utfcode
utf8.sty 3.10 UTF-8 input encoding 13.06.2000
scanner for code UTF-8 installed.

Ultimate Statistics & Probability Cheat Sheet

Essential Formulas, Distributions, Testing & Inference

1. Foundations of Probability

Axioms of Probability:

- $0 \leq P(A) \leq 1$ for any event A.
- $P(\Omega) = 1$ (Sample Space).
- If $A_1, A_2, \ldots$ disjoint: $P(\bigcup A_i) = \sum P(A_i)$.

Key Rules:

- Addition: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Complement: $P(A^c) = 1 - P(A)$
- Conditional: $P(A|B) = \frac{P(A \cap B)}{P(B)}$ $(P(B) > 0)$
- Multiplication: $P(A \cap B) = P(A|B)P(B)$

Independence: A, B independent $\iff$ $P(A \cap B) = P(A)P(B)$

Law of Total Probability:

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Bayes' Theorem:

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

2. Random Variables & Moments

Cumulative Distribution Function (CDF):

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

Expectation $E[X]$:

- Discrete: $E[X] = \sum_x x\, P(X = x)$
- Continuous: $E[X] = \int_{-\infty}^{\infty} x f(x) dx$
- Linearity: $E[aX + bY + c] = aE[X] + bE[Y] + c$

Variance & Std Deviation:

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

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

Covariance & Correlation:

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

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}, \quad -1 \leq \rho_{XY} \leq 1$$

3. Discrete Distributions

Bernoulli(p): $X \in \{0, 1\}$

- PMF: $P(X = x) = p^x (1 - p)^{1-x}$
- $E[X] = p$, $\text{Var}(X) = p(1 - p)$

Binomial(n, p): $X = \#$ successes in n trials

- PMF: $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$
- $E[X] = np$, $\text{Var}(X) = np(1 - p)$

Poisson(λ): $X = \#$ events in unit time

- PMF: $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$
- $E[X] = \lambda$, $\text{Var}(X) = \lambda$

Geometric(p): $X =$ trial of 1st success

- PMF: $P(X = k) = (1 - p)^{k-1} p$
- $E[X] = \frac{1}{p}$, $\text{Var}(X) = \frac{1-p}{p^2}$

4. Continuous Distributions

Uniform(a, b):

- PDF: $f(x) = \frac{1}{b-a}$ for $x \in [a, b]$
- $E[X] = \frac{a+b}{2}$, $\text{Var}(X) = \frac{(b-a)^2}{12}$

Normal $\mathcal{N}(\mu, \sigma^2)$:

- PDF: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
- Standard Normal $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$
- $68-95-99.7$ Empirical Rule for $\mu \pm 1\sigma, 2\sigma, 3\sigma$.

Exponential(λ):

- PDF: $f(x) = \lambda e^{-\lambda x}$ $(x \geq 0)$
- $E[X] = \frac{1}{\lambda}$, $\text{Var}(X) = \frac{1}{\lambda^2}$
- Memoryless: $P(X > s + t \mid X > s) = P(X > t)$

5. Sampling & Limit Theorems

Law of Large Numbers (LLN): Sample mean $\bar{X}_n \to \mu$ as $n \to \infty$.
Central Limit Theorem (CLT): For i.i.d. $X_1, \ldots, X_n$ with mean μ and variance σ^2:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{d} \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \quad \text{as } n \to \infty$$

Standardized: $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$
Sample Statistics:

- Sample Mean: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- Sample Variance: $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

6. Estimation & Hypothesis Testing

Confidence Intervals (CI):

- Normal / Known σ: $\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$
- Unknown σ (t-dist, $df = n - 1$): $\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$

Hypothesis Testing Steps:

1. State H_0 (null) and H_1 (alternative).
2. Choose significance level α (e.g., 0.05).
3. Compute test statistic (Z or t).
4. Calculate p-value or compare to critical value.

Test Statistics:

- One-sample Z-test: $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$
- One-sample t-test: $t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$

Errors:

- Type I Error (α): Reject H_0 when H_0 is true.
- Type II Error (β): Fail to reject H_0 when H_0 is false.
- Power $= 1 - \beta$.