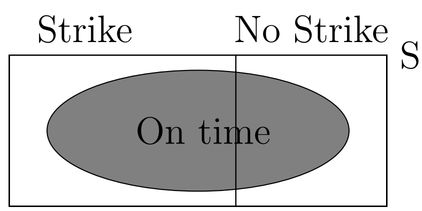
2.4 Rule of Total Probability and Bayes' Theorem

Example 2.4.1 *The completion of a construction job may be delayed because of a strike. The probabilities are* 0.60 *that there will be a strike,* 0.85 *that the construction job will be completed on time if there is no strike, and* 0.35 *that the construction job will be completed on time if there is a strike. What is the probability that the construction job will be completed on time?*



Let A be the event that the job will be completed on time,

B the event of a strike, therefore

B' is the event of no strike.

We want $P(A)$ and are given

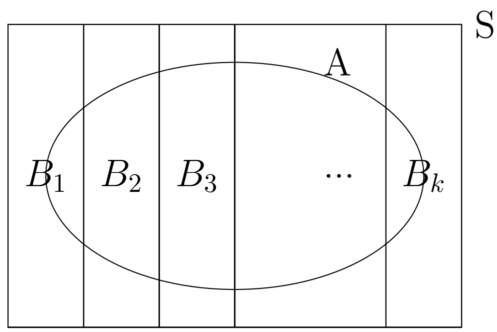
$$P(B) = 0.60, \quad P(A|B) = 0.35, \quad P(A|B') = 0.85.$$

Using the fact that $A = (A \cap B) \cup (A \cap B')$ (union of mutually exclusive events) and the multiplicative rule, we have

$$P(A) = P(A \cap B) + P(A \cap B') = P(B) \cdot P(A|B) + P(B') \cdot P(A|B').$$

Thus $P(A) =$

We can generalize the idea from the last example to obtain a formula for the probability of any event, given that we have a partition of our sample space into events of known probability.



(a partition of a set S is a collection of pairwise disjoint subsets whose union is S)

Theorem 2.4.2 ***(Rule of Total Probability)***

Suppose events $B_1, B_2, \ldots B_k$ form a partition of the sample space S, and $P(B_i) \neq 0$ for $i = 1, \ldots, k$. Then for any event A in S,

$$P(A) = \sum_{i=1}^{k} P(B_i) \cdot P(A|B_i).$$

Example 2.4.3 *Three machines M_1, M_2 and M_3 produce respectively 40, 10 and 50 percent of the items in a factory. The percentage of defective items produced by each respective machine is 2, 3 and 4 percent.*

Find the probability that a randomly selected item from the factory is defective.

Let D denote the event that a randomly selected item is defective.

Then,

2.4.1 Bayes' Theorem

Example 2.4.4 *With reference to the last example: if the randomly selected item is defective, what is the probability that the item was produced by*

(a) machine M_1*,*

(b) machine M_2*, or*

(c) machine M_3*?*

We calculated earlier that the probability of the item being defective is 0.031.

To answer the current question (part (a)), we first ask ourselves

"What is the contribution of the defectives from M_1 to the probability 0.031?"

The idea is captured in the following theorem.

Theorem 2.4.5 ***(Bayes' Theorem)*** *Suppose events $B_1, B_2, \ldots B_k$ form a partition of the sample space S, and $P(B_i) \neq 0$ for $i = 1, \ldots, k$. Then for any event A in S with $P(A) \neq 0$*

$$P(B_r|A) = \frac{P(B_r) \cdot P(A|B_r)}{\sum_{i=1}^{k} P(B_i) \cdot P(A|B_i)}$$

for $r = 1, \ldots, k$.

Proof 2.4.6

$$\begin{aligned} P(B_r|A) &= \frac{P(B_r \cap A)}{P(A)} \quad \text{\textit{(by definition)}} \\ &= \frac{P(B_r) \cdot P(A|B_r)}{P(A)} \quad \text{\textit{(multiplication rule)}} \\ &= \frac{P(B_r) \cdot P(A|B_r)}{\sum_{i=1}^{k} P(B_i) \cdot P(A|B_i)} \quad \text{\textit{(rule of total probability)}} \end{aligned}$$

Back to the example:

Example 2.4.7 *Three machine M_1, M_2 and M_3 produce respectively 40, 10 and 50 percent of the items in a factory. The percentage of defective items produced by each respective machine is 2, 3 and 4 percent. If the randomly selected item is defective, what is the probability that the item was produced by*

(a) machine M_1,

(b) machine M_2, or

(c) machine M_3?

