2.5 Further Statistics: Discrete Distributions Depth

Total Marks

Calculator allowed · about 230 minutes · Write-in edition

Name: _____ Date: ____ / ____ / _____

Answer ALL questions. Show all your working.

Key facts: discrete distributions and probability generating functions

Original content written for Revision Library.

For X ~ Po(lambda), P(X = x) = e^(-lambda) * lambda^x / x!, for x = 0, 1, 2, ..., with E(X) = lambda and Var(X) = lambda. If X ~ Po(lambda1) and Y ~ Po(lambda2) are independent, then X + Y ~ Po(lambda1 + lambda2).

For X ~ Geo(p), the number of independent trials up to and including the first success (write q = 1 - p), P(X = x) = q^(x-1) * p, for x = 1, 2, 3, ..., with E(X) = 1/p and Var(X) = q/p^2. The geometric distribution has the memoryless property: P(X > m+n | X > m) = P(X > n) for positive integers m and n.

For X ~ NB(r, p), the number of independent trials up to and including the rth success, P(X = x) = C(x-1, r-1) * p^r * q^(x-r), for x = r, r+1, r+2, ..., with E(X) = r/p and Var(X) = rq/p^2. X can be written as the sum of r independent Geo(p) random variables.

The probability generating function of a discrete random variable X taking non-negative integer values is G_X(t) = E(t^X) = sum of P(X = x) * t^x. It satisfies G_X(1) = 1, E(X) = G_X'(1), and Var(X) = G_X''(1) + G_X'(1) - [G_X'(1)]^2. If X and Y are independent random variables, then G_(X+Y)(t) = G_X(t) * G_Y(t).

1 Meera is practising penalty kicks. Each attempt is successful, independently of all other attempts, with probability 0.35. Let X be the number of attempts up to and including her first successful kick, so that X ~ Geo(0.35).

(a) Find P(X = 5), giving your answer to 3 significant figures. **(2)**

(b) Find $P(X > 3)$, giving your answer to 3 significant figures. **(2)**

(c) State the value of $E(X)$, giving your answer as an exact fraction, and interpret this value in the context of this scenario. **(1)**

(Total for Question 1 is 5 marks)

2 In parts (a) and (b) below, X ~ Geo(p) models the number of independent trials, each with probability of success p, up to and including the first success; the two parts describe separate, independent scenarios.

(a) Given that E(X) = 12.5, find the value of p. **(2)**

(b) Given instead that Var(X) = 20, form an equation in p and show that it can be written as $20p^2 + p - 1 = 0$. Hence find the value of p, given that $0 < p < 1$. **(5)**

(Total for Question 2 is 7 marks)

3 The discrete random variable Y has the probability distribution shown in the table below.

Y	1	2	3	4
P(Y = y)	0.1	0.4	0.3	0.2

(a) Find the probability generating function $G_Y(t)$. **(2)**

(b) Use $G_Y(t)$ to find E(Y) and Var(Y). **(5)**

(c) Verify your value of E(Y) by direct calculation from the probability distribution. **(2)**

(Total for Question 3 is 9 marks)

---

4 The random variable X has distribution B(100, 0.02).

(a) State two conditions under which a Poisson distribution is a suitable approximation to a binomial distribution, and give the parameter of the approximating distribution. **(2)**

(b) Use the binomial distribution to find P(X = 2), and use the Poisson approximation to find the same probability, each to 4 significant figures. **(4)**

(c) Find the percentage error in using the Poisson approximation here, to 2 significant figures. **(2)**

(Total for Question 4 is 8 marks)

5 The discrete random variable X has probability distribution $P(X = x) = kx^2$ for $x = 1, 2, 3, 4$ and 0 otherwise, where k is a constant.

(a) Show that $k = 1/30$. **(3)**

(b) Find E(X) and Var(X). **(4)**

(c) Find E(5X - 4) and Var(5X - 4). **(4)**

(Total for Question 5 is 11 marks)

6 The independent discrete random variables X and Y each take the values 1, 2 and 3. P(X = 1), P(X = 2), P(X = 3) are 1/5, 3/5, 1/5 respectively, and P(Y = 1), P(Y = 2), P(Y = 3) are 1/4, 5/8, 1/8 respectively.

