PROBABILITY 1 SOLUTIONS

MATH11300

(Paper code MATH-11300J)

January 2018 1 hour 30 minutes

A1 (**8 marks**) [Very standard type of problem, though specific example unseen]

(a) The sample space has $5 \cdot 5 = 25$ elementary outcomes in it.

(b) The outcomes that make up B are $\{Vv, Vw, Vx, Vy, Vz, Wv, Xv, Yv, Zv\}$. There are 9 of these outcomes, the probability is therefore $\mathbb{P}(B) = 9/25$.

(c) $\mathbb{P}(C) = \mathbb{P}(D) = 1/5$ (other draw doesn't matter) and $\mathbb{P}(C \cap D) = 1/25$ (exactly $\{Vv\}$ must be the outcome). Since $B = C \cup D$, we know that

$$\mathbb{P}(B) = \mathbb{P}(C \cup D) = \mathbb{P}(C) + \mathbb{P}(D) - \mathbb{P}(C \cap D) = \frac{5+5-1}{25} = \frac{9}{25}.$$

A2 (**8 marks**) [Very standard type of problem, though specific example unseen]

(a) We write L for the event that Professor X. has a lecture, and T for the event that he is wearing a tie. We are told that $\mathbb{P}(L) = 3/7$ so $\mathbb{P}(L^c) = 4/7$, that $\mathbb{P}(T|L) = 9/10$ and that $\mathbb{P}(T^c|L^c) = 9/10$, so that $\mathbb{P}(T|L^c) = 1/10$. The partition theorem gives

$$\mathbb{P}(T) \quad = \quad \mathbb{P}(L)\mathbb{P}(T|L) + \mathbb{P}(L^c)\mathbb{P}(T|L^c) = \frac{3}{7} \cdot \frac{9}{10} + \frac{4}{7} \cdot \frac{1}{10} = \frac{27+4}{70} = \frac{31}{70}$$

(b) Using Bayes' theorem:

$$\mathbb{P}(L|T) = \frac{\mathbb{P}(L \cap T)}{\mathbb{P}(T)} = \frac{\mathbb{P}(T|L)\mathbb{P}(L)}{\mathbb{P}(T)} = \frac{9/10 \cdot 3/7}{31/70} = \frac{27/70}{31/70} = \frac{27}{31}.$$

A3 (**8 marks**) [Very standard type of problem, though specific example unseen]

We find the conditional distributions: given that no coins are tossed, $\mathbb{P}(Y = 0|X = 0) = 1$. Given that one coin is tossed $\mathbb{P}(Y = 1|X = 1) = \mathbb{P}(Y = 0|X = 1) = 1/2$. Given that two coins are tossed $\mathbb{P}(Y = 0|X = 2) = 1/4$, $\mathbb{P}(Y = 1|X = 2) = 1/2$ and $\mathbb{P}(Y = 2|X = 2) = 1/4$.

Since each of the values of X has probability $1/3$, we can calculate the joint distribution and marginal of Y as

	$Y=0$	$Y=1$	$Y=2$
$X=0$	1/3	0	0
$X=1$	1/6	1/6	0
$X=2$	1/12	1/6	1/12
	7/12	4/12	1/12

The expected value of Y is $\mathbb{E}(Y) = 1 \cdot (4/12) + 2 \cdot (1/12) = 6/12 = 1/2$.

A4 (**8 marks**) [Both parts fairly standard, though care needed with manipulations and to treat sign properly]

(a) From the notes and by independence of X and Y (and hence of X and $-Y$) we know that

$$\begin{aligned}
\mathbb{E}(U) = \mathbb{E}(2X+Y) &= 2\mathbb{E}(X) + \mathbb{E}(Y) = 6 + 3 = 9, \\
\text{Var}\ (U) = \text{Var}\ (2X+Y) &= 4\text{Var}\ (X) + \text{Var}\ (Y) = 36 + 16 = 52, \\
\mathbb{E}(V) = \mathbb{E}(X-2Y) &= \mathbb{E}(X) - 2\mathbb{E}(Y) = 3 - 6 = -3, \\
\text{Var}\ (V) = \text{Var}\ (X-2Y) &= \text{Var}\ (X) + \text{Var}\ (-2Y) = \text{Var}\ (X) + 4\text{Var}\ (Y) = 9 + 64 = 73.
\end{aligned}$$

