

Counting

Theorem: 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: $N + M$.

Theorem: 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 $N_1 \cdot N_2 \cdot \dots \cdot N_m = \prod_{i=1}^m N_i$.

Theorem: Complementary Counting

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

Definition: k -Permutations

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

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

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 + ...

Definition: 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:

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

Definition: Encoding/Stars and Bars Method

The number of ways to distribute n indistinguishable balls into k distinguishable bins is

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

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 $\lceil n/k \rceil$ pigeons.

Discrete Probability

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 *probability space* is a pair (Ω, \mathbb{P}) , where Ω is the sample space $\mathbb{P} : \Omega \rightarrow [0, 1]$ 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

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

Theorem: Bayes Theorem

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

Definition: Partition

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

- **(Exhaustive)** $E_1 \cup E_2 \cup \dots \cup E_n = \bigcup_{i=1}^n E_i = \Omega$ (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, \dots, E_n partition Ω , then for any event F :

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

Theorem: Bayes Theorem with LTP

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

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

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

Theorem: Chain Rule

Let A_1, \dots, A_n be events with nonzero probabilities. Then:

$$\mathbb{P}[A_1 \cap \dots \cap A_n] = \mathbb{P}[A_1] \mathbb{P}[A_2 | A_1] \mathbb{P}[A_3 | A_1 \cap A_2] \dots \mathbb{P}[A_n | A_1 \cap \dots \cap A_{n-1}]$$

Definition: Mutual Independence (Events)

We say n events A_1, A_2, \dots, A_n are **(mutually) independent** if, for any subset $I \subseteq [n] = \{1, 2, \dots, n\}$, 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. $\mathbb{P}[A \cap B | C] = \mathbb{P}[A | C] \mathbb{P}[B | C]$
2. $\mathbb{P}[A | B \cap C] = \mathbb{P}[A | C]$
3. $\mathbb{P}[B | A \cap C] = \mathbb{P}[B | C]$

Random Variables

Definition: Random Variable (RV)

A random variable (RV) X is a numeric function of the outcome $X : \Omega \rightarrow \mathbb{R}$. 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 $p_X : \Omega_X \rightarrow [0, 1]$ 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

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

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 $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.

Zoo

Definition: Bernoulli/Indicator Random Variable

$X \sim \text{Bernoulli}(p)$ ($\text{Ber}(p)$ for short) iff X has PMF:

$$p_X(k) = \begin{cases} p, & k = 1 \\ 1 - p, & k = 0 \end{cases}$$

$\mathbb{E}[X] = p$ and $\text{Var}(X) = p(1 - p)$.

Definition: Binomial Random Variable

$X \sim \text{Binomial}(n, p)$ ($\text{Bin}(n, p)$ for short) iff X has PMF

$$p_X(k) = \binom{n}{k} p^k (1 - p)^{n - k}, \quad k \in \Omega_X = \{0, 1, \dots, n\}$$

$\mathbb{E}[X] = np$ and $\text{Var}(X) = np(1 - p)$.

Definition: Uniform Random Variable (Discrete)

$X \sim \text{Uniform}(a, b)$ ($\text{Unif}(a, b)$ for short), for integers $a \leq b$, iff X has PMF:

$$p_X(k) = \frac{1}{b - a + 1}, \quad k \in \Omega_X = \{a, a + 1, \dots, b\}$$

$\mathbb{E}[X] = \frac{a+b}{2}$ and $\text{Var}(X) = \frac{(b-a)(b-a+1)}{12}$.

Definition: Geometric Random Variable

$X \sim \text{Geometric}(p)$ ($\text{Geo}(p)$ for short) iff X has PMF:

$$p_X(k) = (1 - p)^{k-1} p, \quad k \in \Omega_X = \{1, 2, 3, \dots\}$$

$\mathbb{E}[X] = \frac{1}{p}$ and $\text{Var}(X) = \frac{1-p}{p^2}$.

Definition: Poisson Random Variable

$X \sim \text{Poisson}(\lambda)$ ($\text{Poi}(\lambda)$ for short) iff X has PMF:

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \Omega_X = \{0, 1, 2, \dots\}$$

$\mathbb{E}[X] = \lambda$ and $\text{Var}(X) = \lambda$. If X_1, \dots, X_n are independent Poisson RV's, where $X_i \sim \text{Poi}(\lambda_i)$, then $X = X_1 + \dots + X_n \sim \text{Poi}(\lambda_1 + \dots + \lambda_n)$.