(a) Find P(X = Y). **(2)**

(b) Find P(X > Y). **(2)**

(c) Find E(X + Y) and Var(X - Y). **(4)**

(Total for Question 6 is 8 marks)

7 The number of flaws per metre in a roll of cloth may be modelled by a Poisson distribution with mean 2. Let X be the number in one such interval, so X ~ Po(2).

(a) Find P(X = 5) and P(X ≥ 5), each to 3 significant figures. **(4)**

(b) State the mean and variance of X, and explain why they are equal in a Poisson model. **(2)**

(c) A second independent count Y has distribution Po(1.5). Find P(X + Y = 5), to 3 significant figures. **(3)**

(Total for Question 7 is 9 marks)

8 The independent discrete random variables X and Y each take the values 1, 2 and 3. P(X = 1), P(X = 2), P(X = 3) are 2/7, 3/7, 2/7 respectively, and P(Y = 1), P(Y = 2), P(Y = 3) are 3/14, 3/7, 5/14 respectively.

(a) Find P(X = Y). **(2)**

(b) Find P(X > Y). **(2)**

(c) Find E(X + Y) and Var(X - Y). **(4)**

(Total for Question 8 is 8 marks)

9 A player attempts a task repeatedly until the first success. Each attempt succeeds with probability 0.25, independently of the others. Let X be the number of the attempt on which the first success occurs, so X ~ Geo(0.25).

(a) Find P(X = 6) and P(X > 6), each to 3 significant figures. **(4)**

(b) Find E(X) and Var(X). **(4)**

(c) Find the least number of attempts n such that the probability of at least one success within n attempts exceeds 0.99. **(4)**

(Total for Question 9 is 12 marks)

10 The independent discrete random variables X and Y each take the values 1, 2 and 3. P(X = 1), P(X = 2), P(X = 3) are 1/9, 2/9, 2/3 respectively, and P(Y = 1), P(Y = 2), P(Y = 3) are 1/3, 4/15, 2/5 respectively.

(a) Find P(X = Y). **(2)**

(b) Find P(X > Y). **(2)**

(c) Find E(X + Y) and Var(X - Y). **(4)**

(Total for Question 10 is 8 marks)

11 The independent discrete random variables X and Y each take the values 1, 2 and 3. P(X = 1), P(X = 2), P(X = 3) are 5/12, 1/4, 1/3 respectively, and P(Y = 1), P(Y = 2), P(Y = 3) are 6/13, 6/13, 1/13 respectively.

(a) Find P(X = Y). **(2)**

(b) Find P(X > Y). **(2)**

(c) Find E(X + Y) and Var(X - Y). **(4)**

(Total for Question 11 is 8 marks)

12 The number of calls to a helpline per minute may be modelled by a Poisson distribution with mean 3.5. Let X be the number in one such interval, so X ~ Po(3.5).

(a) Find P(X = 3) and P(X ≥ 3), each to 3 significant figures. **(4)**

(b) State the mean and variance of X, and explain why they are equal in a Poisson model. **(2)**

(c) A second independent count Y has distribution Po(1). Find P(X + Y = 3), to 3 significant figures. **(3)**

(Total for Question 12 is 9 marks)

13 The discrete random variable X has probability distribution $P(X = x) = kx^2$ for $x = 1, 2, 3, 4$ and 0 otherwise, where k is a constant.

(a) Show that $k = 1/30$. **(3)**

(b) Find E(X) and Var(X). **(4)**

(c) Find E(2X - 4) and Var(2X - 4). **(4)**

(Total for Question 13 is 11 marks)

14 The number of calls to a helpline per minute may be modelled by a Poisson distribution with mean 5. Let X be the number in one such interval, so $X \sim Po(5)$.

