Lecture 6

Part 2 Probability and Distributions

Example 4

- There are three buckets that weigh 1lb, 2lb and3lb, respectively. A bucket weighing i lb contains i white balls and $5 - i$ black balls, for $i = 1, 2, 3$. For example, the bucket weighing 1lb contains 1 white ball and 4 black balls. Suppose a bucket is chosen with probability proportional to its weight and two balls are randomly selected (without replacement) from this bucket.

(a) Find the probability that both balls selected are white.

(b) If both balls selected are white, what is the probability that the bucket weighing 3lb was chosen?

Solution - Part (a)

- Let A be the event "both balls selected are white".
- Let B_i be the event that the bucket weighing $i$lb is selected.
- Note that B_1, B_2 and B_3 are a partition of S.
- Therefore, by the Law of Total Probability:

$$
\begin{aligned}
P(A) &= P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3) \\
&= P(A|B_1) \times P(B_1) + P(A|B_2) \times P(B_2) \\
&\quad + P(A|B_3) \times P(B_3)
\end{aligned}
$$

Solution - Part (a)

- The total weight of the buckets is $1+2+3=6$lb.
- Therefore, $P(B_1)=\frac{1}{6}$, $P(B_2)=\frac{1}{3}$ and $P(B_3)=\frac{1}{2}$.
- What about $P(A|B_1)$, $P(A|B_2)$ and $P(A|B_3)$?

$$P(A|B_1)=0$$

$$P(A|B_2)=\frac{2}{5}\times\frac{1}{4}=\frac{1}{10}$$

$$P(A|B_3)=\frac{3}{5}\times\frac{2}{4}=\frac{3}{10}$$

Solution - Part (a)

- Putting this all together, we get:

$$
\begin{aligned}
P(A) &= P(A|B_1) \times P(B_1) + P(A|B_2) \times P(B_2) \\
&\quad + P(A|B_3) \times P(B_3) \\
&= 0 \times \frac{1}{6} + \frac{1}{10} \times \frac{1}{3} + \frac{3}{10} \times \frac{1}{2} \\
&= \frac{11}{60} \\
&= 0.1833
\end{aligned}
$$

Solution - Part (b)

- We want to find $P(B_3|A)$:

$$
\begin{aligned}
P(B_3|A) &= \frac{P(B_3 \cap A)}{P(A)} \\
&= \frac{P(A|B_3) \times P(B_3)}{P(A)} \\
&= \frac{\frac{3}{10} \times \frac{1}{2}}{\frac{11}{60}} \\
&= \frac{9}{11} \\
&= 0.8182
\end{aligned}
$$

Let's talk about **distribution**!

Random Variable

- Suppose we flip a fair coin three times.
- The sample space is:

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

- Each outcome is equally likely to occur.
- Define a new quantity (call it X) which is equal to the number of heads that occur in the three coin flips.

Random Variable

- X can take the value 0, 1, 2 or 3.
- The actual value that X takes is random and depends on the outcome of the experiment.
- X is what we call a random variable.
- Formally, a **random variable** is a function that assigns a numeric value to each simple event in a sample space.

Notation

- Denote random variables using uppercase letters, e.g., X, Y, Z.
- Denote the actual observed or *realised* value of the random variable by lowercase letters, e.g., x, y, z.
- Back to coin flipping example:
  - X is the random variable that can take values 0, 1, 2 or 3.
  - If we actually perform the experiment and observe the outcome HHT, then the realized value of X is $x = 2$.

Discrete Random Variable

- A **discrete random variable** is one that can take on a countable number of possible values.
- For example:
  - Flip a coin five times and let X be the number of heads that occurs. The possible values are $X = 0, 1, 2, 3, 4$ or 5.
  - Flip a coin until it comes up tails and let X be the total number of flips needed. The possible values are $X = 1, 2, 3, 4, 5, 6, 7, \ldots$

Continuous Random Variable

- A **continuous random variable** is one that can take on an uncountable number of possible values - the number of possible values is infinite as a result of continuous variation.
- For example:
  - Let X be the time taken to finish a three hour exam.
  - Let X be the weight of a boxer.

