

《数理统计》知识点总结

创作声明：

大家好！我是一名来自西北工业大学电子信息学院的 2022 级硕士研究生，2022-2023 年第一学期选修了《数理统计》课程（M11G11004），感谢西北工业大学数学与统计学院《数理统计》教学团队各位老师的辛勤付出。

在学习过程中，我总结了本门课程第一章至第六章的主要知识点，参考教材为师义民、徐伟等老师编写的《数理统计》一书。本资料由我一人编写，旨在为同学们复习提供一些帮助。

但由于我的编写能力和时间有限，难免存在一些不足和错误。如果读者在阅读中发现了任何错误或疏漏，敬请联系我，联系方式为：869465624（QQ）。

此外，希望读者朋友多多关注我的微信公众号“知识贩卖机”和 B 站账号“多多〇知识贩卖机”（UID:405870457），我将不定期分享知识干货，希望能够为大家提供更多有益的帮助。

特此声明。

王华元

2023 年 3 月于西北工业大学长安校区

第 0 章 常见分布及其分布函数

类型 1：均匀分布

$$X \sim U(a, b)$$
$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其他} \end{cases}$$
$$EX = \frac{a+b}{2}, \quad DX = \frac{(b-a)^2}{12}$$

类型 2：指数分布

$$X \sim E(\lambda)$$
$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$
$$EX = \frac{1}{\lambda}, \quad DX = \frac{1}{\lambda^2}$$

类型 3：泊松分布

$$X \sim P(\lambda)$$

离散

类型 4: 几何分布

$$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots$$

$$EX = \lambda, \quad DX = \lambda$$

类型 5: 两点分布

$$X \sim GE(p)$$

$$P\{X = k\} = (1-p)^{k-1}p, \quad k = 1, 2, \dots$$

$$EX = \frac{1}{p}, \quad DX = \frac{1-p}{p^2}$$

类型 6: 二项分布

$$X \sim B(1, p)$$

$$P\{X = k\} = p^k(1-p)^{1-k}, \quad k = 0, 1, \dots$$

$$EX = p, \quad DX = p(1-p)$$

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

$$P\{X = k\} = C_N^k p^k (1-p)^{N-k}, \quad k = 0, 1, \dots$$

$$EX = Np, \quad DX = Np(1-p)$$

正态分布

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

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

σ^2 的充分完备统计量:

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

一般题中
为已知

第 1 章 统计量与抽样分布

1. 常用统计量——样本矩

(1) 样本均值

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

(2) 样本方差

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

(3) 修正样本方差

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

2. 样本矩的期望与方差

$$E\bar{X} = EX, \quad D\bar{X} = \frac{1}{n}DX, \quad ES_n^2 = \frac{n-1}{n}DX, \quad ES_n^{*2} = DX$$

3. 利用因子分解定理判定充分统计量

(1) 连续型情况

设总体 X 具有分布密度 $f(x; \theta)$, $(X_1, X_2, \dots, X_n)^T$ 是一个样本, $T(X_1, X_2, \dots, X_n)$ 是一个统计量, 则 T 是 θ 的充分统计量的充要条件是: 样本的联合分布密度函数可以分解为

$$E(aX \pm bY) = aEX \pm bEY$$

$$D(aX \pm bY) = a^2DX + b^2DY$$

$$EX^2 = DX + (EX)^2$$

3. 利用因子分解定理判定充分统计量

(1) 连续型情况

设总体 X 具有分布密度 $f(x; \theta)$, $(X_1, X_2, \dots, X_n)^T$ 是一个样本, $T(X_1, X_2, \dots, X_n)$ 是一个统计量, 则 T 是 θ 的充分统计量的充要条件是: 样本的联合分布密度函数可以分解为

$$L(\theta) \triangleq \prod_{i=1}^n f(x_i; \theta) = h(x_1, x_2, \dots, x_n) g[T(x_1, x_2, \dots, x_n); \theta]$$

其中 h 是 $x_1, x_2, \dots, x_n$ 的非负函数且与 θ 无关, g 仅通过 T 依赖于 $x_1, x_2, \dots, x_n$ 。

(2) 离散型情况

设总体 X 的分布律 $P\{X = x^{(i)}\} = p(x^{(i)}; \theta) (i = 1, 2, \dots)$, $(X_1, X_2, \dots, X_n)^T$ 是一个样本, $T(X_1, X_2, \dots, X_n)$ 是一个统计量, 则 T 是 θ 的充分统计量的充要条件是: 样本的联合分布律可以表示为

$$\prod_{i=1}^n P\{X = x_i\} = h(x_1, x_2, \dots, x_n) g[T(x_1, x_2, \dots, x_n); \theta]$$