(a) Find $P(X = 2)$ and $P(X \geq 2)$, each to 3 significant figures. **(4)**

(b) State the mean and variance of X, and explain why they are equal in a Poisson model. **(2)**

(c) A second independent count Y has distribution Po(2). Find $P(X + Y = 2)$, to 3 significant figures. **(3)**

(Total for Question 14 is 9 marks)

15 In a factory in Swansea, each electronic component produced is defective, independently of all other components, with probability 0.08. Components are inspected one at a time until the 3rd defective component is found. Let X be the total number of components inspected up to and including the 3rd defective one found, so that X follows a negative binomial distribution, $X \sim NB(3, 0.08)$, where $P(X = x) = C(x-1, 2) * (0.08)^3 * (0.92)^{x-3}$, for $x = 3, 4, 5, \ldots$.

(a) Write down the value of E(X). **(1)**

(b) Find the probability that the 3rd defective component found is the 10th component inspected, giving your answer to 3 significant figures. **(3)**

(c) Find Var(X). **(1)**

(Total for Question 15 is 5 marks)

16 The discrete random variable X has the probability distribution shown in the table below.

X	0	1	2
P(X = x)	0.2	0.3	0.5

(a) Write down the probability generating function $G_X(t)$ of X. **(2)**

(b) Verify that $G_X(1) = 1$, and explain why this must always be true for the probability generating function of any discrete random variable. **(2)**

(c) Use $G_X(t)$ to find E(X). **(2)**

(Total for Question 16 is 6 marks)

17 The discrete random variable X follows a Poisson distribution with parameter λ, so that $X \sim Po(\lambda)$, where $P(X = x) = e^{-\lambda} * \lambda^x / x!$, for $x = 0, 1, 2, \ldots$.

(a) Show that the probability generating function of X is given by $G_X(t) = e^{\lambda(t-1)}$. **(4)**

(b) Using your result from part (a), show that $E(X) = \lambda$. **(3)**

(c) Using your result from part (a), show further that $Var(X) = \lambda$. **(3)**

(Total for Question 17 is 10 marks)

18 Let X ~ Po(lambda1) and Y ~ Po(lambda2) be independent random variables, and let W = X + Y.

(a) State the probability generating functions of X and of Y. **(1)**

(b) Show that $G_W(t) = G_X(t) * G_Y(t)$, and hence show that W ~ Po(lambda1 + lambda2). **(4)**

(c) The number of £10 notes given as change by the till of a small shop in Dundee during the morning session follows a Poisson distribution with mean 2.4, independently of the number given as change during the afternoon session, which follows a Poisson distribution with mean 3.1. Find the probability that exactly 5 £10 notes are given as change over the whole day, giving your answer to 3 significant figures. **(3)**

(Total for Question 18 is 8 marks)

19 Let $X \sim \text{Geo}(p)$ be the number of trials up to and including the first success, and write $q = 1 - p$.

(a) Show that, for positive integers m and n, $P(X > m+n \mid X > m) = P(X > n)$. **(4)**

(b) A component is tested repeatedly until it fails; the probability that it fails on any given test is 0.04, independently of the outcome of every previous test. Given that the component has already survived (not failed) 10 tests, find the probability that it survives at least a further 15 tests before it fails, giving your answer to 3 significant figures. **(2)**

(Total for Question 19 is 6 marks)

20 The discrete random variable X follows a geometric distribution with parameter p, so that X ~ Geo(p), where $P(X = x) = (1-p)^{x-1} * p$, for x = 1, 2, 3, ... Write q = 1 - p.

(a) Show that the probability generating function of X is $G_X(t) = pt / (1 - qt)$, for $|t| < 1/q$. **(4)**

(b) Hence show that E(X) = 1/p. **(3)**

(c) Hence show further that $Var(X) = q/p^2$. **(4)**

(Total for Question 20 is 11 marks)

21 A sequence of independent trials is carried out, in each of which the probability of success is p (write q = 1 - p). Let X ~ NB(r, p) be the number of trials up to and including the rth success. X can be written as X = X1 + X2 + ... + Xr, where X1 is the number of trials up to and including the 1st success, and, for i = 2, 3, ..., r, Xi is the number of further trials after the (i-1)th success up to and including the ith success. X1, X2, ..., Xr are independent, and each has the Geo(p) distribution.

