EXAMINATION

September 2005

Subject CT4 — Models (includes both 103 and 104 parts) Core Technical

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.

M Flaherty
Chairman of the Board of Examiners

15 November 2005

EXAMINERS' COMMENTS

Comments on solutions presented to individual questions for this September 2005 paper are given below:

103 Part

Question A1 *This was not well answered.*
There was a lot of repetition in some of the solutions offered - for example several different instances of parameter error may have been mentioned.

Question A2 *This was well answered overall, even by the weaker candidates.*
Credit was not given in part (ii)(b) if the examples cited were not likely to be encountered by an actuary working in a professional capacity.

Question A3 *This was well answered overall.*
Some candidates lost marks by not explaining why the chains were not irreducible and were aperiodic. Many candidates did not correctly identify the state space of the chain C_n and most did not realise that the chain will escape to infinity as the value increases without barrier.

Question A4 *This was very well answered overall, with the majority of candidates scoring highly.*
One common mistake was the omission of the constant term from the likelihood function in part (i).

Question A5 *This was very poorly answered by all but a few candidates.*
Some candidates offered general explanations in parts (i) and (iii), which, if clear enough, were given some credit.

Question A6 *Overall this was not well answered.*
In part (i), few candidates gave the full, correct Kolmogorov equations.
Many candidates lost marks in part (ii) because of insufficient or inaccurate working.

Question A7 *Overall this was not well answered.*
However, part (i) was well answered. Some candidates reached the correct answer via a different solution and received full credit.
Many candidates struggled with part (ii), failing to identify the correct integrand required.
In part (iii), many candidates described the shape of the function, but few explained it, as required by the question.

104 Part

Question B1 *This was not well answered.*
Some candidates commented on the advantages/disadvantages of graduation in general, rather than concentrating on the parametric formula method.

Question B2 *Part (i) was well answered.*
In part (ii), many candidates clearly did not understand the meaning of non-informative censoring.

Question B3 *This was well answered overall.*
In part (ii), the question asked candidates to "derive an expression" and therefore we were looking for clearly set out steps here. Many candidates lost marks by not providing sufficient explanation of their working.

Question B4 *This was very well answered, even by the weaker candidates.*
The main areas where candidates lost marks were: not stating the null hypothesis, or not stating it clearly enough; failure to identify the correct degrees of freedom to be used in the test; and insufficient or insufficiently clear descriptions of the shortcomings.
In part (iii), the majority of candidates seemed confused between two issues in connection with bias. There are two distinct problems. Firstly, if the consistent bias is only small, the chi-squared test may fail to detect it because the resulting number (i.e. the sum of the squared deviations) is not large enough to exceed the critical value. The signs test, which ignores the magnitude of the bias and looks only at how consistent it is across the ages, can be used to identify this. The second problem is that even if the consistent bias is larger and the chi-squared test leads us to reject the null hypothesis, the test gives no indication of whether the graduated rates are too high or too low. This is because the deviations are squared and the test statistic always positive. The signs test is <u>not</u> a solution to this second problem.

Question B5 *This was well answered overall.*
Some parts of the question required candidates to "show" a result; candidates lost marks if their working was not sufficiently clear or complete.

Question B6 *This was not well answered.*
Surprisingly few candidates correctly answered part (i).
In parts (ii) and (iii), very few candidates recognised that the expectation of life was an average of the future lifetimes of those bulbs still shining. As a result, although many candidates correctly calculated the expectation of life for a one-day old bulb, few managed to do so for a new bulb. In part (iii), most candidates commented on the higher force of failure in the first day.

103 Part

A1

Items to be mentioned include:

- Models will be chosen which it is felt give a reasonable reflection of the underlying real world processes, but this may not turn out to be the case. (Model error.)

- The model may be very sensitive to parameters chosen, and the parameters are estimates because the true underlying parameters cannot be observed. (Parameter error.)

