Statistics

1 Probability

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

$$\Pr(A \cup B) = 0 \qquad \text{(mutually exclusive)}$$

2 Conditional probability

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} \quad \text{where } \Pr(B) \neq 0$$

$$\Pr(A) = \Pr(A|B) \cdot \Pr(B) + \Pr(A|B') \cdot \Pr(B') \qquad \text{(law of total probability)}$$

$$\Pr(A \cap B) = \Pr(A|B) \times \Pr(B) \qquad \text{(multiplication theorem)}$$

For independent events:

- $\Pr(A \cap B) = \Pr(A) \times \Pr(B)$
- $\Pr(A|B) = \Pr(A)$
- $\Pr(B|A) = \Pr(B)$

2.1 Discrete random distributions

Any experiment or activity involving chance will have a probability associated with each result or *outcome*. If the outcomes have a reference to **discrete numeric values** (outcomes that can be counted), and the result is unknown, then the activity is a *discrete random probability distribution.*

2.1.1 Discrete probability distributions

If an activity has outcomes whose probability values are all positive and less than one ($\implies 0 \leq p(x) \leq 1$), and for which the sum of all outcome probabilities is unity ($\implies \sum p(x) = 1$), then it is called a *probability distribution* or *probability mass* function.

- **Probability distribution graph -** a series of points on a cartesian axis representing results of outcomes. $\Pr(X = x)$ is on y-axis, x is on x axis.
- **Mean** μ **-** measure of central tendency. Also known as *balance point* or *expected value* of a distribution. Centre of a symmetrical distribution.
- **Mode -** most popular value (has highest probability of X values). Multiple modes can exist if > 1 X value have equal-highest probability.
- **Median** m **-** the value of x such that $\Pr(X \leq m) = \Pr(X \geq m) = 0.5$. If $m > 0.5$, then value of X that is reached is the median of X. If $m = 0.5 = 0.5$, then m is halfway between this value and the next.
- **Variance** σ^2 **-** measure of spread of data around the mean. Not the same magnitude as the original data. Represented by $\sigma^2 = \text{Var}(x) = \sum(x = \mu)^2 \times p(x) = \sum(x - \mu)^2 \times \Pr(X = x)$. Alternatively: $\sigma^2 = \text{Var}(X) = \sum x^2 \times p(x) - \mu^2$
- **Standard deviation** σ **-** measure of spread in the original magnitude of the data. Found by taking square root of the variance: $\sigma = \text{sd}(X) = \sqrt{\text{Var}(X)}$

2.1.2 Expectation theorems

$$\overline{x} = \frac{\Sigma(xf)}{\Sigma(f)} = \Sigma(xp(x)) \qquad \text{(expected value)}$$

$$\begin{aligned} E(aX \pm b) &= aE(X) \pm b \\ E(z) &= z \\ E(X+Y) &= E(X) + E(Y) \\ E(X)^n &= \Sigma x^n \cdot p(x) \\ &\neq [E(X)]^2 \end{aligned}$$