THE ROYAL STATISTICAL SOCIETY

2008 EXAMINATIONS – SOLUTIONS

HIGHER CERTIFICATE

(MODULAR FORMAT)

MODULE 5

FURTHER PROBABILITY AND INFERENCE

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. $\log_{10}$.

Higher Certificate, Module 5, 2008. Question 1

(i) $P(X = x, Y = y) = P(Y = y | X = x)P(X = x)$.

Table of $P(X = x,\ Y = y)$.

		Values of Y				
		1	2	3	4	Total
Values of X	1	1/16	1/16	1/16	1/16	1/4
	2		1/12	1/12	1/12	1/4
	3			1/8	1/8	1/4
	4				1/4	1/4
	Total	3/48	7/48	13/48	25/48	1

(ii) The marginal probability distribution of Y is as follows, copied from the margin of the table above.

$$P(Y=1) = \frac{3}{48}\left(=\frac{1}{16}\right); \quad P(Y=2) = \frac{7}{48}; \quad P(Y=3) = \frac{13}{48}; \quad P(Y=4) = \frac{25}{48}.$$

$$E(Y) = \left(1\times\frac{3}{48}\right)+\left(2\times\frac{7}{48}\right)+\left(3\times\frac{13}{48}\right)+\left(4\times\frac{25}{48}\right) = \frac{156}{48} = \frac{13}{4}.$$

$$E(Y^2) = \left(1\times\frac{3}{48}\right)+\left(4\times\frac{7}{48}\right)+\left(9\times\frac{13}{48}\right)+\left(16\times\frac{25}{48}\right) = \frac{548}{48} = \frac{137}{12}.$$

$$\therefore \text{Var}(Y) = \frac{137}{12}-\left(\frac{13}{4}\right)^2 = \frac{41}{48}.$$

(iii) $$E(XY) = \left(1\times\frac{1}{16}\right)+\left(2\times\frac{1}{16}\right)+\left(3\times\frac{1}{16}\right)+\left(4\times\frac{1}{16}\right)+\left(4\times\frac{1}{12}\right)+\left(6\times\frac{1}{12}\right)$$

$$+\left(8\times\frac{1}{12}\right)+\left(9\times\frac{1}{8}\right)+\left(12\times\frac{1}{8}\right)+\left(16\times\frac{1}{4}\right)$$

$$= \frac{1}{48}(3+6+9+12+16+24+32+54+72+192) = \frac{420}{48}.$$

Solution continued on next page

$$E(X) = (1+2+3+4)\times\frac{1}{4} = \frac{10}{4}$$ (and $E(Y)$ = 13/4, see above).

$$\therefore \mathrm{Cov}(X,Y) = \frac{420}{48} - \left(\frac{10}{4}\times\frac{13}{4}\right) = \frac{1}{48}(420-390) = \frac{30}{48} = \frac{5}{8}.$$

(iv) $U = X + Y$.

$$P(U=2) = P(X=1, Y=1) = \frac{1}{16}$$

$$P(U=3) = P(X=1, Y=2) = \frac{1}{16}$$

$$P(U=4) = P(X=1, Y=3) + P(X=2, Y=2) = \frac{1}{16} + \frac{1}{12} = \frac{7}{48}$$

$$P(U=5) = P(X=1, Y=4) + P(X=2, Y=3) = \frac{1}{16} + \frac{1}{12} = \frac{7}{48}$$

$$P(U=6) = P(X=2, Y=4) + P(X=3, Y=3) = \frac{1}{12} + \frac{1}{8} = \frac{10}{48} = \frac{5}{24}$$

$$P(U=7) = P(X=3, Y=4) = \frac{1}{8}$$

$$P(U=8) = P(X=4, Y=4) = \frac{1}{4}$$

No other values of U have non-zero probability.

Higher Certificate, Module 5, 2008. Question 2

Probability generating function, $\pi(t) = E\left(t^X\right)$.

Moment generating function, $m(t) = E\left(e^{tX}\right)$.

