THE ROYAL STATISTICAL SOCIETY

2007 EXAMINATIONS – SOLUTIONS

GRADUATE DIPLOMA

STATISTICAL THEORY AND METHODS

PAPER I

The Society provides these solutions to assist candidates preparing for the examinations in future years and for the information of any other persons using the examinations.

The solutions should NOT be seen as "model answers". Rather, they have been written out in considerable detail and are intended as learning aids.

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.

Note. In accordance with the convention used in the Society's examination papers, the notation log denotes logarithm to base *e*. Logarithms to any other base are explicitly identified, e.g. $\log_{10}$.

Graduate Diploma, Statistical Theory & Methods, Paper I, 2007. Question 1

(i)(a) $P(X=x, Y=y) = \dfrac{e^{-\theta}\theta^x}{x!}.\dfrac{e^{-\lambda}\lambda^y}{y!}$ by independence (x = 0, 1, …, y = 0, 1, …).

(b) There is only one way that $X + Y = 0$ can happen, when $X = 0$ and $Y = 0$. Hence $P(X+Y=0) = e^{-(\theta+\lambda)}$.

(c) $P(X+Y=1) = P(X=0, Y=1) + P(X=1, Y=0) = e^{-\theta}\lambda e^{-\lambda} + \theta e^{-\theta}e^{-\lambda}$

$= (\theta+\lambda)e^{-(\theta+\lambda)}$.

(d) $P(X+Y=z) = \sum_{x=0}^{z} P(X=x, Y=z-x) = \sum_{x=0}^{z} \dfrac{e^{-\theta}\theta^x}{x!}.\dfrac{e^{-\lambda}\lambda^{z-x}}{(z-x)!}$

$= \sum_{x=0}^{z} \dfrac{e^{-(\theta+\lambda)}\theta^x\lambda^{z-x}}{x!(z-x)!} = \dfrac{e^{-(\theta+\lambda)}}{z!}\sum_{x=0}^{z}\dfrac{z!\theta^x\lambda^{z-x}}{x!(z-x)!} = \dfrac{e^{-(\theta+\lambda)}}{z!}(\theta+\lambda)^z$.

Hence $X + Y$ has the Poisson distribution with mean $(\theta + \lambda)$.

(ii) For y = 0, 1, …, z,

$$P(Y=y \mid X+Y=z) = \frac{P(X=z-y, Y=y)}{P(X+Y=z)} = \frac{\dfrac{e^{-\theta}\theta^{z-y}}{(z-y)!}.\dfrac{e^{-\lambda}\lambda^y}{y!}}{\dfrac{e^{-(\theta+\lambda)}(\theta+\lambda)^z}{z!}}$$

$$= \binom{z}{y}\frac{\theta^{z-y}\lambda^y}{(\theta+\lambda)^z} = \binom{z}{y}\left(\frac{\theta}{\theta+\lambda}\right)^{z-y}\left(\frac{\lambda}{\theta+\lambda}\right)^y = \binom{z}{y}\left(\frac{\lambda}{\theta+\lambda}\right)^y\left(1-\frac{\lambda}{\theta+\lambda}\right)^{z-y}.$$

Hence this conditional distribution is binomial with parameters z and $\dfrac{\lambda}{\theta+\lambda}$.

(iii) Since all $\{U_i\}$ are independent, and using the result of (i)(d) (i.e. the additive property of the Poisson distribution), $U_1 + U_2 + \ldots + U_{n-1}$ has the Poisson distribution with parameter $(n-1)\theta$.

Now using (ii) with $X = U_1 + U_2 + \ldots + U_{n-1}$ and $Y = U_n$, it follows that the conditional distribution of U_n is binomial with parameters z and $\theta/(n\theta) = 1/n$.

Graduate Diploma, Statistical Theory & Methods, Paper I, 2007. Question 2

(i) There are $\binom{20}{5}$ different ways of selecting the committee if there are no constraints.

In order to meet the constraints we require X out of 12 Blues, Y out of 4 Reds and $5 - X - Y$ out of 4 Yellows. The respective numbers of ways for these selections are $\binom{12}{x}, \binom{4}{y}$ and $\binom{4}{5-x-y}$.

