

一、随机事件与概率

1. 事件的关系与运算

事件包含与吸收律

若 $A \subset B$, 则

$$A \cup B = B, \quad AB = A.$$

其中 AB 表示 $A \cap B$, 以上结论称为吸收律。

对偶律 (德摩根律)

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}.$$

可推广到 n 个事件:

$$\overline{\bigcup_{i=1}^n A_i} = \bigcap_{i=1}^n \overline{A_i}, \quad \overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \overline{A_i}.$$

事件独立性

两事件独立:

$$P(AB) = P(A)P(B).$$

三事件两两独立:

$$P(AB) = P(A)P(B), \quad P(AC) = P(A)P(C), \quad P(BC) = P(B)P(C).$$

三事件相互独立: 两两独立, 且

$$P(ABC) = P(A)P(B)P(C).$$

2. 概率核心公式

加法公式

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

推广到三个事件:

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(AB) - P(AC) - P(BC) + P(ABC). \end{aligned}$$

减法公式

$$P(A\overline{B}) = P(A) - P(AB).$$

若 $B \subset A$, 则

$$P(A - B) = P(A) - P(B).$$

条件概率与乘法公式

条件概率:

$$P(A | B) = \frac{P(AB)}{P(B)}, \quad P(B) > 0.$$

乘法公式：

$$P(AB) = P(A | B)P(B) = P(B | A)P(A).$$

全概率公式与贝叶斯公式

若 $A_1, A_2, \dots, A_n$ 为样本空间的完备事件组，则

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

贝叶斯公式：

$$P(A_k | B) = \frac{P(A_k)P(B | A_k)}{\sum_{i=1}^n P(A_i)P(B | A_i)}, \quad P(B) > 0.$$

3. 经典概型

1. **古典概型**：有限等可能样本空间，

$$P(A) = \frac{A \text{ 包含的基本事件数}}{\text{样本空间总基本事件数}}.$$

2. **几何概型**：无限等可能样本空间，

$$P(A) = \frac{A \text{ 的几何测度}}{\text{样本空间总几何测度}}.$$

几何测度可以是长度、面积、体积等。

3. **伯努利概型**： n 重独立重复 0-1 试验，对应二项分布。

二、一维随机变量及其分布

1. 分布函数与分类

分布函数定义：

$$F(x) = P\{X \leq x\}, \quad x \in (-\infty, +\infty).$$

核心性质：非降、右连续，且

$$F(-\infty) = 0, \quad F(+\infty) = 1.$$

连续型随机变量：

$$F(x) = \int_{-\infty}^x f(t) dt,$$

其中 $f(x)$ 为概率密度函数，满足

$$\int_{-\infty}^{+\infty} f(x) dx = 1.$$

随机变量分类：离散型、连续型。

2. 常见离散型分布

分布名称	分布记号	期望 $E(X)$	方差 $D(X)$	核心备注
0-1 分布 (伯努利分布)	$X \sim B(1, p)$	p	$p(1 - p)$	单次 0-1 试验
二项分布	$X \sim B(n, p)$	np	$np(1 - p)$	n 重伯努利试验
泊松分布	$X \sim P(\lambda)$	λ	λ	稀有事件计数, $\lambda > 0$
几何分布	$X \sim Ge(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	首次成功的试验次数

3. 常见连续型分布

分布名称	分布记号	期望 $E(X)$	方差 $D(X)$	核心备注
均匀分布	$X \sim U(a, b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	区间 (a, b) 内等可能取值
指数分布	$X \sim Exp(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	无记忆性, $\lambda > 0$ 为失效率
正态分布 (高斯分布)	$X \sim N(\mu, \sigma^2)$	μ	σ^2	标准化: $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$

4. 标准正态分布核心性质

分布函数记为 $\Phi(x)$, 概率密度记为 $\varphi(x)$ 。

$$\Phi(-a) = 1 - \Phi(a), \quad \varphi(-x) = \varphi(x).$$

三、二维随机变量及其分布

1. 联合分布

联合分布函数:

$$F(x, y) = P\{X \leq x, Y \leq y\}, \quad x, y \in (-\infty, +\infty).$$

离散型联合分布:

$$P\{X = x_i, Y = y_j\} = p_{ij}, \quad \sum_i \sum_j p_{ij} = 1.$$

连续型联合分布: 联合概率密度为 $f(x, y)$, 满足

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du,$$

且

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1.$$

2. 边缘分布

离散型边缘分布律

X 的边缘分布:

$$p_{i\cdot} = P\{X = x_i\} = \sum_j p_{ij}.$$

Y 的边缘分布:

$$p_{\cdot j} = P\{Y = y_j\} = \sum_i p_{ij}.$$

连续型边缘概率密度

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy, \quad f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx.$$

