Faculty of Actuaries **Institute of Actuaries**

EXAMINATIONS

April 2004

Subject 103 — Stochastic Modelling

EXAMINERS' REPORT

Introduction

The attached subject report has been written by the Principal Examiner with the aim of helping candidates. The questions and comments are based around Core Reading as the interpretation of the syllabus to which the examiners are working. They have however given credit for any alternative approach or interpretation which they consider to be reasonable.

J Curtis
Chairman of the Board of Examiners

22 June 2004

1 First method (direct method):

Let $f(X_t) = \ln X_t$. Using Itô's Lemma we have

$$\begin{aligned} df &= f'\left(\rho X_t dt + \sigma X_t dB_t\right) + \tfrac{1}{2} f''(\sigma^2 X_t^2) dt \\ &= \left(\rho X_t \frac{1}{X_t} - \tfrac{1}{2}.\frac{1}{X_t^2}\sigma^2 X_t^2\right) dt + \frac{1}{X_t}\sigma X_t dB_t \\ &= \left(\rho - \tfrac{1}{2}\sigma^2\right) dt + \sigma dB_t \end{aligned}$$

Integrating from t_0 to t we have

$$\ln X_t - \ln X_{t_0} = \left(\rho - \tfrac{1}{2}\sigma^2\right)\left(t - t_0\right) + \sigma\left(B_t - B_{t_0}\right)$$

and using the initial condition we get

$$X_t = x \exp\left[\sigma\left(B_t - B_{t_0}\right) + \left(\rho - \tfrac{1}{2}\sigma^2\right)\left(t - t_0\right)\right] \text{ as required.}$$

Alternative method (working backwards from the solution)

Let $f(t, B_t) = x \exp\left[\sigma\left(B_t - B_{t_0}\right) + \left(\rho - \tfrac{1}{2}\sigma^2\right)\left(t - t_0\right)\right]$. Then

$$\frac{\partial f}{\partial t} = \left(\rho - \frac{1}{2}\sigma^2\right) f, \quad \frac{\partial f}{\partial B} = \sigma f, \quad \frac{\partial^2 f}{\partial B^2} = \sigma^2 f \text{ , so that}$$

$$df_t = \left(\rho - \frac{1}{2}\sigma^2\right) f_t dt + \sigma f_t dB_t + \frac{1}{2}\sigma^2 f_t dt = \rho f_t dt + \sigma f_t dB_t$$

This implies that $X_t = f(t,B_t)$ satisfies the stochastic differential equation.

We need to verify that the initial condition holds for this solution. (It clearly does, but the check needs to be performed.)

Considering that this stochastic differential equation is the most frequently used of all, this question was on average very poorly answered.

2 (i) The solution to $1 - 0.63z = 0$ is $z = 1/0.63$, which is greater than 1. Therefore the model is stationary.

For invertibility, we should check that the roots of $1 + 0.45z - 0.34z^2 = 0$ are outside the unit circle. They are (–0.45 ± 1.25)/(–2*0.34) = 2.5 or –1.18, both OK.

(ii) Let $\psi_0 = \text{Cov}(X_t, e_t)$, $\psi_1 = \text{Cov}(X_t, e_{t-1})$, $\psi_2 = \text{Cov}(X_t, e_{t-2})$. Then

$\gamma_0 = 0.63\gamma_1 + \psi_0 + 0.45\psi_1 - 0.34\psi_2$
$\gamma_1 = 0.63\gamma_0 + 0.45\psi_0 - 0.34\psi_1$
$\gamma_2 = 0.63\gamma_1 - 0.34\psi_0$
$\psi_0 = \sigma^2$ (this may be regarded as obvious and not stated explicitly)
$\psi_1 = 0.63\psi_0 + 0.45\sigma^2 = 1.08\sigma^2$
$\psi_2 = 0.63\psi_1 - 0.34\sigma^2 = 0.3404\sigma^2$

An alternative expression for γ_0 may be obtained using

$$\begin{aligned}\gamma_0 &= \text{Var}(0.63X_{t-1} + e_t + 0.45e_{t-1} - 0.34e_{t-2}) \\ &= 0.63^2\gamma_0 + (1 + 0.45^2 + 0.34^2)\sigma^2 + 2\times 0.63[0.45\psi_0 - 0.34\psi_1]\end{aligned}$$

which removes the need to calculate ψ_2.