Relationship: $m(t) = \pi\left(e^t\right)$.

(i) $\pi(t) = \sum_h t^h P(X = h) = \sum_{h=0}^{n} t^h \binom{n}{h} p^h (1-p)^{n-h}$

$= \sum_{h=0}^{n} \binom{n}{h} (pt)^h (1-p)^{n-h} = (pt + 1 - p)^n$ (using the binomial theorem).

(ii) $E(X) = \left.\frac{d\pi}{dt}\right|_{t=1}$. $\frac{d\pi}{dt} = np(pt + 1 - p)^{n-1}$, $\therefore E(X) = np$.

$E(X(X-1)) = \left.\frac{d^2\pi}{dt^2}\right|_{t=1}$. $\frac{d^2\pi}{dt^2} = n(n-1)p^2(pt + 1 - p)^{n-2}$,

$\therefore E(X(X-1)) = n(n-1)p^2$.

$\therefore E\left(X^2\right) = n(n-1)p^2 + E(X) = n(n-1)p^2 + np$.

$\therefore \text{Var}(X) = n(n-1)p^2 + np - n^2p^2 = n^2p^2 - np^2 + np - n^2p^2 = np(1-p)$.

[Alternatively, could directly use Var(X) = $\pi''(1) + E(X)(1 - E(X))$.]

(iii) $E(X(X-1)(X-2)) = \left.\frac{d^3\pi}{dt^3}\right|_{t=1}$. $\frac{d^3\pi}{dt^3} = n(n-1)(n-2)p^3(pt + 1 - p)^{n-3}$,

$\therefore E(X(X-1)(X-2)) = n(n-1)(n-2)p^3$.

$\therefore E\left(X^3\right) - 3E\left(X^2\right) + 2E(X) = n(n-1)(n-2)p^3$,

$\therefore E\left(X^3\right) = n(n-1)(n-2)p^3 + 3n(n-1)p^2 + 3np - 2np$

$= n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np$.

(iv) $\pi_{X_i}(t) = (pt + 1 - p)^{n_i}$ $(i = 1, 2, ..., m)$.

$\therefore \pi_Y(t) = \prod_{i=1}^{m} (pt + 1 - p)^{n_i} = (pt + 1 - p)^{\Sigma n_i}$, which is the pgf of $\text{B}\left(\sum n_i, p\right)$.

$\therefore$ by the 1-1 correspondence between pgfs and distributions, $Y \sim \text{B}\left(\sum n_i, p\right)$.

<u>Higher Certificate, Module 5, 2008. Question 3</u>

(i) $E(X) = \lambda^{-2}\int_0^\infty x^2 e^{-x/\lambda}dx$

$= \lambda^{-2}\left[-\lambda x^2 e^{-x/\lambda}\right]_0^\infty + 2\lambda\int_0^\infty \lambda^{-2} x e^{-x/\lambda}dx$

Note that the second integral is simply the integral of the pdf

$= [0 - 0] + (2\lambda \times 1) = 2\lambda.$

$\therefore$ the method of moments estimator $\hat{\lambda}$ satisfies $2\hat{\lambda} = \bar{X}$. $\therefore \hat{\lambda} = \frac{1}{2}\bar{X}$.

(ii) $E(\bar{X}) = E(X) = 2\lambda$.

$\therefore E(\hat{\lambda}) = \frac{1}{2}E(\bar{X}) = \lambda$ for all λ, i.e. $\hat{\lambda}$ is unbiased for λ.

$E(X^2) = \lambda^{-2}\int_0^\infty x^3 e^{-x/\lambda}dx = \lambda^{-2}\left[-\lambda x^3 e^{-x/\lambda}\right]_0^\infty + 3\lambda\int_0^\infty \lambda^{-2}x^2 e^{-x/\lambda}dx$

$= 3\lambda E(X) = 6\lambda^2$.

$\therefore \text{Var}(X) = E(X^2) - (E(X))^2 = 6\lambda^2 - 4\lambda^2 = 2\lambda^2$.

