CT4 – P XS – 16

Series X Solutions

ActEd Study Materials: 2016 Examinations

Subject CT4

Contents

Series X Solutions

If you think that any pages are missing from this pack, please contact ActEd's admin team by email at **ActEd@bpp.com**.

How to use the Series X Solutions

Guidance on how and when to use the Series X Solutions is set out in the *Study Guide for the 2016 exams*.

Important: Copyright Agreement

ISBN 178-1-4727-0090-0

1 781472 700900

Assignment X1 – Solutions

Markers: This document sets out one approach to solving each of the questions. Please give credit for other valid approaches.

Solution X1.1

Comment
Reference: Chapter 2

(a) ***Strict stationarity***

A stochastic process $\{X_t : t \in J\}$ is said to be strictly stationary if the joint distributions of $X_{t_1}, X_{t_2}, ..., X_{t_n}$ and $X_{t_1+k}, X_{t_2+k}, ..., X_{t_n+k}$ are identical for all $t_1, t_2, ..., t_n$ and $t_1+k, t_2+k, ..., t_n+k$ in the time set J and all positive integers n. [1½]

Markers: award 1 mark for the less precise answer of "the statistical properties of the process do not change over time".

(b) ***Weak stationarity***

A stochastic process $\{X_t : t \in J\}$ is said to be weakly stationary if:

- $E(X_t)$ is constant, *ie* independent of t, and [½]
- $\text{cov}(X_t, X_{t+k})$ depends only on the lag, k. [1]

[Total 3]

Note that $\text{var}(X_t) = \text{cov}(X_t, X_t)$*, so this will be constant if the process is weakly stationary.*

Solution X1.2

> ***Comment***
>
> *Reference: Chapter 2*

(i) ***Value of λ***

Together the teams score an average of 5 goals in $3 \times 20 = 60$ minutes. So:

$$\lambda = \tfrac{1}{2} \times 5 = 2.5 \qquad [1]$$

(ii) ***Probability of each team scoring exactly one goal in 20 minutes***

The expected number of goals scored by each team in 20 minutes is:

$$\frac{20}{60} \times \lambda = \tfrac{1}{3} \times 2.5 = \tfrac{5}{6} \qquad [½]$$

So the probability that one team scores exactly 1 goal in 20 minutes is:

$$\frac{e^{-5/6} \times \left(\frac{5}{6}\right)^1}{1!} = 0.36217 \qquad [1]$$

Since the teams score goals independently, the probability that each team scores exactly 1 goal in 20 minutes is:

$$0.36217^2 = 0.13116 \qquad [½]$$

[Total 2]

(iii) ***Probability that more than 2 goals are scored in the match***

Method 1 (considering both teams combined)

Let N denote the combined number of goals scored by both teams in a match. Then N has a $Poisson(5)$ distribution. [½]

$$\begin{aligned} P(N > 2) &= 1 - P(N=0) - P(N=1) - P(N=2) \\ &= 1 - e^{-5} - \frac{5e^{-5}}{1!} - \frac{5^2 e^{-5}}{2!} \\ &= 1 - 18.5e^{-5} = 0.87535 \end{aligned} \qquad [1½]$$

Method 2 (considering each team separately)

Let N_1 denote the number of goals scored by the home team and N_2 denote the number of goals scored by the away team in a match. Then the probability that more than 2 goals are scored in a match is:

$$1-\big[P(N_1=0,N_2=0)+P(N_1=1,N_2=0)+P(N_1=2,N_2=0)$$
$$+P(N_1=0,N_2=1)+P(N_1=0,N_2=2)+P(N_1=1,N_2=1)\big] \qquad [½]$$

Since N_1 and N_2 are independent, this is:

$$1-\left[e^{-2.5}e^{-2.5}+\left(2.5e^{-2.5}\right)e^{-2.5}+\left(\frac{2.5^2\,e^{-2.5}}{2!}\right)e^{-2.5}+e^{-2.5}\left(2.5e^{-2.5}\right)\right.$$
$$\left.+e^{-2.5}\left(\frac{2.5^2\,e^{-2.5}}{2!}\right)+\left(2.5e^{-2.5}\right)\left(2.5e^{-2.5}\right)\right] \qquad [1]$$

$$=1-18.5e^{-5}$$

$$=0.87535 \qquad [½]$$

[Total 2]

Solution X1.3

Comment
Reference: Chapter 2

(i) ***White noise***

To say that the process $\{Z_n : n=0,1,2,...\}$ is a white noise process means that the random variables Z_n are independent and identically distributed. [1]

To say that the process has a discrete state space means that the Z_n are discrete random variables. [1]

[Total 2]

(ii)(a) ***Independent increments?***

The increments of the process $\{Z_n\}$ are of the form $Z_{n+m} - Z_n$. The process is said to have independent increments if, for all n and every $m > 0$, $Z_{n+m} - Z_n$ is independent of the past values of the process $\{Z_k : k = 0, 1, ..., n\}$. [1]

We have:

$$\operatorname{cov}\left(Z_{n+m} - Z_n, Z_n\right) = -\operatorname{var}\left(Z_n\right) \neq 0$$

So $Z_{n+m} - Z_n$ and Z_n are not independent and the process does not have independent increments. [1]

(ii)(b) ***Markov property?***

Since the random variables Z_n are independent:

$$P\left(Z_n = z_n \mid Z_{n-1} = z_{n-1}, Z_{n-2} = z_{n-2}, ..., Z_0 = z_0\right) = P\left(Z_n = z_n\right) \quad [½]$$

Also:

$$P\left(Z_n = z_n \mid Z_{n-1} = z_{n-1}\right) = P\left(Z_n = z_n\right) \quad [½]$$

So:

$$P\left(Z_n = z_n \mid Z_{n-1} = z_{n-1}, Z_{n-2} = z_{n-2}, ..., Z_0 = z_0\right) = P\left(Z_n = z_n \mid Z_{n-1} = z_{n-1}\right)$$

and the process is Markov. [1]
[Total 4]

From Chapter 2 we know that a process with independent increments has the Markov property. However, even if a process doesn't have independent increments, it may still have the Markov property, as we see from this example.

Solution X1.4

Comment
Reference: Chapter 2

(i)(a) ***State space***

The state space of the stochastic process $\{X_t : t \in J\}$ is the set of values that the random variables X_t can take. The state space can be discrete or continuous. [½]

(i)(b) ***Time set***

The time set for this stochastic process is J, which contains all points at which the value of the process can be observed. The time set can be discrete or continuous. [½]

(i)(c) ***Sample path***

A sample path is a joint realisation of the random variables X_t for all $t \in J$. [1]

[Total 2]

(ii)(a) ***Examples of stochastic processes***

Discrete state space, discrete time set

Examples include Markov chains, simple random walks and white noise processes that have discrete state spaces. [½]

Discrete state space, continuous time set

Examples include Markov jump processes (of which the Poisson process is a special case) and counting processes. [½]

Continuous state space, discrete time set

Examples include general random walks and time series. [½]

Continuous state space, continuous time set

Examples include Brownian motion, diffusion processes and compound Poisson processes where the state space is continuous. [½] (ii)(b) ***Examples of problems an actuary may wish to study***

Discrete state space, discrete time set

An example of this is a no claims discount system. The random variable X_t represents the discount level given to a policyholder in year t, $t = 1, 2, \ldots$. [½]

Discrete state space, continuous time set

An example of this is the health, sickness, death model, which can be used to value sickness benefits. The random variable X_t takes one of the values healthy, sick or dead for each $t \geq 0$. [½]

Continuous state space, discrete time set

An example of this is a company's share price at the end of each trading day. Another example is the annual UK inflation rate. [½]

Continuous state space, continuous time set

An example of this is the cumulative claim amount incurred on a portfolio of policies up to time t. Another example is a company's share price at time t, where t denotes time since trading began. [½]
[Total 4]

Markers: please award marks for other sensible examples.

Solution X1.5

Comment
Reference: Chapter 3

(i) ***Definition of stationary distribution***

Let S be the state space. We say that $\{\pi_j \mid j \in S\}$ is a stationary probability distribution for a Markov chain with transition matrix P if the following hold for all $j \in S$:

- $\pi_j = \sum_{i \in S} \pi_i P_{ij}$
- $\pi_j \geq 0$
- $\sum_{i \in S} \pi_i = 1$ [Total 2]

(ii) ***Stationary distributions for this process***

We need to solve:

$$(\pi_1, \pi_2, \pi_3)\begin{pmatrix} 2/9 & 1/3 & 4/9 \\ 1/3 & 2/3 & 0 \\ 1/9 & 1/9 & 7/9 \end{pmatrix} = (\pi_1, \pi_2, \pi_3)$$ [1]

Note that we can get rid of the fractions by multiplying the equation through by 9. We have the three equations:

$$2\pi_1 + 3\pi_2 + \pi_3 = 9\pi_1$$
$$3\pi_1 + 6\pi_2 + \pi_3 = 9\pi_2$$
$$4\pi_1 + 7\pi_3 = 9\pi_3$$

Rearranging we get:

$$-7\pi_1 + 3\pi_2 + \pi_3 = 0$$
$$3\pi_1 - 3\pi_2 + \pi_3 = 0$$
$$4\pi_1 - 2\pi_3 = 0$$ [1]

We will use π_1 as the working variable. One of the equations is always redundant so we ignore the first equation. The third is easily solved to get $\pi_3 = 2\pi_1$. We can substitute this into the middle equation to get $5\pi_1 - 3\pi_2 = 0$ and hence $\pi_2 = \frac{5}{3}\pi_1$. [2]

This gives the solution $\left(1, \frac{5}{3}, 2\right)$ but since we need $\pi_1 + \pi_2 + \pi_3 = 1$ we must multiply by $\frac{1}{\left(1+\frac{5}{3}+2\right)}$ to get $\left(\frac{3}{14}, \frac{5}{14}, \frac{6}{14}\right)$. [1]

[Total 5]

Solution X1.6

> ***Comment***
>
> *Reference: Chapter 1*

(i) ***Benefits of modelling in actuarial work***

The benefits of modelling in actuarial work are:

- systems with long time frames, such as the operation of a pension fund, can be studied quickly [1]
- systems with stochastic elements, such as the operation of a life insurance company, can be modelled more realistically to capture the randomness [1]
- different future policies or possible actions can be compared to see which best suits the requirements or constraints of a user [1]
- in a model of a complex system, we can usually get much better control over the experimental conditions so that we can reduce the variance of the output results without changing their mean values. [1]

[Total 4]

(ii) ***Difference between a deterministic and a stochastic model***

A deterministic model uses one set of input parameters and gives the results of the relevant calculations for this single scenario. [1]

A stochastic model involves at least one input parameter varying according to an assumed probability distribution. As such, the output will vary along with the input and the model produces distributions of the relevant results for a distribution of scenarios. [1]

Often the output from a stochastic model is in the form of many thousands of simulated outcomes of the process. We can study the distributions of these outcomes. [1]

[Total 3]

Solution X1.7

Comment

Reference: Chapter 3

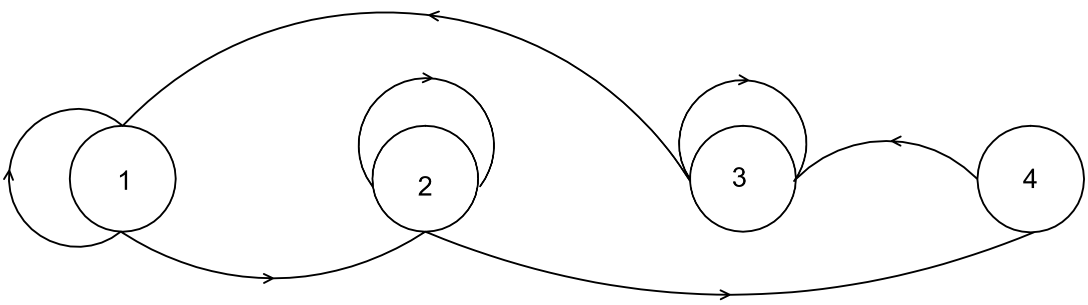
(i) ***Definitions***

A state is said to be periodic with period $d > 1$ if a return to the state is possible only in a number of steps that is a multiple of d. [1]

A Markov chain (or its transition matrix) is said to be irreducible if every state can eventually be reached from every other. [1]

[Total 2]

(ii) ***Matrix P***



From the transition diagram we can see that the process will "cycle" through the states in the order $1 \to 2 \to 4 \to 3 \to 1 \to \cdots$. In the long term, the proportion of time spent in each state (π_i) can then be found from the equation $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$, *ie* the simultaneous equations:

$$\pi_1 = \tfrac{1}{2}(\pi_1 + \pi_3)$$

$$\pi_2 = \tfrac{1}{2}(\pi_1 + \pi_2)$$

$$\pi_3 = \tfrac{1}{2}\pi_3 + \pi_4$$

$$\pi_4 = \tfrac{1}{2}\pi_2$$ [1]

By symmetry, we see that $\pi_1 = \pi_2 = \pi_3$. [1]

So, if we express everything in terms of π_1, these equations reduce to:

$$\pi_1 = \pi_1$$

$$\pi_1 = \pi_1$$

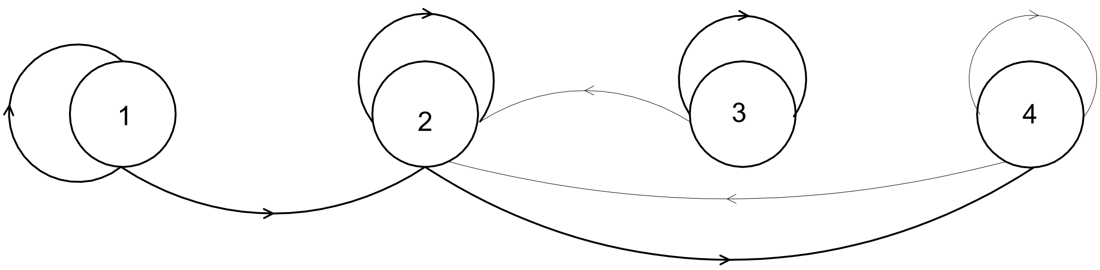
$$\pi_1 = \pi_1$$

$$\pi_4 = \tfrac{1}{2}\pi_1$$

The last one is the only useful equation. Since we also know that $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$, *ie* $3\pi_1 + \pi_4 = 1$, we find that the equilibrium probabilities are:

$$\pi_1 = \pi_2 = \pi_3 = \tfrac{2}{7} \text{ and } \pi_4 = \tfrac{1}{7} \quad [1]$$

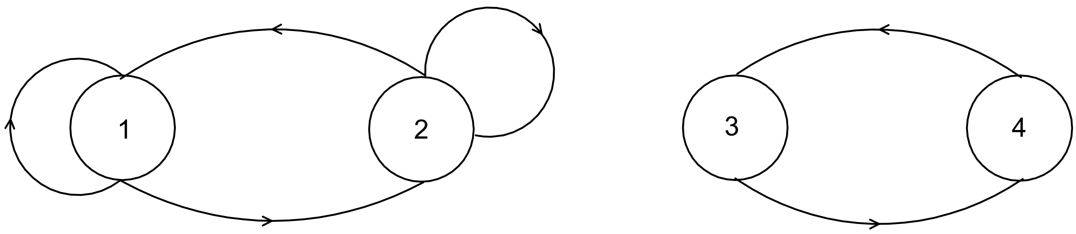
Matrix Q



Once this process has left State 1 or State 3, it can never return to these states. So, in the long term it will end up oscillating at random between State 2 and State 4, swapping on average every two moves. Therefore $\pi_1 = \pi_3 = 0$ and $\pi_2 = \pi_4 = ½$. [2]

In this case, these probabilities are both the stationary distribution and the equilibrium distribution because the process will always converge to this distribution, irrespective of the starting distribution.