- Sampling error may result from running insufficient simulations. (It should be possible to give a confidence interval for the error that could result from this source.)

- The management actions assumed may not match what would happen in extreme circumstances.

- Policyholder behaviour, such as take-up rates for options, may differ in practice.

- There may be future events, such as legislative changes which affect the interpretation of the policy conditions, which have not been anticipated in the modelling.

- There may be errors in the coding of the model. The model is likely to be complex and difficult to verify completely.

- The model relies on input data, which may be grouped rather than being able to run every policy. Any errors in the data could cause the output to be inaccurate.

A2

(i) (a) The state space is the set of values which it is possible for each random variable X_t to take.

(b) The time set is the set J, the times at which the process contains a random variable X_t.

(c) A sample path is a joint realisation of the variables X_t for all t in J, that is a set of values for X_t (at each time in the time set) calculated using the previous values for X_t in the sample path.

(ii) Discrete State Space, Discrete Time

(a) Simple random walk, Markov chain, or any other suitable example

(b) Any reasonable example. For example: No Claims Discount systems, Credit Rating at end of each year

Discrete State Space, Continuous Time

(a) Poisson process, Markov jump process, for example

(b) Any reasonable example. For example: Claims received by an insurer, Status of pension scheme member

Continuous State Space, Discrete Time

(a) General random walk, time series, for example

(b) Any reasonable example. For example: Share prices at end of each trading day, Inflation index

Continuous State Space, Continuous Time

(a) Brownian motion, diffusion or Itô process, for example.
Compound Poisson process if the defined state space is continuous.

(b) Any reasonable example. For example: Share prices during trading period, Value of claims received by insurer

A3 (a) Given the current state (the largest outcome or the number of sixes) up to the nth roll, no additional information is required to predict the status of the chain after the next roll. Therefore both B_n and C_n have the Markov property.

(b) B_n has state space {1, 2, 3, 4, 5, 6},
the state space for C_n is the set of non-negative integers.

(c) For B_n, and $1 \le i, j \le 6$,

$$P\left(B_{n+1} = j \mid B_n = i\right) = \frac{i}{6} \quad \text{for } j = i,$$

$$P\left(B_{n+1} = j \mid B_n = i\right) = \frac{1}{6} \quad \text{for each } j > i$$

and $$P\left(B_{n+1} = j \mid B_n = i\right) = 0 \quad \text{for } i > j$$

For C_n, and for $k = 0,1,2,\ldots,$

$$P\left(C_{n+1} = k+1 \mid C_n = k\right) = \frac{1}{6},$$

$$P\left(C_{n+1} = k \mid C_n = k\right) = \frac{5}{6},$$

and $$P\left(C_{n+1} = j \mid C_n = k\right) = 0 \quad \text{for all other } j \ne k, k+1$$

(d) The chain B_n is clearly aperiodic; if currently at state i, it can remain there if the next outcome is at most i.
It is *not* irreducible, as it cannot be reached from j for $i < j$.

C_n is again aperiodic; if currently at state i, it can remain there if the next outcome is not a 6.
It is not irreducible; state k cannot be reached from m if $k < m$.

(e) In the long run, B_n will reach state 6 and will remain there; hence in equilibrium $P(B_n = 6) = 1$ for sufficiently large n.

C_n cannot decrease and has an infinite state space; therefore, it is certain that it will escape to infinity with probability one.

A4 (i) The likelihood is

$$L = K \times \exp(-652(\sigma + \mu))\exp(-44(\rho + \upsilon))\sigma^{23}\rho^{15}\mu^{3}\upsilon^{5}$$

(ii) $l = \ln L = -652\sigma + 23 \ln \sigma$ + constant with respect to σ

Differentiating with respect to σ gives

$$\frac{\partial l}{\partial \sigma} = -652 + \frac{23}{\sigma}$$

and setting equal to zero gives

$$0 = -652 + \frac{23}{\hat{\sigma}}$$

$$\Rightarrow \hat{\sigma} = \frac{23}{652} = 0.0353 \text{ p.a.}$$

Differentiating again gives

$$\frac{\partial^2 l}{\partial \sigma^2} = -\frac{23}{\sigma^2} < 0$$

therefore $\hat{\sigma}$ is the maximum likelihood estimate

(iii) The variance of $\tilde{\sigma}$ is $-\left(\frac{\partial^2 l}{\partial \sigma^2}\right)^{-1} = \frac{\sigma^2}{23}$,

which we can estimate by $\frac{\hat{\sigma}^2}{23}$.

