Revision 2

Question 1 (taken from Question 5 Exam 2011)

(i) Two football teams, Team X and Team Y, are about to play a match against each other. Let X and Y denote the number of goals scored by Team X and Team Y, respectively. The joint probability function of X and Y is:

	Goals	Y 0	1	2	3
X	0	0.25	0.12	0.05	0.02
	1	0.10	0.09	0.06	0.02
	2	0.07	0.06	0.04	0.03
	3	0.02	0.02	0.03	0.02

(a) What is the probability that Team X wins the match? What is the probability of a draw?

(b) Find the marginal distributions of X and Y, and compute $E(X)$, $E(Y)$.

(c) Compute $\text{Var}(X)$, $\text{Var}(Y)$ and $\text{Cov}(X, Y)$. Are X and Y uncorrelated?

(d) Write down the conditional distribution of Y given that Team X does not score. Are X and Y independent?

(ii) A system consists of k independent components in parallel, i.e. it functions as long as at least one component functions. The components are unreliable: they fail with probability 0.20.

(a) What is the probability that the system functions?

(b) What is the minimum number of components required to ensure that the probability the system fails is less than 1 in 1000 (i.e. 0.1%)?

(iii) Let T be a random variable with range $[1, \infty)$ and hazard rate

$$h(t) = \frac{\alpha}{t}$$

for some $\alpha > 0$.

Find the cumulative hazard function, the cumulative distribution function (CDF) and the probability density function (PDF) of T.

Question 2 (taken from Question 4 Exam 2011)

The continuous random variable Y has probability density function (PDF) given by

$$f_Y(y) = \begin{cases} \theta^{-2} y e^{-y/\theta} & y > 0 \\ 0 & \text{otherwise} \end{cases}.$$

(i) Verify that $f_Y(y)$ is a valid PDF. *(Hint: You need to integrate by parts.)*

(ii) Find the moment generating function of Y and hence, or otherwise, compute $E(Y)$ and $\text{Var}(Y)$.

(iii) Suppose that we draw a random sample $Y_1, Y_2, \ldots, Y_n$ from this distribution. Find the method of moments estimator of θ. Is it unbiased?

(iv) Show that the maximum-likelihood estimator of θ is $\hat{\theta} = \bar{Y}/2$ and compute its mean square error.

Solution 2

Question 1

(a) i.

$$\begin{aligned} P(\text{‘Team X wins’}) &= \sum_{x>y} p_{X,Y}(x,y) \\ &= 0.10 + 0.07 + 0.02 + 0.06 + 0.02 + 0.03 = 0.3 \end{aligned}$$

$$\begin{aligned} P(\text{‘draw’}) &= \sum_{x=y} p_{X,Y}(x,y) \\ &= 0.25 + 0.09 + 0.04 + 0.02 = 0.4 \end{aligned}$$

Unseen — 2 MARKS

ii.

$$\begin{aligned} p_X(0) &= \sum_{y} p_{X,Y}(0,y) \\ &= 0.25 + 0.12 + 0.05 + 0.02 = 0.44 \end{aligned}$$

and similarly for the other of the values of $p_X(x)$, $p_Y(y)$. The resulting marginals are

x	0	1	2	3
$p_X(x)$	0.44	0.27	0.20	0.09

and

y	0	1	2	3
$p_Y(y)$	0.44	0.29	0.18	0.09

Directly from the marginal:

$$E(Y) = 0.44 \times 0 + 0.29 \times 1 + 0.18 \times 2 + 0.09 \times 3 = 0.92$$

and, similarly, $E(X) = 0.94$.