Any of the selections of Blues can go with any of those for Reds and any of those for Yellows. All possible committees are equally likely as the choice is made at random. So the required probability is the ratio (number of ways of observing the constraints) / (total number of ways), which gives the stated result.

This is valid for $x = 0, 1, \ldots, 5$ and $y = 0, 1, \ldots, 4$, with $1 \le x + y \le 5$.

(ii) $$P = \frac{\binom{12}{3}\binom{4}{1}\binom{4}{1}}{\binom{20}{5}} = \frac{\frac{12\times11\times10}{3\times2\times1}\times4\times4}{\frac{20\times19\times18\times17\times16}{5\times4\times3\times2\times1}} = \frac{220\times16}{19\times17\times48} = 0.2270.$$

(iii) Selection of the committee is random sampling without replacement from the finite population of 20 members. Hence, to study X alone, the members are effectively 12 Blues and 8 others, so X has a hypergeometric distribution with

$$P(X = x) = \frac{\binom{12}{x}\binom{8}{5-x}}{\binom{20}{5}}.$$

Solution continued on next page

(iv) By the same argument as in (iii), labelling Yellows as Z we have

$$P(Z=z)=\binom{4}{z}\binom{16}{5-z}/\binom{20}{5}.$$

Thus $P(Z>1) = 1 - P(Z = 0 \text{ or } 1)$

$$= 1-\left\{\frac{\binom{4}{0}\binom{16}{5}}{\binom{20}{5}}+\frac{\binom{4}{1}\binom{16}{4}}{\binom{20}{5}}\right\} = 1-\left(\frac{4368+7280}{15504}\right) = 1-0.7513 = 0.2487.$$

(v) Given the value of Y (y = 0, 1, 2, 3, or 4), there remain $5 - y$ places to be filled by non-Reds. So the situation is similar to that in part (iii): we must select X Blues from the total of 12 Blues and 4 Yellows.

Given that $Y = 2$, there are only 3 remaining places to allocate. So

$$P(X=x\,|\,Y=2)=\frac{\binom{12}{x}\binom{4}{3-x}}{\binom{16}{3}}, \qquad \text{for } x = 0, 1, 2, \text{ or } 3.$$

(vi) If all 4 Yellows are chosen, the remaining 1 person must be chosen from the 12 Blues and 4 Reds.

Thus $P(Y=1\,|\,Z=4)=\dfrac{\binom{12}{0}\binom{4}{1}}{\binom{16}{1}}=\dfrac{1\times 4}{16}=\dfrac{1}{4}$.

[Alternatively, this can be considered as obvious because the fifth person is simply chosen at random from the 12 Blues and 4 Reds, so it is not necessary to make explicit use of hypergeometric probabilities.]

Graduate Diploma, Statistical Theory & Methods, Paper I, 2007. Question 3

(i) $f(y \mid X = x) = \dfrac{f(x,y)}{f(x)}$, where $f(x)$ is the marginal pdf of X. Thus we have

$$E_Y\left(Y \mid X = x\right) = \int_{-\infty}^{\infty} y\, f\left(y|x\right) dy = \int_{-\infty}^{\infty} y \frac{f(x,y)}{f(x)} dy\,.$$

For $h(x)$ any function of X, we have

$$E_X\left[h(X).E_Y\left(Y \mid X\right)\right] = \int_{-\infty}^{\infty} \left\{ h(x) \int_{-\infty}^{\infty} y \frac{f(x,y)}{f(x)} dy \right\} f(x) dx$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x).y.f(x,y) dy dx = E\left(h(X).Y\right).$$