Matrix R



If the process starts in State 1 or State 2, it will oscillate at random between these two states, so that $\pi_1 = \pi_2 = ½$ and $\pi_3 = \pi_4 = 0$. If, on the other hand, it starts in State 3 or State 4, it will oscillate periodically between these two states, swapping states on each move. In this case the process will never settle down to an equilibrium state. [2]

In this case, these probabilities are a stationary distribution, but they are not an equilibrium distribution, because the process will not always converge to this distribution.

[Total 7]

Solution X1.8

> ***Comment***
>
> *Reference: Chapter 4*

(i) ***Maximum likelihood estimate***

We need to consider the survivors and the deaths separately.

Survivors

The ith life is observed from age $x+a_i$ to age $x+t_i$. The contribution of this individual to the likelihood is the probability that the life is still alive at age $x+t_i$, which is ${}_{t_i-a_i}p_{x+a_i}$. However, since the force of mortality is assumed to be constant between the ages of x and $x+1$, we can write this as:

$$ {}_{t_i-a_i}p_{x+a_i} = e^{-(t_i-a_i)\mu} \qquad [1]$$

Deaths

Again, the i th life is observed from age $x+a_i$ to age $x+t_i$. But now the i th life dies at age $x+t_i$. The contribution of this individual to the likelihood is:

$$_{t_i-a_i}p_{x+a_i}\,\mu_{x+t_i} = \mu\, e^{-(t_i-a_i)\mu} \qquad [1]$$

Suppose that a total of d out of the original N lives die. Then, assuming that the lives are independent, the likelihood function is given by:

$$\begin{aligned} L &= \prod_{\text{survivors}} e^{-(t_i-a_i)\mu} \prod_{\text{deaths}} \mu\, e^{-(t_i-a_i)\mu} \\ &= \mu^d \prod_{i=1}^{N} e^{-(t_i-a_i)\mu} \\ &= \mu^d \exp\left\{-\mu \sum_{i=1}^{N}(t_i-a_i)\right\} \\ &= \mu^d e^{-\mu v} \end{aligned} \qquad [2]$$

where $v=\sum_{i=1}^{N}(t_i-a_i)$ denotes the total observed waiting time.

Maximising the likelihood function is equivalent to maximising the log-likelihood:

$$\ln L = d\ln\mu - \mu v$$

Differentiating with respect to μ:

$$\frac{d\ln L}{d\mu} = \frac{d}{\mu} - v$$

Setting this equal to 0 and rearranging, we obtain:

$$\mu = \frac{d}{v} \qquad [1]$$

Checking the second derivative:

$$\frac{d^2\ln L}{d\mu^2} = -\frac{d}{\mu^2} < 0$$

So the maximum likelihood estimate of μ is $\hat{\mu} = \dfrac{d}{v}$. [1]

[Total 6]

(ii) ***Asymptotic sampling distribution of maximum likelihood estimator***

The asymptotic sampling distribution of $\tilde{\mu}$ is $N\left(\mu, \dfrac{\mu}{E(V)}\right)$, where $E(V)$ denotes the expected total waiting time for the lives under observation. [2]

Alternatively, you could say that the variance of $\tilde{\mu}$ is given by the Cramér-Rao lower bound:

$$\frac{-1}{E\left(\dfrac{d^2 \ln L}{d\mu^2}\right)}$$

(iii) ***Confidence interval***

An approximate 95% confidence interval for μ is:

$$\hat{\mu} \pm 1.96\sqrt{\text{var}(\tilde{\mu})}$$ [½]

We have:

$$\hat{\mu} = \frac{20}{780} = 0.02564$$ [½]

and $\text{var}(\tilde{\mu})$ can be estimated by:

$$\frac{\hat{\mu}}{v} = \frac{20}{780^2} = 0.00003287$$ [½]

So an approximate 95% confidence interval for μ is:

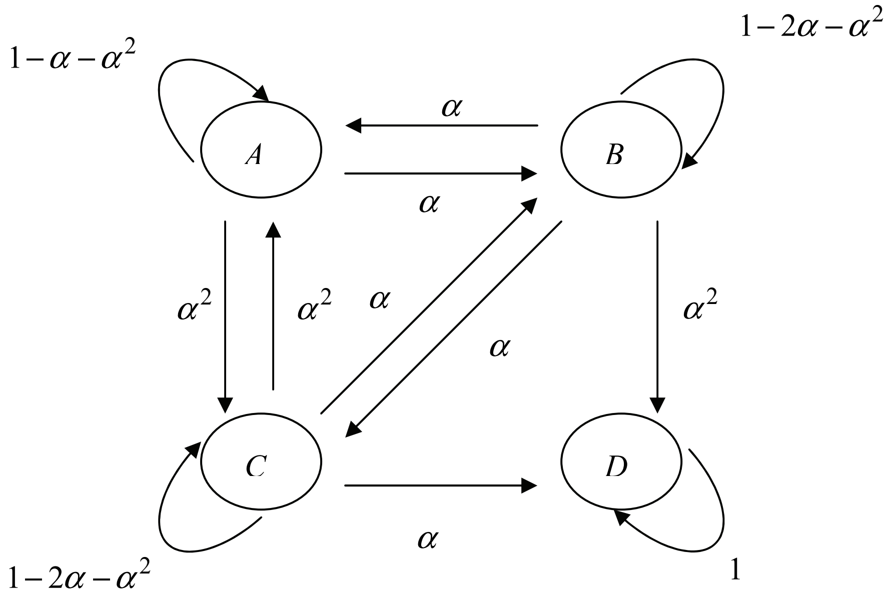
$$\frac{20}{780} \pm 1.96\sqrt{\frac{20}{780^2}} = 0.02564 \pm 0.01124 = (0.01440, 0.03688)$$ [½]

[Total 2]

Solution X1.9

Comment   *Reference: Chapter 3*

(i) ***Transition graph***



[2]

(ii) ***Range of values***

For the matrix to be valid we require all probabilities to lie in the range [0,1]. This immediately imposes α to lie in [0,1], but we also require:

$$0 \le 1-\alpha-\alpha^2 \le 1$$

and:

$$0 \le 1-2\alpha-\alpha^2 \le 1$$

The quadratic $1-\alpha-\alpha^2$ has roots $-0.5 \pm 0.5\sqrt{5}$ and the quadratic $1-2\alpha-\alpha^2$ has roots $-1 \pm \sqrt{2}$. [1]

It follows that α must lie in the interval $[0, -1+\sqrt{2}]$. Indeed, if it does lie in this range, then the matrix does make sense. [1]

[Total 2]

(iii) ***Irreducible and aperiodic?***

The chain is not irreducible since there is no way out of state *D*. It is aperiodic, however, since, for example, you can remain for an arbitrary number of steps in any state if $\alpha < -1+\sqrt{2}$. In the case that $\alpha = -1+\sqrt{2}$ you can still see that the process is aperiodic by inspection, since you can go backwards and forwards between states *A, B* and *C*. [2]

(iv) ***Stationary distribution***

Assume that $\alpha > 0$. Intuitively, no matter where you begin, you must end up in state *D* eventually, with probability 1. The probability of then being in state *A, B* or *C* must be 0. For a stationary distribution, these probabilities cannot change, so the probability of being in any of those 3 states in the stationary distribution must be 0. The only stationary distribution if $\alpha > 0$ is therefore (0,0,0,1). [1]

If $\alpha = 0$, then you always stay where you are; the transition matrix is the identity matrix. It follows in this (unrealistic) case that every possible distribution is stationary.

[1]

[Total 2]

(v) ***Probabilities of having rating D in the third quarter***

If $\alpha = 0.1$, then the transition is:

$$P = \begin{pmatrix} 0.89 & 0.1 & 0.01 & 0 \\ 0.1 & 0.79 & 0.1 & 0.01 \\ 0.01 & 0.1 & 0.79 & 0.1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(a) ***Given rating A in the first quarter***

The initial distribution is $(1,0,0,0)$ and repeated post-multiplication by the matrix P gives:

$$\begin{aligned}(1,0,0,0) &\to (0.89,\ 0.1,\ 0.01,\ 0)\\ &\to (*,\ *,\ *,\ 0.002)\end{aligned}$$

So the probability of being in state *D* in the third quarter is 0.002. [1]

(b) ***Given rating B in the first quarter***

The initial distribution is $(0,1,0,0)$ and repeated post-multiplication by the matrix P gives:

$$\begin{aligned}(0,1,0,0) &\rightarrow (0.1,\, 0.79,\, 0.1,\, 0.01)\\ &\rightarrow (*,\, *,\, *,\, 0.0279)\end{aligned}$$

So the probability of being in state D in the third quarter is 0.0279. [1]

(c) ***Given rating C in the first quarter***

The initial distribution is $(0,0,1,0)$ and repeated post-multiplication by the matrix P gives:

$$\begin{aligned}(0,0,1,0) &\rightarrow (0.01,\, 0.1,\, 0.79,\, 0.1)\\ &\rightarrow (*,\, *,\, *,\, 0.18)\end{aligned}$$

So the probability of being in state D in the third quarter is 0.18. [1]

(d) ***Given rating D in the first quarter***

If you start in state D, then you must stay there. So the probability of being in state D in the third quarter is 1. [1]

[Total 4]

Solution X1.10

Comment
Reference: Chapter 3

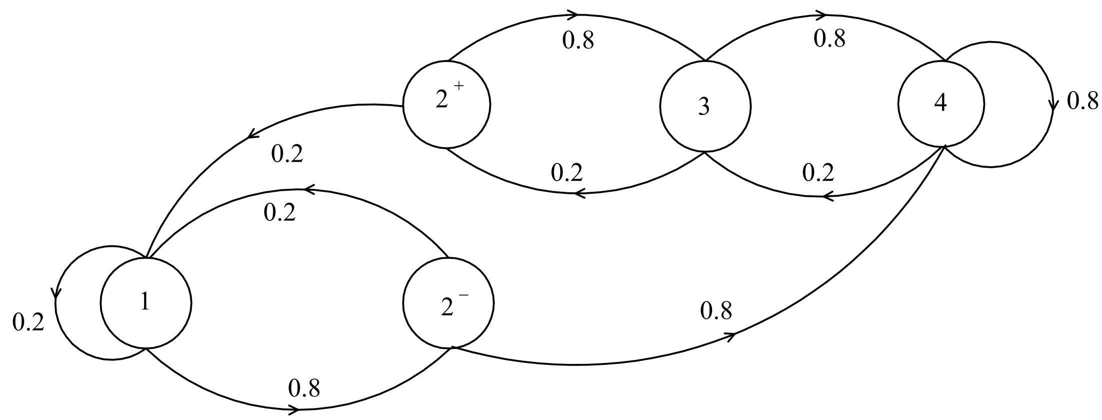
(i) ***Why process is not Markov***

If a policyholder is on Level 2, the probability of moving to Level 4 depends on the level the policyholder was on last year. Hence the process is not Markov. [2]

(ii) ***New process***

Subdivide Level 2 into two levels say 2^- (coming from Level 1 before) and 2^+ (coming from Level 3 before). The new process is Markov and has 5 states. [2]

(iii) ***Transition graph for Y(t)***



[2]

(iv) ***Transition matrix***

One-step transition matrix is:

$$
\begin{array}{c c}
 & \begin{array}{ccccc} 1 & 2^- & 2^+ & 3 & 4 \end{array} \\
\begin{array}{c} 1 \\ 2^- \\ 2^+ \\ 3 \\ 4 \end{array} & \begin{bmatrix} 0.2 & 0.8 & 0 & 0 & 0 \\ 0.2 & 0 & 0 & 0 & 0.8 \\ 0.2 & 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0.2 & 0 & 0.8 \\ 0 & 0 & 0 & 0.2 & 0.8 \end{bmatrix}
\end{array}
$$

[1]

(v) ***Sufficient conditions for unique stationary distribution***

The chain has a finite number of states, 5 …. [½]

… so it has at least one stationary distribution. [½]

The chain is also irreducible as every state can be reached from every other state … [½]

… so it has a unique stationary distribution. [½]

Because the chain is irreducible, the states will either all be aperiodic or will all have the same period. State 1 is aperiodic since it is possible to stay in state 1 in successive time periods. So all the states are aperiodic. [½]

Since the chain is aperiodic as well as having a finite state space and being irreducible, the process with settle down to its unique stationary distribution in the long run. [½]
[Total 3]

(vi) ***Probability of being on Level 2 in the long run***

Solve $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$ to get:

(1) $0.2\pi_1 + 0.2\pi_{2^-} + 0.2\pi_{2^+} = \pi_1$

(2) $0.8\pi_1 = \pi_{2^-}$

(3) $0.2\pi_3 = \pi_{2^+}$

(4) $0.8\pi_{2^+} + 0.2\pi_4 = \pi_3$

(5) $0.8\pi_{2^-} + 0.8\pi_3 + 0.8\pi_4 = \pi_4$ [1]

We can discard (1) and use it for cross checking later. Then:

(3) $\pi_3 = 5\pi_{2^+}$

(4) $\pi_4 = 5\left(\pi_3 - 0.8\pi_{2^+}\right)$

$$= 5\left(5\pi_{2^+} - 0.8\pi_{2^+}\right)$$

$$= 21\pi_{2^+}$$

(5) $$0.8\pi_{2^-} = -0.8\pi_3 + 0.2\pi_4$$

$$= \frac{1}{5}\pi_{2^+}$$

$$\Rightarrow \pi_{2^-} = \frac{1}{4}\pi_{2^+}$$

(2) $$\pi_1 = \frac{5}{4}\pi_{2^-} = \frac{5}{16}\pi_{2^+}$$ [2]

Then normalising the probabilities, we have:

$$\left(\frac{5}{16} + \frac{1}{4} + 1 + 5 + 21\right)\pi_{2^+} = 1$$

$$\Rightarrow \pi_{2^+} = \frac{16}{441}$$

$$\Rightarrow \pi_{2^-} = \frac{4}{441}$$ [1]

Check using equation (1):

Both the LHS and the RHS equal $\frac{5}{441}$.

Then required probability is:

$$\pi_{2^-} + \pi_{2^+} = \frac{20}{441}$$ [1]

[Total 5]

Assignment X2 – Solutions

Markers: This document sets out one approach to solving each of the questions. Please give credit for other valid approaches.

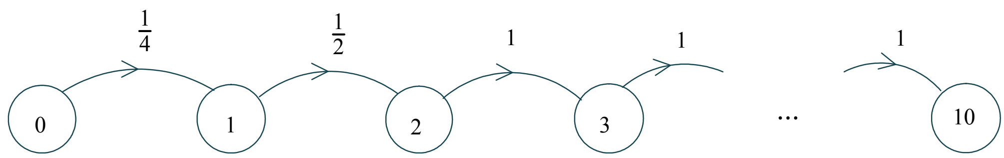
Solution X2.1

Comment
Reference: Chapter 5

(i) ***Probability***

Let $N(t)$ denote the number of breakdowns up to time t. Then the state space of $N(t)$ is the set $\{0, 1, 2, ..., 10\}$.

