If A_1 and A_2 are independent, are A_1^c and A_2^c also independent?

Let A_i be the event that Component i fails. Let A be the event that the system fails. Note that $A = A_1 \cup A_2$

To apply independence, we need an intersection, so instead can look at $A^c = (A_1 \cup A_2)^c = A_1^c \cap A_2^c$

$$\begin{aligned} P(A) &= P(A_1 \cup A_2) \\ &= P(A_1) + P(A_2) - P(A_1 \cap A_2) \\ &= 0.01 + 0.01 - \mathbb{P}(A_1)\mathbb{P}(A_2) \\ &= 0.02 - (0.01)^2 = 0.0199. \end{aligned}$$

$$\begin{aligned} \mathbb{P}(A^c) &= \mathbb{P}(A_1^c \cap A_2^c) = P(A_1^c)P(A_2^c) \\ &= (1 - P(A_1))(1 - P(A_2)) \\ \mathbb{P}(A) &= 1 - P(A^c) \\ &= 1 - (1 - 0.01)(1 - 0.01) \\ &\approx 0.0199 \end{aligned}$$

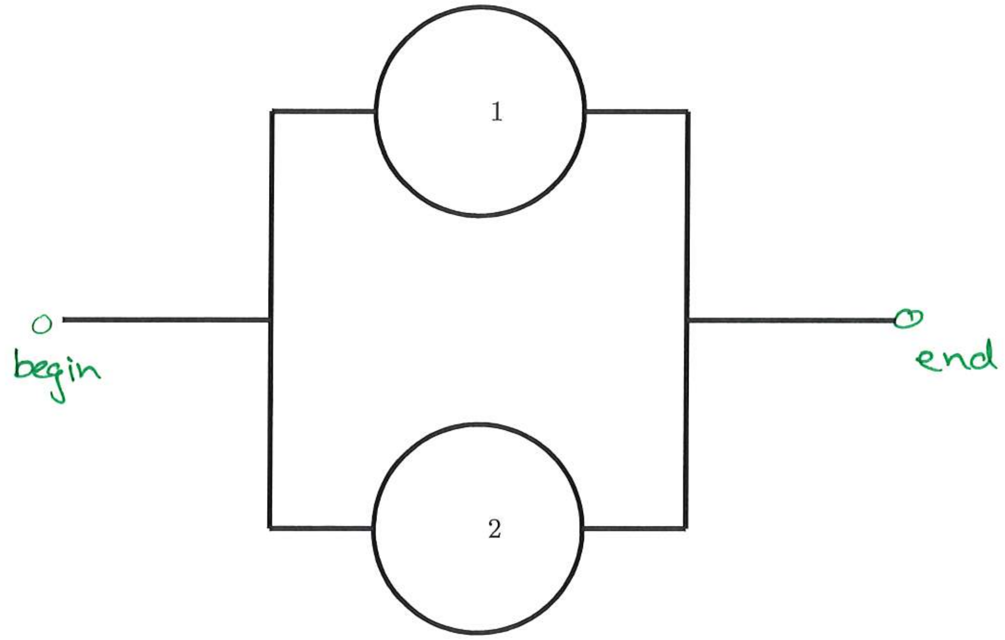

Figure 3.4: The system fails if both components fail.

Example: Consider the system of Figure 3.4, with two components (1 and 2), but this time, the system fails only if *both* of the components fails. Again, each component fails independently of the other with probability 0.01. What is the probability of a system failure?

Let A_i be the event that Component i fails. Let A be the event that the system fails. Note that $A = A_1 \cap A_2$, this time.

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(A_1 \cap A_2) \\ &= \mathbb{P}(A_1)\mathbb{P}(A_2) \quad \text{(by independence)} \\ &= 0.01 \times 0.01 \\ &= 0.0001 \end{aligned}$$

4

Discrete Random Variables

By the end of this chapter you should:

- Know what a random variable is.
- Identify some common distributions for discrete random variables.
- Compute the expectation and variance of a discrete random variable.
- Be able to determine if two discrete random variables are independent from their joint distribution.
- Be able to manipulate joint, marginal and conditional pmf's.
- Compute the expectation and variance of a sum of discrete random variables.
- To use moment generating functions to calculate moments and identify the distribution of a random variable.