(b) Using the fact that the covariance is linear in each argument, and since independence of X and Y means that their covariance is 0:

$$\begin{aligned}
\text{Cov}\ (U,V) &= \text{Cov}\ (2X+Y, X-2Y) \\
&= 2\text{Cov}\ (X,X) + \text{Cov}\ (Y,X) + \text{Cov}\ (2X,-2Y) + \text{Cov}\ (Y,-2Y) \\
&= 2\text{Var}\ (X) + \text{Cov}\ (X,Y) - 4\text{Cov}\ (X,Y) - 2\text{Var}\ (Y) \\
&= 18 - 32 = -14.
\end{aligned}$$

Hence, since their covariance is non-zero, U and V cannot be independent.

A5 (**8 marks**) [Fairly standard, though care needed with manipulations]

(a) Since the random variables are independent, the mean and variance are additive. Hence we know $X+Y$ has mean 6 and variance 25, hence standard deviation 5.

(b) We know that

$$\mathbb{P}(X+Y \geq 14) = \mathbb{P}(X+Y-6 \geq 8) \ \leq \ \mathbb{P}(|X+Y-6| \geq 8) \leq \frac{5^2}{8^2} = \frac{25}{64} = 0.3906.$$

(c) Using the table

$$\begin{aligned}
\mathbb{P}(X+Y \geq 14) &= \mathbb{P}\left(\frac{X+Y-6}{5} \geq \frac{8}{5}\right) \\
&= \mathbb{P}\left(N(0,1) \geq 1.6\right) = 1 - \Phi(1.6) = 1 - 0.9452 = 0.0548.
\end{aligned}$$

B1 [Part (a)(i),(ii) are standard bookwork, a(iii)–(iv) are similar to questions in the notes and problem sheets. MGF calculation is in the notes, though calculation and interpretation of the derivative requires accurate work and good understanding. Example in part (b) is entirely unseen, though calculating the joint and marginal mass functions is very standard. The use of the tower law to calculate mean is standard, the formula for $\mathbb{E}(YZ)$ is unseen, and calculations are stretching in (b)iii)]

(a) i. (**3 marks**) We know

$$\sum_{z=0}^{\infty} p_\lambda(z) = \sum_{z=0}^{\infty} \frac{e^{-\lambda}\lambda^z}{z!} = e^{-\lambda} \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} = e^{-\lambda} e^{\lambda} = 1.$$

Simplifying: $\mathbb{P}(Z=2) = e^{-\lambda}\lambda^2/2$ and $\mathbb{P}(Z \geq 1) = 1 - \mathbb{P}(Z=0) = 1 - e^{-\lambda}$.

ii. (**4 marks**) Similar calculations give

$$\begin{aligned}\mathbb{E}(Z) &= \sum_{z=0}^{\infty} \frac{e^{-\lambda}\lambda^z}{z!} z = e^{-\lambda}\lambda \sum_{z=1}^{\infty} \frac{\lambda^{z-1}}{(z-1)!} = e^{-\lambda}\lambda e^{\lambda} = \lambda. \\ \mathbb{E}(Z(Z-1)) &= \sum_{z=0}^{\infty} \frac{e^{-\lambda}\lambda^z}{z!} z(z-1) = e^{-\lambda}\lambda^2 \sum_{z=2}^{\infty} \frac{\lambda^{z-2}}{(z-2)!} = e^{-\lambda}\lambda^2 e^{\lambda} = \lambda^2.\end{aligned}$$

iii. (**3 marks**) Hence $\mathbb{E}(Z^2) = \mathbb{E}(Z(Z-1)) + \mathbb{E}(Z) = \lambda^2 + \lambda$, so that $\text{Var}\,(Z) = \mathbb{E}(Z^2) - (\mathbb{E}(Z))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

iv. (**5 marks**) The MGF

$$M_Z(t) = \sum_{z=0}^{\infty} \frac{e^{-\lambda}\lambda^z}{z!} e^{tz} = e^{-\lambda} \sum_{z=0}^{\infty} \frac{(\lambda e^t)^z}{z!} = e^{\lambda(e^t-1)}.$$

