1 Probability

Probability theorems

Union:	$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$
Multiplication theorem:	$\Pr(A \cap B) = \Pr(A\|B) \times \Pr(B)$
Conditional:	$\Pr(A\|B) = \dfrac{\Pr(A \cap B)}{\Pr(B)}$
Law of total probability:	$\Pr(A) = \Pr(A\|B) \cdot \Pr(B) + \Pr(A\|B') \cdot \Pr(B')$

Mutually exclusive $\implies \Pr(A \cup B) = 0$

Independent events:

$$\begin{aligned}\Pr(A \cap B) &= \Pr(A) \times \Pr(B)\\ \Pr(A|B) &= \Pr(A)\\ \Pr(B|A) &= \Pr(B)\end{aligned}$$

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*.

Discrete probability distributions

If an activity has outcomes whose probability values are all positive and less than one ($\implies 0 \le p(x) \le 1$), 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.

$$\begin{aligned}\overline{x} = \mu = E(X) &= \frac{\Sigma\left[x \cdot f(x)\right]}{\Sigma f} && \text{(where } f = \text{absolute frequency)}\\ &= \sum_{i=1}^{n}\left[x_i \cdot \Pr(X = x_i)\right] && \text{(for } n \text{ values of } x)\\ &= \int_{-\infty}^{\infty} (x \cdot f(x))\, dx && \text{(for pdf } f)\end{aligned}$$

- **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 $\Pr(X \le m) = \Pr(X \ge m) = 0.5$. 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. To find m, add values of X from smallest to alrgest until the sum reaches 0.5.

$$m = X \text{ such that } \int_{-\infty}^{m} f(x)dx = 0.5$$

- **Variance** σ^2 - measure of spread of data around the mean. Not the same magnitude as the original data. For distribution $x_1 \mapsto p_1, x_2 \mapsto p_2, \ldots, x_n \mapsto p_n$:

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

- **Standard deviation** σ - measure of spread in the original magnitude of the data. Found by taking square root of the variance:

$$\begin{aligned}\sigma &= \text{sd}(X)\\ &= \sqrt{\text{Var}(X)}\end{aligned}$$

Expectation theorems

For some non-linear function g, the expected value $E(g(X))$ is not equal to $g(E(X))$.

$$\begin{aligned}E(X^n) &= \Sigma x^n \cdot p(x) && \text{(non-linear function)}\\ &\neq [E(X)]^n\\ E(aX \pm b) &= aE(X) \pm b && \text{(linear function)}\\ E(b) &= b && \text{(for constant } b \in \mathbb{R}\text{)}\\ E(X+Y) &= E(X) + E(Y) && \text{(for two random variables)}\end{aligned}$$

2 Binomial Theorem

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

Patterns

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 $n+1$

Combinatorics

$$\text{Binomial coefficient:} \quad {}^n\text{C}_r = \binom{N}{k}$$

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

On CAS: (soft keyboard) [↓] → [Advanced] → `nCr(n,cr)`

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

3 Binomial distributions

(aka Bernoulli distributions)

$$\begin{aligned}\text{Defined by} \quad & X \sim \text{Bi}(n, p) \\ \implies \Pr(X = x) &= \binom{n}{x} p^x (1-p)^{n-x} \\ &= \binom{n}{x} p^x q^{n-x}\end{aligned}$$

where:

n is the number of trials

There are two possible outcomes: S or F

$\Pr(\text{success}) = p$

$\Pr(\text{failure}) = 1 - p = q$

Conditions for a binomial variable/distribution

1. Two possible outcomes: **success** or **failure**
2. Pr(success) is constant across trials (also denoted p)
3. Finite number n of independent trials

Solve on CAS

Main → Interactive → Distribution → `binomialPDf`
Input `x` (no. of successes), `numtrial` (no. of trials), `pos` (probbability of success)

Properties of $X \sim \text{Bi}(n, p)$

$$\begin{aligned}\textbf{Mean} \quad & \mu(X) = np \\ \textbf{Variance} \quad & \sigma^2(X) = np(1-p) \\ \textbf{s.d.} \quad & \sigma(X) = \sqrt{np(1-p)}\end{aligned}$$

Applications of binomial distributions

$$\Pr(X \geq a) = 1 - \Pr(X < a)$$