

THE ROYAL STATISTICAL SOCIETY

2010 EXAMINATIONS – SOLUTIONS

HIGHER CERTIFICATE

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, 2010. Question 1

- (i)
- $$P(X = 0, Y = 0) = P(Y = 0 | X = 0)P(X = 0) = 1 \times 0.4 = 0.4$$
- $$P(X = 1, Y = 0) = P(Y = 0 | X = 1)P(X = 1) = 0.6 \times 0.3 = 0.18$$
- $$P(X = 1, Y = 1) = P(Y = 1 | X = 1)P(X = 1) = 0.4 \times 0.3 = 0.12$$
- $$P(X = 2, Y = 0) = P(Y = 0 | X = 2)P(X = 2) = 0.6^2 \times 0.2 = 0.072$$
- $$P(X = 2, Y = 1) = P(Y = 1 | X = 2)P(X = 2) = 2 \times 0.6 \times 0.4 \times 0.2 = 0.096$$
- $$P(X = 2, Y = 2) = P(Y = 2 | X = 2)P(X = 2) = 0.4^2 \times 0.2 = 0.032$$
- $$P(X = 3, Y = 0) = P(Y = 0 | X = 3)P(X = 3) = 0.6^3 \times 0.1 = 0.0216$$
- $$P(X = 3, Y = 1) = P(Y = 1 | X = 3)P(X = 3) = 3 \times 0.6^2 \times 0.4 \times 0.1 = 0.0432$$
- $$P(X = 3, Y = 2) = P(Y = 2 | X = 3)P(X = 3) = 3 \times 0.6 \times 0.4^2 \times 0.1 = 0.0288$$
- $$P(X = 3, Y = 3) = P(Y = 3 | X = 3)P(X = 3) = 0.4^3 \times 0.1 = 0.0064$$

All other entries in the two-way table are zero. Of course, it is not necessary to calculate all the above explicitly – many can be deduced by subtraction. The table is as follows.

		Values of X				Total
		0	1	2	3	
Values of Y	0	0.4	0.18	0.072	0.0216	0.6736
	1	0	0.12	0.096	0.0432	0.2592
	2	0	0	0.032	0.0288	0.0608
	3	0	0	0	0.0064	0.0064
Total		0.4	0.3	0.2	0.1	

- (ii) The marginal distribution of Y is in the Total column in the table in part (i):
 $P(Y = 0) = 0.6736$ $P(Y = 1) = 0.2592$ $P(Y = 2) = 0.0608$ $P(Y = 3) = 0.0064$.

$$\therefore E(Y) = (0 \times 0.6736) + (1 \times 0.2592) + (2 \times 0.0608) + (3 \times 0.0064) = 0.4.$$

- (iii) $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$. From the table, we have

$$E(X) = (0 \times 0.4) + (1 \times 0.3) + (2 \times 0.2) + (3 \times 0.1) = 1.0,$$

$$E(XY) = (0 \times 0 \times 0.4) + \dots + (3 \times 3 \times 0.0064) = 0.8.$$

$$\therefore \text{Cov}(X, Y) = 0.8 - (1.0 \times 0.4) = 0.4.$$

- (iv) We need to find the probability distribution of $X - Y$, which we here call U .

$$P(U = 0) = P(X = Y) = 0.4 + 0.12 + 0.032 + 0.0064 = 0.5584$$

$$P(U = 1) = P(X - Y = 1) = 0.18 + 0.096 + 0.0288 = 0.3048$$

$$P(U = 2) = P(X - Y = 2) = 0.072 + 0.0432 = 0.1152$$

$$P(U = 3) = P(X - Y = 3) = 0.0216$$

$$[P(U = k) = 0 \text{ for all other values of } k.]$$

Higher Certificate, Module 5, 2010. Question 2

$$f(x, y) = \frac{1}{2\pi} \exp\left(-\frac{1}{4}(x-1)^2 - (y - \frac{1}{4}(1+x))^2\right), \quad -\infty < x < \infty, \quad -\infty < y < \infty.$$

