10 Modeling randomness

Try section 1 of this worksheet after you have completed Exercise 10E,
section 2 of this worksheet after you have completed Section 10.4, and
section 3 of this worksheet after you have completed Chapter 10.

1 Other discrete distributions related to the binomial distribution

The binomial distribution can be generalized to include k mutually exclusive events $E_1, \ldots, E_k$ with constant probabilities of success $p_1, \ldots, p_k$ such that $p_1 + \ldots + p_k = 1$. This distribution is called the **multinomial distribution** and is defined as follows.

Let X_i be the number of times the event E_i with probability of success p_i occurs in n Bernoulli trials, for $i \in \{1, 2, \ldots, k\}$ where $n = X_1 + \ldots + X_k$ and $p_1 + \ldots + p_k = 1$. Then $P((X_1, X_2, \ldots, X_k) = (x_1, x_2, \ldots, x_k)) = \frac{n!}{x_1!x_2!\ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$.

EXAMPLE 1

A fair die is tossed 12 times. Find the probability of getting each face up exactly twice.

Answer

If a fair die is tossed 12 times, the probability of getting each face up exactly twice is given by $\frac{12!}{2!2!2!2!2!2!}\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2 = 0.00344 \text{ (3sf)}$

Exercise 1

1 A box contains 5 red balls, 4 yellow balls and 3 white balls. A ball is selected at random from the box, its colour noted and then put back in the box. The process is repeated six times under the same conditions. Find the probability that out of the six balls selected in this manner, 2 are red, 3 are yellow and 1 is white.

2 A fair die is tossed 6 times. Find the probability that each face shows on the up side exactly once.

3 A tetrahedral die is tossed 8 times. Find the probability that it lands on each face exactly twice.

The binomial distribution gets its name because the expression of the probability for each event corresponds to one of the terms of the expansion

$$(p+q)^n = \binom{n}{0}p^0q^n + \binom{n}{1}p^1q^{n-1} + \ldots + \binom{n}{n-1}p^{n-1}q^1 + \binom{n}{n}p^nq^0$$

Similarly, the multinomial probabilities correspond to the terms of the expansion

$$(p_1 + \ldots + p_k)^n = \sum_{x_1+\ldots+x_k=n} \frac{n!}{x_1!x_2!\ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}, \; n \geqslant 2, \; n \geqslant 2.$$

Another distribution related to the binomial is the **hypergeometric distribution**. This distribution is studied in detail as part of the statistics option. While the binomial distribution models the probability of events for trials with replacement (i.e. situations where the probability of success remains constant), the hypergeometric distribution models the probability of events for trials without replacement.

EXAMPLE 2

A box contains 6 strawberry candies and 4 blueberry candies. Kathy takes 2 candies from the box simultaneously. Find the probability that

a she takes one candy of each type

b both candies she takes are strawberry candies.

Answers

Let S represent a strawberry candy and B a blueberry candy.

a You have two options: a strawberry candy followed by a blueberry candy (SB) or a blueberry candy followed by a strawberry candy (BS).

$$\mathrm{P(SB, BS)} = \frac{6}{10}\times\frac{4}{9}+\frac{4}{10}\times\frac{6}{9}=\frac{48}{90}=\frac{8}{15}$$

or

$$\mathrm{P(SB, BS)} = \frac{\binom{6}{1}\binom{4}{1}}{\binom{10}{2}}=\frac{8}{15}$$

b $\mathrm{P(SS)} = \frac{6}{10}\times\frac{5}{9}=\frac{30}{90}=\frac{1}{3}$

or

$$\mathrm{P(SS)} = \frac{\binom{6}{2}\binom{4}{0}}{\binom{10}{2}}=\frac{15}{45}=\frac{1}{3}$$

Exercise 2

1 If a box contains n objects of type A and m objects of type B and we remove t of these objects simultaneously (or sequentially without replacing them at any stage), show that the probability that we take exactly x type A objects and $t-x$ objects of type B is given by

$$\frac{\binom{n}{x}\binom{m}{t-x}}{\binom{n+m}{t}}.$$

State any restrictions you need to impose to the values of the variables and parameters in this expression.

Another interesting distribution associated with Bernoulli trials is the **geometric distribution**. Suppose that you are considering performing a series of independent trials each with a constant probability of success p and a probability of failure $q = 1-p$, $0 < p, q < 1$. The random variable X that denotes the 'number of trials needed until the first success occurs' is said to follow a geometric distribution with parameter p. We write $X \sim \mathrm{Geo}(p)$.

Exercise 3

1 Show that an expression for $P(X = x)$ is $p(1-p)^{x-1} = pq^{x-1}$.

2 Show that $f(x) = p(1-p)^{x-1}$, $x = 1, 2, \ldots$ defines a probability distribution function. Hence give a reason for the name 'geometric distribution' and explain why it is enough to know the value of p to describe this distribution.