(a) State the probability generating function of a single Xi. **(1)**

(b) Hence show that the probability generating function of X is $G_X(t) = [pt / (1 - qt)]^r$. **(3)**

(c) Using $G_X(t)$, show that E(X) = r/p. **(4)**

(d) Show further that $Var(X) = rq/p^2$. **(4)**

(Total for Question 21 is 12 marks)

Mark scheme · 2.5 Further Statistics: Discrete Distributions Depth

Question 1

(a) correct use of $P(X=5) = q^4 * p$ with q = 0.65, p = 0.35, oe — **M1**

(a) awrt 0.0625 — **A1**

(a) Answer: P(X = 5) = awrt 0.0625

(b) recognises $P(X>n) = q^n$, i.e. the probability the first 3 attempts all fail — **M1**

(b) 0.274625, awrt 0.275 — **A1**

(b) Answer: P(X > 3) = 0.274625, awrt 0.275

(c) E(X) = 1/0.35 = 20/7, awrt 2.86, with a valid interpretation in context, oe — **B1**

(c) Answer: E(X) = 20/7 = awrt 2.86; on average, Meera needs about 2.86 attempts up to and including her first success.

Question 2

(a) uses E(X) = 1/p = 12.5, oe — **M1**

(a) p = 0.08, cao — **A1**

(a) Answer: p = 0.08

(b) uses $Var(X) = (1-p)/p^2 = 20$, oe — **M1**

(b) rearranges to $1 - p = 20p^2$ — **M1**

(b) $20p^2 + p - 1 = 0$, cso — **A1**

(b) solves the quadratic, e.g. $p = (-1 \pm \sqrt{1+80})/40 = (-1 \pm 9)/40$ — **M1**

(b) p = 0.2 (rejecting p = -0.25 as 0 < p < 1), cao — **A1**

(b) Answer: p = 0.2

Question 3

(a) $G_Y(t)$ = sum of $P(Y=y)*t^y$, using the table — **M1**

(a) $G_Y(t) = 0.1t + 0.4t^2 + 0.3t^3 + 0.2t^4$, oe — **A1**

(a) Answer: $G_Y(t) = 0.1t + 0.4t^2 + 0.3t^3 + 0.2t^4$

(b) differentiates: $G_Y'(t) = 0.1 + 0.8t + 0.9t^2 + 0.8t^3$ — **M1**

(b) $E(Y) = G_Y'(1) = 2.6$, cao — **A1**

(b) differentiates again: $G_Y''(t) = 0.8 + 1.8t + 2.4t^2$, so $G_Y''(1) = 5.0$ — **M1**

(b) uses $Var(Y) = G_Y''(1) + G_Y'(1) - [G_Y'(1)]^2$, oe — **M1**

(b) Var(Y) = 0.84, cao — **A1**

(b) Answer: E(Y) = 2.6, Var(Y) = 0.84

(c) uses E(Y) = sum of y*P(Y=y) = 1(0.1)+2(0.4)+3(0.3)+4(0.2) — **M1**

(c) = 2.6, matching part (b), cao — **A1**

(c) Answer: E(Y) = 0.1+0.8+0.9+0.8 = 2.6, matching part (b).

