The Society is providing these solutions to assist candidates preparing for the examinations in 2017.

The solutions are intended as learning aids and should not be seen as "model answers".

Users of the solutions should always be aware that in many cases there are valid alternative methods. Also, in the many cases where discussion is called for, there may be other valid points that could be made.

While every care has been taken with the preparation of these solutions, the Society will not be responsible for any errors or omissions.

The Society will not enter into any correspondence in respect of these solutions.
The Society is providing these solutions to assist candidates preparing for the examinations in 2017.

The solutions are intended as learning aids and should not be seen as "model answers".

Users of the solutions should always be aware that in many cases there are valid alternative methods. Also, in the many cases where discussion is called for, there may be other valid points that could be made.

While every care has been taken with the preparation of these solutions, the Society will not be responsible for any errors or omissions.

The Society will not enter into any correspondence in respect of these solutions.
1. (i) None of the states communicate so each forms a class, 1, 2, and 3 are transient while 4 is absorbing so is recurrent. Someone with the disease will end up in state 4 with probability 1, i.e., permanent disability.

(ii) 
\[ P^{(2)} = P \times P = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{16} & \frac{7}{16} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{3}{4} \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 0 & 1 \end{bmatrix} \]

(iii) 
(a) \( P_{24} = \frac{1}{4} \) 
(b) \( P_{13}^{(2)} = \frac{1}{16} \)

\[ P^{(4)} = P^{22} \times P^{23} + P^{23} \times P^{33} = \frac{1}{4} \times \frac{1}{4} + \frac{1}{4} \times \frac{1}{4} = \frac{1}{8} \]

\[ P^x = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 1 \end{bmatrix} \]

All four states now communicate so form a single recurrent class.

(vi) \[ \pi = \pi \times P^x \] subject to \[ \sum_{i=1}^{4} \pi_i = 1 \]

\[ \therefore \pi_1 = \frac{1}{2} \pi_1 + \frac{1}{2} \pi_4 \Rightarrow \pi_1 = \pi_4 \]

\[ \pi_2 = \frac{1}{4} \pi_1 + \frac{1}{2} \pi_2 \Rightarrow \pi_2 = \frac{1}{2} \pi_1 \]

\[ \pi_3 = \frac{1}{4} \pi_2 + \frac{1}{2} \pi_3 \Rightarrow \pi_3 = \frac{1}{2} \pi_2 = \frac{1}{4} \pi_1 \]

\[ (\pi_4 = \frac{1}{4} \pi_1 + \frac{1}{4} \pi_2 + \frac{1}{2} \pi_3 + \frac{1}{2} \pi_4) \]

\[ \sum_{i=1}^{4} \pi_i = 1 \Rightarrow (1 + \frac{1}{2} + \frac{1}{4} + 1) \pi_1 = 1 \Rightarrow \pi_1 = \frac{4}{11} \]

\[ \pi_2 = \frac{2}{11}, \pi_3 = \frac{1}{11}, \pi_4 = \frac{4}{11} \]

\[ \pi(Receiving\ treatment) = \pi_4 = \frac{4}{11} \text{ regardless of the initial state.} \]

(vii) \[ E[Cost] = 0 + \frac{7}{11} \cdot c + \frac{1}{11} \cdot 2c + \frac{4}{11} \cdot 8c = \frac{36}{11}c \]
Since \( X_1 \) consists of a single family \( Z \),
\[ G_1(s) = E[s^{X_1}] = G(s). \]

(ii) \( X_{ni} \) is the total population of a (sub) branching process with \((n-1)\) generations, from one individual.
Hence \( E[s^{X_{ni}}] = G_{n-1}(s). \)

(iii) \[ G_n(s) = E[s^{X_n}] = E\left[ E\left(s^{X_n} \mid X_1\right)\right] \]
Now if \( X_1 = x \), \( X_n = \sum_{i=1}^{x} X_{ni} \), so
\[ E\left[s^{X_n} \mid X_1 = x\right] = E\left[\prod_{i=1}^{x} s^{X_{ni}}\right] = \prod_{i=1}^{x} E\left[s^{X_{ni}}\right], \]
by independence
\[ = G_{n-1}(s)^x, \]
by (ii),
and \( G_n(s) = E\left[G_{n-1}(s)^{X_1}\right] = G\left(G_{n-1}(s)\right) \)
since \( E[s^{X_1}] = G(s), \) by (i).