3 If X is a discrete random variable that follows a geometric distribution with parameter p, show that

a $\mathrm{P}(X \le x) = 1-q^{x}$ **b** $\mathrm{P}(X < x) = 1-q^{x-1}$ **c** $\mathrm{P}(X > x) = q^{x}$

4 Given that $X \sim \text{Geo}(p)$, show that

a $\mathrm{E}(X) = \frac{1}{p}$

b $\mathrm{E}(X^2) = \frac{2}{p} \times \mathrm{E}(X) - (1 + q + q^2 + \ldots) = \frac{2-p}{p^2}$

c $\mathrm{Var}(X) = \frac{1-p}{p^2}$

$1 + q + q^2 + \ldots = \frac{1}{1-q} = \frac{1}{p}$

5 Show that the mode of a geometric distribution is always 1. Explain the meaning of this result for small values of p. Give real-life examples of the geometric distribution.

6 Show that geometric distributions are memory-less, i.e.

$\mathrm{P}(X > a + b \mid X > a) = \mathrm{P}(X > b)$

Try this section after you have completed Section 10.4.

2 Relationship between exponential distribution and Poisson processes

The **exponential distribution** is a continuous distribution that is studied in detail as part of the Statistics option. However, as this distribution is closely related to the Poisson distribution and it has interesting properties, it is worth exploring it at this stage. This distribution also appears often in exam questions where previous knowledge of its properties is not required – so it can be useful to be aware of it.

When events occur at random points in time and the number of events in an interval follows a Poisson distribution with mean m per unit of time, the mean time between two consecutive events is naturally $\frac{1}{m}$. The continuous random variable X 'length of the interval between two consecutive Poisson occurrences' follows an exponential distribution with parameter $\frac{1}{m}$ whose CDF is given by $F(x) = 1 - e^{-mx}$, $x \geqslant 0$.

Exercise 4

1 Show that exponential distributions are memory-less, i.e.

$\mathrm{P}(X > a + b \mid X > a) = \mathrm{P}(X > b)$

Explain the meaning of this result in a real-life context.

2 Show that the mode of an exponential distribution is independent of its parameter and state its value.

Try this section after you have completed Chapter 10.

3 Statisticians' favourite example that the obvious can be false

Exercise 5

1 Consider the distribution whose PDF is defined by

$$f(x) = \frac{1}{\pi\left(1 + (x-m)^2\right)}$$

Use a GDC to obtain its graph for different values of m.

Use the graphs to predict the values of the median and mean of the distribution.

Then use the GDC to approximate the value of the mean in each case by considering the integral between $-M$ and M for large values of $M < 0$. What do you observe?

This is a particular case of a family of distributions known as **Cauchy distributions**. You may wish to explore their properties further and find out about their practical applications.

Chapter 10 extension worked solutions

Exercise 1

1 5R + 4Y + 3W = 12 balls

$\mathrm{P(R)} = \frac{5}{12}$ $\mathrm{P(Y)} = \frac{4}{12} = \frac{1}{3}$ $\mathrm{P}(W) = \frac{3}{12} = \frac{1}{4}$

$\mathrm{P(2R, 3Y, 1W)} = \frac{6!}{2!3!1!}\left(\frac{5}{12}\right)^2\left(\frac{1}{3}\right)^3\left(\frac{1}{4}\right)^1 = 0.002$

2 $\mathrm{P(each\ face\ shows\ once)} = \frac{6!}{1!1!1!1!1!1!}\left(\frac{1}{6}\right)^6 = \frac{6!}{6^6} = \frac{5}{324}$

3 $\mathrm{P(each\ of\ the\ four\ faces\ twice)} = \frac{8!}{2!2!2!2!}\left(\frac{1}{4}\right)^2\left(\frac{1}{4}\right)^2\left(\frac{1}{4}\right)^2\left(\frac{1}{4}\right)^2 = \frac{315}{8192}$

Exercise 2

1 Recall the counting techniques studied in Chapter 1. The number of ways of selecting x distinct objects from a set with n distinct objects is given by $\binom{n}{x}$. Similarly, the number of ways we can select the remaining $t - x$ objects from a set with m distinct objects is given by $\binom{m}{t-x}$ and the number of ways of selecting the t objects from the total A and B (with $n + m$ objects) is given by $\binom{n+m}{t}$

Therefore, the probability is given by

$$\frac{\binom{n}{x}\binom{m}{t-x}}{\binom{n+m}{t}}.$$

This formula is valid for $0 \leqslant x \leqslant n$ and $0 \leqslant t - x \leqslant m$ (or $0 \leqslant t \leqslant n + m$), where n, m, x, $t \in \mathbb{Z}$ and n, $m > 0$.

Exercise 3

1 $\mathrm{P}(X = x) = \underbrace{(1-p)(1-p)\ldots(1-p)}_{\substack{x-1 \text{ times corresponding} \\ \text{to } x-1 \text{ failures in the first} \\ x-1 \text{ attempts (indep. trials)}}} \cdot \underbrace{p}_{\text{first success}} = p(1-p)^{x-1}$ or pq^{x-1}

2 $f(x) = p(1-p)^{x-1} \geq 0 \ \forall\ x = 1, 2, \ldots$

$\sum_{x=1}^{\infty} f(x) = \sum_{x=1}^{\infty} p(1-p)^{x-1} = \frac{p}{1-(1-p)} = 1$

This is a geometric series with $a = p$ and $r = 1 - p$

This series converges as $0 \leq 1 - p \leq 1$.