Having derived the equations, we need to solve them.

$\gamma_1 = 0.63\gamma_0 + 0.0828\sigma^2$
$\gamma_0 = 0.63(0.63\gamma_0 + 0.0828\sigma^2) + 1.370\sigma^2 = 0.3969\gamma_0 + 1.422\sigma^2$,

implying that

$\gamma_0 = 2.358\sigma^2$, $\gamma_1 = 1.5684\sigma^2$, $\gamma_2 = 0.6481\sigma^2$.

Candidates were often unsure of the procedure required to derive the equations for the γ_k but did rather better at solving them. In particular, many candidates did not take correct account of the relationship between X_t and past values of e_t. Marks were awarded for correct methodology when deriving the solutions, even if the equations being solved were not the right ones.

3 (i) It is a jump process because it remains in one state for a period of time and then jumps to another state, or alternatively because it is a continuous-time process with a discrete state space.
Given the entire past history of the process, the probability of a member retiring and beginning to receive benefits in the next dt interval is λdt, i.e. independent of the past. The same applies to the probability of death between times t and $t + dt$: it is dependent only on the state (the number of retired members) at time t, and not on anything which happened before that. Thus the Markov property holds.

(ii) $\sigma_{n,n+1} = \lambda$, $\sigma_{n,n-1} = n\mu$

(iii) From the Kolomogorov Forward equations we have that

$$\frac{dp_n}{dt} = -(\sigma_{n,n-1} + \sigma_{n,n+1})\,p_n + \sigma_{n+1,n}\,p_{n+1} + \sigma_{n-1,n}\,p_{n-1},.$$
$$= -(\lambda + n\mu)\,p_n + (n+1)\mu p_{n+1} + \lambda p_{n-1}$$

(iv) The suggested form of p_n does not depend on t, so its derivative is zero. The RHS is

$$-(\lambda + n\mu)e^{-\rho}\frac{\rho^n}{n!} + (n+1)\mu e^{-\rho}\frac{\rho^{n+1}}{(n+1)!} + \lambda e^{-\rho}\frac{\rho^{n-1}}{(n-1)!} = 0,$$

where $\rho = \lambda/\mu$.

(v) The implication is that Poisson(ρ) is a stationary distribution for the Markov jump process. (In this case it is also a limiting distribution, but that is not deducible from the Core Reading.)

Very few candidates suggested that the rate at which deaths take place should be proportional to the number of retired scheme members.

In (iv) examiners awarded marks for correct attempts to carry out the verification procedure even when the differential equation in (iii) was wrong.

In (v), terms such as "limiting distribution" or "equilibrium distribution" were given full credit.

4 (i) Model fitting involves first choosing a suitable family of models, then a suitable state space and finally estimating the values of the model parameters to isolate a single member of the family.

Having performed the model fitting process, we have the best-fitting model from the given family. Model validation involves checking whether this is an adequate explanation of the observations.

Simulation can be used in model validation to test that the paths typically followed by a simulated version of the process are broadly similar to the paths observed in practice.

(ii) (a) Under the given model, the successive increments ∇S_t form a sequence of i.i.d. $N(\mu, \sigma^2)$ random variables, so the parameters can be estimated using the sample mean and sample variance of the observed increments.

(b) The first thing to do is to take the log of the observations, obtaining x_1, …, x_n say. Then apply the same technique as in (a).

(c) This is a time series model, so a time series technique (such as Box-Jenkins) would be appropriate.

Parameters are estimated by Method of Moments or Maximum Likelihood, or simply by applying a computer package to do the estimation for you.

(iii) Any of the tests described here are equally valid.

(a) The model implies that the increments are Normally distributed, so a standard test of normality, such as the Anderson-Darling test, the Kolmogorov-Smirnov test or a chi-squared goodness-of-fit test can be employed to test this.