Random Variables

Many of the situations we have been concerned with have naturally had a number associated with them. For example, the number resulting from the rolling of a fair die. In other situations the connection between the sample space and a real number may not be so obvious, but is usually very useful.

Suppose that the sample space consists of 20 students chosen at random from this class. What are some numbers associated with a particular outcome ω? ← omega element of Sample space

(one student) (20 students)

$X(\omega) =$ height (in cm) — average height

$Y(\omega) =$ gpa

$Z(\omega) =$ age — maximum age

The outcome of a random experiment is often expressed as a *number* or *measurement.*

nothing here

just missed a \newpage command

sample space

Definition. A function X assigning a real number to every outcome $\omega \in \Omega$ is called a *random variable.*

$$X : \Omega \longrightarrow \mathbb{R}$$

Notation: It is often highly useful to have a function which indicates simply whether (or not) an item belongs to a particular set. Recalling that *events* (e.g. A) are sets, an indicator function is a *random variable* $\mathrm{I}_A : \Omega \to \{0,1\}$ that takes on the value 1 if the outcome ω of our *random experiment* is in the set A (i.e. $\omega \in A$). (Zero otherwise)

Example. We toss a coin three times, with the tosses being independent. The sample space is

$$\Omega = \{H, T\}^3 = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

i.e. sequences of length three of Hs (failures) and Ts (successes).

Consider the function $X : \Omega \to \{0, \ldots, 3\}$ which maps $\omega = (\omega_1, \omega_2, \omega_3)$ to $\omega \in \Omega$

$$X(\omega) := \mathrm{I}_{\{\omega:\, \omega_1 = T\}}(\omega) + \mathrm{I}_{\{\omega:\, \omega_2 = T\}}(\omega) + \mathrm{I}_{\{\omega:\, \omega_3 = T\}}(\omega)\,.$$

defining

X is a random variable. The set $\{X = k\}$ corresponds to the set of outcomes with exactly k successes. Hence, we can interpret X as the *total number of successes in three identical coin tosses.*

Let $p \in [0,1]$ be the **probability of success** in a single coin toss. For example,

$$\begin{aligned}\mathbb{P}(\omega_1 = T, \omega_2 = T, \omega_3 = H) &= \mathbb{P}(\omega_1 = T)\,\mathbb{P}(\omega_2 = T)\,\mathbb{P}(\omega_3 = H) \\ &= p \times p \times (1-p) \\ &= p^2(1-p)\end{aligned}$$

by independence.

Then

$$\begin{aligned}\underbrace{\mathbb{P}(\{\omega \in \Omega : X(\omega) = 2\})}_{\text{2 tails out of 3 coin toss}} &= \mathbb{P}(\{\omega = TTH\} \cup \{\omega = THT\} \cup \{\omega = HTT\}) \\ &= \mathbb{P}(\{\omega = \{TTH\}\}) + \mathbb{P}(\{\omega = \{THT\}\}) + \mathbb{P}(\{\omega = \{HTT\}\}) \\ &= \\ &= p^2(1-p) + p^2(1-p) + p^2(1-p) \\ &= 3p^2(1-p) =: \binom{3}{2} p^2(1-p)\end{aligned}$$

Important Remarks

- Random variables are usually the *most convenient way to describe random ex-*

periments; they allow us to use intuitive notations for certain events, such as $\{X > 1000\}$, $\{\max(X, Y) \leqslant Z\}$, etc.

- Although mathematically a random variable is neither random nor a variable (it is a function), in practice we may interpret a random variable as the *measurement on a random experiment* which we will carry out *"tomorrow"*. However, all the *thinking about the experiment is done "today"*.
- We denote random variables by *upper case Roman letters*, X, Y, ...
- Numbers we get when we make the measurement (the outcomes of the random variables) are denoted by the *lower case* letter, such as x_1, x_2, x_3 for three values for X.

