Chapter 5

Discrete Distributions

Review of last week

Equally Likely Rule: Suppose S contains equally likely outcomes. Then

$$P(E) = \frac{|E|}{|S|}.$$

- Rules for counting: don't skip, don't double count
- **Multiplication principle**: one stage at a time
- **Binomial coefficient**: $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$ is the number of ways to choose k items from n items.

Review of last week

Conditional probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

- **Chain rule for probabilities**

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

- **Bayes' rule for inverse probabilities**

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

- **Independent events**

$$P(A \cap B) = P(A) \times P(B).$$

Random Variable

A random variable is a number that is obtained as or from the result of a random experiment.

- Say we flip a coin 3 times, then our sample space has $2^3 = 8$ possible outcomes:

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

  Typically, the particular sequence of heads or tails is of little interest; what does matter is the number of heads that result.

- If we define $X =$ number of heads in 3 tosses, we have captured the essence of the problem. We call X a random variable.

- Note: X defines a mapping (a function) from the original sample space S to a set of numbers.

  $$X(HHH) = 3, X(HHT) = 2, X(HTH) = 2, X(HTT) = 1$$

  $$X(THH) = 2, X(THT) = 1, X(TTH) = 1, X(TTT) = 0$$

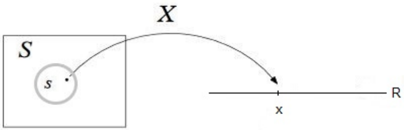
Random variable

Definition 5.1 Let S be a sample space associated with a random experiment, $\mathbb{R}$ is the real line and let

$$X : S \to \mathbb{R}.$$

Then X is called a **random variable** and

$$P(X = x) = P(s \in S : X(s) = x).$$



In defining a random variable, we have also created a new sample space (the range of the random variable).

Random variables often create a dramatically simpler sample space.

They also allow us to describe certain kinds of events very succinctly. In the coin rolling example, we can now write $P(X = 2)$ instead of $P(\text{there are two heads and one tail})$.

Example 5.1

Independent trials consisting of the flipping of a coin having probability $\frac{1}{3}$ of coming up heads are continually performed until a head occurs. The sample space is

$$S = \{(H); (T, H); (T, T, H); (T, T, T, H) \dots\}$$

Suppose we define the random variable X as the number of tosses for the first head to appear.

a. What is the *range*(X)?

Example 5.1

Independent trials consisting of the flipping of a coin having probability $\frac{1}{3}$ of coming up heads are continually performed until a head occurs. The sample space is

$$S = \{(H); (T, H); (T, T, H); (T, T, T, H) \dots\}$$

Suppose we define the random variable X as the number of tosses for the first head to appear.

b. Express the following event in random variable notation and calculate its probability: at least three tosses must be made for the first head to be observed.

We will be concerned with two main types of random variables: discrete and continuous.

A **discrete** random variable is one that can only take on a finite or countably infinite set of values.

- X, the number of heads in 3 flips of a coin is clearly a discrete random variable since $range(X) = \{0, 1, 2, 3\}$.

A **continuous** random variable can take on all the values in an interval.

- Y, the life length of a randomly selected bulb is theoretically a continuous random variable since $range(Y) = [0, \infty)$.

Example 5.2

In each case, state whether the random variable is discrete or continuous.

a. The distance traveled by a football when thrown.

b. Toss a coin repeatedly until the first head appears and record the number of tails.

Probability Mass Function

One useful way to describe the distribution of a discrete random variable, especially one with a finite range, is by way of a table.

X	Number of heads in 3 tosses			
x	0	1	2	3
probability	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

This is an example of a **Probability Mass Function** or PMF. As the name suggests, the PMF is associated with "point probabilities". Note that the table only shows the values which have positive probability.

Probability Mass Function

Definition 5.2 For a discrete random variable X, we define the **Probability Mass Function (PMF)** $f(x)$ by:

$$f(x) = P(X = x),\ \forall\ x.$$

We will write f_X when we want to emphasize the random variable.

The PMF $f(x)$ of a discrete random variable can be positive for at most a countable number of values. That is, if X can take values x_1, x_2,x_3, ... then

$$\begin{aligned} f(x_i) &> 0, \ i = 1, 2, 3, \ldots \\ &= 0 \text{ otherwise.} \end{aligned}$$