Unseen — 4 MARKS

iii. From the marginal distribution of Y:

$$E(Y^2) = 0.44 \times 0^2 + 0.29 \times 1^2 + 0.18 \times 2^2 + 0.09 \times 3^2 = 1.82\,,$$

so the variance is

$$\mathrm{Var}(Y) = E(Y^2) - E(Y)^2 = 1.82 - 0.92^2 = 0.9736\,.$$

Similarly, $E(X^2) = 1.88$ and $\mathrm{Var}(X) = 0.9964$.
We have:

$$\begin{aligned} E(XY) &= \sum_{x,y} xy\,p_{X,Y}(x,y) \\ &= 0 \times 0 \times 0.25 + 1 \times 0 \times 0.10 + \ldots \\ &= 1.15 \end{aligned}$$

and so

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0.2852\,.$$

X,Y are *not* uncorrelated, because $\text{Cov}(X,Y) \neq 0$.

Unseen — 3 MARKS

iv.

$$P(Y=0|X=0) = \frac{p_{X,Y}(0,0)}{p_X(0)} = \frac{0.25}{0.44} = 0.568$$

and similarly for the remaining probabilities. The conditional PMF is given by

y	0	1	2	3
$P(Y=y\|X=0)$	0.568	0.273	0.114	0.045

This is different to the marginal for Y, so the two random variables are not independent. (Or: X,Y are correlated, so they are certainly not independent.)

Unseen — 2 MARKS

(b) A system consists of k components in parallel, i.e. it functions as long as at least one component functions. The components are unreliable: they fail with probability 0.20.

i. The probability is

$$\begin{aligned} P(\text{'system functions'}) &= P(\text{'at least one component functions'}) \\ &= 1 - P(\text{'all components fail'}) = 1 - 0.2^k\,. \end{aligned}$$

Seen similar — 2 MARKS

ii.

$$\begin{aligned} P(\text{'system functions'}) &> 1 - 0.001 \\ 1 - 0.2^k &> 1 - 0.001 \\ 0.2^k &< 0.001 \\ k\log(0.2) &< \log(0.001) \\ -1.609k &< -6.908 \\ k &> \frac{-6.908}{-1.609} = 4.29\,, \end{aligned}$$

so we need $k = 5$ components.

Unseen — 2 MARKS

(c) For $t > 1$, the cumulative hazard is

$$H(t) = \int_0^t h(s)ds = \int_1^t \frac{\alpha}{s} ds = [\alpha \log s]_1^t = \alpha \log t$$

Since

$$\begin{aligned} H(t) &= -\log(R(t)) = -\log(1 - F(t)) \Rightarrow \\ e^{-H(t)} &= 1 - F(t) \Rightarrow \\ F(t) &= 1 - e^{-H(t)}\,, \end{aligned}$$

the CDF is

$$F(t) = 1 - e^{-\alpha \log t} = 1 - t^{-\alpha} ,$$

for $t > 1$ (and $F_X(x) = 0$ otherwise).
The PDF is

$$f(t) = \frac{d}{dt} F(t) = \alpha t^{-\alpha - 1}$$

for $t > 1$ (and $f_X(x) = 0$ otherwise).

Unseen — 5 MARKS

Question 2

(a) Clearly, $f_Y(y) \geq 0$ for all y. We now need to verify that $f_Y(y)$ integrates to 1 over $(-\infty, \infty)$. We have:

$$
\begin{aligned}
\int_{-\infty}^{\infty} f_Y(y)dy &= \int_0^{\infty} \theta^{-2} y e^{-y/\theta} dy \\
&= -\int_0^{\infty} \theta^{-1} y \left(e^{-y/\theta}\right)' dy \\
&= -\left[\theta^{-1} y e^{-y/\theta}\right]_0^{\infty} + \int_0^{\infty} \left(\theta^{-1} y\right)' e^{-y/\theta} dy \\
&= 0 + \int_0^{\infty} \theta^{-1} e^{-y/\theta} dy \\
&= \left[-e^{-y/\theta}\right]_0^{\infty} \\
&= 1,
\end{aligned}
$$

so $f_Y(y)$ is a valid PDF.

Note that theta should be >0 for the integral to converge.