(iv) \[ X_n = 0 \Rightarrow X_{n+1} = 0 \] so \( \Pi_n \leq \Pi_{n+1} \) and
\( \Pi_n \) is increasing.
\[ \Pi_n = G_n(0) = G(G_{n-1}(0)) = G(\Pi_{n-1}) \]

(v) \[ \Pi = \lim_{n \to \infty} \Pi_n = \lim_{n \to \infty} G\left(\Pi_{n-1}\right) = G(\Pi) \]
since \( G(s) = \sum_{i=0}^{\infty} s^i p_i \) is continuous.
\[ G(1) = \sum_{i=0}^{\infty} p_i = 1 \] so 1 is a root.

(vi) \( G(s) = 0.1 + 0.2s + 0.3s^2 + 0.4s^3 \)
so solve \( 0.1 + 0.2s + 0.3s^2 + 0.4s^3 = s \)
or \( 4s^3 + 3s^2 - 8s + 1 = 0 \)
\( s = 1 \) is a root so
\[ (s-1)(4s^2 + 7s - 1) = 0 \]
and the remaining roots are
\[ -7 \pm \sqrt{49 + 16} \]
8
\[ -10.88 \text{ or } 0.13 \]
Hence \( \Pi = 0.13 \) as \( \Pi \) is the smallest non-negative root.
3. (i) Let \( N(t) = N_0 \) in system at time \( t \). Then
\[
\begin{align*}
\mathbb{P}(N(t+s) = i+1 \mid N(t) = i) &= \alpha_i s^i + o(s^i), \\
\mathbb{P}(N(t+s) = i-1 \mid N(t) = i) &= \beta_i s^i + o(s^i), \\
\end{align*}
\]
Since \( \sum_{i=0}^{\infty} \Pi_i = 1 \), \( \Pi_0 = \left[ 1 + \sum_{n=1}^{\infty} \frac{\alpha_n x_1 \cdots x_{i-1} \cdots x_{n-1}}{\beta_1 \beta_2 \cdots \beta_n} \right]^{-1} \)

provides it converges.

(ii) (a) \( x_i = \lambda_i \frac{1}{i+1}, \quad i = 0, 1, 2, \ldots \quad \beta_i = \frac{\mu_i}{i+1}, \quad i = 1, 2, \ldots \)
\[
\begin{align*}
\alpha_0 x_1 \cdots x_{n-1} &= \frac{x_0!}{\beta_1 \beta_2 \cdots \beta_n} = \frac{x_0!}{x_0! x_1 \cdots x_{n-1}} = \left( \frac{n!}{n+1!} \right)^n = (n+1)^n \]

where \( \rho = \frac{\lambda}{\mu} \)
\[
\Pi_0 = 1 + \sum_{n=1}^{\infty} (n+1)^n \rho^n = \sum_{n=1}^{\infty} \rho^n \frac{1}{(1-\rho)^2} = \frac{1}{(1-\rho)^2}
\]

Server is busy a proportion \( 1 - \Pi_0 = 1 - (1 - \rho)^2 = \rho (2 - \rho) \)