Question 4

(a) n is large and p is small — **B1**

(a) X is approximately Po(100 x 0.02) = Po(2) — **B1**

(a) Answer: Po(2)

(b) $100C2 \times 0.02^2 \times 0.98^{98}$ — **M1**

(b) 0.2734 — **A1**

(b) $e^{-2}\, 2^2/2!$ — **M1**

(b) 0.2707 — **A1**

(b) Answer: binomial 0.2734, Poisson 0.2707

(c) |approximate - exact|/exact x 100 — **M1**

(c) 1.0% — **A1**

(c) Answer: 1.0%

Question 5

(a) uses the sum of all probabilities = 1 — **M1**

(a) forms the sum P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) in terms of k — **M1**

(a) k(30) = 1, so k = 1/30 (cso) — **A1**

(a) Answer: k = 1/30

(b) E(X) = sum of x P(X = x) — **M1**

(b) E(X) = 10/3 — **A1**

(b) $Var(X) = E(X^2) - [E(X)]^2$ — **M1**

(b) $E(X^2) = 59/5$, Var(X) = 31/45 — **A1**

(b) Answer: E(X) = 10/3, Var(X) = 31/45

(c) E(aX + b) = aE(X) + b — **M1**

(c) 38/3 — **A1**

(c) $Var(aX + b) = a^2\, Var(X)$ — **M1**

(c) 155/9 — **A1**

(c) Answer: E = 38/3, Var = 155/9

Question 6

(a) sum of P(X = i)P(Y = i) over i = 1, 2, 3, using independence — **M1**

(a) 9/20 — **A1**

(a) Answer: 9/20

(b) adds P(X = i)P(Y = j) over the cases i > j — **M1**

(b) 13/40 — **A1**

(b) Answer: 13/40

(c) E(X + Y) = E(X) + E(Y) — **M1**

(c) 31/8 — **A1**

(c) Var(X - Y) = Var(X) + Var(Y) for independent variables — **M1**

(c) 243/320 — **A1**

(c) Answer: E(X + Y) = 31/8, Var(X - Y) = 243/320

Question 7

(a) $P(X = 5) = e^{-2}\, 2^5/5!$ — **M1**

(a) 0.0361 — **A1**

(a) $P(X \geq 5) = 1 - P(X \leq 4)$ — **M1**

(a) 0.0527 — **A1**

(a) Answer: 0.0361; 0.0527

(b) E(X) = Var(X) = 2 — **B1**

(b) for a Poisson distribution the variance equals the mean — **B1**

(b) Answer: E(X) = Var(X) = 2

(c) X + Y ~ Po(2 + 1.5) = Po(3.5) — **M1**

(c) $P = e^{-3.5}\, 3.5^5/5!$ — **A1**

(c) 0.132 — **A1**

(c) Answer: 0.132

Question 8

(a) sum of $P(X = i)P(Y = i)$ over i = 1, 2, 3, using independence **M1**

(a) 17/49 **A1**

(a) Answer: 17/49

(b) adds $P(X = i)P(Y = j)$ over the cases i > j **M1**

(b) 27/98 **A1**

(b) Answer: 27/98

(c) E(X + Y) = E(X) + E(Y) **M1**

(c) 29/7 **A1**

(c) Var(X - Y) = Var(X) + Var(Y) for independent variables **M1**

(c) 55/49 **A1**

(c) Answer: E(X + Y) = 29/7, Var(X - Y) = 55/49

Question 9

(a) $P(X = 6) = (1 - 0.25)^5 \times 0.25$ **M1**

(a) 0.0593 **A1**

(a) $P(X > 6) = (1 - 0.25)^6$ **M1**

(a) 0.178 **A1**

(a) Answer: 0.0593; 0.178

(b) E(X) = 1/p **M1**

(b) 4 **A1**

(b) $Var(X) = (1 - p)/p^2$ **M1**

(b) 12.00 **A1**

(b) Answer: E(X) = 4.000, Var(X) = 12.00

(c) $P(X \leq n) = 1 - (1 - 0.25)^n > 0.99$ **M1**