(i) $f(x) = \int_{y=-\infty}^{\infty} f(x, y) dy = \frac{1}{2\pi} \exp\left(-\frac{1}{4}(x-1)^2\right) \int_{y=-\infty}^{\infty} \exp\left(-\left(y - \frac{1}{4}(1+x)\right)^2\right) dy$

The method is to express the integral as the integral of a Normal pdf

$$= \frac{1}{2\pi} \exp\left(-\frac{1}{4}(x-1)^2\right) \sqrt{\pi} \int_{y=-\infty}^{\infty} \frac{1}{\sqrt{2\pi \times \frac{1}{2}}} \exp\left(-\frac{\left(y - \frac{1}{4}(1+x)\right)^2}{2 \times \frac{1}{2}}\right) dy$$

Note that the integrand is the pdf of $N\left(\frac{1+x}{4}, \frac{1}{2}\right)$

$$= \frac{1}{\sqrt{2\pi \times 2}} \exp\left(-\frac{(x-1)^2}{2 \times 2}\right) \quad (\text{for } -\infty < x < \infty),$$

which is the pdf of $N(1, 2)$, so we have $X \sim N(1, 2)$ as required.

(ii) The moment generating function of X is

$$\begin{aligned} m_X(t) &= \int_{-\infty}^{\infty} e^{tx} f(x) dx = \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{1}{4}(x-1)^2} dx \\ &= \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{4}(x^2 - 2x + 1 - 4tx)} dx \\ &= \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{4}(x - (1+2t))^2} e^{\frac{1}{4}((1+2t)^2 - 1)} dx \\ &= e^{t+t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \times 2}} e^{-\frac{1}{2 \times 2}(x - (1+2t))^2} dx = e^{t+t^2}. \end{aligned}$$

Note that the integrand is the pdf of $N(1+2t, 2)$

Solution continued on next page

(iii) Differentiating the moment generating function,

$$\frac{dm_X(t)}{dt} = (1+2t)e^{t+t^2},$$

$$\frac{d^2m_X(t)}{dt^2} = (1+2t)^2 e^{t+t^2} + 2e^{t+t^2},$$

$$\frac{d^3m_X(t)}{dt^3} = (1+2t)^3 e^{t+t^2} + 4(1+2t)e^{t+t^2} + 2(1+2t)e^{t+t^2}.$$

$$\therefore E(X^3) = \left. \frac{d^3m_X(t)}{dt^3} \right|_{t=0} = 1+4+2 = 7.$$

[Note. The result may also be obtained in a straightforward way via the power series expansion of the moment generating function.]

Higher Certificate, Module 5, 2010. Question 3

(i) The likelihood is $L(\beta) = \prod_{i=1}^n (\beta(1-x_i)^{\beta-1}) = \beta^n \prod_{i=1}^n (1-x_i)^{\beta-1}$.

\therefore Log likelihood is $\log L(\beta) = n \log \beta - (\beta-1) \sum \log(1-x_i)$.

$$\frac{d \log L}{d \beta} = \frac{n}{\beta} + \sum \log(1-x_i) \text{ which on setting equal to zero gives solution}$$

$$\hat{\beta} = -\frac{n}{\sum \log(1-x_i)}.$$

To investigate whether this is a maximum, consider $\frac{d^2 \log L}{d \beta^2} = -\frac{n}{\beta^2} < 0$, so this is the maximum likelihood estimator of β .

(ii) We have $E\left(-\frac{d^2 \log L}{d \beta^2}\right) = \frac{n}{\beta^2}$.

Thus, using the usual asymptotic result, $\text{Var}(\hat{\beta}) \approx \frac{1}{n/\beta^2} = \frac{\beta^2}{n}$.

Thus, for large n , $\hat{\beta} \sim N\left(\beta, \frac{\beta^2}{n}\right)$, approximately.

Thus an approximate 95% confidence interval for β is given by $\hat{\beta} \pm 1.96 \frac{\hat{\beta}}{\sqrt{n}}$.

[Note. The pivotal quantity method could also be used to find an alternative.]