Drawing a transition diagram might help you see what's going on. We have:



Method 1

Since we are considering a new boiler, $N(0) = 0$. We require $P(N(5) > 1)$, which can be calculated using the equation:

$$P(N(5) > 1) = 1 - P(N(5) = 0) - P(N(5) = 1)$$

First of all, we have:

$$P(N(5) = 0) = e^{-5\times\frac{1}{4}} = 0.28650 \qquad [1]$$

Then, adopting the usual notation:

$$p_{ij}(t) = P(N(t) = j \mid N(0) = i)$$

We can write $P\left(N(5)=1\right)$ in the following integral form:

$$P\left(N(5)=1\right)=\int_0^5 \underbrace{p_{00}(t)}_{\substack{\text{stay in state}\\ \text{0 up to time } t}} \quad \underbrace{\mu_{01}}_{\substack{\text{move from}\\ \text{state 0 to}\\ \text{state 1 at}\\ \text{time } t}} \quad \underbrace{p_{11}(5-t)}_{\substack{\text{stay in state 1}\\ \text{from time } t \text{ to}\\ \text{time 5}}}\ dt \qquad [1]$$

Note that $p_{ii}(t)=p_{\overline{ii}}(t)$ for this model since, once state i has been left, a return to it is impossible.

Now: $p_{00}(t)=e^{-\frac{t}{4}}$ [½]

and $p_{11}(5-t)=e^{-\left(\frac{5-t}{2}\right)}$ [½]

So:

$$P\left(N(5)=1\right)=\int_0^5 e^{-\frac{t}{4}}\tfrac{1}{4}e^{-\left(\frac{5-t}{2}\right)}\,dt$$

$$=\tfrac{1}{4}e^{-\frac{5}{2}}\int_0^5 e^{\frac{t}{4}}\,dt$$

$$=e^{-\frac{5}{2}}\left[e^{\frac{t}{4}}\right]_0^5$$

$$=e^{-\frac{5}{2}}\left(e^{\frac{5}{4}}-1\right)$$

$$=0.20442 \qquad [1½]$$

So: $P\left(N(5)>1\right)=1-0.28650-0.20442=0.50908$ [½]

[Total 5]

Method 2

Alternatively, $P\left(N(5)=1\right)=p_{01}(5)$ can be calculated by solving the appropriate differential equation as follows:

$$\frac{d}{dt}p_{01}(t)=\frac{1}{4}p_{00}(t)-\frac{1}{2}p_{01}(t)$$

$$\Rightarrow \quad \frac{d}{dt}p_{01}(t)+\frac{1}{2}p_{01}(t)=\frac{1}{4}p_{00}(t)=\frac{1}{4}e^{-\frac{t}{4}} \qquad [1]$$

Multiplying through by the integrating factor $e^{\frac{t}{2}}$ gives:

$$e^{\frac{t}{2}}\frac{d}{dt}p_{01}(t)+\frac{1}{2}e^{\frac{t}{2}}p_{01}(t)=\frac{1}{4}e^{\frac{t}{4}} \qquad [½]$$

Then integrating both sides with respect to t, we get:

$$e^{\frac{t}{2}}p_{01}(t)=e^{\frac{t}{4}}+C \qquad [1]$$

where C is a constant of integration. Setting $t=0$ gives:

$$0=1+C$$

ie: $C=-1$ [1]

Hence:

$$p_{01}(t)=e^{-\frac{t}{2}}\left(e^{\frac{t}{4}}-1\right) \qquad [½]$$

and $P\big(N(5)=1\big)=p_{01}(5)=e^{-\frac{5}{2}}\left(e^{\frac{5}{4}}-1\right)=0.20442$ [1]

[Total 5]

Method 3

Another alternative is to use a 3-state model, where state 0 = never broken down, state 1 = broken down once, state 2+ = broken down more than once. Then it's easy to write down and solve the FDE:

$$\frac{\partial}{\partial t}p_{02+}(t)=p_{00}(t)\mu_{02}+p_{01}(t)\mu_{12}+p_{02+}(t)\mu_{2+2+}$$

$$=\frac{1}{2}p_{01}(t)=\frac{1}{2}\big(1-p_{00}(t)-p_{02+}(t)\big)$$

$$\Rightarrow \quad \frac{\partial}{\partial t}p_{02+}(t)+\frac{1}{2}p_{02+}(t)=\frac{1}{2}\left(1-e^{-t/4}\right)$$

This can then be solved using the integrating factor method.

[Total 5]

(ii) ***Expected lifetime of a boiler***

For $i = 0, 1, ..., 9$, the expected holding time in state i is $\dfrac{1}{\lambda_i}$, and state i must be followed by state $i+1$. So the expected lifetime of a boiler is:

$$\sum_{i=0}^{9} \frac{1}{\lambda_i} = 4 + 2 + (8 \times 1) = 14 \text{ years} \qquad [1]$$

Solution X2.2

Comment

Reference: Chapter 5

(i) ***Likelihood function***

The likelihood function is:

$$L = e^{-904(\sigma+\mu)}\, e^{-112(\rho+\upsilon)}\, \sigma^{34}\, \rho^{26}\, \mu^{2}\, \nu^{7} \qquad [2]$$

Full marks should be awarded if a constant factor has been included.

(ii) ***Maximum likelihood estimate***

The log-likelihood function is:

$$\ln L = \ln C - 904(\sigma + \mu) - 112(\rho + \upsilon) + 34 \ln \sigma + 26 \ln \rho + 2 \ln \mu + 7 \ln \nu \qquad [½]$$

Differentiating with respect to ρ:

$$\frac{\partial \ln L}{\partial \rho} = -112 + \frac{26}{\rho} \qquad [1]$$

Setting this equal to 0 gives:

$$\hat{\rho} = \frac{26}{112} = 0.23214 \qquad [½]$$

[Total 2]

(iii) ***Confidence interval***

Let $\tilde{\rho}$ denote the maximum likelihood estimator of ρ. Since $\tilde{\rho}$ is asymptotically normally distributed, an approximate 95% confidence interval for ρ is:

$$\hat{\rho} \pm 1.96\sqrt{\text{var}\left(\tilde{\rho}\right)}$$ [1]

Asymptotically, the variance of the estimator is given by:

$$\text{var}\left(\tilde{\rho}\right) = \frac{-1}{E\left(\dfrac{\partial^2 \ln L}{\partial \rho^2}\right)}$$

This is estimated by:

$$\frac{-1}{\left.\dfrac{\partial^2 \ln L}{\partial \rho^2}\right|_{\rho=\hat{\rho}}} = \frac{\hat{\rho}^2}{26} = \frac{\left(\frac{26}{112}\right)^2}{26} = 0.0020727$$ [1]

Hence an approximate 95% confidence interval for ρ is:

$$0.23214 \pm 1.96\sqrt{0.0020727} = \left(0.14291,\ 0.32138\right)$$ [1]

[Total 3]

Solution X2.3

Comment

Reference: Chapter 6

(i) ***Derivation of differential equation***

Consider the interval from age x to age $x+t+h$. Using the Markov assumption… [½]

… we can write:

$$(*) \qquad {}_{t+h}p_x^{21} = {}_t p_x^{21}\,{}_h p_{x+t}^{11} + {}_t p_x^{22}\,{}_h p_{x+t}^{21} + {}_t p_x^{23}\,{}_h p_{x+t}^{31} \qquad [1]$$

Recall that the Markov assumption states that the probabilities that a life at any given age will be found in any given state at any subsequent age depend only on the ages involved and the state currently occupied.

We now assume that for any two distinct states i and j, and $t \geq 0$:

$${}_h p_{x+t}^{ij} = h\mu_{x+t}^{ij} + o(h)$$

and the probability that a life makes two or more transitions in a short time interval of length h is $o(h)$. [1]

So we can write:

$${}_h p_{x+t}^{21} = h\mu_{x+t}^{21} + o(h)$$

$${}_h p_{x+t}^{31} = h\mu_{x+t}^{31} + o(h)$$

and:

$${}_h p_{x+t}^{11} = 1 - {}_h p_{x+t}^{12} - {}_h p_{x+t}^{13} - {}_h p_{x+t}^{14} = 1 - h\mu_{x+t}^{12} - h\mu_{x+t}^{14} + o(h) \qquad [1]$$

Since the probability of more than one transition in a short time interval of length h is $o(h)$, the ${}_h p_{x+t}^{13}$ term is included in the $o(h)$ term in the equation above.

Substituting these expressions into (*) gives:

$${}_{t+h}p_x^{21} = {}_tp_x^{21}\left(1 - h\mu_{x+t}^{12} - h\mu_{x+t}^{14}\right) + {}_tp_x^{22}h\mu_{x+t}^{21} + {}_tp_x^{23}h\mu_{x+t}^{31} + o(h) \qquad [½]$$

We can rearrange this to get:

$$\frac{{}_{t+h}p_x^{21} - {}_tp_x^{21}}{h} = {}_tp_x^{22}\mu_{x+t}^{21} + {}_tp_x^{23}\mu_{x+t}^{31} - {}_tp_x^{21}\left(\mu_{x+t}^{12} + \mu_{x+t}^{14}\right) + \frac{o(h)}{h} \qquad [½]$$

Finally, letting $h \to 0$, we obtain the result:

$$\frac{\partial}{\partial t}{}_tp_x^{21} = {}_tp_x^{22}\mu_{x+t}^{21} + {}_tp_x^{23}\mu_{x+t}^{31} - {}_tp_x^{21}\left(\mu_{x+t}^{12} + \mu_{x+t}^{14}\right) \qquad [½]$$

since $\lim\limits_{h\to 0}\frac{o(h)}{h} = 0$.

[Total 5]

(ii) ***Other forward equations***

The corresponding differential equations for ${}_tp_x^{23}$ and ${}_tp_x^{32}$ are:

$$\frac{\partial}{\partial t}{}_tp_x^{23} = {}_tp_x^{22}\mu_{x+t}^{23} - {}_tp_x^{23}\left(\mu_{x+t}^{31} + \mu_{x+t}^{34}\right) \qquad [1]$$

and:

$$\frac{\partial}{\partial t}{}_tp_x^{32} = {}_tp_x^{31}\mu_{x+t}^{12} - {}_tp_x^{32}\left(\mu_{x+t}^{21} + \mu_{x+t}^{23} + \mu_{x+t}^{24}\right) \qquad [2]$$

[Total 3]

These differential equations follow the usual pattern. For example, in the first one, we are thinking about going from State 2 to State 3, and we can construct the RHS of the equation as follows:

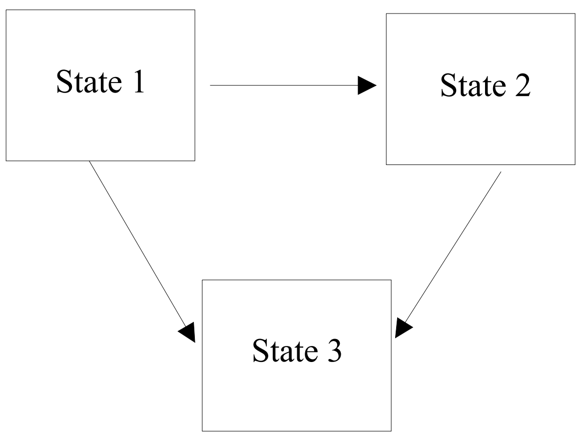
- *Imagine that the life is in State 2 at time 0 (age x).*
- *If the life is in State 2 at time t (the probability of this is ${}_tp_x^{22}$), then to get into State 3 at age $x+t$, he must instantaneously go from State 2 to State 3 at age $x+t$. (The "probability" of this is μ_{x+t}^{23}.)*

- *If the life is in State 3 at time t (probability ${}_tp_x^{23}$), then he must stay there. We need the $-\mu_{x+t}^{31}$ term to ensure that he doesn't move to State 1 at age $x+t$, and the $-\mu_{x+t}^{34}$ to ensure that he doesn't move to State 4.*
- *We don't need a term containing ${}_tp_x^{21}$ since going from State 1 to State 3 requires two transitions, and we are assuming that we can have only one transition in any given instant.*
- *We don't need a term containing ${}_tp_x^{24}$ either since it is impossible to go from State 4 to State 3.*

Solution X2.4

Comment
Reference: Chapters 5 and 6

A picture may help you to see what is going on here. Let's call the states 1, 2 and 3. The transition diagram is as follows:



We need to find the matrix of transition probabilities, $P(t)$, and then calculate:

$$\left(\tfrac{1}{3},\tfrac{1}{3},\tfrac{1}{3}\right)P(1)$$

State 3 in the diagram above resembles the dead state, since once you enter state 3 you cannot leave it.

We have:

$$P_{31}(t) = P_{32}(t) = 0 \text{ and } P_{33}(t) = 1 \qquad [1]$$

Also $P_{21}(t) = 0$. [½]

Since the only path from 2 to 2 is to stay there throughout:

$$P_{22}(t) = P_{\overline{22}}(t) = e^{-\lambda_2 t} = e^{-2t} \qquad [1]$$

It follows that $P_{23}(t) = 1 - e^{-2t}$. [½]

Similarly $P_{11}(t) = P_{\overline{11}}(t) = e^{-3t}$. [1]

To calculate $P_{12}(t)$ we can use the integral form of the Kolmogorov equation. (This is slightly quicker to deal with than the differential form.) If we use the backward form we have:

$$\begin{aligned} P_{12}(t) &= \int_0^t P_{\overline{11}}(w)\mu_{12}P_{22}(t-w)\,dw = \int_0^t e^{-3w} \times 2 \times e^{-2(t-w)}dw \\ &= 2e^{-2t}\left[\frac{e^{-w}}{-1}\right]_0^t = 2e^{-2t}\left(1-e^{-t}\right) \end{aligned} \qquad [2]$$

It follows that:

$$P_{13}(t) = 1 - e^{-3t} - 2e^{-2t}\left(1-e^{-t}\right) = 1 + e^{-3t} - 2e^{-2t} \qquad [1]$$

Finally:

$$\left(\tfrac{1}{3},\tfrac{1}{3},\tfrac{1}{3}\right)P(1) = \left(\tfrac{1}{3},\tfrac{1}{3},\tfrac{1}{3}\right)\begin{pmatrix} e^{-3} & 2e^{-2}\left(1-e^{-1}\right) & 1+e^{-3}-2e^{-2} \\ 0 & e^{-2} & 1-e^{-2} \\ 0 & 0 & 1 \end{pmatrix}$$

$$= (0.0166, 0.1021, 0.8813) \qquad [2]$$

[Total 9]

Solution X2.5

Comment
Reference: Chapter 5

(i) ***Transition rates***

The transition rates are:

$$\mu_{ij} = \begin{cases} 0.83 & \text{for } j = i+1 \\ -0.83 & \text{for } j = i \\ 0 & \text{otherwise} \end{cases}$$ [1]

(ii) ***Distribution of X_t***

X_t has a Poisson distribution with parameter $0.83t$. [1]