Another testable property of the model is independence of increments: it is possible to investigate the lag-1 sample ACF and determine whether it is significantly different from zero, or use the Ljung-Box chi-square statistics to perform a portmanteau test on all lags of the sample ACF simultaneously.

A similar test is to inspect the sample ACF and PACF of the residuals to see whether the observed values fall outside the 5% significance band.

We could test whether there was a relationship between the increment ∇s_t and s_t; this is a test for homoscedasticity, though in the situation described in the question it is unlikely to provide useful results.

A turning points test, a run test or a sign change test could also be used, in each case comparing a calculated statistic with tables of a reference distribution. These are tests for serial independence (usually applied to a sequence of residuals).

(b) Apply the same tests as in (a) to the logarithm of the data.

(c) Tests can be applied to the sequence of residuals arising from the fitting process. Since, if the model is accurate, these should form a sequence of uncorrelated Normal observations, the same tests as in (a) can be applied.

A procedure which might be applied to the model in part (c) but not to the models in (a) or (b) is to inspect the sample partial ACF of the increment process ∇S_t; if the process really is a first-order autoregression as stated in the model, then the sample PACF should have no significant values at lags higher than 1.

This question differentiated well between candidates, with some very good answers indicating that the candidates understood the principles and practice of modelling, and others highlighting deficiencies of understanding.

5 (i) (a) If $X = B(½)$ and $Y = B(1)$, then the marginal distribution of X is N(0,0.5) and the conditional distribution of Y given $X = x$ is N(x, 0.5).

This means that $f_{X,Y}(x,y)$ is given by

$$\frac{1}{\sqrt{2\pi(0.5)}}\exp\left(-\frac{x^2}{2(0.5)}\right).\frac{1}{\sqrt{2\pi(0.5)}}\exp\left(-\frac{(y-x)^2}{2(0.5)}\right)$$

$$=\frac{1}{\pi}\exp(-(y^2-2yx+2x^2))$$

(b) The conditional density function of X given Y is

$$f_{X|Y}(x\,|\,b)=\frac{f_{X,Y}(x,b)}{f_Y(b)}=\frac{\pi^{-1}\exp(-(b^2-2bx+2x^2))}{\frac{1}{\sqrt{2\pi}}\exp\left(-½b^2\right)}$$

$$=\sqrt{\frac{2}{\pi}}\exp\left(-2(x-½b)^2\right)$$

This is the density function of the Normal distribution with mean $½b$ and variance ¼.

(ii) (a) Here we seek $P(Y(1) > 0 \mid Y(½) = 5) = P(B(1) > 0 \mid B(½) = 1)$

$$= \quad P\left[B(\tfrac{1}{2})>-1\right]$$

$$= \quad P\left[\sqrt{2}\left\{B(\tfrac{1}{2})-0\right\}>-\sqrt{2}\right]$$

$$= \quad 1-\Phi\left(-\sqrt{2}\right) = 0.9213.$$

(b) We require $P\left[Y(\tfrac{1}{2})>0 \mid Y(1)=5\right] = P\left[B(\tfrac{1}{2})>0 \mid B(1)=1\right]$

Using part (i)(b), this is $1-\Phi\left(\frac{0-½}{\sqrt{¼}}\right)=\Phi(1)=0.8413.$

This question was designed in such a way that candidates could tackle part (ii) even if part (i) had not worked out right, but many allowed themselves to become discouraged by difficulties in the early part of the question and gave up.

6 (i) Use a linear congruential generator (LCG). Specify three positive integers a,c,m with $m>a, m>c$ and an initial value x_0, then generate a sequence of $x_1,x_2,...,x_n$ in the range $\{0,1,2,...,m-1\}$ by the recursive rule

$$x_n = (ax_{n-1} + c) \pmod m, \qquad n = 1, 2, \ldots, N,$$

Then $u_n = x_n/m$ is a sequence of pseudo-random numbers in $[0,1]$.