Using Bayes' Theorem,

(a)

$$P(M_1|D) = \frac{P(M_1) \cdot P(D|M_1)}{P(D)} = \frac{(0.40)(0.02)}{0.031} \approx 0.2581$$

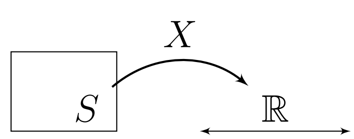
(b)

Chapter 3

Probability Distributions and Probability Densities

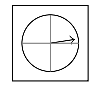
3.1 Random Variables

Let S be a sample space with a probability measure. A **random variable** is a function $X : S \to \mathbb{R}$, which maps the outcomes in the sample space to real numbers. The output of a random variable is something we can measure.



Random variables are defined when we want to focus on a particular property of the outcomes of an experiment. More than one random variable can be defined for a given sample space.

Usually capital letters like "X" are used to denote random variables; and their lower case letter like "x" are used for particular values that X can take.



Example 3.1.1 *Earlier we mentioned the experiment of spinning a probability spinner, and described the sample space as* $\{\theta \text{ degrees} | \theta \in [0, 360)\}$.

However, the actual sample space could include more information, such as multiple rotations, angular velocity at time t, elapsed time, the color it landed on, etc.

A random variable focuses on ***one*** *property of the outcome that can be assigned a real number.*

Some examples of random variables:

- X_1*: resting position (degrees), outputs values in* $[0, 360)$.
- X_2*: resting position (radians), outputs values in* $[0, 2\pi)$.
- X_3*: number of full rotations, can take values* $0, 1, 2, 3, \ldots$.

Example 3.1.2 *Two socks are selected at random and removed in succession from a drawer containing five brown socks and three green socks.*

List the elements of the sample space, the corresponding probabilities, and the corresponding values of the random variable X, where X is the number of brown socks selected.

Element of sample space	*Probability*	*value of X*
BB	$\frac{20}{56}$	*2*
BG		
GB		
GG		

We write: $P(X = 2) = \frac{20}{56}, \quad P(X = 1) = \frac{30}{56}, \quad P(X = 0) = \frac{6}{56}.$

Example 3.1.3 *Three balls are randomly chosen (without replacement) from a bag of 20 balls numbered 1-20. We bet that at least one of the numbers drawn is equal to or greater than 17. What is the probability of winning the bet?*

Outcomes in the sample space are subsets of three numbered balls, and they are all equally likely to occur.

Let random variable X denote the largest number of the three selected. Thus X takes values $3, 4, \ldots, 20$, and we want $P(X \geq 17)$.

Since total probability equals 1, $P(X \geq 17) = 1 - P(X \leq 16)$

Let us calculate $P(X \leq 16)$.

How many 3-element subsets of balls 1-16 are there?

How many 3-element subsets of balls 1-20 are there?

Recall that the set of all possible output values of a function is called its **range**.

If the range of a random variable X is a finite or countably infinite set, then we say that X is a **discrete random variable**.

In contrast, a **continuous random variables** is one whose range is a continuum of values, like an interval or a union of intervals in $\mathbb{R}$.

We will deal with this type later. The important difference to notice is in how the probabilities are assigned.

3.2 Probability Distributions

Example 3.2.1 *Experiment: Rolling two dice*

Let random variable X denote the sum of a roll. The range of X is $\{2, 3, \ldots, 12\}$.

Knowing that each outcome in the sample space has probability $\frac{1}{36}$, we can easily find the probability that X takes on any value in its range. e.g. $P(X = 7) = \frac{6}{36}$, $P(X = 11) = \frac{2}{36}$.

This information can be summarized in a table.

x	$P(X = x)$
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

We would like a rule $f(x)$ which gives $P(X = x)$ for each value x in the range of random variable X.

In this case the probabilities are given by the function

$$f(x) = \ldots.$$

Such a function is called a **probability distribution of X**.

If X is a discrete random variable, the function f given by

$$f(x) = P(X = x)$$

for each x in the range of X, is called the **probability distribution of** X. (Also called the **probability mass function** of X.)