3. 条件分布

离散型条件分布律:

$$P\{X = x_i | Y = y_j\} = \frac{p_{ij}}{p_{\cdot j}}, \quad p_{\cdot j} > 0.$$

连续型条件概率密度:

$$f_{Y|X}(y | x) = \frac{f(x, y)}{f_X(x)}, \quad f_X(x) > 0.$$

4. 随机变量的独立性

离散型:

$$X, Y \text{ 独立} \iff p_{ij} = p_{i\cdot} p_{\cdot j} \quad \text{对所有 } i, j \text{ 成立.}$$

连续型:

$$X, Y \text{ 独立} \iff f(x, y) = f_X(x) f_Y(y) \quad \text{对几乎所有 } x, y \text{ 成立.}$$

5. 二维随机变量函数的分布

和的分布: 设 $Z = X + Y$, 则卷积公式为

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x) dx = \int_{-\infty}^{+\infty} f(z-y, y) dy.$$

最值分布: 设

$$M = \max(X, Y), \quad N = \min(X, Y).$$

常用事件转化:

$$F_M(z) = P(M \leq z) = P(X \leq z, Y \leq z),$$

$$F_N(z) = P(N \leq z) = 1 - P(X > z, Y > z).$$

若 X, Y 独立, 还可进一步化为边缘分布函数的乘积形式。

四、随机变量的数字特征

1. 数学期望 (均值) $E(X)$

定义

离散型:

$$E(X) = \sum_i x_i p_i.$$

连续型:

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx.$$

随机变量函数的期望

一维:

$$E[g(X)] = \sum_i g(x_i) p_i \quad (\text{离散}),$$

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) dx \quad (\text{连续}).$$

二维:

$$E[g(X, Y)] = \sum_i \sum_j g(x_i, y_j) p_{ij} \quad (\text{离散}),$$

$$E[g(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy \quad (\text{连续}).$$

核心性质

$$E(C) = C \quad (C \text{ 为常数}),$$

$$E(CX) = C \cdot E(X),$$

$$E(X + Y) = E(X) + E(Y) \quad (\text{无条件成立}).$$

若 X, Y 相互独立, 则

$$E(XY) = E(X)E(Y).$$

2. 方差 $D(X)$

定义

$$D(X) = E\{[X - E(X)]^2\} = E(X^2) - [E(X)]^2,$$

标准差为

$$\sqrt{D(X)}.$$

核心性质

$$D(C) = 0 \quad (C \text{ 为常数}),$$

$$D(CX) = C^2 D(X), \quad D(aX + b) = a^2 D(X).$$

若 X, Y 相互独立, 则

$$D(X \pm Y) = D(X) + D(Y).$$

通用公式:

$$D(X \pm Y) = D(X) + D(Y) \pm 2 \text{Cov}(X, Y).$$

3. 协方差与相关系数

协方差 $\text{Cov}(X, Y)$

定义：

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

核心性质：

$$\text{Cov}(X, X) = D(X), \quad \text{Cov}(X, Y) = \text{Cov}(Y, X),$$

$$\text{Cov}(aX, bY) = ab \text{Cov}(X, Y),$$

$$\text{Cov}(X_1 + X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y).$$

相关系数 ρ_{XY}

定义：

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}.$$

核心性质：

$$|\rho_{XY}| \leq 1,$$

$$|\rho_{XY}| = 1 \quad \text{表示 } X, Y \text{ 完全线性相关.}$$

五、大数定律与中心极限定理

1. 大数定律

核心概念：依概率收敛。样本均值随样本量增大，依概率收敛于总体均值。

核心定律：

1. **切比雪夫大数定律**：独立随机变量序列，期望和方差存在，样本均值依概率收敛于期望的均值。
2. **伯努利大数定律**： n 重伯努利试验中，事件频率依概率收敛于事件概率。
3. **辛钦大数定律**：独立同分布随机变量序列，期望存在，样本均值依概率收敛于总体期望。

2. 中心极限定理

核心思想：大量独立同分布的随机变量，其和的标准化变量近似服从标准正态分布。

独立同分布中心极限定理

设 $X_1, X_2, \dots, X_n$ 独立同分布，且

$$E(X_i) = \mu, \quad D(X_i) = \sigma^2 > 0,$$

则

$$\lim_{n \rightarrow \infty} P \left\{ \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \leq x \right\} = \Phi(x).$$

棣莫弗-拉普拉斯中心极限定理

二项分布的正态近似。若

$$X \sim B(n, p),$$

当 n 充分大时,

$$X \approx N(np, np(1-p)).$$

六、统计量、三大抽样分布与正态总体抽样分布

1. 基本概念

- **总体**: 研究对象的全体, 对应随机变量 X 。
- **样本**: 独立同分布于总体的随机变量 $X_1, X_2, \dots, X_n$, n 为样本量。
- **统计量**: 不含未知参数的样本函数。

核心统计量:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

样本 k 阶矩、样本 k 阶中心矩:

$$A_k = \frac{1}{n} \sum_{i=1}^n X_i^k, \quad B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k.$$