Therefore the estimated standard deviation of $\tilde{\sigma}$ is $\frac{\hat{\sigma}}{\sqrt{23}} = 0.00736$.

A5 (i) Let N_t denote the number of claims up to time t. Since the Poisson process has stationary increments, we may take t = 0, so that the required conditional distribution is

$$P\left(T_0 \le y \mid N_s = 1\right) = \frac{P\left(T_0 \le y, \quad N_s = 1\right)}{P\left(N_s = 1\right)}$$

$$= \frac{P\left(\ N_y = 1, \quad N_s - N_y = 0\right)}{P\left(N_s = 1\right)}$$

But $N_s - N_y$ is independent of N_y
and has the same distribution as N_{s-y}.

Thus the right hand side above equals

$$\frac{(\lambda y e^{-\lambda y})e^{-\lambda(s-y)}}{\lambda s e^{-\lambda s}} = \frac{y}{s},$$

which is the cdf of the uniform distribution on [0, s].

(ii) Since holding times are independent, each having an exponential distribution, their joint density is

$$\lambda^n e^{-\lambda\left(t_1+t_2+\ldots+t_n\right)} 1_{\left\{t_1,t_2,\ldots,t_n>0\right\}}.$$

(iii) We have, as in part (i),

$$P\left(N_s = k \mid N_t = n\right) = \frac{P\left(N_s = k, \quad N_t = n\right)}{P\left(N_t = n\right)}$$

$$= \frac{P\left(\ N_s = k, \quad N_t - N_s = n-k\right)}{P\left(N_t = n\right)}$$

Using again that the Poisson process has stationary and independent increments, and that the number of claims in an interval [0, t] is Poisson (λt), we derive from above that

$$P\left(N_s = k \mid N_t = n\right) = \frac{\dfrac{e^{-\lambda s}(\lambda s)^k}{k!} \cdot \dfrac{e^{-\lambda(t-s)}\lambda^{n-k}(t-s)^{n-k}}{(n-k)!}}{\dfrac{e^{-\lambda t}(\lambda t)^n}{n!}}$$

$$= \frac{e^{-\lambda t}\lambda^n s^k (t-s)^{n-k}}{k!(n-k)!} \cdot \frac{n!}{e^{-\lambda t}\lambda^n t^n}$$

$$= \frac{n!}{k!(n-k)!} \cdot \frac{s^k (t-s)^{n-k}}{t^k t^{n-k}}$$

$$= \binom{n}{k}\left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}$$

which is binomial with parameters *n* and *s/t*.

A6 (i) The generator matrix is

$$A = \begin{pmatrix} -\lambda & \lambda & & & & & 0 \\ & -\lambda & \lambda & & & & \\ & & \cdot & \cdot & & & \\ & & & \cdot & \cdot & & \\ & & & & \cdot & \cdot & \\ & & & & & \lambda & \\ & & & & & -\lambda & \lambda \\ 0 & & & & & & 0 \end{pmatrix},$$

all other entries being zero

The Kolmogorov equations are $P'(t) = P(t)A$.

