Probability Cheatsheet

A quick reference guide to probability concepts, formulas, and distributions, covering basic probability, conditional probability, random variables, and common distributions.

Basic Probability Concepts

Definitions

Probability:	A measure of the likelihood that an event will occur. It is quantified as a number between 0 and 1, where 0 indicates impossibility and 1 indicates certainty.
Experiment:	A process or action that has observable outcomes.
Sample Space (S):	The set of all possible outcomes of an experiment.
Event (E):	A subset of the sample space, representing a specific outcome or set of outcomes.
Outcome:	A possible result of an experiment.
Mutually Exclusive Events:	Events that cannot occur at the same time (i.e., they have no outcomes in common).

Basic Probability Formula

The probability of an event *E* occurring is defined as:

$P(E) = \frac{\text{Number of favorable outcomes}}{\text{Total number of possible outcomes}} = \frac{n(E)}{n(S)}$

Where:

- $P(E)$ is the probability of event *E*.
- $n(E)$ is the number of outcomes in event *E*.
- $n(S)$ is the number of outcomes in the sample space *S*.

Probability Rules

Rule 1: Probability Range	The probability of any event *E* must be between 0 and 1: $0 \le P(E) \le 1$
Rule 2: Probability of Sample Space	The probability of the entire sample space *S* is 1: $P(S) = 1$
Rule 3: Complement Rule	The probability of an event *E* not occurring is: $P(E') = 1 - P(E)$
Rule 4: Addition Rule	For any two events *A* and *B*: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Rule 5: Addition Rule for Mutually Exclusive Events	If *A* and *B* are mutually exclusive: $P(A \cup B) = P(A) + P(B)$

Conditional Probability and Independence

Conditional Probability

Conditional probability is the probability of an event *A* occurring given that another event *B* has already occurred. It is denoted as $P(A|B)$ and calculated as:

$P(A|B) = \frac{P(A \cap B)}{P(B)}$, where $P(B) > 0$

Independence of Events

Definition	Two events *A* and *B* are independent if the occurrence of one does not affect the probability of the other.
Independence Condition	Events *A* and *B* are independent if and only if: $P(A \cap B) = P(A) \cdot P(B)$
Conditional Probability and Independence	If *A* and *B* are independent, then: $P(A|B) = P(A)$ and $P(B|A) = P(B)$

Bayes' Theorem

Bayes' Theorem describes the probability of an event based on prior knowledge of conditions related to the event. It is given by:

$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$

Where:

- $P(A|B)$ is the posterior probability of *A* given *B*.
- $P(B|A)$ is the likelihood of *B* given *A*.
- $P(A)$ is the prior probability of *A*.
- $P(B)$ is the prior probability of *B*.

In terms of sample space:

$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|A') \cdot P(A')}$

Random Variables and Distributions

Random Variables

Definition:	A random variable is a variable whose value is a numerical outcome of a random phenomenon.
Discrete Random Variable:	A variable whose value can only take on a finite number of values or a countably infinite number of values.
Continuous Random Variable:	A variable whose value can take on any value within a given range.

Probability Mass Function (PMF)

For a discrete random variable *X*, the probability mass function (PMF) gives the probability that *X* takes on a specific value *x*:

$P(X = x)$

Probability Density Function (PDF)

For a continuous random variable *X*, the probability density function (PDF) gives the relative likelihood that *X* will take on a specific value. The probability that *X* falls within a certain interval [a, b] is given by the integral of the PDF over that interval:

$P(a \le X \le b) = \int_{a}^{b} f(x)\, dx$

Where $f(x)$ is the PDF.

Cumulative Distribution Function (CDF)

The cumulative distribution function (CDF) gives the probability that a random variable *X* takes on a value less than or equal to *x*:

$F(x) = P(X \le x)$

Expected Value (Mean)

The expected value (or mean) of a random variable *X* is the weighted average of its possible values:

- For discrete random variable: $E(X) = \sum x \cdot P(X = x)$
- For continuous random variable: $E(X) = \int x \cdot f(x)\, dx$

Variance and Standard Deviation

Variance:	The variance measures the spread of the distribution of a random variable around its mean: $Var(X) = E[(X - E(X))^2]$ Alternative formula: $Var(X) = E[X^2] - (E[X])^2$
Standard Deviation:	The standard deviation is the square root of the variance and provides a measure of the typical deviation of values from the mean: $SD(X) = \sqrt{Var(X)}$

Common Probability Distributions

Discrete Distributions

Bernoulli Distribution

- Represents the probability of success or failure of a single binary event.
- PMF: $P(X = x) = p^x (1-p)^{(1-x)}$, where $x \in \{0, 1\}$ and p is the probability of success.
- $E(X) = p$
- $Var(X) = p(1-p)$

Binomial Distribution

- Represents the number of successes in a fixed number of independent Bernoulli trials.
- PMF: $P(X = k) = \binom{n}{k} p^k (1-p)^{(n-k)}$, where n is the number of trials, k is the number of successes, and p is the probability of success in a single trial.
- $E(X) = np$
- $Var(X) = np(1-p)$

Poisson Distribution

- Represents the number of events occurring in a fixed interval of time or space.
- PMF: $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$, where λ is the average rate of events.
- $E(X) = \lambda$
- $Var(X) = \lambda$

Continuous Distributions

Uniform Distribution

- Represents a constant probability over a given interval.
- PDF: $f(x) = \frac{1}{b-a}$ for $a \le x \le b$, where a and b are the interval endpoints.
- $E(X) = \frac{a+b}{2}$
- $Var(X) = \frac{(b-a)^2}{12}$

Exponential Distribution

- Represents the time until an event occurs in a Poisson process.
- PDF: $f(x) = \lambda e^{-\lambda x}$ for $x \ge 0$, where λ is the rate parameter.
- $E(X) = \frac{1}{\lambda}$
- $Var(X) = \frac{1}{\lambda^2}$

Normal (Gaussian) Distribution

- Represents a symmetric, bell-shaped distribution characterized by its mean and standard deviation.
- PDF: $f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, where μ is the mean and σ is the standard deviation.
- $E(X) = \mu$
- $Var(X) = \sigma^2$