(iii) ***Probabilities***

The first probability given is:

$$\begin{aligned} P[X_3 \le 2] &= P[X_3 = 0] + P[X_3 = 1] + P[X_3 = 2] \\ &= \frac{e^{-3\times 0.83}(3\times 0.83)^0}{0!} + \frac{e^{-3\times 0.83}(3\times 0.83)^1}{1!} + \frac{e^{-3\times 0.83}(3\times 0.83)^2}{2!} \\ &= 0.54638 \end{aligned}$$ [2]

Since Poisson process is discrete and has stationary independent increments, we have $P[X_5 - X_2 < 3] = P[X_3 \le 2]$. [1]

[Total 3]

(iv) ***First holding time***

The first holding time is the time between the start and the first jump. This time is exponentially distributed with parameter 0.83. [1]

(v) ***Proof***

Since the holding times are exponentially distributed, they have the lack of memory property.

$$\begin{aligned}P\left(T_0 > s+t \mid T_0 > s\right) &= P\left(X_{s+t} = 0 \mid X_s = 0\right)\\ &= P\left(X_{s+t} - X_s = 0 \mid X_s = 0\right)\\ &= P\left(X_t - X_0 = 0 \mid X_0 = 0\right)\\ &= P\left(X_t = 0\right)\\ &= P\left(T_0 > t\right)\end{aligned}$$ [2]

Alternatively:

$$\begin{aligned}P(T_0 > s+t \mid T_0 > s) &= \frac{P(T_0 > s+t \text{ and } T_0 > s)}{P(T_0 > s)}\\ &= \frac{P(T_0 > s+t)}{P(T_0 > s)}\\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}\\ &= e^{-\lambda t}\\ &= P(T_0 > t)\end{aligned}$$ [2]

(vi) ***Probability that the remaining time to the first jump is at least 2 hours***

It follows that the remaining time until the first jump is also exponentially distributed with parameter 0.83. So:

$$P\left[T_0 > 6.5 \mid T_0 > 4.5\right] = P\left[T_0 > 2\right] = e^{-2\times 0.83} = 0.190$$ [1]

Solution X2.6

Comment

Reference: Chapter 5

(i) ***Assumptions***

We are assuming that the process is Markov and that the transition rates are constant. [1]

The Markov property of the underlying jump chain can be tested using a chi square test based on triplets of successive observations. [1]

A chi-square test can also be used to test whether the waiting times are exponentially distributed with constant parameter. [1]

[Total 3]

(ii) ***Transition graph***

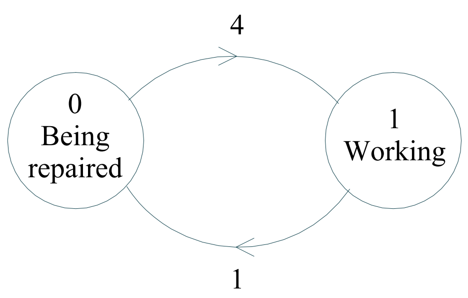
The expected holding time in the broken down state (state 0) is 0.25 days and the expected holding time in the working state (state 1) is 1 day. Since the holding time in state i is $Exp(\lambda_i)$, we have:

$$E(T_0) = \frac{1}{\lambda_0} = 0.25 \Rightarrow \lambda_0 = 4$$

and:

$$E(T_1) = \frac{1}{\lambda_1} = 1 \Rightarrow \lambda_1 = 1 \qquad [1]$$

The transition diagram is as follows:



[1]

[Total 2]

(iii) ***Kolmogorov's differential equations***

The backward differential equation is:

$$\frac{d}{dt}P_{0,0}(t) = -4P_{0,0}(t) + 4P_{1,0}(t)$$ [1]

and the forward differential equation is:

$$\frac{d}{dt}P_{0,0}(t) = P_{0,0}(t)\times(-4) + P_{0,1}(t)\times 1 = -4P_{0,0}(t) + P_{0,1}(t)$$ [1]

[Total 2]

(iv) ***Proof***

Since $P_{0,1}(t) = 1 - P_{0,0}(t)$, we have the differential equation:

$$\frac{d}{dt}P_{0,0}(t) = 1 - 5P_{0,0}(t)$$

together with the boundary condition $P_{0,0}(0) = 1$. [1]

We can solve this equation and boundary condition using an integrating factor of e^{5t} :

$$\frac{d}{dt}P_{0,0}(t)e^{5t} + 5P_{0,0}(t)e^{5t} = e^{5t}$$

$$\Leftrightarrow \quad \frac{d}{dt}\left(P_{0,0}(t)e^{5t}\right) = e^{5t}$$

$$\Rightarrow \quad P_{0,0}(t)e^{5t} = \frac{1}{5}e^{5t} + K$$

$$\Rightarrow \quad P_{0,0}(t) = \frac{1}{5} + Ke^{-5t}$$

Applying the boundary condition $P_{0,0}(0) = 1 \Rightarrow K = \frac{4}{5}$. So we have the required result.

[2]
[Total 3]

Solution X2.7

Comment

Reference: Chapter 6

(i) ***Chapman-Kolmogorov equations***

For any u such that $s < u < t$, we have:

$$P_{ij}(s,t) = \sum_k P_{ik}(s,u)\,P_{kj}(u,t) \qquad [1]$$

(ii) ***Transition rates***

The transition rates for the process are:

$$\mu_{ij}(s) = \left[\frac{\partial}{\partial t}P_{ij}(s,t)\right]_{t=s} = \lim_{h\to 0}\frac{P_{ij}(s,s+h)-P_{ij}(s,s)}{h} = \lim_{h\to 0}\frac{P_{ij}(s,s+h)-\delta_{ij}}{h} \qquad [2]$$

Marks should be awarded for any of these equivalent forms.

(iii) ***Derivation of Kolmogorov forward equations***

The Chapman-Kolmogorov equations can be written in the form:

$$P_{ij}(s,t+h) = P_{ij}(s,t)\,P_{jj}(t,t+h) + \sum_{k\neq j} P_{ik}(s,t)\,P_{kj}(t,t+h)$$

Now we can apply the definition of transition rates. We know that for a small time interval h, we have:

$$P_{jj}(t,t+h) = 1 + h\mu_{jj}(t) + o(h) \qquad [1]$$

where the $o(h)$ term accounts for the fact that there is a very small chance of more than one transition during the period $(t,t+h)$. Similarly:

$$P_{kj}(t,t+h) = h\mu_{kj}(t) + o(h) \qquad [1]$$

for $k \neq j$.

Therefore:

$$P_{ij}(s,t+h) = P_{ij}(s,t)\left(1+h\mu_{jj}(t)\right) + \sum_{k\neq j} P_{ik}(s,t)\,h\mu_{kj}(t) + o(h)$$ [1]

This gives:

$$\frac{P_{ij}(s,t+h)-P_{ij}(s,t)}{h} = \frac{\sum_{k} P_{ik}(s,t)\,h\mu_{kj}(t) + o(h)}{h}$$ [1]

Taking limits $h \to 0$ then gives the desired result. [1]

[Total 5]

(iv) ***Verification of solution***

Substituting $P(s,t) = e^{(t-s)A}$ into the equation from (iii) gives:

$$\frac{\partial}{\partial t}P(s,t) = \frac{\partial}{\partial t}\sum_{k=0}^{\infty} \frac{(t-s)^k}{k!}A^k = \sum_{k=1}^{\infty} \frac{(t-s)^{k-1}}{(k-1)!}A^{k-1}A = P(s,t)\,A$$

as required. [2]

(v) ***Transition probability matrix***

Method 1

We have $A = \begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix}$. Therefore:

$$A^2 = \begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix}\begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix} = \begin{pmatrix} 2\lambda^2 & -2\lambda^2 \\ -2\lambda^2 & 2\lambda^2 \end{pmatrix} = -2\lambda\begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix} = -2\lambda A$$

Repeating this process:

$$A^n = (-2\lambda)^{n-1}A$$ [1]

It follows that:

$$P(s,t) = e^{(t-s)A} = \sum_{k=0}^{\infty} \frac{(t-s)^k}{k!} A^k$$

$$= I + \sum_{k=1}^{\infty} \frac{(t-s)^k}{k!} (-2\lambda)^{k-1} A$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} e^{-2\lambda(t-s)} - 1 & -e^{-2\lambda(t-s)} + 1 \\ -e^{-2\lambda(t-s)} + 1 & e^{-2\lambda(t-s)} - 1 \end{pmatrix}$$

$$= \begin{pmatrix} ½\left(1+e^{-2\lambda(t-s)}\right) & ½\left(1-e^{-2\lambda(t-s)}\right) \\ ½\left(1-e^{-2\lambda(t-s)}\right) & ½\left(1+e^{-2\lambda(t-s)}\right) \end{pmatrix}$$ [2]

[Total 3]

Method 2

As an alternative we could think about the Kolmogorov differential equations:

$$\frac{d}{dt} P(t) = P(t) A$$

In this case, we have:

$$\frac{d}{dt} \begin{pmatrix} p_{11}(t) & p_{12}(t) \\ p_{21}(t) & p_{22}(t) \end{pmatrix} = \begin{pmatrix} p_{11}(t) & p_{12}(t) \\ p_{21}(t) & p_{22}(t) \end{pmatrix} \begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix}$$

So:

$$\frac{d}{dt} p_{11}(t) = -\lambda\, p_{11}(t) + \lambda p_{12}(t)$$

$$= -\lambda\, p_{11}(t) + \lambda\left(1 - p_{11}(t)\right)$$

$$= \lambda - 2\lambda\, p_{11}(t)$$

Rearranging:

$$\frac{d}{dt} p_{11}(t) + 2\lambda\, p_{11}(t) = \lambda$$

We can solve this using the integrating factor method. The integrating factor is $e^{2\lambda t}$. Multiplying though by this, we get:

$$e^{2\lambda t}\frac{d}{dt}p_{11}(t)+2\lambda e^{2\lambda t}\,p_{11}(t)=\lambda e^{2\lambda t}$$

and integrating both sides gives:

$$e^{2\lambda t}\,p_{11}(t)=\frac{1}{2}e^{2\lambda t}+C$$

where C is a constant of integration. Now, setting $t=0$:

$$1=\frac{1}{2}+C\Rightarrow C=\frac{1}{2}$$

So:

$$e^{2\lambda t}\,p_{11}(t)=\frac{1}{2}\left(e^{2\lambda t}+1\right)$$

and hence:

$$p_{11}(t)=\frac{1}{2}\left(1+e^{-2\lambda t}\right)$$

Since $p_{11}(t)+p_{12}(t)=1$, we know that:

$$p_{12}(t)=1-\frac{1}{2}\left(1+e^{-2\lambda t}\right)=\frac{1}{2}\left(1-e^{-2\lambda t}\right)$$

By symmetry:

$$p_{21}(t)=\frac{1}{2}\left(1-e^{-2\lambda t}\right)$$

$$p_{22}(t)=\frac{1}{2}\left(1+e^{-2\lambda t}\right)$$

So:

$$P(t) = \begin{pmatrix} \frac{1}{2}\left(1+e^{-2\lambda t}\right) & \frac{1}{2}\left(1-e^{-2\lambda t}\right) \\ \frac{1}{2}\left(1-e^{-2\lambda t}\right) & \frac{1}{2}\left(1+e^{-2\lambda t}\right) \end{pmatrix}$$

and:

$$P(s,t) = P(t-s) = \begin{pmatrix} \frac{1}{2}\left(1+e^{-2\lambda(t-s)}\right) & \frac{1}{2}\left(1-e^{-2\lambda(t-s)}\right) \\ \frac{1}{2}\left(1-e^{-2\lambda(t-s)}\right) & \frac{1}{2}\left(1+e^{-2\lambda(t-s)}\right) \end{pmatrix}$$

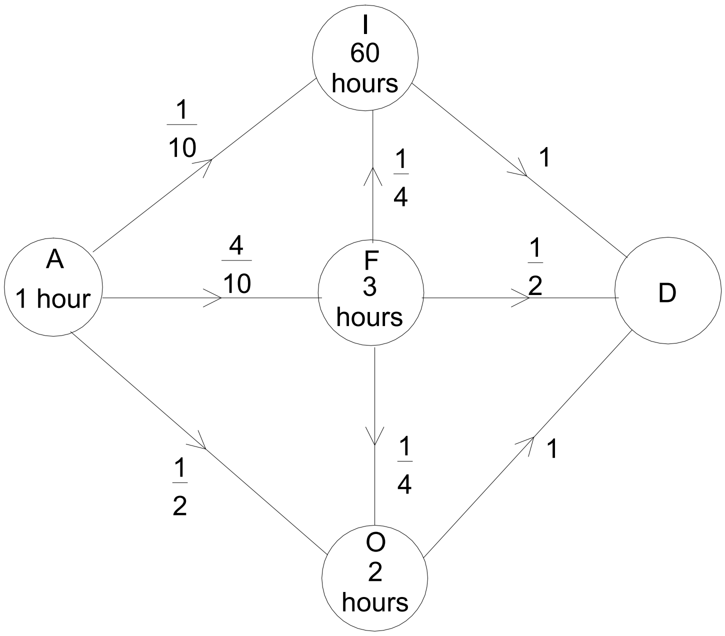
[Total 3]

Solution X2.8

Comment
Reference: Chapter 5

(i) ***Matrix of transition rates***

The situation is represented by the following transition diagram. The numerical values shown are the transition probabilities and the average waiting times.



The transition rates are then found by multiplying the probability of taking each route by the rate for that route, in other words, dividing by the average waiting time. This leads to the following matrix:

$$\begin{array}{c} \\ A \\ F \\ I \\ O \\ D \end{array} \begin{array}{c} \begin{array}{ccccc} A & F & I & O & D \end{array} \\ \begin{bmatrix} -1 & 0.4 & 0.1 & 0.5 & 0 \\ 0 & -\frac{1}{3} & \frac{0.25}{3} & \frac{0.25}{3} & \frac{0.5}{3} \\ 0 & 0 & -\frac{1}{60} & 0 & \frac{1}{60} \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

[2]

(ii) ***Proportion receiving in-patient treatment***

Patients will end up in state I either by going directly from state A or by going via state F. So the probability of eventually reaching state I is:

$$0.1 + 0.4 \times 0.25 = 0.2 \qquad [1]$$

(iii)(a) ***Probability that patient is yet to be classified by a junior doctor***

The forward differential equation for $p_{AA}(t)$, the probability of still being in state A at time t is:

$$\frac{d}{dt} p_{AA}(t) = -p_{AA}(t) \quad \Rightarrow \frac{d}{dt} \log p_{AA}(t) = -1 \quad \Rightarrow p_{AA}(t) = e^{-1\times t} = e^{-t} \qquad [1]$$

(iii)(b) ***Probability that patient is undergoing further investigation***

Here, the forward differential equation is:

$$\frac{d}{dt} p_{AF}(t) = 0.4 p_{AA}(t) - \tfrac{1}{3} p_{AF}(t) \qquad [1]$$

Using the formula derived for $p_{AA}(t)$, and rearranging:

$$\frac{d}{dt} p_{AF}(t) + \tfrac{1}{3} p_{AF}(t) = 0.4e^{-t}$$

Multiplying through by the integrating factor $e^{t/3}$:

$$e^{t/3} \frac{d}{dt} p_{AF}(t) + \tfrac{1}{3} e^{t/3} p_{AF}(t) = 0.4e^{-2t/3}$$

$$\frac{d}{dt}[e^{t/3} p_{AF}(t)] = 0.4e^{-2t/3} \quad \Rightarrow e^{t/3} p_{AF}(t) = -0.6e^{-2t/3} + c \qquad [1]$$

When $t = 0$, $p_{AF}(t) = 0$. So $c = 0.6$, and $p_{AF}(t) = 0.6(e^{-t/3} - e^{-t})$. [1]

[Total 4]

Alternatively, you could use an intuitive approach:

A patient who is in state F at time t must have previously entered state A and then remained in state A until some earlier time s, say, then made the transition to state F (probability 0.4), then remained in state F for the remaining period $t-s$. So, we need to integrate the probability of this sequence of events over the possible times s.

