

《概率论与数理统计》重要知识点及计算方法总结

第一章 随机事件及概率

1. 基本概念

(1) 随机试验、样本空间、随机事件；

(2) 事件的关系：

包含关系： $A \subset B$ 指 A 发生必然导致 B 发生；

互不相容： $AB = \Phi$ ；

对立： $AB = \Phi, A \cup B = S$ ；

独立： $P(AB) = P(A)P(B)$ 。

2. 运算公式

(1) 德摩根律： $\overline{A \cup B} = \bar{A} \cap \bar{B}$ ， $\overline{A \cap B} = \bar{A} \cup \bar{B}$ ；

(2) 条件概率： $P(B|A) = \frac{P(AB)}{P(A)}$ ；

(3) 补集： $P(\bar{A}) = 1 - P(A)$ ， $P(\bar{A}|B) = 1 - P(A|B)$ ；

(4) 加法： $P(A \cup B) = P(A) + P(B) - P(AB)$ ；

$$P[A \cup B | C] = P(A|C) + P(B|C) - P(AB|C)；$$

(5) 减法： $P(A - B) = P(A \cap \bar{B}) = P(A) - P(AB)$ ；

$$P[A - B | C] = P(A|C) - P(AB|C)；$$

(6) 乘法： $P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 A_2) \cdots P(A_n|A_1 A_2 \cdots A_{n-1})$ ；

(7) 全概率公式：

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots + P(B_n)P(A|B_n) = \sum_{j=1}^n P(B_j)P(A|B_j)；$$

(8) 贝叶斯公式： $P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^n P(B_j)P(A|B_j)}$ 。

第二章 一维随机变量及其分布

1. 已知离散型随机变量的分布律，求分布函数 $F(x)$ 的方法。

(1) 以 X 的特殊点（取值点）分割 $-\infty$ 到 $+\infty$ ；

- (2) 按右连续的规则加等号;
- (3) 概率累加。

2. 对分布函数 $F(x)$ 性质的理解。

- (1) $F(-\infty)=0, F(+\infty)=1$;
- (2) $F(x)$ 为单调不减函数;
- (3) $F(x)$ 右连续 $\begin{cases} \text{连续型 } r.v. \text{ --- } F(x) \text{ 始终连续!} \\ \text{离散型 } r.v. \text{ --- } \begin{cases} F(x) \text{ 大部分都连续} \\ F(x) \text{ 在间断点 --- 右连续} \end{cases} \end{cases}$

说明:

- (1) 性质 1—3, 可作为判断分布函数的充要条件;
- (2) $F(b)-F(a)=P\{a < X \leq b\} \quad (b > a)$;
- (3) $P\{X = a\} = F(a) - F(a-0)$; 连续点 $F(a) - F(a-0) = 0$ 。

3. 连续型随机变量的概率、分布函数与概率密度的关系。

概率	分布函数	概率密度
$P\{-\infty < x < +\infty\}$	$= F(+\infty) - F(-\infty)$	$= \int_{-\infty}^{+\infty} f(x) dx = 1$
$P\{X \leq x\}$	$= F(x)$	$= \int_{-\infty}^x f(x) dx$
$P\{a < X < b\}$	$= F(b) - F(a)$	$= \int_a^b f(x) dx$
$P\{X = a\}$	$= F(a) - F(a-0)$	$= \int_a^a f(x) dx = 0$

4. 求一维连续型随机变量函数分布的方法。

已知: (1) X 的概率密度 $f(x)$; (2) $Y = g(X)$,

求: Y 的概率密度 $f_Y(y)$.

