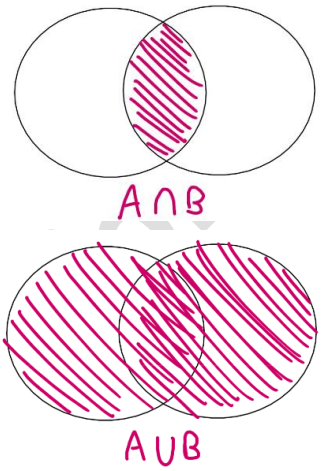
Complete VCE Education

Unit 4 AOS 2 Probability – Cheat Sheet

Probability basics



- **Union:** $A \cup B$ (Event A or Event B or both)
- **Intersection:** $A \cap B$ (both Event A and Event B)
- **Mutually exclusive:** When two events A and B have no outcomes in common. $\Pr(A \cap B) = 0$, so $\Pr(A \cup B) = \Pr(A) + \Pr(B)$
- **Complement:** $Pr(A') = 1 - Pr(A)$, where A' is the complement of A
- **Addition rule:** $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$

Conditional probability and independence

- $Pr(A \mid B)$: probability of event A occurring given that event B has occurred
- **Conditional probability:** $Pr(A \mid B) = \frac{Pr(A \cap B)}{Pr(B)}$
- **Multiplication rule:** $Pr(A \cap B) = Pr(A \mid B)\, Pr(B)$
- **Other useful rules:**
  - $\Pr(A) = \Pr(A \cap B) + \Pr(A \cap B')$
  - **Law of total probability:** $Pr(A) = Pr(A \mid B)Pr(B) + Pr(A \mid B')\ Pr(B')$
- **Independent events:** one event occurring does not affect the probability of the other occurring
  - $\Pr(A \mid B) = \Pr(A)$
  - $\Pr(A \cap B) = \Pr(A) \times \Pr(B)$

Discrete random variables

- **Discrete random variable:** Random variable that can only take a countable number of values (e.g. X=0, 1, 2…)
- **Conditions of a probability mass function:**
  - $0 \leq p(x) \leq 1$ for all x
  - $\sum p(x) = 1$ (sum of probabilities is 1)
- **Expected value:**
  - $E(X) = \sum x \cdot p(x)$
  - $E[g(X)] = \sum g(x) \cdot p(x)$
  - $E(aX + b) = aE(X) + b$
  - $E(X + Y) = E(X) + E(Y)$
- **Variance:**
  - $Var(X) = E(X^2) - [E(X)]^2$
  - $Var(aX + b) = a^2 Var(X)$

- **Standard deviation:**
  - $sd(X) = \sqrt{Var(X)}$
  - $sd(aX + b) = \sqrt{Var(aX + b)} = \sqrt{a^2 Var(X)} = |a|sd(X)$

Binomial probability distribution

- **Features of binomial distribution:**
  - Each trial has only two outcomes; **a success or a failure**
  - The probability of a success on a single trial, p, is constant for all trials
  - The trials are independent
- **Formula:** $\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$, where $\binom{n}{x} = \frac{n!}{x!(n-x)!}$
- <u>On calculator:</u>
  - Binomial Pdf if X is equal to only one value (e.g. $X = 3$ successes)
  - Binomial Cdf when X is equal to multiple values (e.g. $X > 3$ successes)
- Common mistake: if the question is asking for $Pr(X < 3)$, when using binomial Cdf on the CAS, the lower bound will be 0 and the **upper bound** will be 2 (since 3 is not included).
- **Expected value:** $E(X) = np$
- **Variance:** $Var(X) = np(1-p)$
- **Standard deviation:** $sd(X) = \sqrt{Var(X)} = \sqrt{np(1-p)}$
- **Methods of finding the sample size (n):**
  - 1. Use the binomial distribution formula and solve for n.
  - 2. Use a **spreadsheet/table** on your calculator: can look at the table and determine the lowest value for the sample size (n) for which the required probability is achieved.
  - Ti-Nspire only: Use 'inverse binomial N' to find the number of trials. Note: the cumulative probability is the **probability that X is less than or equal to a number.**