Theorem 3.2.2 *A function f is allowable as a probability distribution for X if and only if its values, $f(x)$, satisfy*

1. $f(x) \geq 0$ *for any* x,
2. $\sum_x f(x) = 1$, *(sum taken over all x in the range of X)*

Example 3.2.3 *Let X be the random variable that counts the number of heads obtained in tossing a balanced coin 4 times.*

(a) What is the range of X?

(b) What is $P(X = x)$ for each x in the range of X?

(c) Find formula for the probability distribution of X.

Example 3.2.4 *Return to the dice rolling experiment.*

Let Y *be the maximum that either die shows in a single roll:* $Y(a,b) = \max(a,b)$.

For example, $Y(3,5) = 5$.

(a) What is the range of Y*?*

(b) What is $P(Y = y)$ *for each* y *in the range of* Y*?*

(c) Find a formula for the probability distribution of Y.

Example 3.2.5 *Check whether the function given by*

$$f(x) = \frac{x+2}{25},$$

for $x = 1, 2, 3, 4, 5$ *can serve as the probability distribution of a discrete random variable.*

Probability distributions for a random variable, say X, may be represented graphically by means of a **probability histogram**.

Each rectangle corresponds to a value for X, its height is $P(X = x)$, and its width is 1, so that the area of each rectangle equals $P(X = x)$. *The total area of the histogram is 1.*

The probability histogram below is for the number of heads in 4 coin flips.

3.3 Cumulative Distribution

In many problems we are interested in the probability that the value of a random variable is less than or equal to (or "at most") some real number x. i.e. $P(X \leq x)$.

If X is a discrete random variable with probability distribution f, the function given by

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$$

for $x \in (-\infty, \infty)$, is called the **cumulative distribution of** X(also called the **distribution function**).

Example 3.3.1 *Let X be the random variable that counts the number of heads in 4 coin flips.*

$f(2) = \dfrac{6}{16}$ *while*

$F(2) = f(0) + f(1) + f(2) = \dfrac{1}{16} + \dfrac{4}{16} + \dfrac{6}{16} = \dfrac{11}{16}$

The corresponding columns of the probability histogram are as follows:

Back to a previous example:

Example 3.3.2 *Two socks are selected at random and removed in succession from a drawer containing five brown socks and three green socks.*

Let X be the random variable that counts the number of brown socks selected.

We found these values earlier on:

Element of sample space	*Probability*	x
BB	$\frac{20}{56}$	*2*
BG	$\frac{15}{56}$	*1*
GB	$\frac{15}{56}$	*1*
GG	$\frac{6}{56}$	*0*

The probability distribution f is given by

$$f(x) = \begin{cases} \frac{20}{56} & \textit{for } x = 2 \\ \frac{30}{56} & \textit{for } x = 1 \\ \frac{6}{56} & \textit{for } x = 0 \end{cases}$$

Find $F(0), F(1), F(2)$, and express the cumulative distribution (distribution function) $F(x)$ as a piece-wise defined function.

Example 3.3.3 *Suppose a random variable X has range $\{1, 2, 3, 4\}$. Define f by*

$$f(1) = \frac{1}{4}, \quad f(2) = \frac{1}{2}, \quad f(3) = \frac{1}{8}, \quad f(4) = \frac{1}{8}$$

(a) Show that f is a valid probability distribution for X.

(b) Find the cumulative distribution (distribution function) for X.

Theorem 3.3.4 *The cumulative distribution $F(x)$ satisfies*

1. $\lim_{x\to-\infty} F(x) = 0$ *and* $\lim_{x\to\infty} F(x) = 1$.

2. *If $a < b$ then $F(a) \leq F(b)$ for any $a, b \in \mathbb{R}$.*

Theorem 3.3.5 *If the range of a random variable X consists of the values $x_1 < x_2 < \cdots < x_n$, then $f(x_1) = F(x_1)$ and*

$$f(x_i) = F(x_i) - F(x_{i-1})$$

for $i = 2, 3, \ldots, n$.

Let's see this on a probability histogram.

Example 3.3.6 *The cumulative distribution for a discrete random variable X is given by*

$$F(x) = \begin{cases} 0 & \text{for } x < -2 \\ \frac{4}{18} & \text{for } -2 \leq x < -1 \\ \frac{7}{18} & \text{for } -1 \leq x < 0 \\ \frac{12}{18} & \text{for } 0 \leq x < 1 \\ 1 & \text{for } x \geq 1 \end{cases}$$

Find the probability distribution for X.