Example: Let X be the face value of a fair die, where $\Omega = \{1, 2, \ldots, 5, 6\}$ and $X(\omega) = \omega$ (this is a simple random variable because the sample space already consists of real numbers).

Notation: Events such as $\{\omega \in \Omega : X(\omega) \leqslant x\}$ and $\{\omega \in \Omega : X(\omega) = x\}$, for some real number x, are abbreviated to $\{X \leqslant x\}$ and $\{X = x\}$. The corresponding probabilities of these events, $\mathbb{P}(\{X \leqslant x\})$ and $\mathbb{P}(\{X = x\})$, are abbreviated further to $\mathbb{P}(X \leqslant x)$ and $\mathbb{P}(X = x)$, respectively.

$\{\omega \in \Omega : X(\omega) = 2 \text{ or } 4 \text{ or } 6\}$

Let A be the event that the face value is even, so $A = \{2, 4, 6\}$. Then

$$
\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(\{\omega \in \Omega : X(\omega) = 2 \text{ or } 4 \text{ or } 6\}) \\
&= \mathbb{P}(\{\omega \in \Omega : X(\omega) = 2\} \cup \{\omega \in \Omega : X(\omega) = 4\} \\
&\qquad \cup \{\omega \in \Omega : X(\omega) = 6\}) \\
&= \\
&= \mathbb{P}(X = 2) + \mathbb{P}(X = 4) + \mathbb{P}(X = 6) \\
&= 1/6 + 1/6 + 1/6 = 1/2.
\end{aligned}
$$

Types of Random Variable

Loosely speaking, the *set of all possible values* a random variable X that occur with positive probability is called the **support** of X, often denoted by $\mathrm{supp}(X)$ or Ω_X.

- **Discrete** random variables can only take *isolated* values.

  For example: a count can only take non-negative integer values.
- **Continuous** random variables can take values in an *interval*.

  For example: rainfall measurements, lifetimes of components, lengths, ... are (at least in principle) continuous.

If we know, or can find the probabilities for all events defined by a random variable X, we know the **(probability) distribution** of X.

Question. How do we write down the distribution of a random variable?

Cumulative Distribution Function

The following function is defined for both continuous and discrete random variables.

Definition. The **cumulative distribution function** (cdf) of X is the function $F : \mathbb{R} \to [0,1]$ defined by

$$F(x) = \mathbb{P}(X \leqslant x) \, .$$

$\lfloor \cdot \rfloor$ round down

Example: If X is the face value of a fair, six-sided die, then

$$\mathbb{P}(X \leqslant x) = \begin{cases} 0, & x \leqslant 0 \\ \lfloor x \rfloor / 6 & x \in (0,6) \\ 1, & x \geqslant 6 \end{cases}$$

The cdf is shown in Figure 4.1.

$\mathbb{P}(X \leqslant 3)$

$= \mathbb{P}(X=1 \text{ or } X=2 \text{ or } X=3)$

$= \frac{1}{2}$.

