1 Random variables

A function which associates a unique real number with each outcome in the sample space is called a random variable. In this course we shall mainly focus on discrete random variables – that is experiments where the set of possible outcomes is countable $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$. In the last chapter we look at continuous random variables – that is experiments where the set of possible outcomes is $(-\infty, \infty)$. Every time the experiment is conducted exactly one value of the random variable is observed, this is called a realisation of the random variable.

Introductory examples of random variables

Exercise 3.1

Suppose we ask a student to react to the statement "Lancaster has excellent night life". Find a suitable sample space and sugggest a random variable to quantify the response.

Sol: 3.1

□

Exercise 3.2

Suppose we decide to record the number of children born in the local maternity ward tomorrow as a probability experiment. Find a suitable sample space and random variable.

Sol: 3.2

□

Discrete random variables arise in a variety of ways: From experiments

- with a natural integer valued outcome
  - the number of buses to stop in the hour,
  - the number of goals in a football match.
- with a continuous outcome which is recorded on an integer scale
  - the date in the year when the youngest in a family is born
- with non-integer outcomes to which numerical values are assigned
  - a coin is tossed the outcome is H or T, converted to 1 and 0 respectively.

In some cases numerical values cannot be given directly to outcomes as there is no logical ordering,
- the suit of a playing card drawn at random.
- the results of throwing a coin 3 times. But

Exercise 3.3

A coin is tossed 3 times. The sample space is

$$\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}.$$

Define a rv giving the number of Hs thrown.

Sol: 3.3

$\Omega = \{\omega;\ \omega = HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$. Define $R(\omega) = \#H$ in ω.

$$\begin{aligned} R(TTT) &= 0 \\ R(HTT) = R(THT) = R(TTH) &= 1 \\ R(HHT) = R(HTH) = R(THH) &= 2 \\ R(HHH) &= 3. \end{aligned}$$

Note the equivalence of the events $\{R = 2\} \iff \{HHT, HTH, THH\}$. The induced sample space for R is $\mathcal{S} = \{0, 1, 2, 3\}$ □

Exercise 3.4

Suppose our sample space consists of the outcomes of throwing a fair die, and suppose we gamble on the outcome, lose £1 if outcome is 1, 2 or 3; win nothing if outcome is 4: win £2 if outcome is 5 or 6. Define R to be the random variable profit, and find the induced sample space for R.

Sol: 3.4

□

To summarise

- The outcomes in the sample space, Ω, of the probability experiment may or may not be numerically valued.
- The random variable R is a function that associates a unique real number with each outcome in the sample space, Ω.
- A random variable is not, itself, a number.
- The values taken by random variable R defined on Ω, is known as the induced sample space for R and is sometimes written as $\mathcal{S}$.
- The event $\{R = r\} = \{\omega;\ R(\omega) = r\}$.

We cannot predict the value of R exactly since probability experiments are subject to chance. We can state the values R may take and attach probabilities to these values.

2 Probability mass functions

Here R is a discrete random variable that takes values in the non-negative integers, or a subset of them. The probability mass function of a discrete random variable, R, is defined by

$$p(r) = \mathsf{P}(\{R = r\}) \quad \text{for} \quad r = 0, 1, 2, \ldots.$$

We write pmf for probability mass function. Note that as $p(r)$ is a probability,

$$0 \le p(r) \le 1 \quad \text{for all} \quad r.$$

As $\Omega = \{R = 0, 1, 2, \ldots\}$

$$\begin{aligned} 1 &= \mathsf{P}(\Omega) \quad \text{by Axiom 2} \\ &= \mathsf{P}(\{R = 0, 1, 2, \ldots\} \\ &= \mathsf{P}(\{R = 0\}) + \mathsf{P}(\{R = 1\}) + \ldots \quad \text{by Axiom 3} \\ &= p(0) + p(1) + p(2) + \ldots. \end{aligned}$$

Thus If $p(r)$ is a probability mass function then it satisfies the conditions

$$0 \le p(r) \le 1 \quad \text{for all} \quad r \quad \text{and} \quad \sum_{r=0}^{\infty} p(r) = 1.$$

End 05.18.1

Exercise 3.5

Find the pmf of the number of heads in 3 tosses of a fair coin. Coin tossing sample space is $\Omega = \{\omega;\ \omega = HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

Sol: 3.5

By equi-probability or by independence

$$\begin{aligned} R(TTT) &= 0, \quad p(0) = 1/8 \\ R(HTT) = R(THT) = R(TTH) &= 1, \quad p(1) = 3/8 \\ R(HHT) = R(HTH) = R(THH) &= 2, \quad p(2) = 3/8 \\ R(HHH) &= 3, \quad p(3) = 1/8. \end{aligned}$$

Note $0 \le p(r) \le 1$ for all r and $\sum_{r=0}^{\infty} p(r) = 1$ □

Exercise 3.6

The length of stay in hospital after surgery is modelled as a random variable R. The following table gies the pmf for R.