方法 1 (分布函数法)

(1) 求 Y 分布函数: $F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\} = \int_{g(x) \leq y} f_X(x) dx$

(2) 求导数: $F'_Y(y) = f_Y(y)$

注意:

- (1) 一般要讨论 y 的范围;
- (2) $\int_{g(x) \leq y} f_X(x) dx$ 一般不用计算, 直接用微积分中“变上下限函数”的导数公式求

导, 得到 $f_Y(y)$ 。

方法 2 (公式法)

当 $y = g(x)$ 是分段严格单调的可导函数时, 有:

$$f_Y(y) = \begin{cases} \sum f(h_i(y)) |h'_i(y)|, & y \text{ 有定义区间,} \\ 0, & \text{其他.} \end{cases}$$

其中 $x = h_i(y)$ 是 $y = g(x)$ 的分段区间上的反函数。

5. 几种重要分布的数学期望与方差

分布	分布律或概率密度	数学期望	方差
两点分布 $X \sim B(1, p)$	$P\{X = k\} = p^k (1-p)^{1-k}$ $k = 0, 1$	p	$p(1-p)$
二项分布 $X \sim B(n, p)$	$P\{X = k\} = C_n^k p^k (1-p)^{1-k}$ $k = 0, 1, 2, \dots, n$	np	$np(1-p)$
泊松分布 $X \sim P(\lambda)$	$P\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!}$ $(k=0, 1, 2, \dots, \lambda > 0)$	λ	λ
均匀分布 $X \sim U(a, b)$	$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其他} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布 $X \sim E(\lambda)$	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{其他} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
正态分布 $X \sim N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

第三章 多维随机变量及其分布

1. 已知二维连续型随机变量 (X, Y) 的概率密度 $f(x, y)$ 为分段函数, 求边缘概率密度的方法 (以 $f_X(x)$ 为例)。

(1) 写公式: $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy;$

(2) 画图：平面直角坐标系中画出 $f(x, y) \neq 0$ 的表达式所在的区域；

(3) 作射线：在 (2) 中区域画若干条 y 从 $-\infty$ 到 $+\infty$ 的射线，从而确定 y 穿过 $f(x, y) \neq 0$ 的区域的边界，即确定 y 的积分上下限 $\int_{g(x)}^{h(x)} f(x, y) dy$ ；

(4) 配上 $f_X(x)$ 其他部分：得到 $f_X(x) = \begin{cases} \int_{g(x)}^{h(x)} f(x, y) dy, & x \text{ 的区间,} \\ 0, & \text{其他.} \end{cases}$

2. 已知二维连续型随机变量 (X, Y) 的概率密度 $f(x, y)$ ，求三种概率的方法总结。

$$(1) P\{X, Y \text{ 的不等式}\} = P\{X, Y \text{ 落在某区域 } D\} = \iint_D f(x, y) dx dy$$

$$(2) P\{c < Y < d \mid X = a\} = \int_c^d f_{Y|X}(y|x) dy \quad (\text{带入 } x = a)$$

$$(3) P\{c < Y < d \mid a < X < b\} = \frac{P\{c < Y < d, a < X < b\}}{P\{a < X < b\}}$$

$$\begin{aligned} & \frac{\int \int_{D'} f(x, y) dx dy = \int_a^b dx \int_c^d f(x, y) dy}{\text{方法①} \int_a^b f_X(x) dx} \\ & \text{方法②} \int \int_{D'} f(x, y) dx dy = \int_a^b dx \int_{-\infty}^{+\infty} f(x, y) dy \end{aligned}$$

3. 求二维离散型随机变量函数的分布律的方法。

已知：(1) (X, Y) 的分布律；(2) $Z = g(X, Y)$ ，

求： Z 的分布律。

方法：(1) 列出 Z 的所有可能取值；

(2) 对应算概率 (Z 的取值 $\Leftrightarrow X$ 与 Y 的取值)。

4. 求二维连续型随机变量函数的概率密度的方法。

已知：(1) (X, Y) 的概率密度 $f(x, y)$ ；(2) $Z = g(X, Y)$ ，

求： Z 的概率密度 $f_Z(z)$ 。

方法 1 (分布函数法)

$$(1) \text{ 求 } Z \text{ 的分布函数: } F_Z(z) = P\{Z \leq z\} = P\{g(X, Y) \leq z\} = \iint_{g(x, y) \leq z} f(x, y) dx dy$$

