2.5 Further Statistics: Discrete Distributions Depth: 1-hour Pack

Total Marks

Calculator allowed · about 60 minutes

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 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 1 is 11 marks)

© Revision Library | revisionlibrary.co.uk

2 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 2 is 9 marks)

---

3 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 3 is 8 marks)

---

4 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 4 is 12 marks)

© Revision Library | revisionlibrary.co.uk

5 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 ~ NB(3, 0.08), where $P(X = x) = C(x-1, 2) * (0.08)^3 * (0.92)^{x-3}$, for x = 3, 4, 5, ...

(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 5 is 5 marks)

---

6 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 6 is 6 marks)

---

7 Let X ~ 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 7 is 6 marks)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 2.5 Further Statistics: Discrete Distributions Depth: 1-hour Pack

Question 1

(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 2

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

(a) 0.0361 **A1**

(a) P(X ≥ 5) = 1 - P(X ≤ 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 3

(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 4

(a) P(X = 6) = $(1 - 0.25)^5$ x 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 ≤ 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 5

(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 6

(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

© Revision Library | revisionlibrary.co.uk

Question 7

(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.)

© Revision Library | revisionlibrary.co.uk