$\therefore \text{Var}(\bar{X}) = \frac{\text{Var}(X)}{n} = \frac{2\lambda^2}{n}$.

$\therefore \text{Var}(\hat{\lambda}) = \text{Var}\left(\frac{\bar{X}}{2}\right) = \frac{1}{4}\text{Var}(\bar{X}) = \frac{\lambda^2}{2n}$.

As $\hat{\lambda}$ is unbiased and $\text{Var}(\hat{\lambda}) \to 0$ as $n \to \infty$, $\hat{\lambda}$ is consistent.

(iii) For $n = 3$, $\text{Var}(\hat{\lambda}) = \frac{\lambda^2}{2\times 3} = \frac{\lambda^2}{6}$.

$\text{Var}(\tilde{\lambda}) = \frac{1}{64}\text{Var}(X_1) + \frac{1}{16}\text{Var}(X_2) + \frac{1}{64}\text{Var}(X_3) = \frac{3}{32}\text{Var}(X) = \frac{3\lambda^2}{16}$.

$\therefore$ relative efficiency of $\tilde{\lambda} = \frac{\text{Var}(\hat{\lambda})}{\text{Var}(\tilde{\lambda})} = \frac{\lambda^2}{6}\times\frac{16}{3\lambda^2} = \frac{8}{9}$.

As the relative efficiency is less than one, $\hat{\lambda}$ is preferred.

Higher Certificate, Module 5, 2008. Question 4

(i) $f\left(x_i\right) = \left(2\pi\theta\right)^{-1/2} e^{-x_i^2/2\theta}$ (for $-\infty < x_i < \infty$).

Likelihood $L(\theta) = \prod_{i=1}^{n}\left\{\left(2\pi\theta\right)^{-1/2} e^{-x_i^2/2\theta}\right\} = \left(2\pi\theta\right)^{-n/2} e^{-\Sigma x_i^2/2\theta}$.

(ii) $\log\left(L(\theta)\right) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\theta) - \frac{\Sigma x_i^2}{2\theta}$

$\frac{d\log L}{d\theta} = -\frac{n}{2\theta} + \frac{\Sigma x_i^2}{2\theta^2}$ which on setting equal to zero gives solution $\hat{\theta} = \frac{\Sigma x_i^2}{n}$.

To investigate whether this is a maximum, consider $\frac{d^2\log L}{d\theta^2} = \frac{n}{2\theta^2} - \frac{\Sigma x_i^2}{\theta^3}$.

Inserting $\theta = \hat{\theta}$ gives $\frac{d^2\log L}{d\theta^2} = \frac{n}{2\hat{\theta}^2} - \frac{n\hat{\theta}}{\hat{\theta}^3} = -\frac{n}{2\hat{\theta}^2} < 0$.

$\therefore \hat{\theta} = \Sigma x^2_i / n$ maximises $\log L(\theta)$; thus $\Sigma X_i^2 / n$ is the maximum likelihood estimator of θ.

(iii) $E\left(-\frac{d^2\log L}{d\theta^2}\right) = -\frac{n}{2\theta^2} + \frac{\Sigma E\left(X_i^2\right)}{\theta^3}$.

As the mean is 0, we have $\theta = \text{Var}(X) = E(X^2)$.

$\therefore E\left(-\frac{d^2\log L}{d\theta^2}\right) = -\frac{n}{2\theta^2} + \frac{n\theta}{\theta^3} = \frac{n}{2\theta^2}$

$\therefore$ For large n, $\hat{\theta} \sim \text{N}\left(\theta,\ \frac{2\theta^2}{n}\right)$, approximately.

(iv) $\hat{\theta} = \frac{1000}{100} = 10$.

$\therefore$ approximate 95% confidence interval is given by $10 \pm 1.96\sqrt{\frac{2\times 10^2}{100}}$

i.e. it is $10 \pm 1.96\sqrt{2}$, i.e. (7.23, 12.77).