We know that $M_Z(t) = 1 + t\mathbb{E}(Z) + t^2\mathbb{E}(Z^2)/2 + o(t^2)$, so it should be that $\mathbb{E}(Z) = M_Z'(0)$ and $\mathbb{E}(Z^2) = M_Z''(0)$. We can check this by finding that

$$\begin{aligned}M_Z'(t) &= \lambda(e^t)e^{\lambda(e^t-1)} \quad \text{so that} \quad M_Z'(0) = \lambda(1)e^{\lambda(1-1)} = \lambda. \\ M_Z''(t) &= \lambda(e^t)e^{\lambda(e^t-1)} + \lambda^2(e^t)^2 e^{\lambda(e^t-1)} \quad \text{so that} \quad M_Z''(0) = \lambda + \lambda^2\end{aligned}$$

(b) i. (**3 marks**) If Anna sends z messages, any particular sequence of read-unread with y total messages read has probability $p^y(1-p)^{z-y}$ of occuring, and there are $\binom{z}{y}$ such sequences. Hence

$$\mathbb{P}(Y = y | Z = z) = \binom{z}{y} p^y (1-p)^{z-y} \text{ for any } 0 \leq y \leq z,$$

which we recognise as a Binomial probability mass function.

ii. (**6 marks**) Combining the two expressions, we know that

$$\mathbb{P}(Y = y, Z = z) = \mathbb{P}(Y = y | Z = z)\mathbb{P}(Z = z) = \binom{z}{y} p^y (1-p)^{z-y} \frac{e^{-\lambda}\lambda^z}{z!},$$

for $0 \leq y \leq z$.

We find the marginal for Y by summing over Z (in the range $z \geq y$) to obtain

$$\begin{aligned}\mathbb{P}(Y = y) &= \sum_{z=y}^{\infty} \binom{z}{y} p^y (1-p)^{z-y} \frac{e^{-\lambda}\lambda^z}{z!} = \frac{e^{-\lambda}(p\lambda)^y}{y!} \sum_{z=y}^{\infty} \frac{(\lambda(1-p))^{z-y}}{(z-y)!} \\ &= \frac{e^{-\lambda}(p\lambda)^y}{y!} e^{\lambda(1-p)} = \frac{e^{-p\lambda}(p\lambda)^y}{y!},\end{aligned}$$

which we recognise as the probability mass function of a Poisson with parameter $p\lambda$, so that $\mathbb{P}(Y = 0) = e^{-p\lambda}$.

iii. (**2 marks**) Since $Y|Z = z$ has a binomial distribution with parameters z and p, we know that $A(z) = zp$. Hence the tower law gives

$$\mathbb{E}(Y) = \mathbb{E}\,(Zp) = p\mathbb{E}(Z) = p\lambda,$$

which we recognise as the mean of a Poisson with parameter $p\lambda$, as expected.

iv. (**4 marks**) Again, using the tower law we obtain

$$\mathbb{E}(YZ) = \mathbb{E}\left(Z(Zp)\right) = p\mathbb{E}(Z^2) = p(\lambda + \lambda^2),$$

using the results above. Hence we deduce that

$$\text{Cov }(Y,Z) \quad = \quad \mathbb{E}(YZ) - \mathbb{E}(Y)\cdot\mathbb{E}(Z) = p(\lambda+\lambda^2) - (p\lambda)\lambda = p\lambda$$

The sign of this covariance is positive, which we would expect; the more messages that Anna sends, the more messages Becky is likely to read. Since Y and Z are both Poisson, their variances are equal to their parameters so

$$\rho(Y,Z) = \frac{\text{Cov }(Y,Z)}{\sqrt{\text{Var }(Y)\text{Var }(Z)}} = \frac{p\lambda}{\sqrt{p\lambda}\sqrt{\lambda}} = \sqrt{p}.$$

B2 [Part (a) is very standard and taken closely from notes (though a)iii) is stated and not proved), part b)i) is standard, parts b)ii),iii) are unseen and more challenging]