(ii)(a) $E\left[Y\right] = E_X\left[E_Y\left(Y \mid X\right)\right] = E\left[\alpha + \beta X\right] = \alpha + \beta E\left[X\right].$

$$E\left[XY\right] = E_X\left[X.E_Y\left(Y \mid X\right)\right] = E\left[\alpha X + \beta X^2\right] = \alpha E\left[X\right] + \beta E\left[X^2\right].$$

$$\therefore \rho(X,Y) = \frac{E\left[XY\right] - E\left[X\right]E\left[Y\right]}{\sqrt{\mathrm{Var}\left(X\right)\mathrm{Var}\left(Y\right)}}$$

$$= \frac{\alpha E\left[X\right] + \beta E\left[X^2\right] - E\left[X\right]\left(\alpha + \beta E\left[X\right]\right)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}$$

$$= \frac{\beta\left\{E\left[X^2\right] - \left(E\left[X\right]\right)^2\right\}}{\sqrt{\mathrm{Var}\left(X\right)\mathrm{Var}\left(Y\right)}}$$

$$= \frac{\beta\,\mathrm{Var}\left(X\right)}{\sqrt{\mathrm{Var}\left(X\right)\mathrm{Var}\left(Y\right)}} \;=\; \beta\sqrt{\frac{\mathrm{Var}\left(X\right)}{\mathrm{Var}\left(Y\right)}}\;.$$

(b) The given identity becomes $\mathrm{Var}\left(Y\right) = \mathrm{Var}\left(\alpha + \beta X\right) + E\left[\sigma^2\right] = \beta^2\mathrm{Var}(X) + \sigma^2.$

So $\sigma^2 = \mathrm{Var}(Y) - \beta^2\mathrm{Var}(X) = \mathrm{Var}(Y) - \rho^2 \dfrac{\mathrm{Var}(Y)}{\mathrm{Var}(X)}.\mathrm{Var}(X) = \left(1 - \rho^2\right)\mathrm{Var}(Y).$

Graduate Diploma, Statistical Theory & Methods, Paper I, 2007. Question 4

(i) $f(x,y) = f_X(x) f_Y(y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}$ by independence (for $-\infty < x, y < \infty$).

$U = X + Y$ and $V = X - Y$, so $X = \frac{1}{2}(U+V)$ and $Y = \frac{1}{2}(U-V)$.

Now considering the Jacobian, we have

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = \left| -\frac{1}{4} - \frac{1}{4} \right| = \frac{1}{2}.$$

$$\therefore\ f_{UV}(u,v) = f(x,y)|J| = \frac{1}{2}.\frac{1}{2\pi} e^{-\frac{1}{2}\left[\frac{(u+v)^2}{4} + \frac{(u-v)^2}{4}\right]}$$

$$= \frac{1}{4\pi} e^{-\frac{1}{2}\frac{u^2+2uv+v^2+u^2-2uv+v^2}{4}} = \frac{1}{4\pi} e^{-\frac{u^2+v^2}{4}}, \quad \text{for } -\infty < u, v < \infty.$$

This joint pdf factorises as $\frac{1}{2\sqrt{\pi}} e^{-\frac{u^2}{4}} . \frac{1}{2\sqrt{\pi}} e^{-\frac{v^2}{4}}$.

So by the factorisation theorem U and V are independent. Both have Normal distributions with mean 0 and variance 2 (ie N(0, 2)).

(ii) The sample mean is $M = \frac{1}{2}(W_1 + W_2)$. So the sample variance is

$$S^2 = \frac{1}{2-1}\sum_{i=1}^{2}(W_i - M)^2 = \sum_{i=1}^{2}\left(W_i - \frac{W_1+W_2}{2}\right)^2 = \frac{1}{2}(W_1 - W_2)^2.$$

Now, there are independent N(0, 1) random variables X and Y such that $W_1 = \mu + \sigma X$ and $W_2 = \mu + \sigma Y$. Thus we have $M = \mu + \frac{1}{2}\sigma(X+Y)$ and $S^2 = \frac{1}{2}\sigma^2(X-Y)^2$.

Using (i), $X + Y$ and $X - Y$ are independent. Therefore any function of $X + Y$ is independent of any function of $X - Y$. So this mean and variance are independent.

Graduate Diploma, Statistical Theory & Methods, Paper I, 2007. Question 5

(i) $M_X(t) = E\left(e^{tX}\right)$. $\therefore M_{aX+b}(t) = E\left(e^{t(aX+b)}\right) = e^{bt}E\left(e^{(at)X}\right) = e^{bt}M_X(at)$.