Days stayed	r	4	5	6	7	8	9	10+	total
Probability	$\mathsf{P}(R = r)$	0.038	0.114	0.430	0.300	0.080	0.030	0.008	1

Find the probability of being in hospital for

(a) at most 6 days,

(b) between 5 and 7 days,

(c) at least 7 days.

Sol: 3.6

Days stayed	r	4	5	6	7	8	9	10+	total
Probability	$\mathsf{P}(R = r)$	0.038	0.114	0.430	0.300	0.080	0.030	0.008	1
Cumulative Probability	$\mathsf{P}(R \le r)$	0.038	0.152	0.582	0.882...				

at most 6 days: $\mathsf{P}(R \le 6) = 0.582$

between 5 and7: $\mathsf{P}(5 \le R \le 7) = 0.882 - 0.038 = 0.844$

at least 7: $\mathsf{P}(R \ge 7) = 1 - \mathsf{P}(R \le 6) = 1 - 0.582 = 0.418$ End 04.17.3 □

3 Probability of an event

The probability of an event can be obtained from the pmf directly. Without loss of generality suppose that the event of interest is $\mathsf{E} = \{r_1, \dots, r_k\}$, $E \subseteq \mathcal{S}$ then The probability of event E is $\sum_{r \in E} p(r)$.

Exercise 3.7

Find the probability of an odd number of heads in 3 tosses of a fair coin.

Sol: 3.7

$$\begin{aligned} R(TTT) = 0, &\quad p(0) = 1/8 \\ R(HTT) = R(THT) = R(TTH) = 1, &\quad p(1) = 3/8 \\ R(HHT) = R(HTH) = R(THH) = 2, &\quad p(2) = 3/8 \\ R(HHH) = 3, &\quad p(3) = 1/8. \end{aligned}$$

Then $\sum_{r=1,3} p(r) = 3/8 + 1/8 = 1/2.$ □

4 Expectation

Expectation is a simple measure to calculate the average value taken by a random variable. It measures the central location of the pmf of the random variable.

Exercise 3.8

Recall the gambling exercise above, where the random variable R, profit, is defined by

$$\begin{aligned} R(\omega) &= -1 \quad \text{if} \quad \omega = 1,2,3 \\ &= 0 \quad \text{if} \quad \omega = 4 \\ &= 2 \quad \text{if} \quad \omega = 5,6. \end{aligned}$$

from the throw of a fair die. The induced sample space for R is $\mathcal{S} = \{-1, 0, 2\}$. The pmf of R is $p(-1) = 3/6$, $p(0) = 1/6$, $p(2) = 2/6$. Conjecture how to define expected profit.

Sol: 3.8

□

These two calculations are the same because of the way the pmf is defined. We use the second as this is based on the pmf which makes the numerical values of the random variable explicit. The expected value of a discrete random variable R is

$$\mathsf{E}[R] = \sum_{r=0}^{\infty} r\ p(r).$$

The value $m = \mathsf{E}[R]$ is numerical and is known as the exepctation of R. [It is a population mean and is not to be confused with the sample mean.] The first calculation translates mathematically to $\sum_{\omega\in\Omega} r(\omega)\ \mathsf{P}(\omega)$.

Exercise 3.9

Let R be the random variable representing the number of episodes of middle ear ache (otitis media) in the first two years of life.

gqview weblib/exer/ear1.gif

r	0	1	2	3	4	5	6	7+
$p(r)$	0.129	0.264	0.271	0.185	0.095	0.039	0.017	0

Find the expected number of episodes of otitis media in the first two years of life.

Sol: 3.9

□

Exercise 3.10

Find the expected number of heads in 3 tosses of a fair coin, and the expected number of tails.

Exercise 3.11

Find the expected value of the score on a die.

Sol: 3.11

□

Exercise 3.12

Generalise the definition of expectation to obtain the expected value of the random variable $g(R)$, where g is a (mathematical) function.