(b)
\[
\Pi_n = (1-\rho)^2 (n+1)^n \rho^n, \quad \text{for } n \geq 0
\]
\[
\begin{align*}
\text{Pgf is } P(z) &= \sum_{n=0}^{\infty} (1-\rho)^2 (n+1)^n \rho^n z^n \\
&= (1-\rho)^2 \sum_{j=1}^{\infty} \frac{j^j \rho^j}{(\rho^2)^{j-1}} \\
&= (1-\rho)^2 \sum_{j=1}^{\infty} j \rho^j \\
&= \frac{2 \rho^2}{(1-\rho)^2}
\end{align*}
\]
\[
P'(z) = (1-\rho)^2 2 \rho (1-\rho z)^{-3} \Rightarrow \mathbb{E}[N] = P'(1) = \frac{2 \rho}{1-\rho}
\]
\[
P''(z) = (1-\rho)^2 6 \rho^2 (1-\rho z)^{-4} \Rightarrow \mathbb{E}[N(N-1)] = P''(1) = \frac{6 \rho^2}{(1-\rho)^2}
\]
\[
\text{Var}[N] = \frac{6 \rho^2 + 2 \rho - (2 \rho^2)}{(1-\rho)^2} = \frac{2 \rho}{(1-\rho)^2}
\]

(c)
\[
\mathbb{E}[N^{\circ} \text{ lost } \mid n \text{ in system}] = \lambda x \left[ 1 - \frac{1}{n+1} \right] = \frac{\lambda n}{n+1}
\]
\[
\mathbb{E}[N^{\circ} \text{ lost }] = \sum_{n=0}^{\infty} \frac{\lambda n}{n+1} x (1-\rho)^2 (n+1) \rho^n
\]
\[
= \lambda (1-\rho)^2 \rho \sum_{j=1}^{\infty} j \rho^j = \frac{\lambda \rho}{(1-\rho)^2}
\]
Let $N(t) = N^n$ be the number of particles in $[0, t]$. 

\[
P(N(t) = n) = \frac{1 - \lambda(t)St + o(St)}{(1 + \lambda(t)St + o(St))^n} \quad \text{for } n = 0, 1, 2, \ldots
\]

\[
P(N(t) = n+1 | N(t) = n) = \lambda(t)St + o(St)
\]

\[
P(N(t) + St > n+2 | N(t) = n) = o(St)
\]

\[
P_n(t+St) = (1 - \lambda(t)St + o(St))^2 \left( (1 + \lambda(t)St + o(St))^n \right) - o(St)
\]

\[
P_n(t+St) - P_n(t) = -\lambda(t)P_n(t) + \lambda(t)P_{n-1}(t) + o(St)
\]

\[
\frac{d}{dt}P_n(t) = -\lambda(t)P_n(t) + \lambda(t)P_{n-1}(t), \text{ letting } St \to 0.
\]

Also

\[
\frac{d}{dt}P_0(t) = \lim_{St \to 0} \frac{(1 - \lambda(t)St + o(St))^2 P_0(t) - P_0(t) + o(St)}{St} = -\lambda(t)P_0(t)
\]

Multiplying by $S^n$ and summing $\sum_{n=0}^{\infty} dP_n(t)/dt$ we get

\[
\sum_{n=0}^{\infty} \frac{dP_n(t)}{dt} S^n = -\lambda(t) \sum_{n=0}^{\infty} P_n(t) S^n + \lambda(t) \sum_{n=0}^{\infty} P_{n-1}(t) S^n
\]

\[
= -\lambda(t)G(s, t) + \lambda(t) \sum_{n=0}^{\infty} P_n(t) S^n
\]

\[
= -\lambda(t)G(s, t) + \lambda(t) S G(s, t)
\]

\[
\frac{d}{dt} G(s, t) = \lambda(t)(S-1) G(s, t)
\]

Since $P_0(0) = 1$, $G(s, 0) = \sum_{n=0}^{\infty} P_n(0) = S \times 1 = 1$

\[
\int_0^t \frac{dG(s, u)}{G(s, u)} du = \int_0^t \frac{1}{u} du (S-1) = \Delta(t)(S-1)
\]

\[
\ln G(s, t) = \Delta(t)(S-1) + \text{constant}
\]

\[
G(s, t) = C(s) e^{\Delta(t)(S-1)}
\]

