2.5 Further Statistics: Discrete Distributions Depth: 30-minute Pack

Total Marks

Calculator allowed · about 30 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 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)

© Revision Library | revisionlibrary.co.uk

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 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 3 is 8 marks)

---

4 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 4 is 8 marks)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 2.5 Further Statistics: Discrete Distributions Depth: 30-minute Pack

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) 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 4

(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

© Revision Library | revisionlibrary.co.uk