Sol: 3.12

□

Exercise 3.13

If $p(r) = \frac{1}{3}$ for $r = 0, 1, 2$, find $\mathsf{E}(R^2)$. Here the function $g(r)$ is $g(r) = r^2$.

Sol: 3.13

□

Exercise 3.14

Find the expected value of a random digit R and its square for which $p(r) = 1/6$ for $r = 1, 2, \ldots, 6$.

Sol: 3.14

□

Expectation obeys two important rules of linearity. For arbitrary functions g and h, and a constant c:

$$\begin{aligned} E[g(R) + h(R)] &= \mathsf{E}[g(R)] + \mathsf{E}[h(R)] \\ \mathsf{E}[cg(R)] &= c\,\mathsf{E}[g(R)]. \end{aligned}$$

A special case is that $\mathsf{E}[c] = c$. These results can be verified using the definition of the expectation of a function.

Exercise 3.15

Find $\mathsf{E}[R]$ if it is known that $\mathsf{E}[R(R-1)] = 4$ and $\mathsf{E}[R^2] = 3$.

Sol: 3.15

$$\begin{aligned}
4 = \mathsf{E}[R(R-1)] &= \mathsf{E}[R^2 - R] \\
&= \mathsf{E}[R^2 + (-1)R] \\
&= \mathsf{E}[R^2 + (-1)R] \\
&= \mathsf{E}[R^2] + \mathsf{E}[(-1)R] \\
&= \mathsf{E}[R^2] + (-1)\,\mathsf{E}[R] \\
&= 3 - \mathsf{E}[R] \quad \text{so} \quad \mathsf{E}[R] = -1.
\end{aligned}$$

□

Variance

Expectation is a weighted average, and consequently is a measure of the location of the pmf. The spread, or dispersion, of a random variable is measured by the variance: the expected squared deviation about the expectation. The variance of the random variable R, $\text{var}(R)$, is defined as

$$\mathsf{var}(R) = \mathsf{E}[(R - \mathsf{E}[R])^2],$$

The standard deviation of R, $\mathsf{sd}(R)$, is defined to be the square root of the variance. Clearly the variance is the expectation of the function of the random variable $g(R) = (R-m)^2$, where $m = \mathsf{E}[R]$ is a number.

Example: Find the variance and standard deviation for R, the random variable, giving number of ear aches. Recall $\mathsf{E}(R) = 2.038$ and

r	0	1	2	3	4	5	6
$p(r)$	0.129	0.264	0.271	0.185	0.095	0.039	0.017

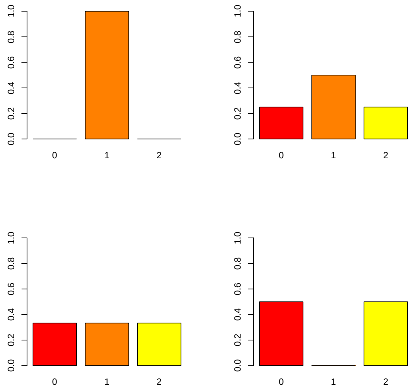
$\mathsf{E}[(R - \mathsf{E}(R))^2]$

$$\begin{aligned}
&= \mathsf{E}[(R-2.038)^2] \\
&= (0-2.038)^2 \times 0.129 + (1-2.038)^2 \times 0.264 + \ldots + (6-2.038)^2 \times 0.017 \\
&= 1.967.
\end{aligned}$$

Standard deviation $\mathsf{sd}(R) = \sqrt{1.967} = 1.402$. Aside: To get a feel for standard deviations experience suggests that for many random variables approximately 95% of the probability mass falls within ± 2 standard deviations of the mean of the random variable.

Exercise 3.16

Suppose that four random variables R_1, R_2, R_3 and R_4 on $\mathcal{S} = \{0, 1, 2\}$ have pmfs



	0	1	2
$p_1(r)$	0	1	0
$p_2(r)$	1/4	1/2	1/4
$p_3(r)$	1/3	1/3	1/3
$p_4(r)$	1/2	0	1/2

respectively.

Note for each pmf the sum of the probs is 1. The expectations are the same so that $\mathsf{E}[R_1] = \mathsf{E}[R_2] = \mathsf{E}[R_3] = \mathsf{E}[R_4] = 1$. Find the variances. Show that the variance ranks the random variables in order of increasing dispersion in the pmf.

