Probability Theory Summary

Mathematical Notes

October 19, 2025

Contents

1 Measure Theory Foundations

1.1 Sigma-Algebras

Definition 1.1. A **σ-algebra** on a set Ω is a collection $\mathcal{F}$ of subsets of Ω such that:

1. $\Omega \in \mathcal{F}$
2. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
3. If $A_1, A_2, \ldots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

1.2 Measures

Definition 1.2. A **measure** on a measurable space $(\Omega, \mathcal{F})$ is a function $\mu : \mathcal{F} \to [0, \infty]$ such that:

1. $\mu(\emptyset) = 0$
2. If $A_1, A_2, \ldots$ are disjoint sets in $\mathcal{F}$, then $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$

1.3 Probability Measures

Definition 1.3. A **probability measure** P on $(\Omega, \mathcal{F})$ is a measure such that $P(\Omega) = 1$.

1.4 Measurable Functions

Definition 1.4. A function $f : \Omega \to \mathbb{R}$ is **measurable** if for every Borel set $B \subseteq \mathbb{R}$, $f^{-1}(B) \in \mathcal{F}$.

2 Random Variables

2.1 Definition

Definition 2.1. A **random variable** is a measurable function $X : \Omega \to \mathbb{R}$.

2.2 Distribution Functions

Definition 2.2. The **cumulative distribution function** (CDF) of random variable X is:

$$F_X(x) = P(X \leq x)$$

2.3 Probability Density Functions

Definition 2.3. For a continuous random variable X, the **probability density function** (PDF) is a function f_X such that:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$$

2.4 Expected Value

Definition 2.4. The **expected value** of random variable X is:

$$E[X] = \int_{\Omega} X(\omega)\, dP(\omega) = \int_{-\infty}^{\infty} x\, dF_X(x)$$

2.5 Variance

Definition 2.5. The **variance** of random variable X is:

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

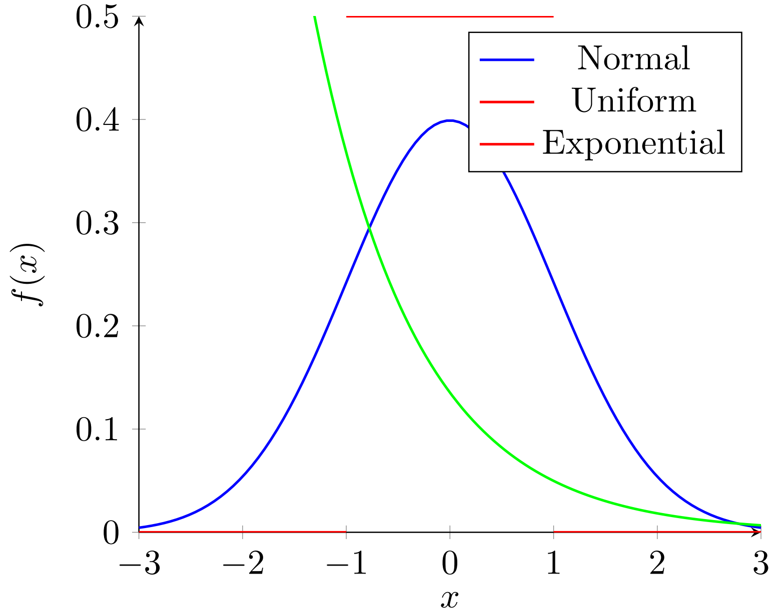

Figure 1: Common probability density functions

3 Common Distributions

3.1 Discrete Distributions

3.1.1 Bernoulli Distribution

Definition 3.1. $X \sim \text{Bernoulli}(p)$ if $P(X = 1) = p$ and $P(X = 0) = 1 - p$.

$$E[X] = p, \quad \text{Var}(X) = p(1 - p)$$

3.1.2 Binomial Distribution

Definition 3.2. $X \sim \text{Binomial}(n, p)$ if $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

$$E[X] = np, \quad \text{Var}(X) = np(1 - p)$$

3.1.3 Poisson Distribution

Definition 3.3. $X \sim \text{Poisson}(\lambda)$ if $P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$.