Continuous Random Variables

- A **continuous random variable** is one that can take any value in an interval of the real number line
- $\Pr(a < X < b) = \int_a^b f(x)\,dx$ (area under the graph)
- <u>Requirements for a probability density function:</u>
  - $f(x) \geq 0$ for all x
  - $\int_{-\infty}^{\infty} f(x)\,dx = 1$ (total area under graph is 1, as the total probability is 1)
- **Expected value:**
  - $E(X) = \int_{-\infty}^{\infty} x\, f(x)\,dx$
  - $E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\,dx$
  - $E(aX + b) = aE(X) + b$

- **Percentiles:** $\int_{-\infty}^{p} f(x)\,dx = q$ (i.e. 75th percentile means q = 0.75)
- **Median:** $\int_{-\infty}^{m} f(x)\,dx = 0.5$ (i.e. 50th percentile)
- **Variance:**
  - $Var(X) = E(X^2) - [E(X)]^2 = \int_{-\infty}^{\infty} x^2\, f(x)\,dx - (\int_{-\infty}^{\infty} x\, f(x)\,dx)^2$
  - $Var(aX + b) = a^2 Var(X)$
- **Standard deviation:**
  - $sd(X) = \sqrt{Var(X)}$
  - $sd(aX + b) = \sqrt{Var(aX + b)} = \sqrt{a^2 Var(X)} = |a| sd(X)$
- **Interquartile range:** difference between 75th percentile and 25th percentile

Normal Distribution

- Key features of the normal distribution:
  - Symmetrical about the mean
  - Total area under graph = 1
  - 68-95-99.7 rule:
    - Approx. 68% of the values lie within one standard deviation of the mean
    - Approx. 95% of the values lie within two standard deviations of the mean
    - Approx. 99.7% of the values lie within three standard deviations of the mean
- **Standard normal distribution (Z):** $\mu = 0, \sigma = 1$
- The standard normal distribution can undergo transformations to produce any normal distribution **with mean μ and standard deviation σ**.

Transformations

	Standard normal distribution→general normal distribution	**General normal distribution→standard normal distribution**
Transformations (in order)	• Dilation by a factor of σ from the y-axis • Dilation by factor of 1/σ from the x-axis • Translation μ units in the positive x-direction	• Translation μ units in the negative x-direction • Dilation by factor σ from the x-axis • Dilation by factor 1/σ from the y-axis
Mapping notation	(x, y) → $(\sigma x+\mu, \frac{y}{\sigma})$	(x, y) → $(\frac{x-\mu}{\sigma}, \sigma y)$

- **Standardisation**: $z = \frac{x-\mu}{\sigma}$ (useful when mean and/or standard deviation unknown)
- On calculator:
  - Calculating probabilities: **Normal Cdf**
  - Probability is given, need to find x-value: **Inverse Normal** (note: on Ti-Nspire, need the probability that X is less than the value).

Populations and samples

- **Population proportion:** p (fixed, whole population)
- **Sample proportion:** $\hat{p}$ (varies between samples)

Sampling from a small population

- Without replacement (cannot use binomial)
- Can use tree diagram or combinatorics:
  - E.g. 13 marbles in a bag (7 blue, 6 green) and 5 are selected. The probability of choosing 3 blue marbles and 2 green marbles is:

$$\Pr(3B, 2G) = \frac{\binom{7}{3}\binom{6}{2}}{\binom{13}{5}}$$

Sampling from a large population

- Can use binomial distribution, since population is very large (p assumed to be constant)
- **Sample proportion:** $\hat{P} = \frac{X}{n}$ (note: X is a binomial random variable and n is the sample size)
- **Expected value:** $E(\hat{P}) = p$
- **Variance:** $Var(\hat{P}) = \frac{p(1-p)}{n}$
- **Standard deviation:** $sd(\hat{P}) = \sqrt{\frac{p(1-p)}{n}}$
- **Confidence intervals:**
  - **C% confidence:** $(\hat{p} - k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}})$, where k is such that $\Pr(-k < Z < k) = \frac{C}{100}$
  - **Interpretation:** expect approximately C% of such intervals to contain the population proportion p.
  - **95% confidence:** $(\hat{p} - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}})$
  - **Margin of error:** $M = k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ (as sample size increases, margin of error decreases)
- On calculator:
  - Ti-Nspire: Menu→Statistics→Confidence Intervals→1-Prop z Interval
  - Casio Classpad: Statistics→Calc→Interval→One-Prop Z Int