$$(2) \text{ 求导数: } F'_Z(z) = f_Z(z)$$

方法2 (公式法) 当 $Z = kX + Y, X + kY, XY, \frac{X}{Y}$ 时, 一般才用公式法, 公式为:

$$(1) Z = kX + Y \text{ 时: } f_Z(z) = \int_{-\infty}^{+\infty} f(x, z - kx) dx \stackrel{\text{若独立}}{=} \int_{-\infty}^{+\infty} f_X(x) \cdot f_Y(z - kx) dx;$$

$$(2) Z = X + kY \text{ 时: } f_Z(z) = \int_{-\infty}^{+\infty} f(z - ky, y) dy \stackrel{\text{若独立}}{=} \int_{-\infty}^{+\infty} f_X(z - ky) \cdot f_Y(y) dy;$$

$$(3) Z = XY \text{ 时: } f_Z(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f\left(x, \frac{z}{x}\right) dx = \int_{-\infty}^{+\infty} \frac{1}{|y|} f\left(\frac{z}{y}, y\right) dy;$$

$$(4) Z = \frac{X}{Y} \text{ 时: } f_Z(z) = \int_{-\infty}^{+\infty} |y| f(yz, y) dy;$$

(通过下面的具体例子讲解方法)

例: $X \sim U[0, a]$, X 与 Y 独立同分布, 求 $Z = X + Y$ 的概率密度 $f_Z(z)$ 。

解: (1) 写公式: $f_Z(z) = \int_{-\infty}^{+\infty} f(x, z - x) dx \stackrel{\text{独立}}{=} \int_{-\infty}^{+\infty} f_X(x) \cdot f_Y(z - x) dx;$

$$(2) \text{ 为使被积函数} \neq 0, \text{ 要求: } \begin{cases} 0 < x < a \\ 0 < z - x < a \rightarrow z - a < x < z \end{cases};$$

(3) 画数轴 (关于积分变量 x 的数轴): x 要落在 $(0, a)$ 之间, x 还要落在 $(z - a, z)$ 之间, 则需讨论 z 的区间;

$$(4) \text{ 讨论 } z \text{ 的区间: } z < 0: f_Z(z) = 0,$$

$$0 \leq z < a: f_Z(z) = \int_0^z \frac{1}{a} \cdot \frac{1}{a} dx = \frac{z}{a^2},$$

$$a \leq z < 2a: f_Z(z) = \int_{z-a}^a \frac{1}{a} \cdot \frac{1}{a} dx = \frac{2a - z}{a^2},$$

$$z \geq 2a: f_Z(z) = 0.$$

第四章 数字特征

1. 计算数学期望的常用思路:

(1) 利用性质: 例如 $E(aX - bY + Z) = aEX - bEY + EZ$

(2) 利用常见分布已知的期望方差结果, 及方差公式变型: $EX^2 = DX + (EX)^2$

例如: $X \sim P(\lambda)$, 则

$$E(X^2 + 3X - 1) = E(X^2) + 3EX - 1 = DX + (EX)^2 + 3EX - 1 = \lambda^2 + 4\lambda - 1$$

(3) 利用随机变量函数的期望公式:

例: $E(g(X)) = \int_{-\infty}^{+\infty} g(x)f(x)dx$; $E(g(X,Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy$ 。

