Reference Sheet

Theorem: Binomial Theorem

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

Theorem: 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 + ...

Theorem: 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.

Definition: 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 $E \cap F = \emptyset$.

Definition: Probability space

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

Definition: Conditional Probability

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$$

Theorem: Bayes Theorem

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[B \mid A]\,\mathbb{P}[A]}{\mathbb{P}[B]}$$

Definition: Partition

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

- **(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)

Theorem: Law of Total Probability (LTP)

If events $E_1, \ldots, E_n$ partition Ω, then for any event F:

$$\mathbb{P}[F] = \sum_{i=1}^{n} \mathbb{P}[F \cap E_i] = \sum_{i=1}^{n} \mathbb{P}[F \mid E_i]\,\mathbb{P}[E_i]$$

Theorem: Bayes Theorem with LTP

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

$$\mathbb{P}[E_1 \mid F] = \frac{\mathbb{P}[F \mid E_1]\,\mathbb{P}[E_1]}{\sum_{i=1}^{n} \mathbb{P}[F \mid E_i]\,\mathbb{P}[E_i]}$$

Definition: Independence (Events)

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. $\mathbb{P}[B \mid A] = \mathbb{P}[B]$

Theorem: Chain Rule

Let $A_1, \ldots, A_n$ be events with nonzero probabilities. Then:
$\mathbb{P}[A_1 \cap \cdots \cap A_n] =$
$\mathbb{P}[A_1]\,\mathbb{P}[A_2 \mid A_1]\,\mathbb{P}[A_3 \mid A_1 \cap A_2] \cdots \mathbb{P}[A_n \mid A_1 \cap \cdots \cap A_{n-1}]$

Definition: Mutual Independence (Events)

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 $[n]$.

Definition: 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. $\mathbb{P}[B \mid A \cap C] = \mathbb{P}[B \mid C]$

Definition: 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.

Definition: 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: $p_X(k) = \mathbb{P}[X = k]$.

Definition: Expectation

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

Theorem: Linearity of Expectation (LoE)

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

Theorem: Law of the Unconscious Statistician (LOTUS)

For a discrete RV X and function g, $\mathbb{E}[g(X)] = \sum_{b \in \Omega_X} g(b) \cdot p_X(b)$.

Definition: Variance

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

Theorem: Property of Variance

$\text{Var}(aX + b) = a^2 \text{Var}(X)$.

Definition: Independence (Random Variables)

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

Theorem: Variance Adds for Independent RVs

If X, Y are independent, then $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.