$$E[X] = \lambda, \quad \text{Var}(X) = \lambda$$

3.2 Continuous Distributions

3.2.1 Uniform Distribution

Definition 3.4. $X \sim \text{Uniform}(a, b)$ has PDF $f(x) = \frac{1}{b-a}$ for $x \in [a, b]$.

$$E[X] = \frac{a+b}{2}, \quad \text{Var}(X) = \frac{(b-a)^2}{12}$$

3.2.2 Normal Distribution

Definition 3.5. $X \sim \mathcal{N}(\mu, \sigma^2)$ has PDF $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

$$E[X] = \mu, \quad \text{Var}(X) = \sigma^2$$

3.2.3 Exponential Distribution

Definition 3.6. $X \sim \text{Exponential}(\lambda)$ has PDF $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$.

$$E[X] = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}$$

4 Convergence of Random Variables

4.1 Almost Sure Convergence

Definition 4.1. $X_n \to X$ **almost surely** if $P(\lim_{n\to\infty} X_n = X) = 1$.

4.2 Convergence in Probability

Definition 4.2. $X_n \to X$ **in probability** if for every $\epsilon > 0$:

$$\lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0$$

4.3 Convergence in Distribution

Definition 4.3. $X_n \to X$ **in distribution** if $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ for all continuity points of F_X.

4.4 Convergence in L^p

Definition 4.4. $X_n \to X$ **in** L^p if $\lim_{n\to\infty} E[|X_n - X|^p] = 0$.

5 Limit Theorems

5.1 Law of Large Numbers

Theorem 5.1 (Strong Law of Large Numbers)**.** If $X_1, X_2, \ldots$ are i.i.d. with $E[X_i] = \mu < \infty$, then:

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mu \text{ almost surely}$$

5.2 Central Limit Theorem

Theorem 5.2 (Central Limit Theorem)**.** If $X_1, X_2, \ldots$ are i.i.d. with $E[X_i] = \mu$ and $\mathrm{Var}(X_i) = \sigma^2 < \infty$, then:

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} \mathcal{N}(0, 1)$$

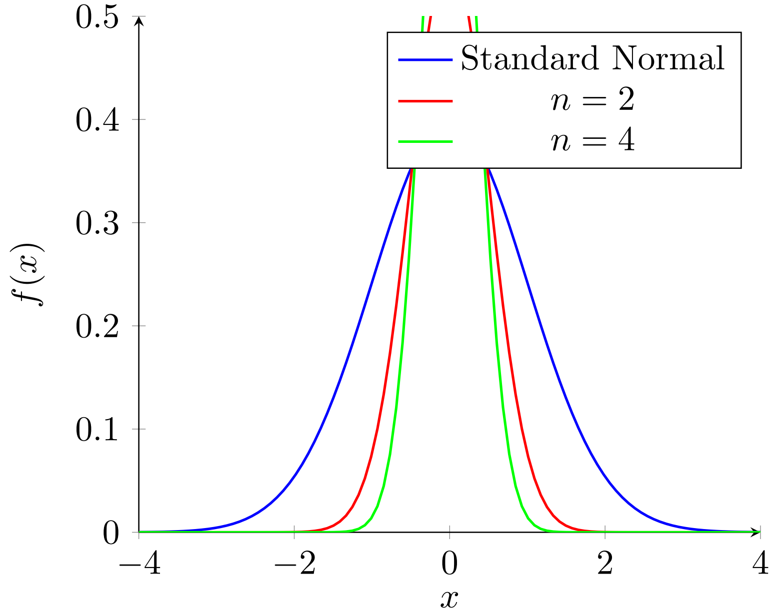

Figure 2: Central Limit Theorem: convergence of sample means to normal distribution

6 Conditional Probability and Independence

6.1 Conditional Probability

Definition 6.1. The **conditional probability** of A given B is:

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

provided $P(B) > 0$.

6.2 Independence

Definition 6.2. Events A and B are **independent** if $P(A \cap B) = P(A)P(B)$.

6.3 Conditional Expectation

Definition 6.3. The **conditional expectation** $E[X|Y]$ is the random variable that equals $E[X|Y = y]$ when $Y = y$.

7 Characteristic Functions

7.1 Definition

Definition 7.1. The **characteristic function** of random variable X is:

$$\phi_X(t) = E[e^{itX}] = \int_{-\infty}^{\infty} e^{itx}\, dF_X(x)$$

7.2 Properties

Theorem 7.1.