In a component form the forward equations read

$$p'_{ii}(t) = -\lambda p_{ii}(t) \qquad \text{for } 0 \le i \le N-1$$

$$p'_{ij}(t) = -\lambda p_{ij}(t) + \lambda p_{i,j-1}(t) \qquad \text{for } i < j < N$$

$$p'_{iN}(t) = \lambda p_{i,N-1}(t).$$

(ii) Differentiating the function given in the question, we get first for $i = j$,

$$p'_{ii}(t) = -\lambda e^{-\lambda t},$$

while for $i < j \le N$,

$$p'_{ij}(t) = -\lambda e^{-\lambda t}\frac{(\lambda t)^{j-i}}{(j-i)!} + \lambda e^{-\lambda t}\frac{(\lambda t)^{j-i-1}}{(j-i-1)!}$$

We can then check that the above satisfy the forward equations.

(iii) For $i = j(<N)$, the solution in (ii) implies that $p_{ii}(t) = e^{-\lambda t}$, so that the distribution of the holding times $T_0, T_1, \ldots, T_{N-1}$ is exponential with parameter λ.

For $i = N$, this is obviously not true; once the chain reaches state N, it stays there forever.

A7 (i) $$\frac{d}{dt}P_{\overline{AA}}(t) = -2t \times P_{\overline{AA}}(t)$$

$$\Rightarrow \frac{d}{dt}\left[\ln P_{\overline{AA}}(t)\right] = -2t$$

$$\Rightarrow \ln P_{\overline{AA}}(s) = -s^2 + \text{constant}$$

We know $P_{\overline{AA}}(0) = 1$, hence constant $= 0$

Hence, $P_{\overline{AA}}(s) = \exp^{-s^2}$

(ii) P(in first visit to B at time $T \mid$ in state A at $t = 0$)

$$= \int_0^T P(\text{remains in A to time } s)$$

$$\times P(\text{transition to B in time } s, s + ds)$$

$$\times P(\text{remains in B to time } T)\, ds$$

$$= \int_{s=0}^{T} P_{\overline{AA}}(s) \times 2s \times P_{\overline{BB}}(s,T)\,ds$$

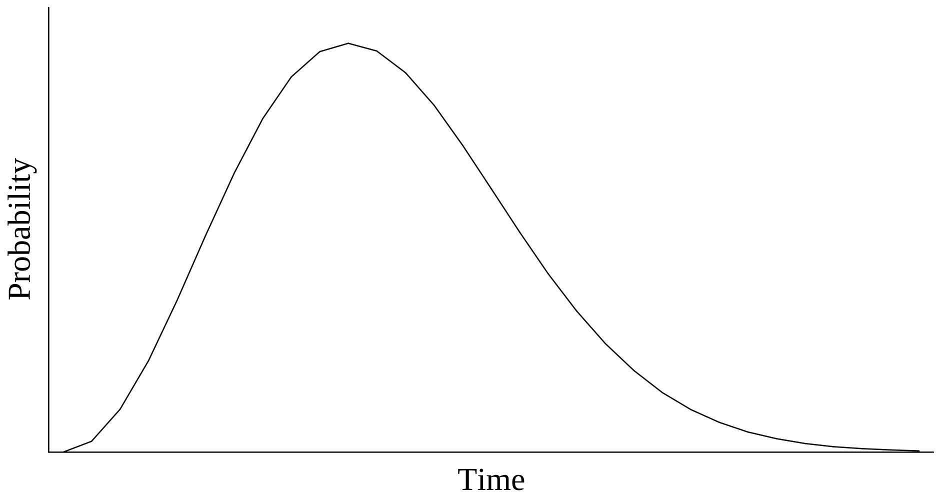
Using the result from part (i) and the similar result for P_{BB} with boundary condition $P_{BB}(s, s) = 1$, this gives us:

$$= \int_{s=0}^{T} e^{-s^2} \times 2s \times e^{-T^2+s^2}\,ds$$

$$= \int_{s=0}^{T} 2s \times e^{-T^2}\,ds$$

$$= e^{-T^2} \times T^2$$

(iii) (a) The sketch should be shaped like:



(b) Commentary:

- Initially probability increases from 0 at $T = 0$, and accelerates as the transition rate from A to B increases.