Discrete Probability Distribution

- For a discrete random variable X, how can we determine $P(X = x)$ for any given value of x?
- The probability that a discrete random variable X takes the value x is denoted by $p(x)$ and is equal to the sum of all the probabilities of the simple events for which $X = x$.

Discrete Probability Distribution

- A discrete probability distribution is a table or formula listing all possible values that a discrete random variable can take, together with the corresponding probability for each value.
- A discrete probability distribution must satisfy two requirements:

$$\begin{aligned} &1. \quad 0 \le p(x) \le 1 \text{ for all } x. \\ &2. \quad \sum_{\text{all } x} p(x) = 1 \end{aligned}$$

Example

- Flip a coin three times, let X be the number of heads.

$$p(0) = P(X = 0) = P(\{TTT\}) = \frac{1}{8}$$

$$p(1) = P(X = 1) = P(\{HTT, THT, TTH\}) = \frac{3}{8}$$

$$p(2) = P(X = 2) = P(\{HHT, HTH, THH\}) = \frac{3}{8}$$

$$p(3) = P(X = 3) = P(\{HHH\}) = \frac{1}{8}$$

Example

x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

- What is the probability of at most one head?

$$P(X \leq 1) = p(0) + p(1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}$$

- What is the probability of at least one head?

$$\begin{aligned} P(X \geq 1) &= p(1) + p(2) + p(3) = \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = \frac{7}{8} \\ &= 1 - p(0) = 1 - \frac{1}{8} = \frac{7}{8} \end{aligned}$$

Probability Distributions and Populations

- Probability distributions represent populations.
- Rather than recording every observation in the population, a probability distribution summaries the population by listing only the possible values that appear in the population, together with their corresponding probabilities.
- We can calculate population parameters such as the population mean and population variance from a probability distribution.

Expected Value

- Let X be a discrete random variable with probability distribution $p(x)$. The **expected value** (or **population mean**) of X is defined to be:

$$\mu = \mathbb{E}(X) = \sum_{\text{all } x} (x \times p(x))$$

- Compare this to the formula for the population mean given in topic 1:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} X_i = \sum_{i=1}^{N} \left(x_i \times \frac{1}{N} \right)$$

Expected Value

- It is straightforward to calculate the expected value of any function of a discrete random variable X.
- Let $g(X)$ be some function of X. Then the expected value of $g(X)$ is defined to be:

$$\mathbb{E}(g(X)) = \sum_{\text{all } x} (g(x) \times p(x))$$

Example

x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$\begin{aligned}\mathbb{E}(X) &= \sum_{\text{all } x}(x \times p(x)) \\ &= 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} \\ &= \frac{12}{8} \\ &= 1.5\end{aligned}$$

Example

x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$\begin{aligned}
\mathbb{E}(X^2) &= \sum_{\text{all } x} (x^2 \times p(x)) \\
&= 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} \\
&= \frac{24}{8} \\
&= 3
\end{aligned}$$

Laws of Expected Value

- If X and Y are random variables (discrete or continuous) and c is any constant, then:

$$
\begin{aligned}
&1. \quad \mathbb{E}(c) = c \\
&2. \quad \mathbb{E}(cX) = c\mathbb{E}[X] \\
&3. \quad \mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y) \\
&4. \quad \mathbb{E}(X-Y) = \mathbb{E}(X) - \mathbb{E}(Y)
\end{aligned}
$$

- And if X and Y are independent, then:

$$5. \quad \mathbb{E}(XY) = \mathbb{E}(X) \times \mathbb{E}(Y)$$

Example

- Let $Z = 3X + 2Y - 2XY + 3$ with $\mathbb{E}(X) = 3, \mathbb{E}(Y) = 5$ and X and Y independent. Then:

$$\begin{aligned}
\mathbb{E}(Z) &= \mathbb{E}(3X + 2Y - 2XY + 3) \\
&= \mathbb{E}(3X) + \mathbb{E}(2Y) - \mathbb{E}(2XY) + \mathbb{E}(3) \\
&= 3\mathbb{E}(X) + 2\mathbb{E}(Y) - 2\mathbb{E}(X)\mathbb{E}(Y) + 3 \\
&= 3 \times 3 + 2 \times 5 - 2 \times 3 \times 5 + 3 \\
&= -8
\end{aligned}$$