This gives:

$$\int_0^t e^{-1\times s} \times 0.4 \times e^{-\frac{1}{3}(t-s)} ds = 0.4e^{-\frac{1}{3}t} \int_0^t e^{-\frac{2}{3}s} ds$$
$$= 0.4e^{-\frac{1}{3}t} \times \tfrac{3}{2}(1-e^{-\frac{2}{3}t}) = 0.6(e^{-\frac{1}{3}t} - e^{-t})$$

(iv)(a) ***Explanation of equation***

In order for a patient in a typical state i to be discharged, they must first make a transition out of state i. The average time taken to do this is $\dfrac{1}{\lambda_i}$ where λ_i is the sum of the transition rates out of state i.

This transition must take them to a state other than i (which is therefore excluded from the sum). The probability that the transition takes them to a particular state j equals $\dfrac{\mu_{ij}}{\lambda_i}$. If this state is not state D (which can therefore also be excluded from the sum), we must add on the expected time taken to reach state D from state j, weighted by the corresponding probability. [2]

(iv)(b) ***Expected time until discharge***

We know that $m_I = 60$ and $m_O = 2$.

The equation in part (iv)(a), applied to state F, then tells us that:

$$m_F = 3 + 0.25m_I + 0.25m_O = 3 + 0.25\times 60 + 0.25\times 2 = 18.5$$

When applied to state A, it tells us that:

$$m_A = 1 + 0.4m_F + 0.1m_I + 0.5m_O = 1 + 0.4\times 18.5 + 0.1\times 60 + 0.5\times 2 = 15.4 \quad [2]$$

[Total 4]

Unfortunately, since A is the first state, we need to work out the m's for all the other states along the way!

(v) ***Distribution of the time spent in each state visited***

The distribution of the time spent in a typical state i is exponential with mean $\dfrac{1}{\lambda_i}$. [1]

You might be wondering why we've been asked this, since we've already used this result in part (iii). It's actually a hint to help you with the next part.

(vi) ***Checking the model***

The key assumption of the model is that the transition rates are constant over time. We could check whether this was a realistic assumption by carrying out a statistical test to see whether the observed waiting times in each state are consistent with an exponential distribution (with constant mean). [1]

One way to do this would be to make use of the fact that the standard deviation of an exponential distribution is equal to the mean. So, if the observed standard deviations of the waiting times in each state do not appear to be consistent with the mean waiting times, we would have cause to question this assumption. [1]
[Total 2]

(vii)(a) ***Calculations needed to compare options***

The management committee will be interested in reducing the average time (m_A) taken to discharge a new patient (and also the average cost of treating a patient, which will be linked to this).

We could investigate the first option using the existing model, but replacing the average waiting times and the transition probabilities by values appropriate for a more senior doctor. We could then recalculate m_A.

We could investigate the second option using the existing model, but reducing the average time for outpatients (currently 2 hours) to an appropriate value. We could then recalculate m_A.

The option with the lower value of m_A would be preferable. [2]

(vii)(b)***Is current model suitable?***

From a theoretical viewpoint, our model only looks at the average waiting times. We have not considered the variation in waiting times for individual patients, which should also be taken into account. Nor have we taken into account explicitly the financial information, such as the doctors' rates of pay and details of the cost of treatment for inpatients and outpatients (and any diagnostic tests required).

From a practical viewpoint, the "appropriate values" referred to in part (a) might be difficult to estimate without actually trying out each of the options during a trial period.

[2]

[Total 4]

Assignment X3 – Solutions

Markers: This document sets out one approach to solving each of the questions. Please give credit for other valid approaches.

Solution X3.1

Comment
Reference: Chapter 7

I is false. e_{x+1} and e_x are related by the equation: $e_x = p_x(1+e_{x+1})$. If we take, as an extreme example, $e_{x+1} = 49$ and $p_x = 0.5$, then $e_x = 0.5(1+49) = 25 < e_{x+1}$. [1]

II is false. If $t=1$ (say), then it is perfectly possible to have p_{x+1} greater than p_x, if mortality rates are decreasing over that age range. [1]

III is true. T_x must lie in the range $K_x \le T_x < K_x + 1$. The statement follows from the right hand inequality. [1]

[Total 3]

Solution X3.2

Comment
Reference: Chapter 7

(i) ***Complete expectation of life at age 45***

The complete expectation of life at age 45 is:

$$\mathring{e}_{45} = \int_0^{65} \left(1-\frac{t}{65}\right)^{½} dt \qquad [1]$$

Making the substitution $u = 1-\dfrac{t}{65}$, this becomes:

$$\mathring{e}_{45} = \int_0^1 65u^{½}\, du = \left[65\times\frac{2}{3}u^{3/2}\right]_0^1 = 43.33 \text{ years} \qquad [2]$$

[Total 3]

(ii) ***Force of mortality at age 45***

The force of mortality at age x is given by:

$$\mu_x = -\frac{S'(x)}{S(x)} = -\frac{d}{dx}\ln S(x)$$ [1]

But:

$$S(x) = {}_x p_0 = \left(1-\frac{x}{110}\right)^{½}$$ [½]

So:

$$\mu_x = -\frac{d}{dx}\left[½\ln\left(1-\frac{x}{110}\right)\right] = -\frac{1}{2}\times\frac{-\frac{1}{110}}{1-\frac{x}{110}} = \frac{1}{2(110-x)}$$ [1]

and:

$$\mu_{45} = \frac{1}{2(110-45)} = \frac{1}{130}$$ [½]

[Total 3]

Alternatively, we can evaluate:

$$\mu_{45} = \lim_{h\to 0}\frac{{}_h q_{45}}{h}$$

Now:

$${}_h p_{45} = \left(1-\frac{h}{110-45}\right)^{½} = \left(1-\frac{h}{65}\right)^{½}$$

So:

$$\lim_{h\to 0}\frac{{}_h q_{45}}{h} = \lim_{h\to 0}\frac{1-\left(1-\frac{h}{65}\right)^{½}}{h}$$

Because the denominator tends to 0 as h tends to 0, we could use L'Hopital's rule to evaluate this limit:

$$\lim_{h\to 0}\frac{1-\left(1-\frac{h}{65}\right)^{1/2}}{h}=\lim_{h\to 0}\frac{\frac{\partial}{\partial h}\left[1-\left(1-\frac{h}{65}\right)^{1/2}\right]}{\frac{\partial}{\partial h}h}$$

$$=\lim_{h\to 0}\frac{-\frac{1}{2}\left(1-\frac{h}{65}\right)^{-1/2}\times-\frac{1}{65}}{1}$$

$$=\lim_{h\to 0}\frac{1}{130\sqrt{1-\frac{h}{65}}}=\frac{1}{130}$$

or we could use the binomial expansion:

$$\left(1-\frac{h}{65}\right)^{1/2}=1+\tfrac{1}{2}\left(-\frac{h}{65}\right)+\frac{\tfrac{1}{2}\left(-\tfrac{1}{2}\right)}{2!}\left(-\frac{h}{65}\right)^2+\cdots$$

which gives:

$$\lim_{h\to 0}\frac{1-\left(1-\frac{h}{65}\right)^{1/2}}{h}=\lim_{h\to 0}\frac{-\left[\tfrac{1}{2}\left(-\frac{h}{65}\right)+\frac{\tfrac{1}{2}\left(-\tfrac{1}{2}\right)}{2!}\left(-\frac{h}{65}\right)^2+\cdots\right]}{h}=\frac{1}{130}$$

Solution X3.3

Comment

Reference: Chapter 8

(i) ***Differences between random censoring and Type I censoring***

Both random censoring and Type I censoring are examples of right censoring. Right censoring occurs when a life exits the investigation for a reason other than death. [1]

With random censoring, the censoring times are not known in advance – they are not chosen by the investigator and are random variables. An example of random censoring in life insurance is the event of a policyholder choosing to surrender a policy. [1]

Type I censoring occurs when the censoring times are known in advance, *ie* the censoring times are chosen by the investigator. [1]

An example of Type I censoring is when observation ceases for all those still alive at the end of the period of investigation. [1]
[Total 4]

(ii) ***Non-informative censoring***

Censoring is non-informative if it gives no information about the future patterns of mortality by age for the censored lives. [1]

In the context of this investigation, non-informative censoring occurs if at any given time, lives are equally likely to be censored regardless of their subsequent force of mortality. This means that you cannot tell anything about a person's mortality after the date of the censoring event from the fact that they have been censored. [1]

In this investigation withdrawals might be informative, since lives that are in better health may be more likely to surrender their policies than those in a poor state of health. Lives that are censored are therefore likely to have lighter mortality than those that remain in the investigation. [1]
[Total 3]

Markers, please give credit for any suitable example.

Solution X3.4

Comment
Reference: Chapter 8

(i) ***Nelson- Aalen estimate of cumulative hazard***

We first have to work out the length of time for which each patient was observed. These figures are given in the table below.

Patient number	Length of observation period (months)	Reason for exit
1	11	Censored
2	6	Death
3	9	Censored
4	12	Censored
5	5	Censored
6	2	Death
7	12	Censored
8	8	Death
9	6	Death
10	7	Censored

[1]

So we have:

Death times: 2, 6 (two deaths) and 8

Censoring times: 5, 7, 9, 11, 12 (two lives)

No life is observed for more than 12 months.

The calculations for the cumulative hazard function are summarised in the following table:

t_j	n_j	d_j	$\hat{\lambda}_j = \dfrac{d_j}{n_j}$
2	10	1	0.1
6	8	2	0.25
8	5	1	0.2

[2]

The Nelson-Aalen estimate of the cumulative hazard function is then given by:

$$\hat{\Lambda}(t) = \sum_{t_j \le t} \hat{\lambda}_j = \begin{cases} 0 & \text{for } 0 \le t < 2 \\ 0.1 & \text{for } 2 \le t < 6 \\ 0.35 & \text{for } 6 \le t < 8 \\ 0.55 & \text{for } 8 \le t \le 12 \end{cases}$$ [1]

[Total 4]

(ii) ***Estimate of survival function***

The Nelson-Aalen estimate of the survival function is:

$$\hat{S}(t) = \exp\left(-\hat{\Lambda}(t)\right) = \begin{cases} 1 & \text{for } 0 \le t < 2 \\ 0.90484 & \text{for } 2 \le t < 6 \\ 0.70469 & \text{for } 6 \le t < 8 \\ 0.57695 & \text{for } 8 \le t \le 12 \end{cases}$$ [2]

(iii) ***Confidence interval for survival probability***

The variance of the Nelson-Aalen estimator of the integrated hazard function is given by:

$$\text{var}\left[\tilde{\Lambda}(t)\right] = \sum_{t_j \le t} \frac{d_j\left(n_j - d_j\right)}{n_j^3}$$

So:

$$\text{var}\left[\tilde{\Lambda}(10)\right] = \frac{1\times 9}{10^3} + \frac{2\times 6}{8^3} + \frac{1\times 4}{5^3} = 0.064438$$ [1]

An approximate 95% confidence interval for $\Lambda(10)$ is:

$$\begin{aligned}\hat{\Lambda}(10) \pm 1.96\sqrt{\text{var}\left[\tilde{\Lambda}(10)\right]} &= 0.55 \pm 1.96\sqrt{0.064438} \\ &= 0.55 \pm 0.49754 \\ &= \left(0.05246, 1.04754\right)\end{aligned}$$ [2]

So an approximate 95% confidence interval for $S(10)$ is:

$$\left(e^{-1.04754}, e^{-0.05246}\right) = \left(0.3508, 0.9489\right) \qquad [1]$$

[Total 4]

(iv) ***Comment***

As 0.9 lies within the confidence interval constructed in (iii), there is insufficient evidence to dispute, at the 2.5% significance level, the hypothesis that at least 90% of patients survive for 10 months or more after the operation. [2]

The significance level here is 2.5% because the 95% confidence interval in part (iii) has 2.5% in each tail. The hypothesis here refers to "10 months <u>or more</u>", which is one-sided.

Solution X3.5

Comment
Reference: Chapter 9

(i) ***Model of force of mortality***

The model for the force of mortality is:

$$\mu(t,\mathbf{Z}) = \mu_0(t)\exp\left(-0.20Z_1 + 0.12Z_2 - 0.05Z_3 - 0.06Z_4\right)$$

where:

$t =$ time since patient underwent procedure

$\mu_0(t) =$ baseline hazard at time t

$\mathbf{Z} = \left(Z_1, Z_2, Z_3, Z_4\right)$

$$Z_1 = \begin{cases} 1 & \text{if patient is female} \\ 0 & \text{if patient is male} \end{cases}$$

$$Z_2 = \begin{cases} 1 & \text{if patient received Treatment B} \\ 0 & \text{if patient did not receive Treatment B} \end{cases}$$

$$Z_3 = \begin{cases} 1 & \text{if patient received Treatment C} \\ 0 & \text{if patient did not receive Treatment C} \end{cases}$$

$$Z_4 = \begin{cases} 1 & \text{if patient attended Hospital B} \\ 0 & \text{if patient attended Hospital A} \end{cases}$$ [3]

Markers: please give credit for alternative correct solutions. Deduct ½ mark for each omission, subject to a minimum of 0.

(ii)(a) ***Semi-parametric model***

The model:

$$\mu(t,\mathbf{Z}) = \mu_0(t)\exp\left(-0.20Z_1 + 0.12Z_2 - 0.05Z_3 - 0.06Z_4\right)$$

as defined above is a semi-parametric model because half of it is parametric and half of it isn't. [½]

The parameters are present in the exponential term but not in the baseline hazard term. [½]

(ii)(b) ***Proportional hazards model***

The model is a proportional hazards model since the hazards of different lives with covariate vectors $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ are in the same proportion at all times, *ie* the ratio $\dfrac{\mu\left(t,\mathbf{Z}^{(1)}\right)}{\mu\left(t,\mathbf{Z}^{(2)}\right)}$ does not depend on t. [1]

[Total 2]

(iii)(a) ***Baseline hazard group***

The baseline hazard refers to the lives whose Z values are all 0, *ie* to male patients on Treatment A who attended Hospital A. [1]

(iii)(b) ***Group with lowest force of mortality***

Here we must make the power in the exponential as negative as possible.