(c) $(1 - 0.25)^n < 0.01$, and takes logs (changing the inequality) **M1**

(c) n > 16.008 **A1**

(c) n = 17 **A1**

(c) Answer: n = 17

Question 10

(a) sum of $P(X = i)P(Y = i)$ over i = 1, 2, 3, using independence **M1**

(a) 49/135 **A1**

(a) Answer: 49/135

(b) adds $P(X = i)P(Y = j)$ over the cases i > j **M1**

(b) 64/135 **A1**

(b) Answer: 64/135

(c) E(X + Y) = E(X) + E(Y) **M1**

(c) 208/45 **A1**

(c) Var(X - Y) = Var(X) + Var(Y) for independent variables **M1**

(c) 2426/2025 **A1**

(c) Answer: E(X + Y) = 208/45, Var(X - Y) = 2426/2025

Question 11

(a) sum of $P(X = i)P(Y = i)$ over i = 1, 2, 3, using independence **M1**

(a) 1/3 **A1**

(a) Answer: 1/3

(b) adds $P(X = i)P(Y = j)$ over the cases i > j **M1**

(b) 11/26 **A1**

(b) Answer: 11/26

(c) E(X + Y) = E(X) + E(Y) **M1**

(c) 551/156 **A1**

(c) Var(X - Y) = Var(X) + Var(Y) for independent variables **M1**

(c) 27587/24336 **A1**

(c) Answer: E(X + Y) = 551/156, Var(X - Y) = 27587/24336

Question 12

(a) $P(X = 3) = e^{-3.5}\ 3.5^3/3!$ **M1**

(a) 0.216 **A1**

(a) $P(X \geq 3) = 1 - P(X \leq 2)$ **M1**

(a) 0.679 **A1**

(a) Answer: 0.216; 0.679

(b) E(X) = Var(X) = 3.5 **B1**

(b) for a Poisson distribution the variance equals the mean **B1**

(b) Answer: E(X) = Var(X) = 3.5

(c) X + Y ~ Po(3.5 + 1) = Po(4.5) **M1**

(c) $P = e^{-4.5}\ 4.5^3/3!$ **A1**

(c) 0.169 **A1**

(c) Answer: 0.169

Question 13

(a) uses the sum of all probabilities = 1 **M1**

(a) forms the sum P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) in terms of k **M1**

(a) k(30) = 1, so k = 1/30 (cso) **A1**

(a) Answer: k = 1/30

(b) E(X) = sum of x P(X = x) **M1**

(b) E(X) = 10/3 **A1**

(b) $Var(X) = E(X^2) - [E(X)]^2$ **M1**

(b) $E(X^2) = 59/5$, Var(X) = 31/45 **A1**

(b) Answer: E(X) = 10/3, Var(X) = 31/45

(c) E(aX + b) = aE(X) + b **M1**

(c) 8/3 **A1**

(c) $Var(aX + b) = a^2\ Var(X)$ **M1**

(c) 124/45 **A1**

(c) Answer: E = 8/3, Var = 124/45

Question 14

(a) $P(X = 2) = e^{-5} 5^2/2!$ **M1**

(a) 0.0842 **A1**

(a) $P(X \geq 2) = 1 - P(X \leq 1)$ **M1**

(a) 0.960 **A1**

(a) Answer: 0.0842; 0.960

(b) E(X) = Var(X) = 5 **B1**

(b) for a Poisson distribution the variance equals the mean **B1**

(b) Answer: E(X) = Var(X) = 5

(c) X + Y ~ Po(5 + 2) = Po(7) **M1**

(c) $P = e^{-7} 7^2/2!$ **A1**

(c) 0.0223 **A1**

(c) Answer: 0.0223

Question 15

(a) E(X) = r/p = 3/0.08 = 37.5, cao **B1**

(a) Answer: E(X) = 37.5

(b) identifies x = 10, r = 3 and uses $P(X=x) = C(x-1,r-1) * p^r * q^{x-r}$, oe **M1**