(4) 涉及正态分布的特殊情形: 利用相互独立的正态分布的线性组合仍服从正态分布, 将多个正态分布的问题转化为一维正态分布的相关问题。

例: X, Y 相互独立, 且均服从正态分布 $N(\mu, \sigma^2)$, 求 $P\{X - Y > 0\}$, $E(|X - Y|)$ 。

解: 设 $Z = X - Y$, 则 $Z \sim N(0, 2\sigma^2)$, 则

$$P\{X - Y > 0\} = P\{Z > 0\} = 0.5.$$

$$\begin{aligned} E(|X - Y|) &= E(|Z|) = \int_{-\infty}^{+\infty} |z|f_z(z)dz = \int_{-\infty}^{+\infty} |z| \frac{1}{\sqrt{2\pi} \cdot \sqrt{2\sigma}} e^{-\frac{z^2}{2 \cdot 2\sigma^2}} dz \\ &= \frac{2}{2\sqrt{\pi}\sigma} \int_0^{+\infty} ze^{-\frac{z^2}{4\sigma^2}} dz = \frac{1}{\sqrt{\pi}\sigma} \cdot 2\sigma^2 = \frac{2\sigma}{\sqrt{\pi}}. \end{aligned}$$

(5) 涉及离散型随机变量求期望的特殊情形: 利用分解思想,

$$EX = E(X_1 + X_2 + \cdots + X_n)$$

——适用于离散型随机变量的分布律较难求的情况。

例: 求二项分布随机变量 X 的期望和方差时, 将 X 分解为若干个相互独立的 0-1 分布 X_i 的和, 即 $X = X_1 + X_2 + \cdots + X_n$, 得到 $EX = nE(X_i)$, $DX = nD(X_i)$ 。

2. 本章常用公式

(1) 方差相关公式

$$DX = E(X - EX)^2 = EX^2 - (EX)^2, \quad EX^2 = DX + (EX)^2$$

$$D(aX + b) = a^2 \cdot DX$$

$$D(X \pm Y) = DX + DY \pm 2\text{cov}(X, Y)$$

$$D(X + Y \pm Z) = DX + DY + DZ + 2\text{cov}(X, Y) \pm 2\text{cov}(Y, Z) \pm 2\text{cov}(X, Z)$$

(2) 协方差、相关系数相关公式

$$\text{cov}(X, Y) = E[(X - EX)(Y - EY)] = E(XY) - EX \cdot EY = \rho_{XY} \cdot \sqrt{DX} \cdot \sqrt{DY}$$

$$\text{cov}(X, X) = DX, \quad \text{cov} X, C = 0$$

$$\text{cov}(aX + bY, Z) = a \cdot \text{cov}(X, Z) + b \cdot \text{cov}(Y, Z) \quad (\text{“记得用”})$$

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{DX} \cdot \sqrt{DY}}$$

例: $EX=1, EY=2, DX=1, DY=4, \rho_{XY}=0.6$, 则 $E(2X-Y+1)^2 = \underline{4.2}$.

3. 独立与不相关的关系、不相关的若干充要条件

X 与 Y 独立(无任何关系) \iff X 与 Y 不相关(无线性关系) $\left\{ \begin{array}{l} \Leftrightarrow \rho_{XY}=0 \\ \Leftrightarrow \text{cov}(X,Y)=0 \\ \Leftrightarrow E(XY)=EX \cdot EY \\ \Leftrightarrow D(X \pm Y)=DX+DY \end{array} \right.$

4. 正态分布的若干结论

(1) $X \sim N(\mu, \sigma^2)$, 则 $\frac{X-\mu}{\sigma} \sim N(0,1)$

(2) $X \sim N(\mu, \sigma^2)$, 则 $aX+b \sim N(a\mu+b, a^2\sigma^2)$

(3) 相互独立的正态分布的线性组合仍服从正态分布, 即 $X_i \sim N(\mu_i, \sigma_i^2)$,

$(i=1,2,\dots,n)$, 且 X_1, X_2, \dots, X_n 相互独立, 则

$$a_1X_1+a_2X_2+\dots+a_nX_n \sim N\left(\sum_{i=1}^n a_i\mu_i, \sum_{i=1}^n a_i^2\sigma_i^2\right)$$

(4) 若 (X,Y) 服从二维正态分布, 即 $(X,Y) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$, 则

① $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$; (反之不成立)

② X 与 Y 独立 \Leftrightarrow X 与 Y 不相关; (反之不成立)

③ $aX+bY$ 服从一维正态分布; (反之不成立)