The lives with the lowest force of mortality according to this model are those for which $Z_1 = 1$, $Z_2 = 0$, $Z_3 = 1$ and $Z_4 = 1$, *ie* female patients on Treatment C who attended Hospital B. [1]

[Total 2]

(iv) ***Comparison of Hospital A with Hospital B***

Suppose that β_4 is the parameter associated with covariate Z_4. We want to test the null hypothesis:

$H_0 : \beta_4 = 0$ (*ie* hospital is not significant)

against the alternative hypothesis:

$H_1 : \beta_4 < 0$ (*ie* Hospital B is better) [1]

The estimated value of β_4 is –0.06, and the standard error of the estimator is 0.04. So the value of our test statistic is:

$$TS = \frac{-0.06 - 0}{0.04} = -1.5$$ [1]

Comparing this with the lower 5% point of the standard normal distribution (–1.6449), we find that it does not fall into the rejection region. So there is insufficient evidence to conclude that attending Hospital B improves the chances of survival. [1]
[Total 3]

Alternatively, you could construct an approximate one-sided 95% confidence interval for β_4 and check whether or not 0 lies within the interval.

The interval is:

$$\left(-\infty, \hat{\beta}_4 + 1.645\, se\left(\tilde{\beta}_4\right)\right) = (-\infty, -0.06 + 1.645 \times 0.04) = (-\infty, 0.0058)$$

Since 0 lies in this interval, there is insufficient evidence to reject the null hypothesis (as stated above).

(v) ***Proportion***

According to the model, the force of mortality at time t since the procedure for a male patient on Treatment B who attended Hospital A is:

$$\mu(t,0,1,0,0) = \mu_0(t)\exp(0.12) \qquad [½]$$

Also, the force of mortality at time t since the procedure for a female patient on Treatment C who attended Hospital B is:

$$\mu(t,1,0,1,1) = \mu_0(t)\exp(-0.20-0.05-0.06) = \mu_0(t)\exp(-0.31) \qquad [½]$$

Dividing the first of these expressions by the second, we obtain:

$$\frac{\mu_0(t)\exp(0.12)}{\mu_0(t)\exp(-0.31)} = e^{0.43} = 1.5373$$

So the force of mortality for the male patient exceeds the force of mortality for the female patient by 53.73%. [1]

[Total 2]

Solution X3.6

Comment
Reference: Chapter 7

(i) ***Meaning of the rates of mortality and the relationship between them***

q_x is the initial rate of mortality. It represents the probability that a life, currently aged exactly x years, dies within the next year. [1]

m_x is the central rate of mortality. It is the rate at which deaths are occurring over the year of age x to $x+1$ exact, relative to the expected amount of time that a life, initially alive at age x, will spend alive in the that year of age. [1]

The rates are related by:

$$m_x = \frac{q_x}{\int_0^1 {}_tp_x \, dt} \qquad [1]$$

[Total 3]

Note that $\int_0^1 {}_tp_x \, dt$ *is the expected time spent alive in the year beginning at exact age* x.

(ii) ***Weighted average***

Since:

$$q_x = \int_0^1 {}_tp_x \, \mu_{x+t} \, dt$$

we can write:

$$m_x = \frac{\int_0^1 {}_tp_x \, \mu_{x+t} \, dt}{\int_0^1 {}_tp_x \, dt} \qquad [1]$$

This is a weighted average of μ_{x+t}, the force of mortality at each age between x and $x+1$, with the probability of survival, ${}_tp_x$, as weights. [1]

[Total 2]

(iii)(a) ***UDD assumption***

Under the UDD assumption:

$$\int_0^1 {}_tp_x\,dt = \int_0^1 (1-tq_x)\,dt = \left[t-\frac{1}{2}t^2q_x\right]_0^1 = 1-\frac{1}{2}q_x \qquad [1]$$

Since $q_x = 0.4$, we now have:

$$m_x = \frac{0.4}{1-0.2} = \frac{0.4}{0.8} = 0.5 \qquad [½]$$

(iii)(b) ***Constant force assumption***

Under the constant force of mortality assumption:

$$q_x = 1-e^{-\mu} \qquad [1]$$

and:

$$\int_0^1 {}_tp_x\,dt = \int_0^1 e^{-\mu t}\,dt = \left[-\frac{1}{\mu}e^{-\mu t}\right]_0^1 = \frac{1}{\mu}\left(1-e^{-\mu}\right) = \frac{q_x}{\mu} \qquad [1]$$

So:

$$m_x = \frac{q_x}{\frac{1}{\mu}q_x} = \mu = -\ln(1-q_x) = -\ln 0.6 = 0.51083 \qquad [½]$$

(iii)(c) ***Balducci assumption***

Under the Balducci assumption:

$${}_{1-t}q_{x+t} = (1-t)q_x$$

So:

$${}_tp_x = \frac{p_x}{{}_{1-t}p_{x+t}} = \frac{p_x}{1-{}_{1-t}q_{x+t}} = \frac{p_x}{1-(1-t)q_x} = \frac{p_x}{p_x+t\,q_x} \qquad [1]$$

Hence:

$$\int_0^1 {}_tp_x\,dt = \int_0^1 \frac{p_x}{p_x + tq_x}\,dt = \int_0^1 \frac{0.6}{0.6+0.4t}\,dt \qquad [½]$$

Making the substitution $u = 0.6 + 0.4t$, we have:

$$\int_0^1 {}_tp_x\,dt = \int_{0.6}^1 \frac{0.6}{u} \times \frac{1}{0.4}\,du = \int_{0.6}^1 \frac{1.5}{u}\,du = 1.5\ln 1 - 1.5\ln 0.6 = -1.5\ln 0.6 \qquad [1]$$

So:

$$m_x = \frac{0.4}{-1.5\ln 0.6} = 0.52203 \qquad [½]$$

[Total 7]

(iv) ***Comment***

Since m_x is a weighted average of μ_{x+t}, the calculated value of m_x will be affected by the way in which μ_{x+t} is assumed to vary over the age range. [½]

Assuming μ_{x+t} is constant gives an m_x value of 0.51083.

The UDD assumption produces a lower value, of 0.5. UDD implies that the force of mortality is rising over the year of age. (If we consider each month, then we have the same number of deaths each month, but the number of survivors at the beginning of each month decreases over the age range). [½]

As the weights, ${}_tp_x$, are higher at the start of the year and decrease over the year, the greatest weights will be applied to μ_{x+t} towards the start of the year, where mortality rates are lowest. This will tend to reduce the weighted average mortality rate compared to (b). [1]

The Balducci assumption yields a value of 0.52203 for m_x. This is higher than when a constant force of mortality is assumed. By comparison with (a), this must indicate that the Balducci assumption implies a reducing force of mortality over the year of age. [1]

[Total 3]

Solution X3.7

Comment

Reference: Chapter 9

(i) ***Model***

The general model is: $\lambda(t;z_i) = \lambda_0(t)\exp(\beta z_i^T)$

where:

$\lambda(t;z_i)$ is the hazard function (or force of mortality) at time t for the i th life

$\lambda_0(t)$ is the baseline hazard function at time t

z_i is a row vector with the values of the covariates for the i th life

β is a row vector of parameters [2]

Initially, there is just one covariate (smoker status), so z_i and β are just scalars. [1]

Our aim is to estimate the value of the parameter β based on the data given. We can estimate β by maximising the partial likelihood, which equals:

$$L(\beta) = \prod_{j=1}^{k} \frac{\exp\left(\beta z_j^T\right)}{\sum_{i \in R(t_j)} \exp\left(\beta z_i^T\right)}$$

where k is the number of deaths (assumed to occur at distinct times), t_j is the j th lifetime, and $R(t_j)$ denotes the set of lives at risk just before time t_j. [1]

[Total 4]

(ii) ***Partial likelihood***

From the data given, we see that Patient 1 (a smoker) dies first, and at time 3. Since there were 2 smokers and 4 non-smokers in the at-risk group just before time 3, the contribution to the partial likelihood from the first death is:

$$\frac{e^{\beta}}{2e^{\beta}+4}$$ [1]

The second life to die is Patient 6 (a non-smoker). Just before this death, there were 1 smoker and 4 non-smokers at risk. So the contribution to the partial likelihood from the second death is:

$$\frac{1}{e^{\beta}+4}$$ [1]

The third life to die is Patient 3 (a non-smoker). Just prior to this death there are 1 smoker and 2 non-smokers at risk. Note that Patient 5 is censored at time 8, so is no longer part of the at-risk group. The contribution to the partial likelihood from this death is therefore:

$$\frac{1}{e^{\beta}+2}$$ [1]

So the partial likelihood is:

$$L(\beta)=\frac{e^{\beta}}{2e^{\beta}+4}\times\frac{1}{e^{\beta}+4}\times\frac{1}{e^{\beta}+2}=\frac{e^{\beta}}{2\left(e^{\beta}+2\right)^{2}\left(e^{\beta}+4\right)}$$ [1]

[Total 4]

(iii) ***Parameter estimate***

The partial log-likelihood is:

$$\ln L=\beta-\ln 2-2\ln\left(e^{\beta}+2\right)-\ln\left(e^{\beta}+4\right)$$ [½]

Differentiating this with respect to β:

$$\frac{d\ln L}{d\beta}=1-\frac{2e^{\beta}}{e^{\beta}+2}-\frac{e^{\beta}}{e^{\beta}+4}$$ [1]

Setting this equal to 0 and rearranging gives:

$$
\begin{aligned}
&\left(e^{\beta}+2\right)\left(e^{\beta}+4\right)=2e^{\beta}\left(e^{\beta}+4\right)+e^{\beta}\left(e^{\beta}+2\right) \\
&\Rightarrow e^{2\beta}+6e^{\beta}+8=3e^{2\beta}+10e^{\beta} \\
&\Rightarrow 2e^{2\beta}+4e^{\beta}-8=0 \\
&\Rightarrow e^{2\beta}+2e^{\beta}-4=0 \\
&\Rightarrow e^{\beta}=\frac{-2\pm\sqrt{4+(4\times 4)}}{2}=\frac{-2\pm\sqrt{20}}{2}=1.23607
\end{aligned}
$$

So:

$$\beta=\ln 1.23607=0.21194 \qquad [1\tfrac{1}{2}]$$

Differentiating the partial log-likelihood a second time gives:

$$
\begin{aligned}
\frac{d^2\ln L}{d\beta^2}&=-2\left[\frac{e^{\beta}\left(e^{\beta}+2\right)-e^{\beta}e^{\beta}}{\left(e^{\beta}+2\right)^2}\right]-\left[\frac{e^{\beta}\left(e^{\beta}+4\right)-e^{\beta}e^{\beta}}{\left(e^{\beta}+4\right)^2}\right] \\
&=-\frac{4e^{\beta}}{\left(e^{\beta}+2\right)^2}-\frac{4e^{\beta}}{\left(e^{\beta}+4\right)^2}<0\Rightarrow \max \qquad [1]
\end{aligned}
$$

So the maximum partial likelihood estimate of the model parameter β is $\hat{\beta}=0.21194$.

[Total 4]

(iv) ***Extended partial likelihood***

Since we have ties, we need to use Breslow's approximation:

$$L(\beta)\approx\prod_{j=1}^{k}\frac{\exp(\beta s_j^T)}{\left(\sum_{i\in R(t_j)}\exp(\beta z_i^T)\right)^{d_j}}$$

where s_j is the sum of the covariates of the lives that die at time t_j.

There were 3 smokers and 5 non-smokers at risk just before time 3. The two patients that died at time 3 were both smokers. So the contribution to the partial likelihood from these deaths is:

$$\frac{e^{2\beta}}{\left(3e^{\beta}+5\right)^2}$$ [1½]

The next death (Patient 6 at time 7) contributes $\dfrac{1}{e^{\beta}+5}$ and the last death (Patient 3 at time 9) contributes $\dfrac{1}{e^{\beta}+3}$. [1]

So the partial likelihood is now:

$$L=\frac{e^{2\beta}}{\left(3e^{\beta}+5\right)^2\left(e^{\beta}+5\right)\left(e^{\beta}+3\right)}$$ [½]

[Total 3]

(v) ***Adapted model***

The model would be of similar form, but with two covariates (smoking status and sex). Specifically:

$$\lambda(t;z_i)=\lambda_0(t)\exp(\beta_1\times smoker\,status\,code+\beta_2\times gender\,code)$$ [1]

Gender could be coded as 1 for males and 0 for females (say). [1]

With this coding, the baseline hazard refers to female non-smokers. [1]
[Total 3]

Note to markers: if the gender is coded as 0 for males and 1 for females, then baseline hazard refers to male non-smokers. Give full credit for this.

Solution X3.8

Comment

Reference: Chapter 8

(i)(a) ***Estimates of discrete hazard function***

First re-order the observations to: 7, 8, 9, 10, 11*,12, 13*, 14, 14, alive.

j	t_j	d_j	n_j	$\hat{\lambda}_j$
1	7	1	10	$1/10$
2	8	1	9	$1/9$
3	9	1	8	$1/8$
4	10	1	7	$1/7$
5	12	1	5	$1/5$
6	14	2	3	$2/3$

[3]

So the discrete hazard function is estimated by:

$$\hat{\lambda}(7) = \frac{1}{10}$$

$$\hat{\lambda}(8) = \frac{1}{9}$$

$$\hat{\lambda}(9) = \frac{1}{8}$$

$$\hat{\lambda}(10) = \frac{1}{7}$$

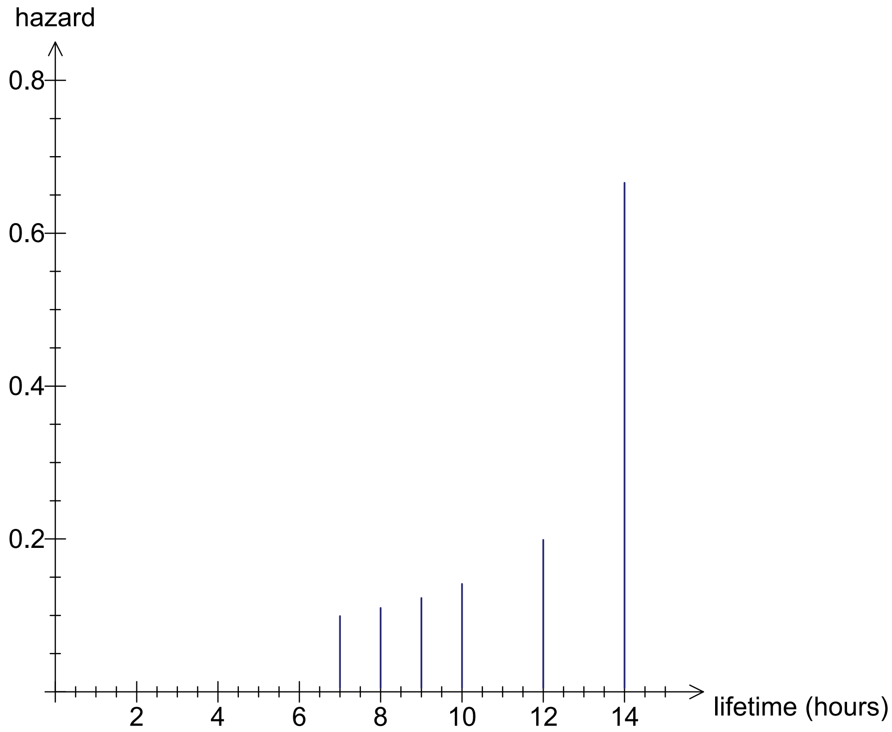
$$\hat{\lambda}(12) = \frac{1}{5}$$

$$\hat{\lambda}(14) = \frac{2}{3}$$

$\hat{\lambda}(t) = 0$ otherwise [1]

