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 or **expected value** $E(X)$ - measure of central tendency. Also known as *balance point.* Centre of a symmetrical distribution.

$$\begin{aligned}\overline{x} = \mu = E(X) &= \frac{\Sigma(xf)}{\Sigma(f)} \\ &= \sum_{i=1}^{n}(x_i \cdot P(X = x_i))\end{aligned}$$

- **Mode** - most popular value (has highest probability of X values). Multiple modes can exist if > 1 X value have equal-highest probability. Number must exist in distribution.
- **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. For distribution $x_1 \mapsto p_1, x_2 \mapsto p_2, \dots, x_n \mapsto p_n$:

$$\begin{aligned}\sigma^2 = \text{Var}(x) &= \sum_{i=1}^{n} p_i(x_i - \mu)^2 \\ &= \sum (x-\mu)^2 \times \Pr(X = x) \\ &= \sum x^2 \times p(x) - \mu^2\end{aligned}$$

- **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

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

3 Binomial Theorem

$$\begin{aligned}(x+y)^n &= \binom{n}{0} x^n y^0 + \binom{n}{1} x^{n-1} y^1 + \binom{n}{2} x^{n-2} y^2 + \cdots + \binom{n}{n-1} x^1 y^{n-1} + \binom{n}{n} x^0 y^n \\ &= \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \\ &= \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}\end{aligned}$$

1. powers of x decrease $n \to 0$
2. powers of y increase $0 \to n$
3. coefficients are given by nth row of Pascal's Triangle where $n = 0$ has one term
4. Number of terms in $(x+a)^n$ expanded & simplified is $n + 1$

Combinations: ${}^n\text{C}_r = \binom{N}{k}$ (binomial coefficient)

- Arrangements $\binom{n}{k} = \frac{n!}{(n-r)}$
- Combinations $\binom{n}{k} = \frac{n!}{r!(n-r)!}$
- Note $\binom{n}{k} = \binom{n}{k-1}$

3.0.1 Pascal's Triangle

$n =$													
0							1						
1						1		1					
2					1		2		1				
3				1		3		3		1			
4			1		4		6		4		1		
5		1		5		10		10		5		1	
6	1		6		15		20		15		6		1