# Chapter 6 - Statistics

# Statistical distributions

## Chapter Overview

1. General probability distributions
2. Binomial distribution
3. Cumulative binomial probabilities

| | | | |
|---|---|---|---|
| **4**<br><br>**Statistical distributions** | 4.1 | **Understand and use simple, discrete probability distributions (calculation of mean and variance of discrete random variables is excluded), including the binomial distribution, as a model; calculate probabilities using the binomial distribution.** | **Students will be expected to use distributions to model a real-world situation and to comment critically on the appropriateness.**<br><br>**Students should know and be able to identify the discrete uniform distribution.**<br><br>**The notation $X \sim B(n, p)$ may be used.**<br><br>**Use of a calculator to find individual or cumulative binomial probabilities.** |

# Probability Distributions

You are already familiar with the concept of **variable** in statistics: a collection of values (e.g. favourite colour of students in the room):

| $\boldsymbol{x}$ | red | green | blue | orange |
|---|---|---|---|---|
| $\boldsymbol{P(X=x)}$ | 0.3 | 0.4 | 0.1 | 0.2 |

If each is assigned a probability of occurring, it becomes a **random variable**.

✎ A **random variable $X$ represents a single experiment/trial.** It consists of outcomes with a probability for each.

i.e. $X$ is a random variable (capital letter), but $x$ is a particular outcome.

$$P(X = x)$$

A shorthand for $P(X = x)$ is ✎ $p(x)$ (note the lowercase $p$).
It's like saying "the probability that the outcome of my coin throw was heads" ($P(X = heads)$) vs "the probability of heads" ($p(heads)$). In the latter the coin throw was implicit, so we can skip the '$X =$ '.

## Probability Distributions vs Probability Functions

There are two ways to write the mapping from outcomes to probabilities:

1. As a function

$$p(x) = \begin{cases} 0.1x, & x = 1,2,3,4 \\ 0, & otherwise \end{cases}$$

2. As a table

| $\boldsymbol{x}$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\boldsymbol{p(x)}$ | | | | |

Example

The random variable $X$ represents the **number of heads when three coins are tossed**.

**Sample Space**

**Distribution as a table**

| $\boldsymbol{x}$ | | | | |
|---|---|---|---|---|
| $\boldsymbol{p(X = x)}$ | | | | |

**Distribution as a function**

Example

**1.** A discrete random variable $X$ has the probability function

$$P(X=x)=\begin{cases}k(1-x)^2 & x=-1,\ 0,\ 1 \text{ and } 2\\ 0 & \text{otherwise.}\end{cases}$$

(*a*) Show that $k=\frac{1}{6}$.

**(3)**

| $x$ | | | | |
|---|---|---|---|---|
| $p(X=x)$ | | | | |

Remember:

## Probability of a Range

We may need to consider the probability of a range of solutions.

Example: Find the given probabilities for this probability distribution

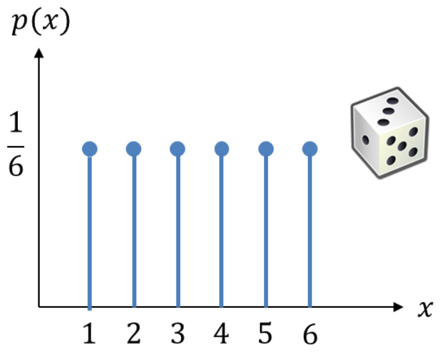
| $x$ | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| $p(x)$ | 0.1 | 0.3 | 0.2 | 0.4 |

$a)\ P(X>3)$

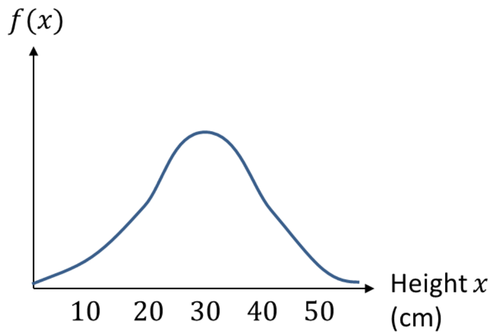
b) $P(2\leq X<4)$

c) $P(2X+1\geq 6)$

## We can also represent a probability distribution graphically:




- The throw of a die is an example of a **discrete uniform distribution** because the probability of each outcome is the same.
- $p(x)$ for discrete random variables is known as a **probability mass function**, because the probability of each outcome represents an actual 'amount' (i.e. mass) of probability.




- We can also have probability distributions for **continuous** variables, e.g. height
- However, the probability that something has a height of say **exactly** 30cm, is infinitely small (effectively 0).

$p(x)$ (written $f(x)$) for continuous random variables is known as a **probability density function**. $p(30)$ wouldn't give us the probability of being 30cm tall, but the amount of probability **per unit height**, i.e. the density. This is similar to histograms where frequency density is the "frequency per unit value". Just as an area in a histogram would then give a frequency, and area under a probability density graph would give a probability (mass).

You will encounter the **Normal Distribution** in Year 2, which is an example of a continuous probability distribution.

Textbook Exercise 6A Pg 86