To generate a sequence over the range (a,b), a linear transformation $v_n = (b-a)u_n + a$ is needed. Thus the procedure is: generate a variable $U \sim U(0,1)$ using a LCG as above; return $V = (b-a)U + a$.

(ii) The main advantage of using pseudo-random numbers is reproducibility. When testing a model, this almost certainly depends on a number of parameters and assumptions, so it is very often desirable to examine the sensitivity of the model to these assumptions and the values of the parameters used. It is thus necessary to reduce the random element to a minimum and rerun the experiment using the same "data".

(iii) (a) We use the inverse transform method. The cdf of the Pareto is

$$F(x) = \int_0^x \frac{a}{(1+x)^{a+1}} = 1 - \frac{1}{(1+x)^a},$$

so the inverse is

$$F^{-1}(y) = (1-y)^{-1/a} - 1.$$

Thus the procedure is

- Generate an observation U from the Uniform (0,1) distribution.
- Return $X = (1-U)^{-1/a} - 1$ (or, equally good, $X = U^{-1/a} - 1$) as an observation from the Pareto.

(b) Since the number of possible values for X is $n+1$, we need to split the interval $(0,1)$ into $n+1$ subintervals, the length of which has to be proportional to the corresponding probabilities. Given a uniform variable U, one such procedure is:

- For $i = 1, 2, \ldots, n$, if $(i-1)/(n+2) < U \le i/(n+2)$, then set $X = i$
- If $U > n/(n+2)$, then set $X = 0$.

(iv) From (iii), we see that the tail of the Pareto distribution is $1 - F(x) = (1+x)^{-a}$, which decreases to zero much more slowly than e.g. the exponential or normal distributions.
Alternatively, it is called "fat-tailed" because of the relatively high probability of producing values which are a long way from the mean/median.

In general, the Pareto offers a suitable choice for portfolios where there is a non-negligible chance of very large claims, which is reasonable in various forms of general insurance, in particular insurance that deals with natural catastrophes (floods, earthquakes, hurricanes etc).

This question was generally well answered and caused few difficulties.

7 (i) (a) We will use the matrix form of the Kolmogorov DE, which states that $P'(t) = AP(t)$.

It follows that $\frac{d}{dt}\pi^T P(t) = \pi^T \frac{dP}{dt} = \pi^T AP(t) = 0^T P(t) = 0^T$.

[The other Kolmogorov DE, $P'(t) = P(t)A$, cannot be used to complete the proof]

(b) If X_0 is random with distribution π, then the distribution of X_t is given by $\pi^T P(t)$.

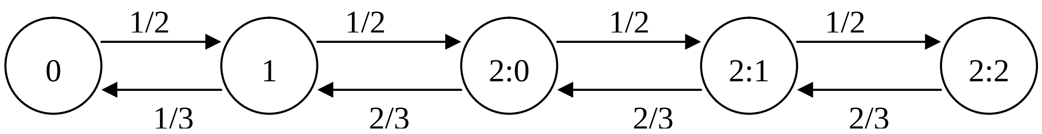
The fact that the differential of this is equal to zero implies that $\pi^T P(t) = \pi^T P(0) = \pi^T$ for all t, in other words X_t has the same distribution for all t.

For an irreducible Markov chain on a finite state space, the stationary distribution is also the limiting (equilibrium, long-term, steady state) distribution.

(ii) (a) The states can be labelled as 0, 1, 2:0, 2:1, 2:2, where 0 and 1 represent the number of operators occupied and 2:*j* means that both operators are occupied and *j* calls are on hold.

Candidates with a different collection of states can earn 0.5 marks, as long as their states *are* states. For example, "Call arrives" is not a state, but an event triggering a transition from one state to another.

(b)



The labels on the arcs represent the transition rates.