(i)(b) ***Sketch***

A sketch of the estimate of the discrete hazard function is given below:



[3]
[Total 7]

(ii) ***Estimate of distribution function***

Since $\hat{F}(t) = 1 - \prod_{t_j \le t} \left(1 - \hat{\lambda}_j\right)$ and:

j	t_j	$1-\hat{\lambda}_j$	$\hat{F}(t)$
1	7	$1 - \frac{1}{10} = \frac{9}{10}$	$1 - \frac{9}{10} = \frac{1}{10} = 0.1$
2	8	$1 - \frac{1}{9} = \frac{8}{9}$	$1 - \frac{9}{10} \times \frac{8}{9} = \frac{2}{10} = 0.2$
3	9	$1 - \frac{1}{8} = \frac{7}{8}$	$1 - \frac{9}{10} \times \frac{8}{9} \times \frac{7}{8} = \frac{3}{10} = 0.3$
4	10	$1 - \frac{1}{7} = \frac{6}{7}$	$1 - \frac{9}{10} \times \frac{8}{9} \times \frac{7}{8} \times \frac{6}{7} = \frac{4}{10} = 0.4$
5	12	$1 - \frac{1}{5} = \frac{4}{5}$	$1 - \frac{9}{10} \times \frac{8}{9} \times \frac{7}{8} \times \frac{6}{7} \times \frac{4}{5} = 1 - \frac{24}{50} = \frac{26}{50} = 0.52$
6	14	$1 - \frac{2}{3} = \frac{1}{3}$	$1 - \frac{9}{10} \times \frac{8}{9} \times \frac{7}{8} \times \frac{6}{7} \times \frac{4}{5} \times \frac{1}{3} = 1 - \frac{8}{50} = \frac{42}{50} = 0.84$

we have:

$$\hat{F}(t) = \begin{cases} 0 & \text{for } 0 \le t < 7 \\ 0.1 & \text{for } 7 \le t < 8 \\ 0.2 & \text{for } 8 \le t < 9 \\ 0.3 & \text{for } 9 \le t < 10 \\ 0.4 & \text{for } 10 \le t < 12 \\ 0.52 & \text{for } 12 \le t < 14 \\ 0.84 & \text{for } 14 \le t \le 16 \end{cases}$$

[3]

(iii) ***Probability of battery failing within 11 hours***

For the interval between time 10 and time 12, the probability is estimated to be 0.4. [1]

(iv) ***Variance of $\tilde{F}(t)$***

The variance is calculated using Greenwood's formula:

$$\operatorname{var}\left(\tilde{F}(t)\right) \approx \left[1 - \hat{F}(t)\right]^2 \sum_{t_j \le t} \frac{d_j}{n_j (n_j - d_j)}$$

Therefore we get:

j	t	d_j	n_j	$\hat{F}(t)$	$\dfrac{d_j}{n_j(n_j-d_j)}$	$\text{var}[\tilde{F}(t)]$
1	7	1	10	0.1	1/90	0.009
2	8	1	9	0.2	1/72	0.016
3	9	1	8	0.3	1/56	0.021
4	10	1	7	0.4	1/42	0.024
5	12	1	5	0.52	1/20	0.027
6	14	2	3	0.84	2/3	0.020

ie:

$$\text{var}[\tilde{F}(7)] \approx 0.009$$
$$\text{var}[\tilde{F}(8)] \approx 0.016$$
$$\text{var}[\tilde{F}(9)] \approx 0.021$$
$$\text{var}[\tilde{F}(10)] \approx 0.024$$
$$\text{var}[\tilde{F}(12)] \approx 0.027$$
$$\text{var}[\tilde{F}(14)] \approx 0.020$$

[5]

(v) ***Long-life battery***

Reordering the observations to: 7, 11, 12*, 13*, 14, 14, 15, alive, alive, alive.

j	t_j	d_j	n_j	$\hat{\lambda}_j$	$1-\hat{\lambda}_j$	$\hat{F}(t)$
1	7	1	10	$1/10$	$9/10$	0.1
2	11	1	9	$1/9$	$8/9$	0.2
3	14	2	6	$2/6$	$4/6$	0.467
4	15	1	4	$1/4$	$3/4$	0.6

So:

$$\hat{F}(t) = \begin{cases} 0 & \text{for } 0 \le t < 7 \\ 0.1 & \text{for } 7 \le t < 11 \\ 0.2 & \text{for } 11 \le t < 14 \\ 0.467 & \text{for } 14 \le t < 15 \\ 0.6 & \text{for } 15 \le t \le 16 \end{cases}$$

[3]

For the variance we get:

j	t	d_j	n_j	$\hat{F}(t)$	$\dfrac{d_j}{n_j(n_j-d_j)}$	$\text{var}[\tilde{F}(t)]$
1	7	1	10	0.100	1/90	0.009
2	11	1	9	0.200	1/72	0.016
3	14	2	6	0.467	2/24	0.031
4	15	1	4	0.600	1/12	0.031

ie:

$$\text{var}[\tilde{F}(7)] \approx 0.009$$
$$\text{var}[\tilde{F}(11)] \approx 0.016$$
$$\text{var}[\tilde{F}(14)] \approx 0.031$$
$$\text{var}[\tilde{F}(15)] \approx 0.031$$

[3]

[Total 6]

(vi) ***Comment***

The long-life battery appears to last longer. $F(t)$ for the standard battery at time 14 is 0.840 compared to 0.467 for the long-life battery. Looking at the variance, this difference looks significant. [1]

Based on the data, however, it is not easy to tell whether the long-life batteries last *three times* as long (or more). However, it may well be worth paying more for a long-life battery. [1]

The results for the long-life batteries are distorted by the battery that failed very early. Is this a fluke or not? [½]

Ideally we should investigate further the three batteries still running at the end. When did they fail? [½]

[Total 3]

(vii) ***Kaplan-Meier vs Nelson-Aalen***

The Kaplan-Meier approach estimates the survival function directly. [1]

The Nelson-Aalen method, on the other hand, first estimates the integrated hazard Λ_j, using:

$$\hat{\Lambda}_j = \sum_{t_j \le t} \frac{d_j}{n_j}$$

Then the survival function is estimated as $e^{-\hat{\Lambda}_j}$. [1]

[Total 2]

Assignment X4 – Solutions

Markers: This document sets out one approach to solving each of the questions. Please give credit for other valid approaches.

Solution X4.1

Comment

Reference: Chapter 13

It would be appropriate to graduate the results of a mortality investigation using a mathematical function if:

- a suitable mathematical formula can be found that can describe mortality rates adequately over the entire age range of interest [1]
- the expected number of deaths is great enough at all ages to give reliable answers [1]
- the data can be considered to be complete and accurate, and is adequately subdivided with respect to age, sex and other relevant categories [1]
- an analytic method or computer software that can determine the optimal parameter values is available. [1]

[Total 4]

An example of such a situation would be an investigation of the mortality of a large group of life office policyholders or a national mortality investigation.

Solution X4.2

Comment

Reference: Chapter 12

(i)(a) ***Outliers***

The probability of an individual standardised deviation falling outside the range –3 to +3 should be approximately $2 \times 0.00135 = 0.27\%$. The number of such values will have a *Bin*(100, 0.00270) distribution. The probability of 2 or more such values is therefore $p = 1 - 0.76310 - 0.20660 = 3.03\%$. [1]

(i)(b) ***Signs***

The probability of a positive/negative individual standardised deviation should be ½. The number of such values will have a $B(100,½)$ distribution, which can be approximated using a $N(50,25)$ distribution. The probability of 60 or more values with the same sign corresponds (applying a continuity correction) to a *z*-value of:

$$(z = 59.5 - 50)\big/\sqrt{25} = 1.90$$

ie a probability of $p = 2 \times 0.0287 = 5.74\%$. [1]

This is a two-sided test as overgraduation results from a small or a large number of positive deviations.

(i)(c) ***Serial correlation***

The T ratio in this case will be $t_1 = \sqrt{100} \times 0.18 = 1.8$. The T ratio has a $N(0,1)$ distribution. For a one-sided test the corresponding probability is $p = 3.59\%$. [1]

(i)(d) ***Absolute deviations***

The probability of an absolute individual standardised deviation exceeding $\tfrac{2}{3}$ should be almost ½. The number of such values will have a $B(100,½)$ distribution, which can be approximated using a $N(50,25)$ distribution. The probability of 55 or more values on exceeding $\tfrac{2}{3}$ corresponds (applying a continuity correction) to a *z*-value of:

$$z = (54.5 - 50)\big/\sqrt{25} = 0.90$$

So the *p*-value is 18.4%. [1]

[Total 4]

(ii) ***Which result?***

Test (a) has the smallest p-value and hence most strongly indicates that the rates might be overgraduated. [1]

Solution X4.3

Comment
Reference: Chapters 12 and 13

The reasoning behind the suggestion is firstly, that the graduated rates lie on a smooth curve. [1]

Secondly, under the binomial model, the distribution of the number of deaths at age x last birthday is approximately $Bin(E_x, q_x)$. So the standard deviation is $\sqrt{E_x q_x (1-q_x)}$, or approximately $\sqrt{E_x q_x}$. [1]

The true value of this is unknown, but it can be estimated as $\sqrt{E_x \hat{q}_x} = \sqrt{\theta_x}$. [1]

Deviations are unlikely to exceed twice the standard deviation, so the graduated rate at each age should lie within the range $\hat{q}_x \pm 2\sqrt{\theta_x}\big/E_x$. [1]

This would be quite a helpful procedure if the graduation were to be performed graphically, although the smooth curve should pass outside the limits given approximately once every twenty ages. [1]

However, once a smooth curve has been drawn, it will be necessary to apply all the usual tests to it, and until this has been done it will not be possible to adopt the graduation. [1]

The conclusion is that the suggestion could be a satisfactory first step in graduating crude data, but would be unlikely in itself to produce a satisfactory graduation. [1]

[Maximum 5]

Solution X4.4

Comment

Reference: Chapter 11

(i) ***Central exposed to risk***

X contributes from 1 July 2010 to 30 June 2011, *ie* 1 year. [½]
Y contributes from 1 January 2011 to 30 June 2011, *ie* 6 months. [½]
Z contributes from 1 October 2012 to 31 December 2012, *ie* 3 months. [½]

So the total contribution is 1 year and 9 months. [½]
[Total 2]

(ii) ***Central exposed to risk versus initial exposed to risk***

The central exposed to risk is the total time that lives in the population are observed during the year of age under consideration (within the investigation period). [1]

The initial exposed to risk includes an additional adjustment for the lives who die during the year under consideration. For these individuals the period from the date of death to the end of the year of age is added on. [1]

The central exposed to risk is the population measure for a Poisson model. The initial exposed to risk is the measure for a binomial model. [1]
[Total 3]

Solution X4.5

Comment

Reference: Chapter 11

(i) ***Central exposed to risk***

Let $P_x(t)$ denote the number of lives at time t aged x next birthday and suppose that time is measured in years from 1 January 2014. [½]

We know the values of $P_x(0)$, $P_x(1)$ *and* $P_x(1\tfrac{1}{2})$ *for all* x.

Also, let d_x denote the number of deaths during the investigation aged x last birthday.

Since the death data and the census data don't match, define $P'_x(t)$ to be the number of lives at time t aged x last birthday. [½]

Then: $E_x^c = \int_0^{1½} P'_x(t)\,dt$ [½]

Assuming that $P'_x(t)$ varies linearly between time 0 and time 1, and also between time 1 and time 1½ … [1]

$$E_x^c = \frac{1}{2}\left[P'_x(0) + P'_x(1)\right] + \frac{1}{4}\left[P'_x(1) + P'_x(1½)\right]$$ [½]

Now: $P'_x(0)$ = number of lives at time 0 aged x last birthday

Then: $P'_x(0)$ = number of lives at time 0 aged $x+1$ next birthday = $P_{x+1}(0)$ [½]

Similarly:

$$P'_x(1) = P_{x+1}(1)$$

$$P'_x(1½) = P_{x+1}(1½)$$ [½]

So: $$E_x^c = \frac{1}{2}\left[P_{x+1}(0) + P_{x+1}(1)\right] + \frac{1}{4}\left[P_{x+1}(1) + P_{x+1}(1½)\right]$$ [1]

$$= \frac{1}{2}P_{x+1}(0) + \frac{3}{4}P_{x+1}(1) + \frac{1}{4}P_{x+1}(1½)$$

[Total 5]

(ii) ***Value of f***

Since deaths are classified according to age last birthday, the rate interval starts at exact age x and ends at exact age $x+1$. So the age in the middle of the rate interval is $x+½$, *ie* $f = ½$. [1]

Solution X4.6

Comment
Reference: Chapter 12

The null hypothesis for this test is:

H_0: the graduated rates are the true mortality rates experienced by the assured lives [1]

The serial correlation coefficient is given by:

$$r_1 = \frac{\frac{1}{10}\sum_{x=50}^{59}\left(z_x - \bar{z}\right)\left(z_{x+1} - \bar{z}\right)}{\frac{1}{11}\sum_{x=50}^{60}\left(z_x - \bar{z}\right)^2}$$

where z_x denotes the standardised deviation at age x.