2. 三大抽样分布

χ^2 (卡方) 分布

定义: 若 $X_1, X_2, \dots, X_n$ 独立同分布于 $N(0, 1)$, 则

$$\chi^2 = \sum_{i=1}^n X_i^2 \sim \chi^2(n),$$

其中 n 为自由度。

数字特征:

$$E[\chi^2(n)] = n, \quad D[\chi^2(n)] = 2n.$$

可加性: 若

$$\chi_1^2 \sim \chi^2(n_1), \quad \chi_2^2 \sim \chi^2(n_2),$$

且二者独立, 则

$$\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2).$$

t 分布

定义: 若

$$X \sim N(0, 1), \quad Y \sim \chi^2(n),$$

且 X 与 Y 独立, 则

$$T = \frac{X}{\sqrt{Y/n}} \sim t(n),$$

其中 n 为自由度。

性质：关于 0 对称；当 $n \rightarrow \infty$ 时， t 分布趋近于标准正态分布。

F 分布

定义：若

$$U \sim \chi^2(n_1), \quad V \sim \chi^2(n_2),$$

且 U 与 V 独立，则

$$F = \frac{U/n_1}{V/n_2} \sim F(n_1, n_2),$$

其中 (n_1, n_2) 为自由度。

性质：

$$F_{1-\alpha}(n_1, n_2) = \frac{1}{F_{\alpha}(n_2, n_1)}.$$

3. 正态总体的抽样分布

设总体

$$X \sim N(\mu, \sigma^2),$$

$X_1, \dots, X_n$ 为样本， $\bar{X}$ 为样本均值， S^2 为样本方差。

1. 样本均值的分布：

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \quad \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

2. 样本方差的分布：

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1), \quad \bar{X} \text{ 与 } S^2 \text{ 相互独立}.$$

3. t 统计量：

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1).$$

七、参数估计

1. 点估计

1. **矩估计**：用样本矩替换总体矩，求解未知参数的估计量。

2. **最大似然估计**：构造似然函数 $L(\theta)$ ，求使 $L(\theta)$ 或对数似然函数 $\ln L(\theta)$ 取最大值的 $\hat{\theta}$ ，即为最大似然估计量。

2. 估计量的评选标准

无偏性

若

$$E(\hat{\theta}) = \theta,$$

则 $\hat{\theta}$ 为 θ 的无偏估计量。

样本均值 $\bar{X}$ 是总体均值 μ 的无偏估计；样本方差 S^2 是总体方差 σ^2 的无偏估计：

$$E(\bar{X}) = \mu, \quad E(S^2) = \sigma^2.$$

有效性

同为无偏估计量的 $\hat{\theta}_1$ 和 $\hat{\theta}_2$, 若

$$D(\hat{\theta}_1) < D(\hat{\theta}_2),$$

则 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效。

相合性 (一致性)

当 $n \rightarrow \infty$ 时, 若 $\hat{\theta}$ 依概率收敛于 θ , 则 $\hat{\theta}$ 为 θ 的相合估计量。

3. 区间估计

置信区间: 给定置信水平 $1 - \alpha$, 构造统计量确定区间 $(\hat{\theta}_L, \hat{\theta}_U)$, 使

$$P\{\hat{\theta}_L < \theta < \hat{\theta}_U\} = 1 - \alpha.$$

核心场景: 正态总体均值、方差的置信区间, 包括单正态总体、双正态总体。

八、假设检验

1. 基本思想与步骤

核心思想: 小概率事件原理, 即小概率事件在一次试验中几乎不发生。

基本步骤:

1. 提出原假设 H_0 和备择假设 H_1 ;
2. 确定显著性水平 α , 构造检验统计量;
3. 确定拒绝域, 计算检验统计量的观测值;
4. 做出决策: 观测值落入拒绝域则拒绝 H_0 , 否则不拒绝 H_0 。

2. 两类错误

1. **第一类错误 (弃真):** H_0 为真时拒绝 H_0 , 概率为 α , 即显著性水平。
2. **第二类错误 (取伪):** H_0 为假时接受 H_0 , 概率为 β 。

3. 常见假设检验场景

正态总体参数的检验:

- u 检验: 方差已知的均值检验;
- t 检验: 方差未知的均值检验;
- χ^2 检验: 方差检验;
- F 检验: 两总体方差比检验。

九、二级结论与心得体会总结

1. 组合数核心公式

$$C_n^k = \frac{n!}{k!(n-k)!}.$$

例：

$$C_5^2 = \frac{5!}{2!3!} = 10.$$

2. 概率计算核心思路

先判断事件关系或概型，优先用公式简化计算；复杂事件用全概率公式或贝叶斯公式拆解。

3. 分布计算技巧

正态分布优先标准化；离散型分布先写分布律；连续型分布注意积分区间与密度函数的非零域。

4. 数字特征计算

优先用性质简化，避免直接积分或求和；方差计算优先用

$$D(X) = E(X^2) - [E(X)]^2.$$

5. 统计量核心结论

样本方差分母为 $n - 1$ ，是为了保证无偏性；三大分布均由正态分布构造，是正态总体统计推断的核心。

6. 参数估计与假设检验的关联

区间估计与假设检验是对偶关系，均基于抽样分布的性质，核心是用样本信息推断总体特征。