To satisfy $G(s, 0) = 1$, $C(s) = 1$ since $\Delta(0) = 0$.

\[
\Delta(t) = \Delta(t)(S-1)
\]

\[
\text{Comparing with } G(s, t) = \sum_{n=0}^{\infty} P_n(t) S^n \text{ we get }
\]

\[
P_n(t) = e^{-\Delta(t)} \frac{\Delta(t)^n}{n!} \text{ for } n = 0, 1, 2, \ldots
\]

\[
\text{i.e. Poisson (}\Delta(t)\text{)}
\]
(i) \( L_t = \alpha Y_t + (1-\alpha)L_{t-1} \)

(ii) \( L_t = L_{t-1} + \alpha (Y_t - L_{t-1}) \)
\[ = L_{t-1} + \alpha e_t \] since \( L_{t-1} = Y_{t-1} - (1) \)

(iii) \( L_t = \alpha Y_t + (1-\alpha)(\alpha Y_{t-1} + (1-\alpha)L_{t-2}) \)
\[ = \alpha Y_t + (1-\alpha)\alpha Y_{t-1} + (1-\alpha)^2(\alpha Y_{t-2} + (1-\alpha)L_{t-3}) \]
\[ = \alpha Y_t + (1-\alpha)Y_{t-1} + \ldots + (1-\alpha)^t Y_{t-t} + (1-\alpha)^{t-1}Y_{t-1} \]
since \( L_1 = Y_1 \)

(iv) \( \alpha = 0.3 \)
\[
\begin{array}{c|c|c|c}
  t & \frac{Y_t}{13} & Y_t - L_{t-1} = e_t & L_t = L_{t-1} + 0.3 e_t \\
  \hline
  1 & 13 & \frac{Y_t}{13} & \frac{Y_t}{13} \\
  2 & 15 & 15 - 13 = 2 & 13 + 0.3 \times 2 = 13.6 \\
  3 & 12 & 12 - 13 = -1.6 & 13.6 - 0.3 \times 1.6 = 13.04 \\
  4 & 17 & 17 - 13 = 3.88 & 13.04 + 0.3 \times 3.88 = 14.28 \\
\end{array}
\]

(b) \( Y_t = X_t - X_{t-1} = (W_t + A_t) - (W_{t-1} + A_{t-1}) \)
\[ = W_{t-1} + E_t + A_t - W_{t-1} - A_t - A_{t-1} \]
\[ = E_t + A_t - A_{t-1} \]

Hence \( \text{Var}(Y_t) = 2E^2 + 2A^2 \) as \( \text{Cov}(A_t, A_{t-1}) = 0 \)
and \( \text{Cov}(Y_t, Y_{t-1}) = \text{Cov}(E_t + A_t - A_{t-1}, E_{t-1} + A_t - A_{t-1} - A_{t-2}) \)
\[ = \text{Var}(A_{t-1}) \quad = -2A^2 \]
since all other terms = 0.
\[ \beta_1(Y_t) = \frac{\text{Cov}(Y_t, Y_{t-1})}{\text{Var}(Y_t)} = \frac{-2A^2}{2E^2 + 2A^2} \]

Also \( \text{Cov}(Y_t, Y_{t-k}) = \text{Cov}(E_t + A_t - A_{t-1}, E_{t-k} + A_t - A_{t-1} - A_{t-k-1}) \)
\[ = 0 \]
so \( \beta_1(k) = 0 \) for \( k > 2 \)

\( Y_t \) will have the acf of an MA(1), so \( X_t \)
will be \( \text{ARIMA}(0, 1, 1) \) i.e. \( p = 0, d = 1 = q \)
The model is $MA(2)$ or $ARMA(0, 2)$: $p = 0, q = 2$. 