(b) correct unsimplified expression $C(9,2) * (0.08)^3 * (0.92)^7$, oe **A1**

(b) awrt 0.0103 **A1**

(b) Answer: P(X = 10) = awrt 0.0103

(c) $Var(X) = r*q/p^2 = 3*0.92/0.0064 = 431.25$, cao **B1**

(c) Answer: Var(X) = 431.25

Question 16

(a) $G_X(t)$ = sum of $P(X=x) * t^x$, using the probabilities given **M1**

(a) $G_X(t) = 0.2 + 0.3t + 0.5t^2$, oe **A1**

(a) Answer: $G_X(t) = 0.2 + 0.3t + 0.5t^2$

(b) $G_X(1) = 0.2 + 0.3 + 0.5 = 1$ **B1**

(b) valid explanation, e.g. $G_X(1)$ = sum of all the probabilities in the distribution, which must total 1 for any valid probability distribution, oe **B1**

(b) Answer: $G_X(1) = 1$, since the probabilities in any distribution must sum to 1.

(c) differentiates: $G_X'(t) = 0.3 + t$, oe **M1**

(c) $E(X) = G_X'(1) = 1.3$, cao **A1**

(c) Answer: E(X) = 1.3

Question 17

(a) writes $G_X(t)$ = sum from x=0 to infinity of $t^x * e^{-\lambda} * \lambda^x / x!$, oe **M1**

(a) factors out $e^{-\lambda}$ and combines terms: $G_X(t) = e^{-\lambda} *$ sum from x=0 to infinity of $(\lambda*t)^x / x!$ **M1**

(a) recognises sum from x=0 to infinity of $(\lambda*t)^x/x! = e^{\lambda*t}$ (the exponential series) **M1**

(a) $G_X(t) = e^{-\lambda} * e^{\lambda*t} = e^{\lambda(t-1)}$, cso **A1**

(a) Answer: $G_X(t) = e^{\lambda(t-1)}$

(b) differentiates: $G_X'(t) = \lambda * e^{\lambda(t-1)}$ **M1**

(b) substitutes t = 1 into $G_X'(t)$ **dM1**

(b) $E(X) = G_X'(1) = \lambda * e^0 = \lambda$, cso **A1**

(b) Answer: E(X) = λ

(c) differentiates again: $G_X''(t) = \lambda^2 * e^{\lambda(t-1)}$, so $G_X''(1) = \lambda^2$ **M1**

(c) uses $Var(X) = G_X''(1) + G_X'(1) - [G_X'(1)]^2 = \lambda^2 + \lambda - \lambda^2$, oe **M1**

(c) $Var(X) = \lambda$, cso **A1**

(c) Answer: Var(X) = λ

Question 18

(a) $G_X(t) = e^{lambda1(t-1)}$ and $G_Y(t) = e^{lambda2(t-1)}$, both stated (ft from Q17, or quoted) **B1**

(a) Answer: $G_X(t) = e^{lambda1(t-1)}$, $G_Y(t) = e^{lambda2(t-1)}$

(b) writes $G_W(t) = E(t^{X+Y}) = E(t^X * t^Y)$, oe **M1**

(b) uses independence of X and Y: $E(t^X * t^Y) = E(t^X)*E(t^Y) = G_X(t)*G_Y(t)$ **M1**

(b) combines: $G_W(t) = e^{lambda1(t-1)} * e^{lambda2(t-1)} = e^{(lambda1+lambda2)(t-1)}$ **M1**

(b) identifies this as the PGF of a Po(lambda1+lambda2) distribution; since a PGF uniquely determines a distribution, W ~ Po(lambda1+lambda2), cso **A1**

(b) Answer: W ~ Po(lambda1 + lambda2)

(c) identifies W ~ Po(2.4+3.1) = Po(5.5), using the result of part (b), oe **M1**

(c) correct use of $P(W=5) = e^{-5.5} * 5.5^5 / 5!$ **M1**

(c) awrt 0.171 **A1**

(c) Answer: P(W = 5) = awrt 0.171

Question 19

(a) states or derives $P(X > k) = q^k$ for a positive integer k, oe **B1**

(a) uses $P(X>m+n \mid X>m) = P(X>m+n \text{ and } X>m)/P(X>m) = P(X>m+n)/P(X>m)$, since {X>m+n} is a subset of {X>m} **M1**

(a) substitutes: q^{m+n}/q^m **M1**

(a) $= q^n = P(X>n)$, as required, cso **A1**

(a) Answer: $P(X > m+n \mid X > m) = q^n = P(X > n)$ (shown)

(b) uses the memoryless property from part (a): required probability = $P(X>15) = q^{15}$, oe **M1**

(b) awrt 0.542 **A1**

(b) Answer: 0.542 (3 s.f.)