④ $aX+bY, cX+dY$ 服从二维正态分布。

第五章 大数定律及中心极限定理

1. 契比雪夫不等式

$$P\{|X-E(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2}, \quad P\{|X-E(X)| < \varepsilon\} \geq 1 - \frac{D(X)}{\varepsilon^2}.$$

2. 大数定律

(1) 2 项内涵: 通过严谨证明保证——频率依概率趋于概率, 平均值依概率趋于真值。

(2) 做题常用公式: $\frac{1}{n} \sum_{i=1}^n X_i^k \xrightarrow{P} E X_i^k$, 其中 X_i 两两相互独立且同分布。

3. 中心极限定理

(1) 勒维—林德贝格(独立同分布)中心极限定理: $\sum_{i=1}^n X_i \rightarrow N(n\mu, n\sigma^2)$, 其

中 X_i 为两两相互独立且同分布, 且 $E(X_k) = \mu$ 、 $D(X_k) = \sigma^2 > 0$ 。

(2) 棣莫弗—拉普拉斯 (二项分布) 中心极限定理: 若随机变量 $X \sim B(n, p)$, 则在 $n \rightarrow \infty$ 时 $X \sim N(np, np(1-p))$ 。

第六章 数理统计的基本概念

1. 本章常用公式与结论

(1) 样本均值 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$; 样本方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$;

(2) $E(\bar{X}) = EX$, $D(\bar{X}) = \frac{DX}{n}$, $E(S^2) = DX$;

(3) χ^2 、 t 、 F 分布的构造原理:

$\chi^2(n) = X_1^2 + X_2^2 + \cdots + X_n^2$, $t(n) = \frac{X}{\sqrt{Y/n}}$, $F(n, m) = \frac{X/n}{Y/m}$;

(4) 正态总体的若干统计量及其分布 (前提: 总体 $X \sim N(\mu, \sigma^2)$), 则

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

2. 求统计量服从分布的一般思路

(1) 猜: 根据问题先猜统计量服从的分布;

(2) 凑: 看已知什么, 按 χ^2 、 t 、 F 分布的构造原理, 自己去凑;

(3) 记得正态分布“标准化”;

(4) 最后整理, 如果猜的方向对, 自然就会出来结果。

第七章 参数估计

1. 点估计是要作什么?

已知: (1) 总体 X 的分布, 其中含有未知参数 θ ;

(2) 样本 X_1, X_2, \cdots, X_n ,

目的: 构造关于样本 X_1, X_2, \cdots, X_n 的函数, 去估计 θ ,

称: $\hat{\theta} = g(X_1, X_2, \cdots, X_n)$ 为 θ 的估计量, $\hat{\theta} = g(x_1, x_2, \cdots, x_n)$ 为 θ 的估计值。

2. 矩估计的计算步骤:

(1) 求出总体矩: $EX, EX^2, \dots, EX^K, \dots$;

(2) 写出样本矩: $A_1 = \frac{1}{n} \sum_{i=1}^n X_i, A_2 = \frac{1}{n} \sum_{i=1}^n X_i^2, \dots, A_K = \frac{1}{n} \sum_{i=1}^n X_i^K \dots$;

(3) 列方程: 令
$$\begin{cases} EX = A_1 \\ EX^2 = A_2 \\ \vdots \\ EX^K = A_K \\ \vdots \end{cases},$$
 即可求出 $\hat{\theta}$ 。

例: $X \sim E \lambda, \lambda > 0$ 未知, 样本 X_1, X_2, \dots, X_n , 求 λ 的矩估计。

解: (1) $EX = \frac{1}{\lambda}$; (2) $A_1 = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$; (3) 令 $\frac{1}{\lambda} = \bar{X} \Rightarrow \hat{\lambda} = \frac{1}{\bar{X}}$ 。

说明:

(1) 一般有几个参数, 列几个方程;

(2) 结果记得加 \wedge ;

(3) 矩估计法不唯一 (由此后面引出了点估计的评价)