(c) From the transition diagram, the generator matrix is as below.

$$A = \begin{pmatrix} -1/2 & 1/2 & 0 & 0 & 0 \\ 1/3 & -5/6 & 1/2 & 0 & 0 \\ 0 & 2/3 & -7/6 & 1/2 & 0 \\ 0 & 0 & 2/3 & -7/6 & 1/2 \\ 0 & 0 & 0 & 2/3 & -2/3 \end{pmatrix}$$

(d) We have

$$\begin{aligned}
\tfrac{1}{2}\pi_0 &= \tfrac{1}{3}\pi_1 & &\Rightarrow & \pi_1 &= \tfrac{3}{2}\pi_0 \\
\tfrac{5}{6}\pi_1 &= \tfrac{1}{2}\pi_0 + \tfrac{2}{3}\pi_{2:0} & &\Rightarrow & \pi_{2:0} &= \tfrac{9}{8}\pi_0 \\
\tfrac{7}{6}\pi_{2:0} &= \tfrac{1}{2}\pi_1 + \tfrac{2}{3}\pi_{2:1} & &\Rightarrow & \pi_{2:1} &= \tfrac{27}{32}\pi_0 \\
\tfrac{7}{6}\pi_{2:1} &= \tfrac{1}{2}\pi_{2:0} + \tfrac{2}{3}\pi_{2:2} & &\Rightarrow & \pi_{2:2} &= \tfrac{81}{128}\pi_0
\end{aligned}$$

To solve this we add the condition that the π_i must sum to 1. Therefore the stationary distribution is

$$\pi = \begin{pmatrix} \tfrac{128}{653} & \tfrac{192}{653} & \tfrac{144}{653} & \tfrac{108}{653} & \tfrac{81}{653} \end{pmatrix}$$

(in decimal form this is
$\pi = \begin{pmatrix} 0.1960 & 0.2940 & 0.2205 & 0.1654 & 0.1240 \end{pmatrix}$)

Many candidates made a good attempt at part (i), though not many scored full marks. In part (ii) a surprising number of candidates were unable to distinguish between states *– in which the process stays for a certain length of time – and* events*, which occur instantaneously and trigger transitions from one state to another.*

8 (i) Suppose at time t X has just arrived in state i. The probability that X remains in state i until time $t + k$ and then leaves, giving a duration of $k + 1$ steps in state i, is $p_{ii}^k(1-p_{ii})$. In other words, $P(D_{i,n} = d) = p_{ii}^{d-1}(1-p_{ii})$, which is the probability function of a geometric random variable.

The fact that $D_{i,n}$ is independent of previous durations follows from the Markov property: What happens to X after arriving in state i is independent of anything that happened before that moment.

(ii) A sequence of successive observations of a geometric random variable is likely to produce 1 more often than any other value. The durations of the restorer's stays in the various locations are always at least 2 days and often much more. There is an indication that the geometric distribution is inappropriate for the data set provided.

An alternative indication that the Markov property is dubious is the fact that the restorer appears to return to Bath at regular intervals, i.e. regardless of the time spent in each state, the path taken appears to lack randomness.

(iii) $\hat{p}_{ij} = n_{ij} / n_{i+}$, where n_{ij} is the number of transitions from state i to state j and n_{i+} is the total number of transitions out of state i. For example, of the 8 days spent in Warwick, one is followed by a trip to Caernarvon, one by a trip to Bath and the remaining 6 are followed by another day in Warwick, so that n_{WC} = 1, n_{WB} = 1 and n_{WW} = 6. We therefore have (using the order B, C, S, W)

$$\hat{P} = \begin{pmatrix} \frac{9}{12} & 0 & \frac{2}{12} & \frac{1}{12} \\ \frac{2}{15} & \frac{13}{15} & 0 & 0 \\ 0 & \frac{1}{10} & \frac{8}{10} & \frac{1}{10} \\ \frac{1}{8} & \frac{1}{8} & 0 & \frac{6}{8} \end{pmatrix}$$

(iv) The model is irreducible. Starting from Bath it is possible to visit Stratford, Warwick, Caernarvon and return to Bath, completing the circuit.

It is also aperiodic, since $p_{ii} > 0$ for some (in fact for all) state i.

(v) We need the entry of $\hat{P}^3$ corresponding to the row and column of Warwick (4th row, 4th column in this case).