(ii) (a) $M_X(t) = \int_{-\infty}^{\infty} e^{tx}.\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx$.

The exponent in this integral is

$$tx - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 = tx - \frac{1}{2\sigma^2}\left(x^2 - 2\mu x + \mu^2\right)$$

$$= -\frac{1}{2\sigma^2}\left[\left\{x - \left(\mu + \sigma^2 t\right)\right\}^2 - 2\mu\sigma^2 t - \sigma^4 t^2\right]$$

$$= -\frac{1}{2\sigma^2}\left[x - \left(\mu + \sigma^2 t\right)\right]^2 + \mu t + \frac{1}{2}\sigma^2 t^2.$$

Hence $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}\int_{-\infty}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}\left[x - (\mu + \sigma^2 t)\right]^2\right\}dx$

$= e^{\mu t + \frac{1}{2}\sigma^2 t^2}$ since the integral = 1 (pdf of N($\mu + \sigma^2 t$, σ^2)).

(b) For Z, $a = \frac{1}{\sigma}$ and $b = -\frac{\mu}{\sigma}$ in the notation of part (i).

Hence $M_Z(t) = \exp\left(-\frac{\mu t}{\sigma}\right)\exp\left(\frac{\mu t}{\sigma} + \frac{\sigma^2}{2}\frac{t^2}{\sigma^2}\right) = e^{\frac{t^2}{2}}$.

This is the same as $M_X(t)$ above when μ = 0 and σ^2 = 1, so by the uniqueness of mgfs the distribution is Normal with mean 0 and variance 1, ie N(0, 1).

(c) $E\left[Y^k\right] = E\left[e^{kX}\right] = M_X(k)$.

$\therefore E[Y] = M_X(1) = e^{\mu + \frac{1}{2}\sigma^2}$.

Also, $E\left[Y^2\right] = M_X(2) = e^{2\mu + 2\sigma^2}$, and so

$$\text{Var}(Y) = e^{2\mu+2\sigma^2} - \left(e^{\mu+\frac{1}{2}\sigma^2}\right)^2 = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right).$$

(i) For U(0, 1), the pdf and cdf are $f(x) = 1$ and $F(x) = x$ (for $0 < x < 1$).

If $X_{(j)}$ is the jth order statistic, $j - 1$ of the X values are below it and $n - j$ are above it. Hence its pdf $g(x_{(j)})$ is (for the general case of pdf $f(x)$ and cdf $F(x)$)

$$g(x_{(j)}) = \frac{n!}{(j-1)!(n-j)!}\{F(x)\}^{j-1} f(x) \{1-F(x)\}^{n-j},$$

using the multinomial distribution.

Thus for U(0, 1) we have

$$g\left(x_{(j)}\right) = \frac{n!}{(j-1)!(n-j)!}x^{j-1}(1-x)^{n-j} \qquad \text{(for } 0 < x < 1\text{)}.$$

$$\therefore E\left[X_{(j)}\right] = \frac{n!}{(j-1)!(n-j)!}\int_0^1 x.x^{j-1}(1-x)^{n-j}\,dx$$

$$= \frac{n!}{(j-1)!(n-j)!}\frac{j!(n-j)!}{(n+1)!} \qquad \text{(using the result given in the question)}$$

$$= \frac{j}{n+1}.$$

Similarly,

$$E\left[\left(X_{(j)}\right)^2\right] = \frac{n!}{(j-1)!(n-j)!}\int_0^1 x^2 x^{j-1}(1-x)^{n-j}\,dx$$

$$= \frac{n!}{(j-1)!(n-j)!}\frac{(j+1)!(n-j)!}{(n+2)!} = \frac{j(j+1)}{(n+1)(n+2)}.$$

Hence

$$\text{Var}\left(X_{(j)}\right) = \frac{j(j+1)}{(n+1)(n+2)} - \frac{j^2}{(n+1)^2} = \frac{j\left(n-j+1\right)}{\left(n+1\right)^2\left(n+2\right)}.$$

