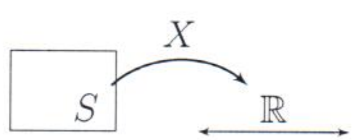
Chapter 3

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.*

5B 3G

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.

prob. that 1st sock is brown

prob. that 2nd sock is brown

	Element of sample space	Probability	value of X
$\frac{5}{8} \cdot \frac{4}{7} = \frac{20}{56}$	BB	$\frac{20}{56}$	2
$\frac{5}{8} \cdot \frac{3}{7}$	BG	$\frac{15}{56}$	1
$\frac{3}{8} \cdot \frac{5}{7}$	GB	$\frac{15}{56}$	1
$\frac{3}{8} \cdot \frac{2}{7}$	GG	$\frac{6}{56}$	0

random variable

$P(X=1) = \frac{15}{56} + \frac{15}{56} = \frac{30}{56}$

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

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, $P(X \geq 17) = 1 - P(X \leq 16)$

the event that the largest of the 3 balls is 16 or less.

the complement of the event $X \geq 17$

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

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

$$\binom{16}{3} = \frac{16 \cdot 15 \cdot 14}{3 \cdot 2 \cdot 1} = 560$$

# of 3-elt subsets of $\{1, \ldots, 16\}$

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

$$\binom{20}{3} = \frac{20 \cdot 19 \cdot \cancel{18}^{3}}{\cancel{3} \cdot \cancel{2} \cdot 1} = 1140$$

# of 3-elt subsets of $\{1, \ldots, 20\}$

$$P(X \leq 16) = \frac{560}{1140} = \frac{28}{57}$$

$$\Rightarrow P(X \geq 17) = 1 - \frac{28}{57} = \frac{29}{57}$$

Note that this probability is a bit larger than 50%.

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

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

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

$$f(x) = \frac{6-|x-7|}{36}$$

where $x = 2, \ldots, 12$

absolute value

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

$$f(3) = \frac{6-|3-7|}{36} = \frac{2}{36}$$

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?*

$\{0, 1, 2, 3, 4\}$

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

"probability that 0 heads occur"

$\leftarrow P(X=0) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{16}$ TTTT

HTTT
THTT
TTHT
TTTH

$P(X=1) = \frac{4}{16}$

$P(X=2) = \frac{6}{16}$

HHTT THHT
HTHT THTH
HTTH TTHH

$P(X=3) = \frac{4}{16}$

$P(X=4) = \frac{1}{16}$

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

$$P(X=x) = \frac{\binom{4}{x}}{2^4}$$

Example 3.2.4 *Return to the dice rolling experiment.* two dice are rolled

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

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

$Y(2,6) = 6$

$Y(4,4) = 4$

(a) What is the range of Y? $\{1, 2, 3, 4, 5, 6\}$

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

$P(Y=1)$

$P(Y=2)$

$P(Y=3)$

$P(Y=4)$

$P(Y=5)$

$P(Y=6)$

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

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