Cheat Sheet

The Sum Rule: If an experiment can either end up being one of N outcomes, or one of M outcomes (where there is no overlap), then the total number of possible outcomes is: $N + M$.

The Product Rule: If an experiment has N_1 outcomes for the first stage, N_2 outcomes for the second stage, ..., and N_m outcomes for the m^{th} stage, then the total number of outcomes of the experiment is $N_1 \times N_2 \cdot \cdots \cdot N_m = \prod_{i=1}^{m} N_i$.

Complementary Counting: Let $\mathcal{U}$ be a (finite) universal set, and S a subset of interest. Then, $| S |=| \mathcal{U} | - | \mathcal{U} \backslash S |$.

Permutation: The number of orderings of N **distinct** objects is $N! = N \cdot (N-1) \cdot (N-2) \cdot \ldots 3 \cdot 2 \cdot 1$.

k-Permutations: If we want to *pick* (**order matters**) only k out of n distinct objects, the number of ways to do so is:

$$P(n,k) = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

Multinomial Coefficients: If we have k distinct types of objects (n total), with n_1 of the first type, n_2 of the second, ..., and n_k of the k-th, then the number of arrangements possible is

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1! n_2! \ldots n_k!}$$

k-Combinations/Binomial Coefficients: If we want to *choose* (**order doesn't matter**) only k out of n distinct objects, the number of ways to do so is:

$$C(n,k) = \binom{n}{k} = \frac{P(n,k)}{k!} = \frac{n!}{k!(n-k)!}$$

Encoding/Stars and Bars Method: The number of ways to distribute n indistinguishable balls into k distinguishable bins is

$$\binom{n+k-1}{k-1} = \binom{n+k-1}{n}$$

Binomial Theorem: Let $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$ a positive integer. Then: $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.

Principle of Inclusion-Exclusion (PIE):
2 events: $|A \cup B| = |A| + |B| - |A \cap B|$
3 events: $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$
k events: singles - doubles + triples - quads + ...

Pigeonhole Principle: If there are n pigeons we want to put into k holes (where $n > k$), then at least one pigeonhole must contain at least 2 (or to be precise, $\lceil n/k \rceil$) pigeons.

Combinatorial Proofs: To prove two quantities are equal, you can come up with a combinatorial situation, and show that both in fact count the same thing, and hence must be equal.

Key Probability Definitions: The **sample space** is the set Ω of all possible outcomes of an experiment. An **event** is any subset $E \subseteq \Omega$. Events E and F are **mutually exclusive** if $E \cap F = \emptyset$.

Probability space: A *probablity space* is a pair $(\Omega, \mathbb{P})$, where Ω is the sample space and $\mathbb{P} : \Omega \to [0,1]$ is a *probability measure* such that $\sum_{x \in \Omega} \mathbb{P}(x) = 1$. The probability of an event $E \subseteq \Omega$ is $\mathbb{P}(E) = \sum_{x \in E} \mathbb{P}(x)$.

Equally Likely Outcomes: If Ω is a sample space such that each of the unique outcome elements in Ω are equally likely, then for any event $E \subseteq \Omega$: $\mathbb{P}(E) = |E|/|\Omega|$.

Conditional Probability: $\mathbb{P}(A \mid B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

Bayes Theorem: $\mathbb{P}(A \mid B) = \dfrac{\mathbb{P}(B \mid A) \mathbb{P}(A)}{\mathbb{P}(B)}$

Partition: Non-empty events $E_1, \ldots, E_n$ **partition** the sample space Ω if they are both:

- **(Exhaustive)** $E_1 \cup E_2 \cup \cdots \cup E_n = \bigcup_{i=1}^{n} E_i = \Omega$ (they cover the entire sample space).
- **(Pairwise Mutually Exclusive)** For all $i \neq j$, $E_i \cap E_j = \emptyset$ (none of them overlap)

Note that for any event E, E and E^C always form a partition of Ω.

Law of Total Probability (LTP): If events $E_1, \ldots, E_n$ partition Ω, then for any event F:

$$\mathbb{P}(F) = \sum_{i=1}^{n} \mathbb{P}(F \cap E_i) = \sum_{i=1}^{n} \mathbb{P}(F \mid E_i) \mathbb{P}(E_i)$$

Bayes Theorem with LTP: Let events $E_1, \ldots, E_n$ partition the sample space Ω, and let F be another event. Then:

$$\mathbb{P}(E_1 \mid F) = \frac{\mathbb{P}(F \mid E_1)\mathbb{P}(E_1)}{\sum_{i=1}^{n} \mathbb{P}(F \mid E_i)\mathbb{P}(E_i)}$$

Chain Rule: Let $A_1, \ldots, A_n$ be events with nonzero probabilities. Then:

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1)\mathbb{P}(A_3 \mid A_1 \cap A_2) \cdots \mathbb{P}(A_n \mid A_1 \cap \cdots \cap A_{n-1})$$

Independence: A and B are **independent** if any of the following equivalent statements hold:

1. $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$ 2. $\mathbb{P}(A \mid B) = \mathbb{P}(A)$ 3. $\mathbb{P}(B \mid A) = \mathbb{P}(B)$

Mutual Independence: We say n events $A_1, A_2, \ldots, A_n$ are **(mutually) independent** if, for *any* subset $I \subseteq [n] = \{1, 2, \ldots, n\}$, we have

$$\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i)$$

This equation is actually representing 2^n equations since there are 2^n subsets of $[n]$.

Conditional Independence: A and B are **conditionally independent given an event** C if any of the following equivalent statements hold:

1. $\mathbb{P}(A \cap B \mid C) = \mathbb{P}(A \mid C)\,\mathbb{P}(B \mid C)$ 2. $\mathbb{P}(A \mid B \cap C) = \mathbb{P}(A \mid C)$ 3. $\mathbb{P}(B \mid A \cap C) = \mathbb{P}(B \mid C)$

Random Variable (RV): A random variable (RV) X is a numeric function of the outcome $X : \Omega \to \mathbb{R}$.
The set of possible values X can take on is its **range/support**, denoted Ω_X.
If Ω_X is finite or countable infinite (typically integers or a subset), X is a **discrete RV**.

Probability Mass Function (PMF): For a discrete RV X, assigns probabilities to values in its range. That is $p_X : \Omega_X \to [0, 1]$ where: $p_X(k) = \mathbb{P}(X = k)$.

Linearity of Expectation (LoE): For any random variables X, Y (possibly dependent):
$$\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$$

Expectation: The **expectation** of a discrete RV X is: $\mathbb{E}[X] = \sum_{k \in \Omega_X} k \cdot p_X(k)$.

Law of the Unconscious Statistician (LOTUS): For a discrete RV X and function g, $\mathbb{E}[g(X)] = \sum_{b \in \Omega_X} g(b) \cdot p_X(b)$.

Linearity of Expectation with Indicators: If asked only about the expectation of a RV X which is some sort of "count" (and not its PMF), then you may be able to write X as the sum of possibly dependent **indicator** RVs $X_1, \ldots, X_n$, and apply LoE, where for an indicator RV X_i, $\mathbb{E}[X_i] = 1 \cdot \mathbb{P}(X_i = 1) + 0 \cdot \mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1)$.

Variance: $\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$.

Standard Deviation (SD): $\sigma_X = \sqrt{\mathrm{Var}(X)}$.

Property of Variance: $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$.

Independence: Random variables X and Y are **independent** if for *all* $x \in \Omega_X$ and all $y \in \Omega_Y$:
$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y).$$

Variance Adds for Independent RVs: If X, Y are independent, then $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.