(4) 大数定律: $\frac{1}{n} \sum_{i=1}^n X_i^K \xrightarrow{P} EX^K$, 即 $A_K \xrightarrow{P} EX^K$ 保证了矩估计法的合理性。

3. 极大似然估计的计算步骤:

计算原理: 当极大似然函数 $L(\theta)$ 取最大值时, 求相应的参数 θ 的方法。

常见计算步骤:

(1) 写出极大似然函数:

$$L(\theta) = \begin{cases} P\{X_1 = x_1\}P\{X_2 = x_2\} \cdots P\{X_n = x_n\}, & \text{离散型总体,} \\ f_{X_1}(x_1)f_{X_2}(x_2) \cdots f_{X_n}(x_n), & \text{连续型总体.} \end{cases}$$

(2) 取对数: $\ln L(\theta)$;

(3) 求导数: 令 $\frac{d \ln L(\theta)}{d \theta} = 0$, 求出估计量 $\hat{\theta}$ 。

说明:

(1) 若有多个参数, 则上面第三步中 $\ln L(\theta)$ 应对每个参数求偏导数, 并令其为 0, 求出各个参数的估计量;

(2) 上述求 $L(\theta)$ 最大值的方法并不严格 (未讨论导数不存在的点、未判断是否为极大值点或最大值点), 但因概率都是实际问题, 一般导数为零的结果只有一

个, 就认为这个结果为 θ 的估计了。

(3) 上面的步骤可能失效 (一般未知参数为总连续型总体概率密度的分段点时, 一定失效), 无法求出 $\hat{\theta}$ 。此时, 应从原理上分析, 直接分析 $L(\theta)$, 找到使 $L(\theta)$ 取得最大值时 θ 的估计量, 该情况一般的结论是

$$\hat{\theta} = \max(X_1, X_2, \dots, X_n) \text{ 或 } \hat{\theta} = \min(X_1, X_2, \dots, X_n).$$

4. 点估计的评价

- (1) 无偏性: $E(\hat{\theta}) = \theta$, 称 $\hat{\theta}$ 是参数 θ 的无偏估计。
- (2) 有效性: $E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta$, 且 $D(\hat{\theta}_1) < D(\hat{\theta}_2)$, 称 $\hat{\theta}_1$ 较 $\hat{\theta}_2$ 有效。
- (3) 一致性: $\hat{\theta}$ 依概率收敛于 θ , 称 $\hat{\theta}$ 为 θ 的一致 (相合) 估计。

5. 一个正态总体前提下对参数的双侧区间估计

未知参数		选择的统计量及分布	$1-\alpha$ 置信区间
μ	σ^2 已知	$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$	$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right)$
	σ^2 未知	$T = \frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t(n-1)$	$\left(\bar{X} - \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1), \bar{X} + \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1) \right)$
σ^2		$\chi^2 = \frac{n-1}{\sigma^2} S^2 \sim \chi^2(n-1)$	$\left(\frac{(n-1)S^2}{\chi_{\alpha/2}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n-1)} \right)$

第八章 假设检验

1. 显著性水平 α 是什么?

显著性水平 α = 犯第一类错误 (弃真) 的概率 = $P\{\text{拒绝 } H_0 | H_0 \text{ 为真}\}$

2. 一个正态总体前提下对参数的双侧假设检验

原假设 H_0		选择的统计量及分布	拒绝域
$\mu = \mu_0$	σ^2 已知	$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$	$ z > z_{\alpha/2}$
	σ^2 未知	$T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \sim t(n-1)$	$ t > t_{\alpha/2}(n-1)$
$\sigma^2 = \sigma_0^2$		$\chi^2 = \frac{n-1}{\sigma_0^2} S^2 \sim \chi^2(n-1)$	$\chi^2 \geq \chi_{\alpha/2}^2(n-1) \text{ 或 } \chi^2 \leq \chi_{1-\alpha/2}^2(n-1)$