- However, as transitions increase, it becomes more likely that the process has already visited state B and jumped back to A. Therefore the probability of being in the first visit to B tends (exponentially) to zero.

(c) Differentiate to find turning point:

$$\frac{d}{dt}\left[e^{-t^2} \times t^2\right] = 2t \times e^{-t^2} - 2t^3 \times e^{-t^2}$$

set derivative equal to zero

$$e^{-t^2} \times 2t \times (1 - t^2) = 0$$

implies $t = 1$ for a positive solution
and, from above analysis, this is clearly a maximum.

104 Part

B1 *Advantages:*

The graduated rates will progress smoothly provided the number of parameters is small.

Good for producing standard tables.

Can easily be extended to more complex formulae, provided optimisation can be achieved.

Can fit the same formula to different experiences and compare parameter values to highlight differences between them.

Disadvantages:

It can be hard to find a formula to fit well at all ages without having lots of parameters.

Care is required when extrapolating: the fit is bound to be best at ages where we have lots of data, and can often be poor at extreme ages.

B2 (i) The table below gives the relevant calculations.

Lecture j	n_j	d_j	c_j	λ_j	$1-\lambda_j$	*S(j)*
1	50	1	5	1/50	49/50	0.980
2	44	0	3	0	1	0.980
3	41	3	2	3/41	38/41	0.908
4	36	1	0	1/36	35/36	0.883
5	35	2	0	2/35	33/35	0.833
6	33	1	0	1/33	32/33	0.807
7	32	0	0	0	1	0.807
8	32					

The Index of Lecture Boringness is therefore equal to 0.807.

(ii) Censoring in this case is unlikely to be non-informative.

This is because the students who switched courses were probably less interested in the subject matter of Survival Models than those who remained registered.

Therefore they would have been more likely, had they not switched courses, to cease attending lectures than those who did not switch.

B3 (i) The classification of deaths implies a calendar year rate interval.

A person who dies will be aged x on the birthday in the calendar year of death, which implies that he or she will be aged x next birthday on 1 January in the calendar year of death.

Since 1 January is the start of the rate interval, the age range at the start is $x - 1$ to x.

(ii) A census of those aged x next birthday on 1 January in each year would correspond to the classification of deaths.

But we have lives classified by age x last birthday.

However, the number alive aged x next birthday on any date is equal to the number alive aged $x - 1$ last birthday.

The number alive aged $x - 1$ last birthday on 1 January in year t is given by $P_{x-1}(t)$.

At the end of year t this cohort will be aged x last birthday.

Thus, using the trapezium rule, the correct exposed to risk at age x in year t is given by

$$\frac{1}{2}\left[P_{x-1}(t) + P_x(t+1)\right].$$

Over the three calendar years 2002, 2003 and 2004, we have, therefore, exposed to risk =

$$\frac{1}{2}\left[P_{x-1}(2002) + P_x(2003)\right]$$
$$+\frac{1}{2}\left[P_{x-1}(2003) + P_x(2004)\right]$$
$$+\frac{1}{2}\left[P_{x-1}(2004) + P_x(2005)\right].$$

(iii) Assuming birthdays are uniformly distributed over the calendar year, the average age at the start of the rate interval will be x – ½.

Therefore the average age in the middle of the rate interval is x.

Assuming a constant force of mortality between x – ½ and x + ½, therefore, $f = 0$.

B4 (i) The null hypothesis is that the observed data come from a population in which the graduated rates are the true rates.

The chi-squared statistic is given by the formula:

$$\sum_x \frac{(d_x - E_x \mathring{q}_x)^2}{E_x \mathring{q}_x}.$$

The calculations are shown in the table below.