Sol: 3.16

$$\begin{array}{llll}
(0-1)^2 \times 0 & +(1-1)^2 \times 1 & +(2-1)^2 \times 0 & = 0, \\
(0-1)^2 \times 1/4 & +(1-1)^2 \times 1/2 & +(2-1)^2 \times 1/4 & = 1/2, \\
(0-1)^2 \times 1/3 & +(1-1)^2 \times 1/3 & +(2-1)^2 \times 1/3 & = 2/3, \\
(0-1)^2 \times 1/2 & +(1-1)^2 \times 0 & +(2-1)^2 \times 1/2 & = 1.
\end{array}$$

$\mathsf{var}[R_1] < \mathsf{var}[R_2] < \mathsf{var}[R_3] < \mathsf{var}[R_4]$

□

This formulation of the variance is inconvenient for calculation, so alternative forms have been derived which simplify evaluation. Writing $\mathsf{E}[R]$ as μ a constant, we have

$$\begin{array}{rcll}
\mathsf{var}(R) & = & \mathsf{E}[(R - \mathsf{E}(R))^2], & \text{def} \\
 & = & \mathsf{E}[(R-\mu)^2], & \text{rewrite} \\
 & = & \mathsf{E}[R^2 - 2R\mu + \mu^2], & \text{expand} \\
 & = & \mathsf{E}[R^2] - 2\,\mathsf{E}[\mu R] + \mathsf{E}[\mu^2], & \text{lin} \\
 & = & \mathsf{E}[R^2] - 2\mu\,\mathsf{E}[R] + \mu^2, & \mathsf{E} \quad \text{const} \\
 & = & \mathsf{E}[R^2] - 2\mu\mu + \mu^2, & \text{def} \quad \mu \\
 & = & \mathsf{E}[R^2] - \mu^2, & \text{simplify} \\
 & = & \mathsf{E}[R^2] - (\,\mathsf{E}[R])^2, & \text{rewrite.}
\end{array}$$

End 05.18.3

Exercise 3.17

Find the variance of a random digit R uniformly distributed on the integers $r = 1, 2, \ldots, 6$. From above $\mathsf{E}(R) = 3.5$ and $\mathsf{E}(R^2) = \frac{91}{6}$.

Sol: 3.17

□

The next formula is often used with discrete rvs as $\mathsf{E}[R(R-1)]$ is easier to compute than $\mathsf{E}[R^2]$.

$$\begin{aligned} \mathsf{var}(R) &= \mathsf{E}[R^2] - (\mathsf{E}[R])^2 && \text{fromabove} \\ &= \mathsf{E}[R^2] - \mathsf{E}[R] + \mathsf{E}[R] - (\mathsf{E}[R])^2 && \text{add + subtract} \\ &= \mathsf{E}[R^2 - R] + \mathsf{E}[R] - (\mathsf{E}[R])^2 && \text{linearity} \\ &= \mathsf{E}[R(R-1)] + \mathsf{E}[R] - (\mathsf{E}[R])^2. && \text{simplify} \end{aligned}$$

As with linearity for expectation there is an important result for the variance of linear functions. Suppose a and b are constants

$$\begin{aligned} \mathsf{var}(aR+b) &= \mathsf{E}[(aR+b-\mathsf{E}[aR+b])^2] && \text{def var} \\ &= \mathsf{E}[(aR+b-(a\,\mathsf{E}[R]+b])^2] && \text{lin E} \\ &= \mathsf{E}[(aR-a\,\mathsf{E}[R])^2] && \text{factor} \\ &= \mathsf{E}[a^2(R-\mathsf{E}[R])^2] && \text{factor} \\ &= a^2\,\mathsf{E}[(R-\mathsf{E}[R])^2] && \text{lin E} \\ &= a^2\,\mathsf{var}(R). && \text{def var} \end{aligned}$$

This result shows the important properties of variance

$$\begin{aligned} &\text{if} \quad a=1, \qquad \mathsf{var}(R+b) = \mathsf{var}(R) \\ &\text{if} \quad b=0, \qquad \mathsf{var}(aR) = a^2\,\mathsf{var}(R). \end{aligned}$$

In summary

$$\begin{aligned} \mathsf{var}(R) &= \mathsf{E}[(R-\mathsf{E}(R))^2] \\ &= \mathsf{E}[R^2] - (\mathsf{E}[R])^2, \\ &= \mathsf{E}[R(R-1)] + \mathsf{E}[R] - (\mathsf{E}[R])^2, \end{aligned}$$

and for constants a and b

$$\mathsf{var}(aR+b) = a^2\,\mathsf{var}(R),$$

$$\mathsf{sd}(R) = \sqrt{\mathsf{var}(R)}.$$

End 04.18.3