(a) Mark applies for a series of jobs until he is made an offer; he is offered the ith job with probability p, independently of all other job decisions.

i. (**4 marks**) Clearly for $n = 0$ we can use the fact that $U^0 = 1$ to know that $\mathbb{E}(U^0) = 1$. Since $\mathbb{P}(U = 1) = p$ and $\mathbb{P}(U = 0) = 1 - p$, we know that for $n > 1$

$$\mathbb{E}\left(U^n\right) \quad = \quad (1-p)\cdot 0^n + p\cdot 1^n = p$$

meaning that $\text{Var }(U) = \mathbb{E}U^2 - (\mathbb{E}U)^2 = p - p^2$.

ii. (**5 marks**) For $X = i$, we know that Mark was not offered the first $i - 1$ jobs, and was offered the ith. Since outcomes are independent, the probability of this is

$$\mathbb{P}(X = i) = (1-p)^{i-1}p,$$

which is geometric as we require. Writing $y = x + z$, the probability

$$\begin{aligned}
\mathbb{P}(X > x) &= \sum_{y=x+1}^{\infty} \mathbb{P}(X = y) \\
&= \sum_{y=x+1}^{\infty} (1-p)^{y-1}p \\
&= p(1-p)^x \sum_{z=1}^{\infty} (1-p)^{z-1} \\
&= p(1-p)^x \frac{1}{1-(1-p)} = (1-p)^x,
\end{aligned}$$

by summing a geometric progression.

iii. (**5 marks**) The problem asks us to calculate

$$\begin{aligned}
\mathbb{P}(X > x+6 | X > 6) &= \frac{\mathbb{P}(\{X > x+6\} \cap \{X > 6\})}{\mathbb{P}(X > 6)} = \frac{\mathbb{P}(X > x+6)}{\mathbb{P}(X > 6)} \\
&= \frac{(1-p)^{x+6}}{(1-p)^6} = (1-p)^x.
\end{aligned}$$

This illustrates the lack-of-memory property; having waited 6 applications does not change the waiting time to come.

(b) i. (**5 marks**) The distribution function $F_Z(z) = 0$ for $z \le 0$ and is

$$F_Z(z) = \int_0^z \lambda e^{-\lambda y} dy = -\left[e^{-\lambda y}\right]_0^z = 1 - e^{-\lambda z}$$

for all $z \ge 0$. The mean can be found by integration by parts as

$$\mathbb{E}(Z) = \int_0^\infty y\lambda e^{-\lambda y} dy = -\int_0^\infty (-e^{-\lambda y}) dy = -\left[\frac{1}{\lambda} e^{-\lambda y}\right]_0^\infty = \frac{1}{\lambda}.$$

Using the distribution function, we know that

$$\mathbb{P}(a < Z \le b) = F_Z(b) - F_Z(a) = e^{-\lambda a} - e^{-\lambda b}.$$

ii. (**5 marks**) Since Z only takes positive values, we know that (for $z \ge 0$)

$$F_{Z^2}(z) = \mathbb{P}(Z^2 \le z) = \mathbb{P}(Z \le \sqrt{z}) = F_Z(\sqrt{z}) = 1 - e^{-\lambda\sqrt{z}}.$$

Taking a derivative with respect to z we deduce that

$$f_{Z^2}(z) = \begin{cases} \frac{\lambda}{2\sqrt{z}} e^{-\lambda\sqrt{z}} & z \ge 0, \\ 0 & z < 0. \end{cases}$$

iii. (**6 marks**)The key is to observe that (for any integer n), $Y = n$ if and only if $n - 1 < Z \le n$. Hence using the previous result

$$\mathbb{P}(Y = n) = \mathbb{P}(n - 1 < Z \le n) = e^{-\lambda(n-1)} - e^{-\lambda n} = \left(e^{-\lambda}\right)^{n-1} (1 - e^{-\lambda}).$$

We recognise this as a geometric mass function (as above) with parameter $1 - e^{-\lambda}$. Hence $X = Y$ when $p = 1 - e^{-\lambda}$, or $\lambda = -\log(1 - p)$.

End of examination.