Now $\hat{P}^2 = \begin{pmatrix} .5729 & .0271 & .2583 & .1417 \\ .2156 & .7511 & .0222 & .0111 \\ .0258 & .1792 & .6400 & .1550 \\ .2042 & .2021 & .0208 & .5729 \end{pmatrix}$, so the required answer is

$$\frac{1}{8} \times .1417 + \frac{1}{8} \times .0111 + \frac{0}{8} \times .1550 + \frac{6}{8} \times .5729 = 0.4488.$$

Each part of this question attracted some unexpected answers as well as some good ones. The answers to part (i) were on the whole disappointing, but the final parts were rather better answered in general.

9 (i) (a) In one sense it is reasonable, because in the absence of rainfall gardeners might increase their demand for water. But the relationship does not work in reverse: in particularly rainy weather the demand is likely to be no less than in normally rainy weather.

(b) Yes. It reflects the seasonal pattern of rainfall.

(ii) No. It satisfies an equation which is explicitly dependent on t, so it will not be stationary unless all the μ_t are equal to one another. (Arguments based on whether $|\alpha| > 1$ are only relevant if all the μ_t are equal.)

(iii) (a) Normality: by induction. If R_{t-1} and e_t are Normal, then R_t must be Normal, as it is a linear combination of these two. Since R_0 is Normal, normality follows for all t.

$E(R_t) - \mu_t = \alpha[E(R_{t-1}) - \mu_{t-1}]$. This can be iterated backwards to zero, showing that $E(R_t) = \mu_t$.

An alternative approach is to solve explicitly for R_t and read off the answers from there.

$$R_t - \mu_t = \alpha^t (R_0 - \mu_0) + \sum_{k=0}^{t-1} \alpha^k e_{t-k}$$

Now we can see that R_t is a linear combination of Normal random variables with some constants, hence Normally distributed, and that the mean is μ_t.

(b) $\text{Var}(R_t)$ is equal to $\alpha^2 \text{Var}(R_{t-1}) + \sigma^2$. If R_r has constant variance σ_R^2 then $\sigma_R^2 = \alpha^2 \sigma_R^2 + \sigma^2$, implying that $\sigma_R^2 = \sigma^2 / (1-\alpha^2)$.

If instead the above expression for R_t is used, we can see that

$$\sigma_R^2 = \alpha^{2t}\sigma_R^2 + \sum_{k=0}^{t-1} \alpha^{2k}\sigma^2, \text{ so that } \sigma_R^2 = \frac{1}{1-\alpha^{2t}} \times \frac{1-\alpha^{2t}}{1-\alpha^2}\sigma^2$$

(iv) μ_t represents the mean rainfall in one particular month of the year. Estimate this by calculating the average rainfall in that month over the 10-year period,

$$\hat{\mu}_t = \frac{1}{10}(r_t + r_{t+12} + r_{t+24} + \cdots + r_{t+108}).$$

α is best estimated using the lag-1 autocorrelation of $R-\mu$,

$$\hat{\alpha} = \frac{\sum_{t=2}^{120}(r_t - \hat{\mu}_t)(r_{t-1} - \hat{\mu}_{t-1})}{\sum_{t=1}^{120}(r_t - \hat{\mu}_t)^2}.$$

Noticing that $\text{Cov}(D_t, R_t) = -\theta \text{ Var}(R_t)$, we suggest minus the ratio of the sample covariance of D with R divided by the sample variance of the R, i.e.

$\hat{\theta} = -\dfrac{\sum_{t=1}^{120}(r_t - \overline{r})(d_t - \overline{d})}{\sum_{t=1}^{120}(r_t - \overline{r})^2}$, although in fact a better estimate would be obtained if account was taken of the variable mean of the R_t.

(Less appropriate is using the fact that $\text{Var}(D_t) = \theta^2 \text{Var}(R_t)$ to suggest $\hat{\theta} = \sqrt{s_D^2 / s_R^2}$, since the sign of $\hat{\theta}$ is lost in this operation.)