Solution continued on next page

(ii) When n is odd, the sample median is $X_{\left(\frac{n+1}{2}\right)}$ and so E(median) = $\frac{1}{2}$ and

$$\text{Var(median)} = \frac{\left((n+1)/2\right)\left((n+1)/2\right)}{(n+1)^2(n+2)} = \frac{1}{4(n+2)}.$$

(iii) Using the multinomial distribution, there is one observation below $X_{(2)}$, one each at $X_{(2)}$ and $X_{(3)}$, and one above $X_{(3)}$, so the joint pdf is

$$g(x_2, x_3) = \frac{4!}{1!1!1!1!} x_2 (1 - x_3) = 24 x_2 (1 - x_3), \quad \text{for } 0 < x_2 < x_3 < 1.$$

$$\therefore E\left[X_{(2)} X_{(3)}\right] = 24 \int_0^1 \int_0^{x_3} x_2^{\,2} x_3 (1 - x_3)\, dx_2 dx_3$$

$$= 24 \int_0^1 x_3 (1 - x_3) \int_0^{x_3} x_2^{\,2} dx_2 dx_3$$

$$= 24 \int_0^1 x_3 (1 - x_3) . \frac{x_3^{\,3}}{3} dx_3 = 8 \int_0^1 x_3^4 (1 - x_3)\, dx_3 = 8 . \frac{4!1!}{6!} = \frac{4}{15}.$$

(iv) When n = 4, the sample median is $M = \frac{1}{2}\left(X_{(2)} + X_{(3)}\right)$.

$$E[M] = \frac{1}{2}\left(E\left[X_{(2)}\right] + E\left[X_{(3)}\right]\right) = \frac{1}{2}\left(\frac{2}{5} + \frac{3}{5}\right) = \frac{1}{2}, \quad \text{using part (i).}$$

$$\text{Cov}\left(X_{(2)}, X_{(3)}\right) = E\left[X_{(2)} X_{(3)}\right] - E\left[X_{(2)}\right] E\left[X_{(3)}\right]$$

$$= \frac{4}{15} - \left(\frac{2}{5} \times \frac{3}{5}\right) = \frac{20 - 18}{75} = \frac{2}{75}.$$

Thus

$$\text{Var}(M) = \frac{1}{4}\left[\text{Var}\left(X_{(2)}\right) + \text{Var}\left(X_{(3)}\right) + 2\text{Cov}\left(X_{(2)}, X_{(3)}\right)\right]$$

$$= \frac{1}{4}\left(\frac{1}{25} + \frac{1}{25} + \frac{4}{75}\right) = \frac{10}{4 \times 75} = \frac{1}{30}.$$

Graduate Diploma, Statistical Theory & Methods, Paper I, 2007. Question 7

(i) The pdf and cdf of U are $f_U(u) = 1$ and $F_U(u) = u$ (for $0 < u < 1$).

$X = -(1/\lambda)\log(1 - U)$. So the cdf of X is

$$F_X(x) = P(X \le x) = P\left(-\frac{1}{\lambda}\log(1-U) \le x\right) = P\left(U \le 1 - e^{-\lambda x}\right) = 1 - e^{-\lambda x},$$

for $x > 0$.

So the pdf of X is $f_X(x) = \lambda e^{-\lambda x}$ (for $x > 0$), ie X is exponential with parameter λ. Thus its mean is $1/\lambda$.

(ii) $f_Y(y) = \lambda^2 y e^{-\lambda y}$ for $y > 0$. So the cdf of Y is (for $y > 0$)

$$F_Y(y) = \lambda^2 \int_0^y t e^{-\lambda t} dt = \lambda^2 \left\{ \left[t \frac{e^{-\lambda t}}{-\lambda} \right]_0^y - \int_0^y \frac{e^{-\lambda t}}{-\lambda} dt \right\}$$

$$= \lambda^2 \left\{ -\frac{y e^{-\lambda y}}{\lambda} - \left[\frac{1}{\lambda^2} e^{-\lambda t} \right]_0^y \right\} = 1 - e^{-\lambda y} - \lambda y e^{-\lambda y}.$$

The inverse cdf method (probability integral transform) is not suitable because F_Y^{-1} cannot be found in closed form.