(iii) $P(X_i < 0.5) = \int_0^{0.5} \beta(1-x)^{\beta-1} dx = \left[-(1-x)^\beta\right]_0^{0.5} = -0.5^\beta + 1 = 1 - 0.5^\beta$.

(iv) $Y \sim B(n, 1 - 0.5^\beta)$.

So $P(Y = y) = \binom{n}{y} (1 - 0.5^\beta)^y (0.5^\beta)^{n-y}$ and the likelihood is simply

$$\binom{n}{y} (1 - 0.5^\beta)^y 0.5^{\beta(n-y)}.$$

Higher Certificate, Module 5, 2010. Question 4

Part (a)

The bias of an estimator $\hat{\theta}$ of a parameter θ is defined as $E(\hat{\theta} - \theta)$.

An unbiased estimator is one for which $E(\hat{\theta} - \theta) = 0$ for all θ .

If the bias of an estimator is non-zero and independent estimates are made using it based on different samples, the average of these estimates will not tend to the true value (θ), no matter how many samples are taken. Other things being equal, therefore, unbiased estimators are preferred. However, there are situations where a biased estimator has smaller variance than the best unbiased estimator, and would therefore be preferred (at least if the bias is small). Indeed, in some situations there is no unbiased estimator at all.

Among unbiased estimators, the most accurate is the one with the smallest variance. The relative efficiency of one unbiased estimator, $\hat{\theta}_1$, compared with another, $\hat{\theta}_2$, is

$$\frac{\text{Var}(\hat{\theta}_2)}{\text{Var}(\hat{\theta}_1)}$$

(sometimes multiplied by 100 to make it a percentage). A value greater than 1 (or 100%) suggests that $\hat{\theta}_1$ is better; a value less than 1 suggests that $\hat{\theta}_2$ is better.

It can be shown that, under regularity conditions, the variance of an unbiased estimator cannot be less than

$$-\frac{1}{E\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)} \quad (= V, \text{ say}),$$

where L is the likelihood. The efficiency of an unbiased estimator $\hat{\theta}$ is $\frac{V}{\text{Var}(\hat{\theta})}$

(again sometimes multiplied by 100 to make it a percentage). A value of 1 (or 100%) states that $\hat{\theta}$ is the best unbiased estimator.

Solution continued on next page

Part (b)

(i) For each Y_i we have

$$E(Y) = \left((-2) \times \frac{1}{2} (1-p)^2 \right) + (0 \times p(1-p)) + (1 \times p) + \left(2 \times \frac{1}{2} (1-p)^2 \right) = p.$$

Therefore the method of moments estimator \tilde{p} satisfies $\bar{Y} = \tilde{p}$, i.e. \tilde{p} is simply \bar{Y} .

An obvious unsatisfactory feature is that, although p must lie between 0 and 1, the value of \tilde{p} can be outside this range.

(ii)
$$E(Y_i^2) = \left(4 \times \frac{1}{2} (1-p)^2 \right) + (0 \times p(1-p)) + (1 \times p) + \left(4 \times \frac{1}{2} (1-p)^2 \right)$$
$$= 4(1-p)^2 + p = 4p^2 - 7p + 4.$$

$$\therefore \text{Var}(Y_i) = E(Y_i^2) - \{E(Y_i)\}^2 = 4p^2 - 7p + 4 - p^2 = 3p^2 - 7p + 4.$$

$$\therefore \text{Var}(\tilde{p}) = \text{Var}(\bar{Y}) = \frac{3p^2 - 7p + 4}{n}.$$

(iii) From part (ii), $E\left(\frac{1}{n} \sum Y_i^2\right) = 4p^2 - 7p + 4$ and so $E\left(\frac{1}{4n} \sum Y_i^2\right) = p^2 - \frac{7}{4}p + 1$.

$$\therefore E\left(\frac{1}{4n} \sum Y_i^2 + \frac{7}{4n} \sum Y_i - 1\right) = p^2.$$

So $\frac{1}{4n} \sum Y_i^2 + \frac{7}{4n} \sum Y_i - 1$ is an unbiased estimator of p^2 .