(b) The MGF is

$$
\begin{aligned}
M_Y(t) &= E(e^{tY}) \\
&= \int_{-\infty}^{\infty} e^{ty} f_Y(y) dy \\
&= \int_0^{\infty} e^{ty} \theta^{-2} y e^{-y/\theta} dy \\
&= \theta^{-2} \int_0^{\infty} y e^{-y(1/\theta - t)} dy
\end{aligned}
$$

If we set $\lambda = (1/\theta - t)^{-1}$, we can rearrange the expression inside the integral so that it has the same form as the PDF (with parameter λ) and integrates to 1.

$$
\begin{aligned}
M_Y(t) &= \theta^{-2} \int_0^{\infty} y e^{-y/\lambda} dy \\
&= \theta^{-2} \lambda^2 \int_0^{\infty} \lambda^{-2} y e^{-y/\lambda} dy \\
&= \theta^{-2} (1/\theta - t)^{-2} \\
&= (1 - \theta t)^{-2}
\end{aligned}
$$

Unseen — 4 MARKS

For the integral to converge, we need $\lambda > 0$ or, equivalently, $t < 1/\theta$.

Unseen — 1 MARK

Differentiating the MGF, we find

$$
\begin{aligned}
M_Y'(t) &= \frac{d}{dt}(1 - \theta t)^{-2} = 2\theta(1 - \theta t)^{-3} \\
M_Y''(t) &= \frac{d}{dt} 2\theta(1 - \theta t)^{-3} = 6\theta^2(1 - \theta t)^{-4},
\end{aligned}
$$

so the mean and variance are

$$E(Y) = M'_Y(0) = 2\theta$$

$$\text{Var}(Y) = E(Y^2) - E(Y)^2 = M''_Y(0) - (M'_Y(0))^2 = 6\theta^2 - 4\theta^2 = 2\theta^2$$

Unseen — 3 MARKS

(c) Let $\hat{\theta}_{MM}$ be the method of moments estimator. Set the sample mean equal to the population mean to find

$$\bar{Y} = 2\hat{\theta}_{MM} \Rightarrow \hat{\theta}_{MM} = \frac{\bar{Y}}{2}.$$

Its expectation is

$$E(\hat{\theta}_{MM}) = E(\bar{Y}/2) = \frac{2\theta}{2} = \theta,$$

so it is an unbiased estimator of θ.

Unseen — 4 MARKS

(d) The likelihood function is

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} f_Y(y_i) = \prod_{i=1}^{n} \theta^{-2} y_i e^{-y_i/\theta} \\ &= \theta^{-2n} \left(\prod_{i=1}^{n} y_i \right) e^{-\sum_i y_i/\theta} \\ &= \theta^{-2n} \left(\prod_{i=1}^{n} y_i \right) e^{-n\bar{y}/\theta}. \end{aligned}$$

The log-likelihood is

$$\ell(\theta) = \log(L(\theta)) = -2n \log\theta - \frac{n\bar{y}}{\theta} + C$$

We differentiate with respect to θ and set the derivative equal to 0 to obtain the maximum-likelihood estimator $\hat{\theta}_{ML}$.

$$\frac{d}{d\theta}\ell(\theta) = -\frac{2n}{\theta} + \frac{n\bar{y}}{\theta^2} \Rightarrow -\frac{2n}{\hat{\theta}_{ML}} + \frac{n\bar{Y}}{\hat{\theta}_{ML}^2} = 0 \Rightarrow \hat{\theta}_{ML} = \bar{Y}/2,$$

which is the same as the method of moments estimator. We know that it is unbiased, so its MLE is equal to its variance:

$$\text{MSE}(\hat{\theta}_{ML}) = \text{Var}(\hat{\theta}_{ML}) = \text{Var}(\bar{Y}/2) = \frac{1}{4}\text{Var}(\bar{Y}) = \frac{1}{4}\frac{2\theta^2}{n} = \frac{\theta^2}{2n}$$