(a) $X_t$ is stationary since all pure $MA(q)$ are stationary if $|rac{1}{1+0.4^2 + (-0.45)^2}| < \infty$ or the AR polynomial $\phi(B) = 1$ has no roots, hence none lie outside the unit circle.

(b) Roots of $\Theta(B) = 1 + 0.4B - 0.45B^2 = (1 + 0.9B)(1 - 0.5B)$, both roots $B_1 = -0.9$ and $B_2 = 0.5$ lie outside $|B| = 1$ so invertible.

\[
E[X_t] = 10 + E[A_t] + 0.4E[A_{t-1}] - 0.45E[A_{t-2}] = 10
\]
since $E[A_{t-i}] = 0$ for all $i$.

\[
\text{Var}[X_t] = \text{Var}[A_t] + 0.4^2 \text{Var}[A_{t-1}] + (-0.45)^2 \text{Var}[A_{t-2}] = 1.3625 \cdot \text{ Var }, \text{since Cov}(A_{t-i}, A_{t-j}) = 0 (i \neq j)
\]

(iii) $X_t(k) = \begin{cases} E[X_{t+k} | X_t, X_{t-1}, \ldots ] \\ 10 + E[A_{t+k} + 0.4A_{t+k-1} - 0.45A_{t+k-2}] X_t, X_{t-1}, \ldots \end{cases}$

for $k \geq 3$ since $E[A_{t+i} | X_t, X_{t-1}, \ldots ] = 0$

for $j > 1$ due to independence of $A_{t+i}$ from past $X_t$'s.

$X_t(1) = E[X_{t+1} | X_t, X_{t-1}, \ldots ] = 10 + E[A_{t+1} + 0.4A_t - 0.45A_{t-1} | X_t, X_{t-1}, \ldots ]$

$= 10 + 0 + 0.4A_t - 0.45A_{t-1}$

\[\text{where } a_t = E[A_t | X_t, X_{t-1}, \ldots ] \approx X_t - X_{t-1}(1)\]

and $a_{t-1} = E[A_{t-1} | X_t, X_{t-1}, \ldots ] \approx X_{t-1} - X_{t-2}(1)$

are 1-step ahead forecast errors.

Similarly $X_t(2) = 10 + E[A_{t+2} + 0.4A_{t+1} - 0.45A_t | X_t, X_{t-1}, \ldots ]$

$= 10 + 0 + 0.4 \times 0 - 0.45A_t$

$= 10 - 0.45A_t$

(iv) $V(1) = \text{Var}[A_{t+1}] = 0.2^2, V(2) = \text{Var}[A_{t+2} + 0.4A_{t+1}] = 1.16 \cdot 0.2^2$

$V(k) = \text{Var}[A_{t+k} + 0.4A_{t+k-1} - 0.45A_{t+k-2}] = 1.3625 \cdot 0.2^2$

for $k \geq 3$.

90% confidence interval will be:

\[10 \pm 1.6449 \times \sqrt{1.3625 \cdot 0.2^2} (\text{i.e. } \pm 1.928)\]

since $P(Z > 1.6449) = 0.05$ for $Z \sim N(0, 1)$. Hence constant width $= 3.842$, for all $k \geq 3$. 
(a) Stationary: any roots of $\phi(B)$ must lie outside $|B| = 1$, i.e. $|B_i| > 1$.

(b) Invertible: any roots of $\Theta(B)$ outside $|B| = 1$.

\[ \phi(B) = 1 - 1.8B + 0.8B^2 = (1 - 0.8B)(1 - B) \] so roots are $B_1 = \frac{3}{4} = 0.75 > 1$ and $B_2 = 1$ \Rightarrow not stationary.

$\Theta(B) = 1 + 0.6B$. Roots is $\frac{5}{3}$ and $|\frac{5}{3}| > 1$ so invertible.

