0.0527}$

$P[20\% \to 20\%] = 1 - 0.7623 - 0.0527 = \underline{0.1850}$

$P[40\% \to 0\%] = P[2 \text{ claims}] = 0.1 \times 0.852 \times 0.527$

$= \underline{0.0449}$

$P[40\% \to 60\%] = 0.7 + 0.2 \times 0.148 + 0.1 \times 0.148^2$

$= \underline{0.7318}$

$P[40\% \to 40\%] = 1 - .0449 - .7318 = \underline{0.2233}$

$P[60\% \to 0\%] = 0.1 \times 1 \times 0.449 = \underline{0.0449}$

$P[60\% \to 60\%] = 1 - 0.0449 = 0.9551$

Transition matrix

$$\begin{bmatrix} 0.2377 & 0.7623 & 0 & 0 \\ 0.0527 & 0.1850 & 0.7623 & 0 \\ 0.0449 & 0 & 0.2233 & 0.7318 \\ 0.0449 & 0 & 0 & 0.9551 \end{bmatrix}$$

14 (i) $\int_d^{L+d} xf(x)dx = \int_{\log d}^{\log(L+d)} e^y f(y)dy$

$$= \int_{\log d}^{\log(L+d)} \frac{1}{\sqrt{2ps^2}} e^{-\frac{y^2-2ym+m^2-2ys^2}{2s^2}} dy$$

$$= e^{m+½s^2} \int_{\log d}^{\log(L+d)} \frac{1}{\sqrt{2ps^2}} e^{-\frac{(y-m-s^2)^2}{2s^2}} dy$$

$$= e^{m+½s^2} \left[\Phi\left(\frac{\log(L+d) - m - s^2}{s} \right) - \Phi\left(\frac{\log d - m - s^2}{s} \right) \right]$$

(ii) (a) $X_R = \begin{cases} 0 & x \le d \\ x - d & d < x \le L + d \\ L & x > L + d \end{cases}$

$$E[X_R] = \int_d^{L+d} (x-d) f(x)dx + LP(X > L + d)$$

$$= \int_d^{L+d} xf(x)dx - dP(d < X \le L+d) + LP(X > L+d)$$

$$= e^{m+\frac{1}{2}s^2}\left[\Phi\left(\frac{\log(L+d)-m-s^2}{s}\right) - \Phi\left(\frac{\log d - m - s^2}{s}\right)\right]$$

$$-d\left[\Phi\left(\frac{\log(L+d)-m}{s}\right) - \Phi\left(\frac{\log d - m}{s}\right)\right] + L\left[1-\Phi\left(\frac{\log(L+d)-m}{s}\right)\right]$$

$$e^{m+\frac{1}{2}s^2} = 2000 \text{ and } (e^{m+\frac{1}{2}s^2})^2\,(e^{s^2}-1) = 2500^2$$

$$e^{s^2} - 1 = 1.5625 \qquad \therefore \qquad s^2 = 0.94098$$

$$\therefore \quad \mu = 7.13041$$

$$E[X_R] = 2000[\Phi(1.592) - \Phi(0.460)] - 5000[\Phi(2.562) - \Phi(1.430)]$$

$$+10000[1 - \Phi(2.562)]$$

$$= 2000[0.9443 - 0.6772] - 5000[0.9948 - 0.9236]$$

$$+ 10000[1 - 0.9948]$$

$$= 230.2$$

$$E[X_R \mid X > 5000] = \frac{E[X_R]}{P(X > 5000)} = \frac{230.2}{1 - \Phi(1.430)} = 3013$$

(b) The expected value is

$$\frac{1.1\int_{\frac{5000}{1.1}}^{\frac{15000}{1.1}}\left(x - \frac{5000}{1.1}\right)f(x)dx + 10000P\left[X > \frac{15000}{1.1}\right]}{P\left[X > \frac{5000}{1.1}\right]}$$

$$P\left[X > \frac{5000}{1.1}\right] = P\left[\frac{\log X - 7.13041}{0.97004} > \frac{\log\left(\frac{5000}{1.1}\right) - 7.13041}{0.97004}\right]$$

$$= 1.33$$

$$= 1 - 0.90824 = \underline{0.09176}$$

$$P\left[X > \frac{15000}{1.1}\right] = 1 - \Phi(2.46) = 1 - 0.99305$$

$$= \underline{\underline{0.00695}}$$

$$\int_{\frac{5000}{1.1}}^{\frac{15000}{1.1}} xf(x)dx = 2000[\Phi(1.49 - \Phi(0.36)]$$

$$0.93189 - 0.64058$$

$$= \underline{\underline{582.62}}$$

$\rightarrow$ Expected value is

$$\frac{1.1 \times 582.6 - 5000[0.99305 - 0.90824] + 10000 \times 0.00695}{0.09176}$$

$$= \underline{\underline{£3120}}$$