

Problem Sheet 3

Please send any comments or corrections to martin@stats.ox.ac.uk.

1. Find the mean of the Binomial(n, p) distribution. [*Try this two different ways: (i) directly from the probability mass function; (ii) using the representation of a binomial random variable as a sum of Bernoulli random variables, and linearity of expectation.*]

2. Suppose that X is a discrete random variable taking non-negative integer values. Show that

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} \mathbb{P}(X > k).$$

Hint: you may find it helpful to write the right-hand side in terms of the probability mass function.

3. Let X have geometric distribution with parameter p , i.e. $\mathbb{P}(X = k) = p(1 - p)^{k-1}$ for $k \geq 1$.

- (a) Find $\mathbb{P}(X > k)$ for $k \geq 0$.
 (b) Show that

$$\mathbb{P}(X = k + r | X > k) = \mathbb{P}(X = r) \text{ for all } k \geq 0, r \geq 1.$$

[This is called the *memoryless property* of the geometric distribution.]

4. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be given by $f(x) = e^{\theta x}$ for some $\theta \in \mathbb{R}$ and suppose that $X \sim \text{Po}(\lambda)$ for some $\lambda > 0$. Show that

$$\mathbb{E}[f(X)] = \exp(\lambda(e^{\theta} - 1)).$$

5. Let $X \sim \text{Bin}(n, \lambda/n)$ for some parameter $\lambda > 0$.

- (a) By using the approximation $(1 - \frac{x}{n})^n \approx e^{-x}$ (which is valid for large n), show that for fixed $k \geq 0$ and for n large,

$$\mathbb{P}(X = k) \approx \frac{e^{-\lambda} \lambda^k}{k!}.$$

- (b) A text file contains 1000 characters. When the file is sent by e-mail from one machine to another, each character (independently of all other characters) has probability 0.001 of being corrupted. From part (a), use a Poisson random variable to estimate the probability that the file is transferred without error. Compare this to the answer obtained when you model the number of errors as a binomial random variable.

6. You repeatedly flip a coin, with probability p of heads and $q = 1 - p$ of tails. Flips are independent. Let X denote the number of flips until you get two heads in a row. By considering the partition A_1, A_2, A_3 where A_1 is the event {tail on the first flip}, A_2 is the event {head on the first flip, tail on the second} and A_3 is the event {heads on the first two flips}, find $\mathbb{E}[X]$.

7. **(The coupon collector problem).** Packets of cornflakes contain a plastic toy, of which there are n different types. Each packet is equally likely to contain any of the toys. You want to collect all of them.
- (a) You find the first new toy in the first packet you open. What is the probability that the next packet contains a *different* toy? What is the distribution of the number of packets after the first that you need to buy in order to find the second type of toy? State any assumptions you make.
 - (b) More generally, suppose that you have already found $k - 1$ different types of toy, for some $k \geq 1$. What is the distribution of T_k , the additional number of packets of cornflakes you have to buy in order to find the k th type of toy?
 - (c) Let $T = T_1 + \dots + T_n$ be the total number of packets needed to collect the full set of different toys. What is the expectation of T ? How quickly does this grow as n becomes large?