The name is related to the expression of the PDF which is given by consecutive terms of a geometric sequence. Its values depend just on p.

3 **a** $\mathrm{P}(X \leqslant x) = \mathrm{P}(1) + \mathrm{P}(2) + \ldots + \mathrm{P}(x)$

$= p + p(1-p) + \ldots + p(1-p)^{x-1}$

$= p\frac{1-(1-p)^x}{1-(1-p)} = \cancel{p}\frac{1-(1-p)^x}{\cancel{p}} = 1 - q^x$

b $\mathrm{P}(X < x) = \mathrm{P}(X \leq x) - \mathrm{P}(X = x)$

$= 1 - q^x - p(1-p)^{x-1}$

$= 1 - q^x - pq^{x-1}$

$= 1 - \underbrace{(p+q)}_{1}\ q^{x-1}$

$= 1 - q^{x-1}$

c $\mathrm{P}(X > x) = 1 - \mathrm{P}(X \leq x)$

$= 1 - (1 - q^x) = q^x$

4 **a** $\mathrm{E}(X) = 1\mathrm{P}(x = 1) + 2\mathrm{P}(x = 2) + 3\mathrm{P}(x = 3) + \ldots$

$$= \underbrace{1p + 2p(1-p) + 3p(1-p)^2 + \ldots}$$

$$= (1-q) + 2(1-q)q + 3(1-q)q^2 + \ldots$$

$$= 1 \underbrace{-q + 2q} - \underbrace{2q^2 + 3q^2} - \underbrace{3q^3 + \ldots}$$

$$= 1 + q + q^2 + q^3 + \ldots$$

$$= \frac{1}{1-q} = \frac{1}{p}$$

b $\mathrm{E}(X^2) = 1\mathrm{P}(x = 1) + 4\mathrm{P}(x = 2) + 9\mathrm{P}(x = 3) + \ldots$

$$= 1p + 4p(1-p) + 9p(1-p)^2 + 16p(1-p)^3 + \ldots$$

$$= 1 - q + 4(1-q)q + 9(1-q)q^2 + 16(1-q)q^3 + \ldots$$

$$= 1 - q + 4q - 4q^2 + 9q^2 - 9q^3 + 16q^3 - 16q^4 + \ldots$$

$$= 1 + 3q + 5q^2 + 7q^3 + 9q^4 + 11q^5 + \ldots$$

$$= 2(1 + q + 2q^2 + 3q^3 + 4q^4 + \ldots) - \underbrace{(1 + q + q^2 + \ldots)}$$

$$= \frac{2}{p}\underbrace{(p + pq + 2pq^2 + \ldots)}_{\mathrm{E}(x)} - \frac{1}{p}$$

$$= \frac{2}{p^2} - \frac{1}{p} = \frac{2-p}{p^2}$$

c $\mathrm{Var}(X) = \mathrm{E}(x^2) - (\mathrm{E}(x))^2 = \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{1-p}{p^2}$

5 $f(x) = p(1-p)^{x-1}$
As $1 - p < 1$ the sequence is decreasing and therefore the maximum value occurs for $x = 1$
For example, if a group of people roll a die till they obtain a '6', most of them will get a 6 at first attempt.

6 $\mathrm{P}(X > a + b \mid X > a) = \frac{\mathrm{P}(X > a+b)}{\mathrm{P}(X > a)} = \frac{q^{a+b}}{q^a} = q^b = (1-p)^b = \mathrm{P}(X > b)$

Exercise 4

1 $\mathrm{P}(X > a+b \mid X > a) = \frac{\mathrm{P}(X > a+b)}{\mathrm{P}(X > a)} = \frac{1 - F(a+b)}{1 - F(a)} = \frac{e^{-m(a+b)}}{e^{-ma}} = e^{-mb} = 1 - F(b) = \mathrm{P}(X > b)$

This means that no matter what happened before, the probability that a Poisson event occurs after a certain period of time a is the same.

2 The PDF is given by $f(x) = F'(x) = me^{-mx}$, $x \geq 0$ which is a decreasing function. Therefore its maximum value occurs at $x = 0$.

Exercise 5

1 The **Cauchy distribution**, also known as the Cauchy–Lorentz distribution, was first studied early in the 19th century by French mathematician Augustin-Louis Cauchy. It was later applied by the 19th-century Dutch physicist Hendrik Lorentz to explain forced resonance, or vibrations.

The Cauchy distribution is important as an example of a pathological case. Cauchy distributions look similar to a normal distribution. However, they have much heavier tails. The mean and standard deviation of the Cauchy distribution are undefined. The practical meaning of this is that 1000 data points give no more accurate an estimate of the mean and standard deviation than does a single point.