Age	$E_x q_x$	$E_x \mathring{q}_x$	$(E_x q_x - E_x \mathring{q}_x)^2$	$\frac{(E_x q_x - E_x \mathring{q}_x)^2}{E_x \mathring{q}_x}$
18	6	6.24	0.0576	0.0092
19	8	6.50	2.2500	0.3461
20	12	7.20	23.0400	3.2000
21	8	8.50	0.2500	0.0294
22	9	7.22	3.1684	0.4388
23	6	7.20	1.4400	0.2000
24	8	6.72	1.6384	0.2438

Therefore the calculated chi-squared value is

0.0092 + 0.3461 + 3.2000 + 0.0294 + 0.4388 + 0.2000 + 0.2438 = 4.4673

Since we have 7 ages, we compare this with the tabulated value at the 5% level at, say, 4 degrees of freedom (since we lose 2–3 degrees for every 10 ages graduated graphically).

The tabulated value with 4 degrees of freedom is 9.488.

Since 4.4673 < 9.488 we have no evidence to reject the null hypothesis.

(ii) On the basis of the chi-squared test, the graphical graduation adheres to the data satisfactorily.

However, there is a large deviation at age 20 which requires further investigation.

(iii) Possible shortcomings, and the relevant tests are:

There may be long runs of deviations of the same sign caused by undergraduation.
These can be detected by the grouping of signs test or the serial correlations test.

There may be one or two large deviations at particular ages, balanced by lots of small deviations (as in the example in part (i))
These can be detected by the individual standardised deviations test.

The graduated rates may be too high or too low over the whole of the age range, but by an amount too small for the chi-squared test to detect.
The signs test or the cumulative deviations test will detect this.

The results of the graduation may not be smooth.
This can be detected by looking at the third order differences of the graduated rates $\mathring{q}_x$. If the rates are smooth, these should be small in magnitude compared with the quantities themselves and should progress regularly.

B5 (i) Taking logarithms of the Gompertz hazard produces

$$\log \lambda_x = \log B + x \log c$$

which indicates that the rate of increase of the hazard with age is constant.

Empirically, this is often a reasonable assumption for middle ages and older ages, which include the age range 50–65 years.

(ii) Putting $B = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3)$ into the Gompertz model produces

$$\lambda_x = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3) \,.\, c^x,$$

defining x as duration since 50th birthday.

The hazard can therefore be factorised into two parts:

$\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3)$, which depends only on the values of the covariates, and

c^x, which depends only on duration.

Therefore the ratio between the hazards for any two persons with different characteristics does not depend on duration, and so the model is a proportional hazards model.

(iii) (a) The baseline hazard in this model relates to

a female,
non-smoker,
who drinks less than 21 units of alcohol per week.

(b) For a female cigarette smoker who does not consume alcohol we have

$X_1 = 0,\ X_2 = 1,\ X_3 = 0$ and $x = 5$.

Therefore the hazard is given by

$$\begin{aligned}\lambda_5 &= \exp(\beta_0 + \beta_1 .0 + \beta_2 .1 + \beta_3 .0)\ .\ c^5 \\ &= \exp(-5 + 0.75) \times 1.10^5 \\ &= 0.0230.\end{aligned}$$

(c) The hazard for a non-smoker at duration u is given by the formula

$$\lambda_u = \exp(\beta_0 + \beta_1 X_1 + \beta_3 X_3)\ .\ c^u,$$

The hazard for a smoker at duration v is given by the formula

$$\lambda^*_v = \exp(\beta_0 + \beta_1 X_1 + 0.75 + \beta_3 X_3)\ .\ c^v.$$

If the smoker's and non-smoker's hazards are the same, then

$$\lambda_u = \lambda^*_v,$$

which implies that

$$\exp(\beta_0 + \beta_1 X_1 + \beta_3 X_3).c^u = \exp(\beta_0 + \beta_1 X_1 + 0.75 + \beta_3 X_3)\ .\ c^v.$$

which simplifies to

$$c^u = \exp(0.75)\ .\ c^v,$$

so that

$$c^u/c^v = c^{u-v} = \exp(0.75) = 2.117.$$

Since $c = 1.1$, we have

$$1.1^{u-v} = 2.117.$$

Therefore

$$\begin{aligned}u - v &= \log(2.117)/\log(1.1) \\ &= 0.75/0.0953 = 7.87.\end{aligned}$$

So when the two hazards are equal, the non-smoker is approximately eight years older than the smoker.