From the data, we have:

$\bar{z} = -0.1065$ [1]

Age, x	z_x	$z_x - \bar{z}$	$(z_x - \bar{z})(z_{x+1} - \bar{z})$	$(z_x - \bar{z})^2$
50	–0.5641	–0.4576	0.2211	0.2094
51	–0.5898	–0.4833	0.1504	0.2336
52	–0.4177	–0.3112	–0.2913	0.0968
53	0.8296	0.9361	0.8736	0.8763
54	0.8267	0.9332	0.9582	0.8709
55	0.9203	1.0268	0.8007	1.0544
56	0.6733	0.7798	–0.3058	0.6081
57	–0.4987	–0.3922	0.2294	0.1538
58	–0.6914	–0.5849	0.2344	0.3421
59	–0.5072	–0.4007	0.4192	0.1605
60	–1.1527	–1.0462		1.0945
Total			**3.2899**	**5.7004**

So: $r_1 = \dfrac{\frac{1}{10} \times 3.2899}{\frac{1}{11} \times 5.7004} = 0.635$ [3]

The test statistic for the serial correlation test is:

$$r_1\sqrt{n} = 0.635 \times \sqrt{11} = 2.106$$ [1]

The test is one-tailed, since we are only concerned about positive correlation. [1]

The test statistic exceeds 1.6449, the upper 5% point of the standard normal distribution. So we reject the null hypothesis at the 5% significance level, and conclude that there is evidence of positive correlation, or grouping of deviations of the same sign. [1]

[Total 8]

Solution X4.7

> ***Comment***
>
> *Reference: Chapter 12*

(i) ***Undergraduation and overgraduation***

When graduating a set of crude mortality rates, there is a trade off between close adherence to the crude rates (goodness of fit) and smoothness. [½]

A satisfactory graduation must achieve an appropriate balance between these two extremes. [½]

"Undergraduation" occurs when too much emphasis is given to goodness of fit. Undergraduated rates adhere closely to the crude rates, but the resulting rates do not show a smooth progression from age to age. [1]

"Overgraduation" occurs when too much emphasis is given to smoothness. Overgraduated rates show a smooth progression from age to age, but the resulting rates do not adhere closely to the crude rates. [1]

[Total 3]

(ii) ***The dangers of overgraduation***

Inadequate premium rates

The office may make losses through underestimating mortality for death benefits or overestimating mortality for survival benefits (since the graduated rates do not accurately reflect the true mortality rates at all ages). [1]

Excessive premium rates

The reverse occurs where the office may lose business through setting excessively high premium rates. [½]

Selection

The office may be exposed to selection from other offices whose premium rates more accurately reflect the true mortality rates. [½]

The dangers of undergraduation

Inappropriate premium rates

The office may make losses or lose business if the premium rates at particular ages have been distorted by random sampling errors that were not "smoothed out". [1]

Anomalies

The office may lose business or incur unnecessary alteration expenses if the rates do not show a consistent progression from age to age. (Policyholders may wait a few years because the rates become cheaper, or they may surrender and take out a new policy to take advantage of an anomaly in the rates at a particular age.) [1]

Reserves

Using biased rates can also lead to inappropriate levels of reserves being held. Holding insufficient reserves can endanger the company's solvency, whereas holding excessive reserves will reduce the company's profitability. [1]

[Total 5]

Solution X4.8

> ***Comment***
>
> *Reference: Chapter 11*

(i) ***Why the company may subdivide their mortality data***

Mortality risk varies between individuals for many reasons. [½]

However, mortality models assume that we are dealing with groups of people who have the same mortality characteristics ("identical lives"). [½]

We can subdivide the data in order to try to achieve these homogeneous groups. [½]

This will reduce much of the heterogeneity present. [½]
[Total 2]

(ii) ***Two main problems***

One problem with subdividing data is that some of the subgroups may be very small, containing only a few individuals. [½]

Estimates of mortality rates derived from the small groups will be unreliable, as it could make it difficult to then ascertain the true underlying rates. [½]

The other main problem is that there may be missing data or the data may be inaccurate or may contain mistakes. [½]

This could result in unreliable mortality estimates. [½]
[Total 2]

(iii) ***Factors for subdividing the data***

Markers: There is a maximum of 2 marks for the 4 main factors below (worth ½ mark each) and a maximum of 2 marks for the other factors (worth ¼ mark each).

Factors that a life insurance company may use to subdivide mortality data include:

- sex [½]
- age [½]
- smoker status [½]
- occupation [½]

[Total 2]

- nationality or ethnic group [¼]
- type of policy [¼]
- level of underwriting [¼]
- duration in force [¼]
- sales channel [¼]
- policy size [¼]
- known impairments [¼]
- current state of health [¼]
- disabilities [¼]
- postcode/geographical location [¼]
- residential status (*eg* homeowner, renting) [¼]
- marital status [¼]

[Maximum 2]
[Total for part (iii) 4]

Solution X4.9

Comment
Reference: Chapter 10

(i) ***Likelihood function***

Let D denote the random variable number of deaths during the investigation, and let d denote the observed value of D. Then the likelihood function under the Poisson model is:

$$L(\mu) = P(D = d) = \frac{e^{-\mu E_x^c}\left(\mu E_x^c\right)^d}{d!} \qquad [1]$$

(ii) ***Why Poisson model is not exact***

The Poisson model is not exact as it allows a non-zero probability of more than N deaths. [1]

However, this probability is usually very small. So the Poisson model is often a good approximation. [1]

The Poisson model assumes that μ is constant over each year of age, but in reality it is a function of the exact age. [1]

However, over individual years of age, the value of μ will not vary very much. So the Poisson model is often a good approximation. [1]

[Maximum 2]

(iii) ***Maximum likelihood estimate of force of mortality***

From (i), the log-likelihood is:

$$\ln L(\mu) = -\mu E_x^c + d\ln\mu + d\ln E_x^c - \ln d!$$

Differentiating with respect to μ gives:

$$\frac{\partial}{\partial\mu}\ln L(\mu) = -E_x^c + \frac{d}{\mu} \qquad [1]$$

Setting this equal to 0 and solving for μ, we obtain:

$$\hat{\mu} = \frac{d}{E_x^c} \qquad [1]$$

We need to check that this maximises the likelihood, so we look at the sign of the second derivative:

$$\frac{\partial^2}{\partial\mu^2} \ln L(\mu) = -\frac{d}{\mu^2} < 0 \Rightarrow \max \qquad [1]$$

So the maximum likelihood estimate of μ is $\hat{\mu} = \dfrac{d}{E_x^c}$.

[Total 3]

(iv) ***Confidence interval***

Using the result in (iii), the maximum likelihood estimate of μ is:

$$\hat{\mu} = \frac{52}{8,460} = 0.006147 \qquad [1]$$

The asymptotic distribution of $\tilde{\mu}$, the maximum likelihood estimator of μ, is $N\left(\mu, \dfrac{\mu}{E_x^c}\right)$. So an approximate 95% confidence interval for μ is:

$$\hat{\mu} \pm 1.96\sqrt{\frac{\hat{\mu}}{E_x^c}} = 0.006147 \pm 1.96\sqrt{\frac{0.006147}{8,460}}$$

$$= (0.004476,\ 0.007818) \qquad [2]$$

[Total 3]

Solution X4.10

Comment
Reference: Chapter 10

(i) ***Estimates of the force of mortality***

The most reasonable assumption we can make here is to assume that the average number of policies in force throughout the year can be approximated by the number in force on 1 July each year. [1]

We can then estimate the force of mortality by dividing the number of deaths (θ_x) by the central exposed to risk:

$$\hat{\mu} = \frac{\theta_x}{E_x^c}$$ [½]

Deaths are classified by age last birthday at the date of death. So at the start of the rate interval, all lives are aged exactly x. [½]

In the middle of the year of age $(x, x+1)$ the lives will be aged $x + ½$. So this will give us an estimate of $\mu_{x+½}$. [1]

This leads to the following results:

Age	E_x^c	θ_x	$\hat{\mu}_{x+½}$
63	18,410	430	0.0234
64	17,196	490	0.0285
65	16,960	507	0.0299

[3]

For example, when $x = 63$, we have:

$$E_x^c = 4{,}192 + 4{,}444 + 4{,}885 + 4{,}889 = 18{,}410$$

$$\theta_x = 104 + 100 + 117 + 109 = 430$$

[Total 6]

(ii) ***Relationship between q and μ***

The general formula for deriving survival probabilities from the force of mortality is:

$${}_t p_x = \exp\left(-\int_0^t \mu_{x+s} ds\right) \quad [½]$$

The Poisson model assumes that μ is constant over the year of age $(x, x+1)$. [½]

So we have:

$$q_x = 1 - p_x = 1 - \exp\left(-\int_0^1 \mu ds\right) = 1 - e^{-\mu} \quad [1]$$

[Total 2]

(iii) ***Estimates of the initial rates of mortality***

We can estimate the initial rates of mortality using the formula in part (ii) and the estimated values of μ from part (i):

Age	$\hat{\mu}_{x+½}$	$\hat{q}_x = 1 - \exp(-\hat{\mu}_{x+½})$
63	0.0234	0.0231
64	0.0285	0.0281
65	0.0299	0.0295

[1]

At the start of the year of age $(x, x+1)$ the lives will be aged x. So these will give us estimates of q_x. [1]

[Total 2]

Solution X4.11

> ***Comment***
>
> *Reference: Chapter 13*

(i) ***Considerations in choosing table***

Whatever table is chosen it must satisfy several key criteria, in particular:

- it must be available for all classes of lives, *eg* males and females [1]
- it must relate to a similar class of lives, *eg* assurances and not annuities in this case [1]
- it must be a "benchmark" table, *ie* generally acceptable to all other actuaries [½]
- it should be up-to-date *ie* relate to fairly recent experience [½]
- it must cover the age range for which rates are required. [½]

In addition it should have the correct pattern of rates by age (not necessarily the correct level of rates though). [1]

It should not have any special features that are unlikely to be present in the experience being graduated. [1]

[Maximum 5]

(ii) ***Checking suitability of formula***

The formula implies that the ratio of the rates varies linearly with age. Is there any external evidence to indicate that this is the correct pattern? [1]

A check could be made by plotting $\overset{\circ}{q}_x / q_x^s$ against x. The plot should be roughly linear. [2]

[Total 3]

(iii)(a) ***Weighted least squares estimation***

We need to look at the function: $\sum[\hat{q}_x - q_x^s(ax+b)]^2 \times \dfrac{E_x}{\hat{q}_x}$ [1]

and minimise it by the choice of the parameters a and b. This is a weighted sum of squares, where the weights are inversely proportional to the approximate variance of $\hat{q}_x$. (For ages where $\hat{q}_x = 0$ an appropriate finite weighting should be used in the sum of squares.) [2]

(iii)(b) ***Maximum likelihood estimation***

We have:

θ_x = observed deaths at age x

E_x = initial exposed to risk at age x

and the graduated rates are to be calculated from the relationship:

$$\mathring{q}_x = q_x^s(ax+b)$$

Let L be the likelihood of obtaining the actual observed set of θ_x values based on given parameter values a and b. Then:

$$L = \prod_x \binom{E_x}{\theta_x} \left(\mathring{q}_x\right)^{\theta_x} \left(1-\mathring{q}_x\right)^{E_x-\theta_x}$$

$$= \prod_x \binom{E_x}{\theta_x} \left(q_x^s(ax+b)\right)^{\theta_x} \left(1-q_x^s(ax+b)\right)^{E_x-\theta_x}$$ [2]

The maximum likelihood estimates of a and b are those values of these parameters that maximise L, or equivalently $\log_e L$, which is:

$$\log_e L = \sum_x \left\{ \log_e \binom{E_x}{\theta_x} + \theta_x \log_e \left(q_x^s(ax+b)\right) + (E_x-\theta_x)\log_e\left(1-q_x^s(ax+b)\right) \right\}$$

[1]

The first term in the summation is fixed and so we can reduce the log-likelihood to:

$$\log_e L = \sum_x \left\{ \theta_x \log_e \left(q_x^s(ax+b)\right) + (E_x-\theta_x)\log_e\left(1-q_x^s(ax+b)\right) \right\}$$ [1]

We can obtain simultaneous equations for the maximum likelihood estimates of a and b by solving the simultaneous equations $\frac{\partial}{\partial a}\log_e L = 0$ and $\frac{\partial}{\partial b}\log_e L = 0$. [1]

[Total 8]

Solution X4.12

Comment

Reference: Chapter 12

(i) ***Statistical tests***

The null hypothesis for all the tests is that the true underlying rates of the population are consistent with the rates given in the AM92 Ultimate table. [½]

(i)(a) ***Chi-squared test***

The table shows the figures required for the chi squared and the individual standardised deviations (ISD) test:

Age	z_x	z_x^2
30	–0.6187	0.383
31	–0.6754	0.456
32	–1.2388	1.535
33	–0.4391	0.193
34	1.1837	1.401
35	2.2199	4.928
36	1.6005	2.562
37	–0.2850	0.081
38	–1.5652	2.450
39	–0.2419	0.059
40	2.0349	4.141
41	–1.3719	1.882
42	–2.6433	6.987
43	–0.4396	0.193
44	1.2462	1.553
45	–0.6421	0.412

[3]

We have ignored the rule of combining groups with an expected number less than 5, as the groups for ages 30 to 33 are only just below 5.

The test statistic is:

$$\sum z_x^2 = \sum \frac{(\theta_x - E_x q_x)^2}{E_x q_x p_x} \simeq \sum \frac{(\theta_x - E_x q_x)^2}{E_x q_x} \qquad [1]$$

since here p_x is close to 1. Summing the z_x^2 column gives:

$$\sum z_x^2 = 29.216 \qquad [1]$$

Under the null hypothesis, this test statistic has a χ^2 distribution with 16 degrees of freedom. [½]

This is a one-sided test. From Page 169 of the *Tables*, the upper 5% point of χ^2_{16} is 26.30. [½]

Since the value of the test statistic is greater than 26.30, we reject the null hypothesis at the 5% significance level and conclude that the observed mortality rates are not consistent with the standard table rates. [1]

(i)(b) ***Signs test***

If the null hypothesis is true, we would expect that roughly half the crude rates would be above the expected rates and half below. The number of positive standardised deviations will have a $Bin(16,½)$ distribution. [½]

From the table above there are 5 positive standardised deviations.

$$P\left(\text{Observed number of positive deviations} \le 5\right)$$

$$= P[Binomial(16,½) \le 5]$$

$$= \frac{\binom{16}{0} + \binom{16}{1} + \cdots + \binom{16}{5}}{2^{16}}$$

$$= \frac{6,885}{65,536}$$

$$= 0.10506 \qquad [1]$$

This is a two-sided test. So the p-value is $2 \times 0.10506 = 0.21$. [1]

Since the p-value is greater than 5%, there is no reason to reject the null hypothesis. [½]

Here, there are only 16 ages so the normal approximation to the binomial is not really appropriate. Markers: please award only 1 mark to students who use the following approach.

The observed standard normal value, including the continuity correction, is:

$$\frac{5.5-16\times 0.5}{\sqrt{16\times 0.5\times 0.5}}=-1.25$$

The result is not significant at the 5% level.

(i)(c) ***Grouping of signs test***

The number of positive signs is 5, and the number of negatives 11. There are 3 groups (runs) of positive standardised deviations. [½]

This is a one-sided test. From Page189 of the *Tables*, we see that the critical value is 1. [½]

The observed number of positive groups is greater than 1, so there is no reason to reject the null hypothesis at the 5% level of significance. [½]

(i)(d) ***Individual Standardised Deviations (ISDs)***

The ISDs have been calculated in (i)(a) above. The following table summarises the expected and actual distribution of the ISDs:

Range	$(-\infty,-3)$	$(-3,-2)$	$(-2,-1)$	$(-1,0)$	$(0,1)$	$(1,2)$	$(2,3)$	$(3,\infty)$
Actual	0	1	3	7	0	3	2	0
Expected	0.0	0.3	2.2	5.4	5.4	2.2	0.3	0.0

The distribution of the ISDs is not quite as we would expect. [½]

Overall there are more negatives than positives. However, this result was not significant on the basis of the signs test. [½]

There are 6 deviations in the range $\left(-\frac{2}{3},\frac{2}{3}\right)$. We would expect to have 50% of the 16 deviations (*ie* 8) in this range. [½]

There are 3 standardised deviations with absolute value greater than 2, which is not satisfactory in 16 deviations. [½]

[Total 14]

(ii) ***Conclusion***

There is no statistical evidence that the observed rates are consistently larger or smaller than the AM92 rates. [½]

Nor is there any statistical evidence of grouping of deviations of the same sign. So the pattern of the observed rates seems to match the pattern of the AM92 rates. [½]

However, there is evidence that the AM92 rates are a poor overall fit to the observed rates. This is a result of a poor fit at ages 35, 40 and 42. The observed numbers of deaths at ages 35 and 40 are much higher than expected, while the observed number at age 42 is much lower than expected. [1]

This is perhaps an indication that some deaths may have been allocated to incorrect ages. This should be investigated. [½]

[Maximum 2]