However, this gamma random variable is the sum of two independent exponential random variables, each of which could be simulated as in (i) and the results added.

(iii) Start at a randomly chosen point in the table and use digits in sets of 3. This gives U(0, 1) random variates. Suppose we obtain 566112799517 from the table, then 0.566, 0.112, 0.799, 0.517 are the U(0, 1) variates.

From Table 1, the cdf of B(15, 0.25) is

x	0	1	2	3	4	5	…
$F_X(x)$	0.0134	0.0802	0.2361	0.4613	0.6865	0.8516	…

$u_1 = 0.566$ is above 0.4613 and below 0.6865, so it gives $x_1 = 4$. Similarly, $u_2 = 0.112$ gives $x_2 = 2$; $u_3 = 0.799$ gives $x_3 = 5$; $u_4 = 0.517$ gives $x_4 = 4$. Hence the pseudo-random binomial variates are 2, 4, 4, 5.

Graduate Diploma, Statistical Theory & Methods, Paper I, 2007. Question 8

(i) Take the states of the model as

0: basic premium
1: small reduction
2: larger reduction
3: even larger reduction.

Assuming that the probability of making a claim in any year is θ, independent of all other years, the conditions for a Markov chain apply.

The transition probabilities are

$p_{00} = p_{10} = p_{20} = p_{30} = \theta,$ when a claim is made.

$p_{01} = p_{12} = p_{23} = p_{33} = 1-\theta,$ when no claim is made.

All other transition probabilities are 0.

$$\therefore \mathbf{P} = \begin{bmatrix} \theta & 1-\theta & 0 & 0 \\ \theta & 0 & 1-\theta & 0 \\ \theta & 0 & 0 & 1-\theta \\ \theta & 0 & 0 & 1-\theta \end{bmatrix}.$$

(ii)

$$\mathbf{P}^2 = \begin{bmatrix} \theta & \theta(1-\theta) & (1-\theta)^2 & 0 \\ \theta & \theta(1-\theta) & 0 & (1-\theta)^2 \\ \theta & \theta(1-\theta) & 0 & (1-\theta)^2 \\ \theta & \theta(1-\theta) & 0 & (1-\theta)^2 \end{bmatrix}$$

$$\mathbf{P}^3 = \begin{bmatrix} \theta & \theta(1-\theta) & \theta(1-\theta)^2 & (1-\theta)^3 \\ \theta & \theta(1-\theta) & \theta(1-\theta)^2 & (1-\theta)^3 \\ \theta & \theta(1-\theta) & \theta(1-\theta)^2 & (1-\theta)^3 \\ \theta & \theta(1-\theta) & \theta(1-\theta)^2 & (1-\theta)^3 \end{bmatrix}$$

For both (a) and (b) the probability is $\theta(1 - \theta)$.

It would appear that the probability of going from state 1 back to state 1 is $\theta(1 - \theta)$ for any number of steps greater than 1.

Solution continued on next page

(iii) Denoting the stationary distribution by $\boldsymbol{\pi} = [\pi_0 \quad \pi_1 \quad \pi_2 \quad \pi_3]$, we have

$$\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$$

and so

$$[\pi_0 \quad \pi_1 \quad \pi_2 \quad \pi_3]\begin{bmatrix} \theta & 1-\theta & 0 & 0 \\ \theta & 0 & 1-\theta & 0 \\ \theta & 0 & 0 & 1-\theta \\ \theta & 0 & 0 & 1-\theta \end{bmatrix} = [\pi_0 \quad \pi_1 \quad \pi_2 \quad \pi_3] .$$

Therefore $\pi_0 = \theta(\pi_0 + \pi_1 + \pi_2 + \pi_3) = \theta$ since $\sum \pi_i$ must be 1.

$\pi_1 = \pi_0(1-\theta) = \theta(1-\theta)$.

$\pi_2 = \pi_1(1-\theta) = \theta(1-\theta)^2$.

$\pi_3 = (\pi_2 + \pi_3)(1-\theta)$, giving $\theta\pi_3 = (1-\theta)\pi_2$ so that $\pi_3 = (1-\theta)^3$.