1. $\phi_X(0) = 1$
2. $|\phi_X(t)| \leq 1$
3. $\phi_{aX+b}(t) = e^{itb}\phi_X(at)$
4. If X and Y are independent, then $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$

8 Martingales

8.1 Definition

Definition 8.1. A sequence of random variables $\{X_n\}$ is a **martingale** with respect to filtration $\{\mathcal{F}_n\}$ if:

1. X_n is $\mathcal{F}_n$-measurable
2. $E[|X_n|] < \infty$
3. $E[X_{n+1}|\mathcal{F}_n] = X_n$

8.2 Martingale Convergence Theorem

Theorem 8.1. If $\{X_n\}$ is a martingale with $\sup_n E[|X_n|] < \infty$, then X_n converges almost surely.

9 Stochastic Processes

9.1 Markov Chains

Definition 9.1. A **Markov chain** is a sequence of random variables $\{X_n\}$ such that:

$$P(X_{n+1} = j|X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j|X_n = i)$$

9.2 Brownian Motion

Definition 9.2. Brownian motion $\{B_t\}_{t\geq 0}$ is a stochastic process such that:

1. $B_0 = 0$
2. $B_t - B_s \sim \mathcal{N}(0, t-s)$ for $s < t$
3. Increments are independent
4. Paths are continuous

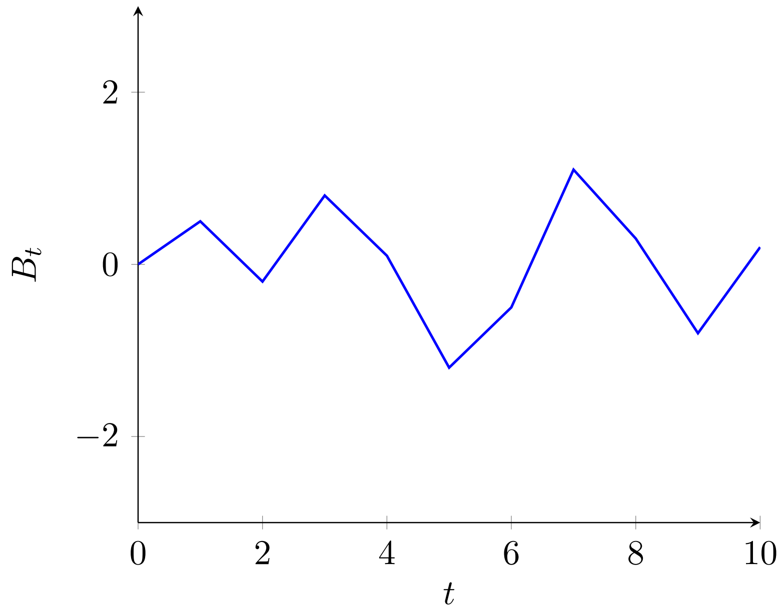

Figure 3: Sample path of Brownian motion

10 Applications

10.1 Finance

Probability theory is fundamental to:

- Option pricing (Black-Scholes model)
- Risk management
- Portfolio optimization
- Credit risk modeling

10.2 Statistics

Applications include:

- Parameter estimation
- Hypothesis testing
- Confidence intervals
- Bayesian inference

10.3 Physics

Used in:

- Statistical mechanics
- Quantum mechanics

- Brownian motion
- Phase transitions

11 Important Theorems

11.1 Radon-Nikodym Theorem

Theorem 11.1. If μ and ν are σ-finite measures with $\nu \ll \mu$, then there exists a measurable function f such that $\nu(A) = \int_A f \, d\mu$.

11.2 Fubini's Theorem

Theorem 11.2. If f is integrable on $X \times Y$, then:

$$\int_{X \times Y} f \, d(\mu \times \nu) = \int_X \left(\int_Y f(x, y) \, d\nu(y) \right) d\mu(x)$$

11.3 Dominated Convergence Theorem

Theorem 11.3. If $f_n \to f$ almost everywhere and $|f_n| \leq g$ for some integrable g, then:

$$\lim_{n \to \infty} \int f_n \, d\mu = \int f \, d\mu$$