Now we have an estimate for θ, we can use $\hat{\nu} = \overline{R}_t + \hat{\theta}\overline{D}_t$.

(v) $X_{121} = X_{120} + R_{121} - D_{121}$. In addition, $R_{121} = \mu_1 + \alpha(R_{120} - \mu_0) + e_{121}$ and $D_{121} = \nu - \theta R_{121}$.

Therefore $X_{121} = X_{120} - \nu + (1+\theta)[\mu_1 + \alpha(R_{120} - \mu_0)] + (1+\theta)e_{121}$, which implies that $\hat{x}_{120}(1) = X_{120} - \nu + (1+\theta)[\mu_1 + \alpha(R_{120} - \mu_0)]$
and $\text{Var}(X_{121} - \hat{x}_{120}(1)) = (1+\theta)^2\sigma^2$.

Most candidates made encouraging progress with this question, though part (iv) proved difficult for most. In part (v), as well, the relationships between the variables were not always well understood.

10 (i) The increment $S_{t+s}–S_t$ is equal to $cs - (N_{t+s}–N_t)$, which is independent of anything that happened before time t and has the same distribution as $cs - (N_s–N_0)$, using the independent increment property of the Poisson process.

Alternatively, it is possible to use the Decomposition Theorem, to state that a Lévy process is any sum of three components: a deterministic linear function of time, a purely continuous random component (a multiple of Brownian motion) and a purely discontinuous random component (a compound Poisson process). In this case the coefficient of the Brownian component is zero and the jumps of the compound Poisson process all have size –1.

(ii) By the independent increments property, $N_{t+s} - N_t \sim Po(\lambda s)$ given N_t.

(iii) (a) $E[Y_{t+s} \mid F_t] = e^{\beta[u+c(t+s)]}E[e^{-\beta N_{t+s}} \mid F_t] = e^{\beta[u+c(t+s)-N_t]+\lambda s[e^{-\beta}-1]}$.

This is equal to Y_t if $\beta c = \lambda[1-e^{-\beta}]$.

(b) If $c < \lambda$, the line $y = c\beta$ increases slowly from 0 as β increases, whereas $\lambda[1–e^{-\beta}]$ increases more rapidly to start with, but never exceeds λ. Therefore there must be a crossing point at a positive value of β.

If, on the other hand, $c > \lambda$, then there is no positive crossing point; the line $y = c\beta$ goes slowly to $-\infty$ as $\beta \to -\infty$, whereas $\lambda[1-e^{-\beta}]$ tends to $-\infty$ exponentially fast as $\beta \to -\infty$.

(c) Yes; this is the condition for the company to remain solvent.

(iv) (a) If S hits K first, then $S_T = K$ exactly, since increases in S occur continuously. But if S becomes negative before hitting K, then the negative movement must have been caused by a downward jump; jumps are of magnitude 1, so in this case S_T can be no less than -1.

(b) Since S_t is between -1 and K for all $0 \le t \le T$, it follows that $Y_{\min(t,T)}$ is bounded above and below, from which we deduce that the optional stopping theorem applies.

(c) $S_T = K$ with probability $1-\psi$, or with probability ψ, S_T takes a value somewhere between -1 and 0. Therefore Y_T takes the value $e^{\beta K}$ with probability $1-\psi$, or with probability ψ it takes some value between 1 and $e^{-\beta}$. Hence the result.

(d) Clearly, $\psi \le \dfrac{E[Y_T]-e^{\beta K}}{1-e^{\beta K}} = \dfrac{e^{\beta u}-e^{\beta K}}{1-e^{\beta K}}$.

As K becomes larger, the fact that $\beta < 0$ means that $e^{\beta K} \to 0$, so in the limit we have $\psi \le e^{\beta u}$.

There was an error in the final part of the question; where the question asked for a lower bound, the quantity which could be derived from the previous part of the question was in fact an upper bound. Candidates who reached the last part and were confused by the error were treated generously.

It appeared that many candidates tried this question when they were short on time. Those candidates who attempted part (iv) often did quite well on it, even if they had omitted earlier parts of the question.