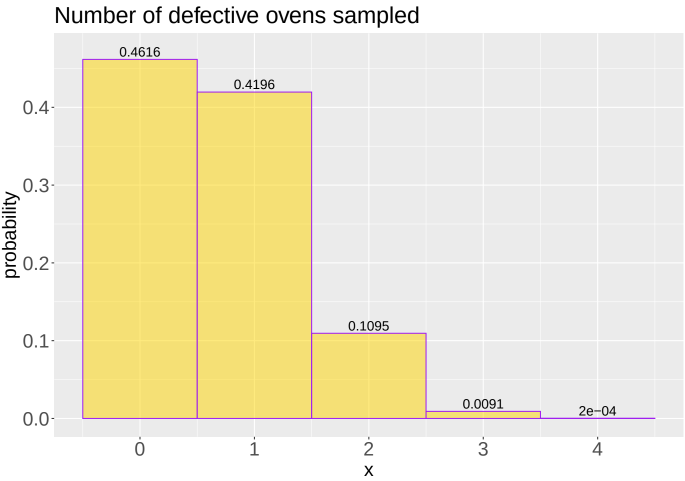
Furthermore, since X must take one of the values x_i, we have

$$\sum_{i=1}^{\infty} f(x_i) = 1.$$

Example 5.3

A store manager receives a shipment of 30 microwave ovens, 5 of which are (unknown to the manager) defective. The store manager takes a random sample of 4 ovens from the shipment and tests them to see if they are defective. Let X denote the number of defective ovens found. Write the PMF of X in a tabular format. (You may assume that every sample of 4 ovens is equally likely to be selected.)

Probability Histogram



Probability histogram: code

```r
# create data frame consisting of xi and f(xi)
library(tidyverse)
ovens <- data.frame(
  x = 0:4,
  f = c(0.4616, 0.4196, 0.1095, 0.0091, 0.0002)
)

# make probability histogram
ggplot(data = ovens, mapping = aes(x = x, y = f)) +
  geom_col(
    width = 1, alpha = 0.5,
    fill = "gold", color = "purple"
  ) +
  geom_text(mapping = aes(label = round(f, 4), y = f + 0.01)) +
  labs(
    x = "x",
    y = "probability",
    title = "Number of defective ovens sampled"
  )
```

For random variables with infinitely many possible values, a formula provides the most concise representation of the PMF.

Example 5.4

Let X denote the number of tosses until the first head when tossing a fair coin. Find the P.M.F. of X. You may assume the outcome on one toss is independent of the outcome on a different toss.

Cumulative Distribution Function

There is yet one more important way to describe the distribution of a discrete random variable, with a **cumulative distribution function (CDF)**

Definition 5.3 The **Cumulative Distribution Function** F of a random variable X is defined by

$$F(x) = P(X \leq x), \ \forall x.$$

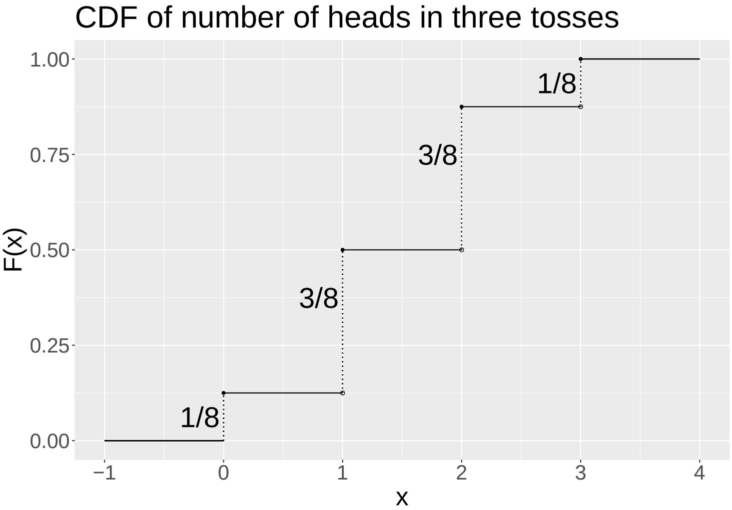
PMF for number of heads in 3 tosses

X	Number of heads in 3 tosses			
x	0	1	2	3
probability	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

CDF for number of heads in 3 tosses

$$
\begin{aligned}
F(x) &= P(X \leq x) \\
&= \begin{cases} 0 & x < 0 \\ \frac{1}{8} & 0 \leq x < 1, \\ \frac{4}{8} & 1 \leq x < 2 \\ \frac{7}{8} & 2 \leq x < 3 \\ 1 & 3 \leq x \end{cases}
\end{aligned}
$$

The CDF for the number of heads in 3 tosses is graphed below:



- The graph of the CDF of a discrete random variable is a step function.
- The step function is non decreasing, has jumps at each of the possible values x and the size of the jump is equal to $P(X = x)$.
- Note that, at the jump point F takes the value at the top of the jump. This is known as *right continuity*.

Example 5.5

A discrete uniform random variable X has a PMF of the form

$$f(x) = \frac{1}{n}, \quad x = 1, 2, \dots, n.$$

Find the CDF of X.

CDF and PMF

There is of course a connection between the PMF and CDF of a given random variable:

- To get the CDF from the PDF, we simply add up the probabilities for all possible values up to and including x.
- To get the PMF from the CDF, we look at how much the CDF has changed from the last jump.