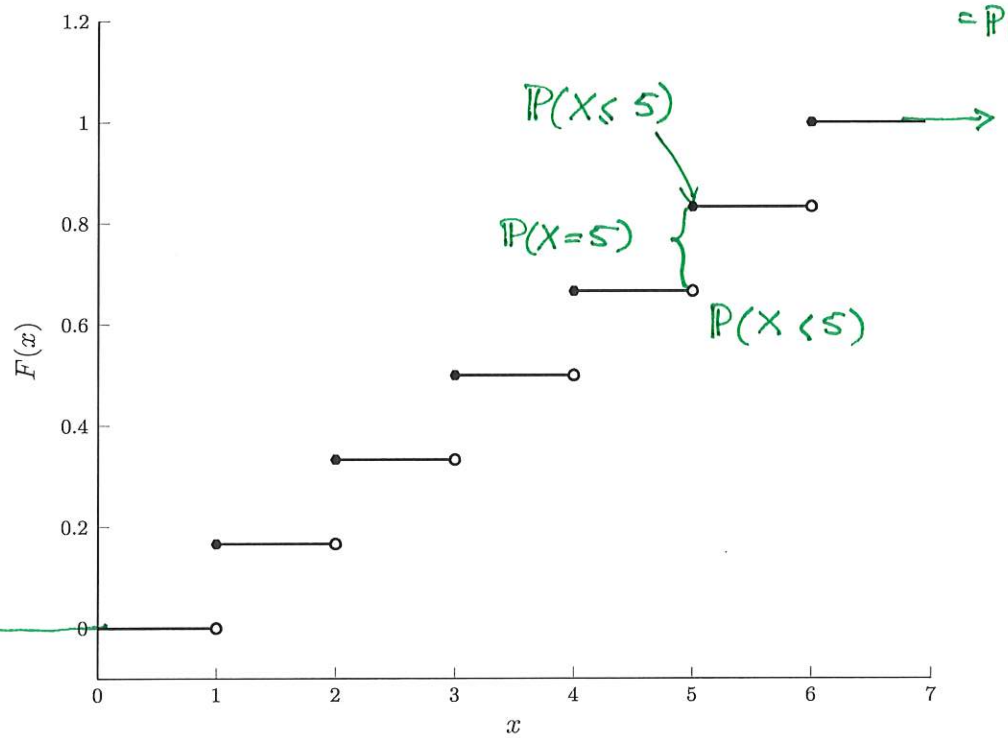

Figure 4.1: The cdf of the face value of a fair, six-sided die.

CDF Properties:

Axiom 1: $\mathbb{P}(A) \geqslant 0$

$\mathbb{P}(A) \leqslant 1$

$F(x) = \mathbb{P}(X \leqslant x)$

- $0 \leqslant F(x) \leqslant 1$.
- $\lim_{x \to \infty} F(x) = 1$ and $\lim_{x \to -\infty} F(x) = 0$.
- F is *non-decreasing*: If $x < y$, then $F(x) \leqslant F(y)$.
- F is *right-continuous*: If $x_n \downarrow x$, then $\lim_{n \to \infty} F(x_n) = F(x)$.

if $A \subseteq B$, $\mathbb{P}(A) \leqslant \mathbb{P}(B)$

$\{X \leqslant x\} \subseteq \{X \leqslant y\}$

For a *discrete* random variable X the cdf F is a **step function** with jumps of size $\mathbb{P}(X = x)$ at all the points $x \in \Omega_X$.

Probability Mass Function

Definition. For a *discrete* random variable X, the function $x \mapsto \mathbb{P}(X = x)$ is called the **probability mass function** (pmf) of X.

For a set B we have

$$f_X(B) = \mathbb{P}(X \in B) = \sum_{x \in B} \mathbb{P}(X = x) .$$

Example: Roll a die and let X be its face value. X is discrete with support $\Omega_X = \{1, 2, 3, 4, 5, 6\}$. If the die is fair, the probability mass function is given by

x	1	2	3	4	5	6	$\sum$
$f_X(x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	1

Note that $\mathbb{P}(X \in \Omega_X) =$ 1

Example: Roll two dice and let M be the largest face value showing. The distribution of M can be found to be

m	1	2	3	4	5	6	$\sum$
$f_M(m)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	1

or, as a formula:

$$f_M(m) = \mathbb{P}(M = m) = \frac{2m - 1}{36}, \text{ for } m = 1, 2, \ldots, 6 .$$

We can now work out the probability of *any* event defined by M so we know the distribution of M.

For example:

$$\boxed{\mathbb{P}(M > 4) = \mathbb{P}(M = 5 \text{ OR } M = 6) = \mathbb{P}(M=5) + \mathbb{P}(M=6) = \frac{9}{36} + \frac{11}{36}} = \frac{20}{36}$$

Common discrete distributions

Discrete Uniform Distribution

In the above example of rolling a die and recording the face value, all outcomes were equally likely. This is a uniform distribution on $\{1, 2, \ldots, 6\}$. In general, we say that a random variable X has a **discrete uniform distribution** on a finite set A if

$$\mathbb{P}(X = x) = \frac{1}{|A|}, \quad x \in A .$$