Question 20

(a) writes $G_X(t)$ = sum from x=1 to infinity of $t^x * q^{x-1} * p$, oe **M1**

(a) takes out a factor of pt: $G_X(t) = pt *$ sum from k=0 to infinity of $(qt)^k$, oe **M1**

(a) recognises sum from k=0 to infinity of $(qt)^k = 1/(1-qt)$ as an infinite geometric series, valid for $|qt| < 1$ **M1**

(a) $G_X(t) = pt/(1-qt)$, cso **A1**

(a) Answer: $G_X(t) = pt/(1-qt)$

(b) differentiates using the quotient rule: $G_X'(t) = p*(1-qt)^{-1} + p*q*t*(1-qt)^{-2}$, oe **M1**

(b) substitutes t = 1, using 1 - q = p **dM1**

(b) $E(X) = G_X'(1) = p/p + pq/p^2 = 1 + q/p = (p+q)/p = 1/p$, cso **A1**

(b) Answer: E(X) = 1/p

(c) differentiates $G_X'(t)$ again to obtain $G_X''(t) = 2pq*(1-qt)^{-2} + 2pq^2*t*(1-qt)^{-3}$, oe **M1**

(c) substitutes t = 1 to get $G_X''(1) = 2q/p + 2q^2/p^2$ **dM1**

(c) uses $Var(X) = G_X''(1) + G_X'(1) - [G_X'(1)]^2$ and combines all terms over a common denominator p^2 **dM1**

(c) $Var(X) = q/p^2$, cso, using p+q=1 to simplify the numerator to q **A1**

(c) Answer: $Var(X) = q/p^2$

Question 21

(a) $G_{Xi}(t) = pt/(1-qt)$, ft from Q20(a) or quoted **B1**

(a) Answer: $G_{Xi}(t) = pt/(1-qt)$

(b) states X is the sum of r independent, identically distributed random variables **M1**

(b) uses the PGF-of-a-sum result (extended to r independent variables): the PGF of a sum of independent variables is the product of their individual PGFs **M1**

(b) $G_X(t) = [pt/(1-qt)]^r$, cso **A1**

(b) Answer: $G_X(t) = [pt/(1-qt)]^r$

(c) lets $g(t) = pt/(1-qt)$, so $G_X(t) = [g(t)]^r$; differentiates using the chain rule: $G_X'(t) = r*[g(t)]^{r-1} * g'(t)$ **M1**

(c) uses g(1) = 1 and g'(1) = 1/p (from Q6) **M1**

(c) substitutes t = 1: $G_X'(1) = r * 1^{r-1} * (1/p)$ **dM1**

(c) $E(X) = G_X'(1) = r/p$, cso **A1**

(c) Answer: E(X) = r/p

(d) differentiates again using the product/chain rule: $G_X''(t) = r(r-1)*[g(t)]^{r-2}*[g'(t)]^2 + r*[g(t)]^{r-1}*g''(t)$ **M1**

(d) substitutes t=1 using g(1)=1, g'(1)=1/p, $g''(1)=2q/p+2q^2/p^2$ (from Q6) to get $G_X''(1) = r(r-1)/p^2 + 2rq/p + 2rq^2/p^2$ **M1**

(d) uses $Var(X) = G_X''(1) + G_X'(1) - [G_X'(1)]^2$ and combines all terms over a common denominator p^2 **dM1**

(d) $Var(X) = rq/p^2$, cso, using p+q=1 to simplify the numerator to rq **A1**

(d) Answer: $Var(X) = rq/p^2$