Alternatively this could be demonstrated by calculating λ_u and λ^*_{u-8} and showing that they are approximately the same.

B6 (i) Let the probability of failure within the first 20 days be ${}_{20}q_0$.

We have:

$$\begin{aligned}{}_{20}q_0 &= 1 - {}_{20}p_0 = 1 - {}_{1}p_0 \cdot {}_{19}p_1 \\ &= 1 - (1 - {}_{1}q_0)\exp(-19\mu) \\ &= 1 - 0.95\exp(-19 \times 0.01) \\ &= 1 - 0.95\exp(-0.19) \\ &= 1 - 0.95(0.82696)\end{aligned}$$

which is 0.21439.

(ii) (a) The complete expectation of life of a one-day old light bulb, $\mathring{e}_1$ is given by

$$\mathring{e}_1 = \int_0^\infty {}_t p_1 dt$$
$$= \int_0^\infty e^{-0.01t} dt$$

Integrating, this gives

$$\mathring{e}_1 = -\frac{1}{0.01}\left[e^{-0.01t}\right]_0^\infty = -\frac{1}{0.01}\left[0 - 1\right]$$

= 100 days.

(b) The complete expectation of life of a new light bulb, $\mathring{e}_0$ is given by

$$\mathring{e}_0 = \int_0^\infty {}_t p_0 dt = \int_0^1 {}_t p_0 dt + \int_1^\infty {}_t p_0 dt. \quad (*)$$

Alternative 1

Assume a uniform distribution of failure times between exact ages 0 and 1,

the first term in (*) is equal to

$$\frac{1}{2}\left(1+{}_1p_0\right)$$

$$=\frac{1}{2}\left[1+(1-{}_1q_0)\right]$$

$$=\frac{1}{2}(1+0.95)=0.975$$

The second term is equal to

$${}_1p_0\int_0^{\infty}{}_tp_1\,dt=0.95(100)$$

(using the result from part (i) above).

Therefore:

$$\overset{\circ}{e}_0=0.975+100\times 0.95=95.975 \text{ days.}$$

Alternative 2

Assume a constant force of failure between exact ages 0 and 1

Let this constant force be δ.

Then

$${}_1p_0=\exp\left[-\int_0^1 \delta\,ds\right]=\exp(-\delta)$$

$$=1-{}_1q_0=0.95.$$

So that

$$\exp(-\delta)=0.95$$

and

$$\delta=-\log(0.95)=0.0513.$$

Thus the first term on the right-hand side of (*) is

$$\int_0^1 {}_t p_0 dt = \int_0^1 \exp(-0.0513t)dt$$

$$= \frac{1}{-0.0513}\left[\exp(-0.0513t)\right]_0^1$$

$$= \frac{1}{-0.0513}\left[\exp(-0.0513) - 1\right]$$

$$= 0.97478,$$

and the second term is equal to

$${}_1 p_0 \int_0^\infty {}_t p_1 dt = 0.95(100)$$

(using the result from part (i) above).

So that

$$\overset{\circ}{e}_0 = 0.97478 + 100 \times 0.95 = 95.97478 \text{ days.}$$

(iii) The complete expectation of life of a light bulb at any age is an average of the future lifetimes of all bulbs which have not failed before that age.

The value of $\overset{\circ}{e}_0$ is lower than $\overset{\circ}{e}_1$ because the average $\overset{\circ}{e}_0$ includes the very short lifetimes of the relatively large proportion of bulbs which fail in the first day, which deflate the average, whereas $\overset{\circ}{e}_1$ excludes these.

END OF EXAMINERS' REPORT