Probability Theory Cheat Sheet

Chapter 10: General Notions in Probability

1. Probability Spaces

Definition 10.1

- Triplet $(\Omega, \mathcal{F}, P)$ where:
  - Ω: Sample space (non-empty set of possible outcomes)
  - $\mathcal{F}$: σ-algebra (collection of measurable events)
  - P: Probability measure satisfying $P(\Omega) = 1$

Key Examples

Type	Formula	Interpretation
Discrete Uniform	$P(A) = \frac{A}{\Omega}$	Equally likely outcomes
Continuous Uniform	$P(A) = \frac{\lambda(A)}{\lambda(\Omega)}$	Lebesgue measure normalized
Poker Probability	$4 \cdot \frac{\binom{13}{5}}{\binom{52}{5}} \approx 0.00198$	Probability of a flush

2. Random Variables & Vectors

Core Definitions

- **Random Variable**: Measurable function $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$

- **Random Vector**: $\mathbf{X} = (X_1, ..., X_k)$ where each X_i is a RV

Distribution Law:
$P_X(B) = P(X^{-1}(B)) = P(\{\omega \in \Omega : X(\omega) \in B\})$

2.1 Discrete Distributions

Properties:

- Countable support $S \subseteq \mathbb{R}$
- Probability mass function $p(x) = P(X = x)$

Common Distributions

Name	PMF	Parameters
Bernoulli	$p^x(1-p)^{1-x}$	$p \in [0,1], x \in \{0,1\}$
Binomial	$\binom{n}{x}p^x(1-p)^{n-x}$	$n \in \mathbb{N}, p \in [0,1]$
Poisson	$e^{-\lambda}\frac{\lambda^x}{x!}$	$\lambda > 0$
Geometric	$(1-p)^{x-1}p$	$p \in (0,1)$

2.2 Continuous Distributions

Properties:

- Density function $f(x) \geq 0$ with $\int_{-\infty}^{\infty} f(x)dx = 1$
- $P(X \in A) = \int_A f(x)dx$

Common Distributions

Name	Density Function	Parameters
Uniform	$\frac{1}{b-a}$ on $[a,b]$	$a < b$
Exponential	$\lambda e^{-\lambda x}$ on $[0,\infty)$	$\lambda > 0$
Normal	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	$\mu \in \mathbb{R}, \sigma > 0$

3. Expectation and Moments

Definition:

$$\mathbb{E}[X] = \int_\Omega X dP = \begin{cases} \sum x_i p(x_i) & \text{if discrete} \\ \int x f(x) dx & \text{if continuous} \end{cases}$$

Properties:

1. Linearity: $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$
2. Monotonicity: $X \leq Y \Rightarrow \mathbb{E}[X] \leq \mathbb{E}[Y]$
3. If $X \perp Y$, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$

Important Inequalities:

- Markov: $P(|X| \geq a) \leq \frac{\mathbb{E}[|X|]}{a}$
- Chebyshev: $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$

4. Variance and Covariance

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

Covariance: $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

Properties:

1. $\text{Var}(aX + b) = a^2\text{Var}(X)$
2. If $X \perp Y$, $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$
3. Correlation: $\rho(X, Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$

5. Distribution Functions

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

Properties:

1. Right-continuous
2. $\lim_{x \to -\infty} F(x) = 0, \lim_{x \to \infty} F(x) = 1$
3. $P(a < X \leq b) = F(b) - F(a)$

Quantile Function: $Q(p) = \inf\{x \in \mathbb{R} : F(x) \geq p\}$

6. Characteristic Functions

Definition: $\varphi_X(t) = \mathbb{E}[e^{itX}]$

Key Properties:

1. $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$
2. $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$ if $X \perp Y$
3. Inversion formula exists

Examples:

- Normal: $\varphi(t) = \exp(i\mu t - \sigma^2 t^2/2)$
- Poisson: $\varphi(t) = \exp(\lambda(e^{it} - 1))$

7. Important Theorems

Law of Large Numbers:
$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mathbb{E}[X]$ almost surely

Central Limit Theorem:
$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \to N(0,1)$ in distribution

Bayes' Theorem:
$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

8. Multivariate Distributions

Joint Density: $f_{X,Y}(x, y)$

Marginal Density: $f_X(x) = \int f_{X,Y}(x, y)dy$

Conditional Density:
$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$

Multivariate Normal:
$f(\mathbf{x}) = (2\pi)^{-k/2}|\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T\Sigma^{-1}(\mathbf{x} - \mu)\right)$

Appendix: Common Notation

- Ω: Sample space
- $\mathcal{F}$: σ-algebra
- μ: Mean
- σ^2: Variance
- λ: Lebesgue measure/rate parameter
- $C(n, k)$: Binomial coefficient
- 1_A: Indicator function