$x_t$ is ARIMA(1,1,1) i.e. $p = 1 = d = q$.

(iii) $a_t = a_t(\theta, \phi) = x_t - 1.8x_{t-1} + 0.8x_{t-2} - 0.6a_{t-1}$
while initially $x_0 = x_{-1} = E[x_t] = 0$ and $a_0 = 0$.

$S(\theta, \phi) = \sum t^2(\theta, \phi)$ is then minimized by searching over $\theta$ and $\phi$.

"Conditional" since conditioning on initial choices of $x_0, x_{-1}, a_0$.

(iv) $\phi(B) = 1 - 2.8B + 2.6B^2 - 0.8B^3 = (1 - B)(1 - 1.8B + 0.8B^2)
\Theta(B) = 1 - 0.4B - 0.6B^2 = (1 - B)(1 + 0.6B)$

Hence $1 - B$ cancels and the model is redundant, simplified model is same as (ii) i.e.

$(1 - 0.8B)(1 - B)x_t = (1 + 0.6B)a_t$ so ARIMA(1,1,1).

(v) If $\phi(B)D^d x_t = \Theta(B) a_t$ where $\phi(B) = (1 - \omega B)\phi_1(B)$
and $\Theta(B) = (1 - \omega B)\Theta_1(B)$ then

\[ a_t(\theta, \phi) = \Theta(B)^{-1}(\phi(B)x_t) \]

\[ = (1 - \omega B)^{-1}(\Theta_1(B)(1 - \omega B)\phi_1(B))x_t \]

\[ = \Theta_1(B)\phi_1(B)x_t \]

The error will not depend on $\omega$. Hence $S(\omega, \phi, \theta)$ will be constant in $\omega$ for any $\phi, \theta$ and the minimization routine will fail to converge.
Values outside the dotted lines in Figures 2 and 3 are acf or pacf values significantly different from 2.50 (at 5%).

a. White noise rejected as acf has many values outside.

b. MA(2): acf should drop to 2.50 after first two so not a plausible candidate.

c. AR(2): pacf drops to 2.50 after first two so this is a possibility.

None of the models require differencing.

1. AR(2): \[ X_t = 9.9283 + 0.1168X_{t-1} + 0.6282X_{t-2} + \Delta_t \]

2. AR(3): \[ X_t = 9.9284 + 0.1145X_{t-1} + 0.6278X_{t-2} + 0.0037X_{t-3} + \Delta_t \]

3. ARMA(2,1): \[ X_t = 9.9285 + 0.1201X_{t-1} + 0.6272X_{t-2} + \Delta_t - 0.0055\Delta_{t-1} \]

(iii) \[ T\text{-ratio} = \frac{\text{Coefficient}}{SE} \text{ Judged significantly different (at 5%)} \text{ if } |T| > 1.96 \text{ (as } n = 187 \text{ is large)} \]

<table>
<thead>
<tr>
<th>Model</th>
<th>AR(1)</th>
<th>AR(2)</th>
<th>AR(3)</th>
<th>MA(1)</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>2.07</td>
<td>11.08</td>
<td>0.05</td>
<td>-</td>
<td>37.23</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.57</td>
<td>10.98</td>
<td>0.05</td>
<td>-</td>
<td>37.07</td>
</tr>
<tr>
<td>Model 3</td>
<td>1.33</td>
<td>10.37</td>
<td>0.05</td>
<td>-</td>
<td>37.07</td>
</tr>
</tbody>
</table>

Conclusion: AR(2) satisfactory but over-fitting to AR(3) or ARMA(2,1) unnecessary.

The log-likelihoods are identical but AIC is lower for AR(2): more evidence that AR(2) should be preferred.

- SD(\(X_t\)) is higher suggesting that we may have over-differenced.
- ACF of \(X_t\) declines exponentially which is characteristic of a stationary series.
- AR(2) provides a satisfactory fit. Apparent "trend" is probably illusory.