其中 h 是 $x_1, x_2, \dots, x_n$ 的非负函数且与 θ 无关, g 仅通过 T 依赖于 $x_1, x_2, \dots, x_n$ 。

4. 利用指数型分布族判定充分完备统计量

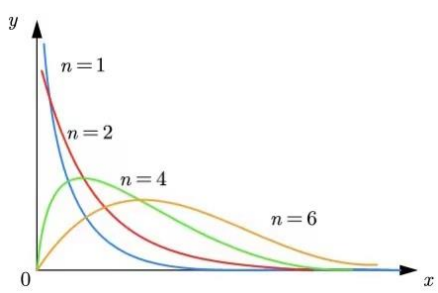
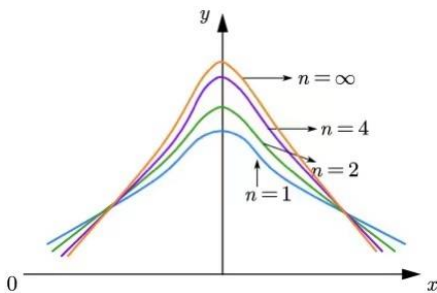
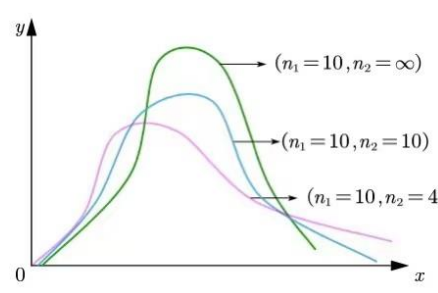
设总体 X 的分布密度为 $f(x; \theta)$ 为指数族分布, 即样本的联合分布密度具有如下形式:

$$\lambda^x = e^{x/\ln \lambda}$$

$$\prod_{i=1}^n f(x_i; \theta) = C(\theta) e^{\sum_{j=1}^m b_j(\theta) T_j(x_1, x_2, \dots, x_n)} h(x_1, x_2, \dots, x_n)$$

其中 $\theta = (\theta_1, \theta_2, \dots, \theta_m)^T$, $\theta \in \Theta$ 。如果 Θ 中包含一个 m 维矩阵, 而且 $B = [b_1(\theta), b_2(\theta), \dots, b_m(\theta)]^T$ 的值域包含有一个 m 维开集, 则 $T = [T_1(X_1, X_2, \dots, X_n), T_2(X_1, X_2, \dots, X_n), \dots, T_m(X_1, X_2, \dots, X_n)]^T$ 是参数 $\theta = (\theta_1, \theta_2, \dots, \theta_m)^T$ 的充分完备统计量。

5. 三大抽样分布

χ^2 分布	t 分布	F 分布
设随机变量$X_1, X_2, \dots, X_n$相互独立且同服从于标准正态分布$N(0,1)$, 则称随机变量 $\chi_n^2 = X_1^2 + X_2^2 + \dots + X_n^2$ 服从自由度为n的χ^2分布, 记为$\chi_n^2 \sim \chi^2(n)$。	设$X \sim N(0,1)$, $Y \sim \chi^2(n)$, 且X与Y相互独立, 则称随机变量 $T = \frac{X}{\sqrt{Y/n}}$ 服从自由度为n的t分布, 记为$T \sim t(n)$。	设$X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$, 且X与Y相互独立, 则称随机变量 $F = \frac{X/n_1}{Y/n_2}$ 服从自由度为(n_1, n_2)的F分布, 记为$F \sim F(n_1, n_2)$。
		
(1) $E\chi_n^2 = n$, $D\chi_n^2 = 2n$。 (2) 若$\chi_1^2 \sim \chi^2(n)$, $\chi_2^2 \sim \chi^2(m)$, 且χ_1^2与χ_2^2相互独立, 则$\chi_1^2 + \chi_2^2 \sim \chi^2(n+m)$。	设$T \sim t(n)$, 当$n > 2$时, $ET = 0$, $DT = \frac{n}{n-2}$。	设$F \sim F(n_1, n_2)$, 则$\frac{1}{F} \sim F(n_2, n_1)$。

6. 概率分布的分位数

设 X 是随机变量, 对于给定的实数 $\alpha (0 < \alpha < 1)$, 若存在 x_α 使

$$P\{X > x_\alpha\} = \alpha$$

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

则称 x_α 为 X 的上侧分位数。

标准正态分布	t 分布	F 分布
$u_\alpha = -u_{1-\alpha}$	$t_\alpha(n) = -t_{1-\alpha}(n)$	$F_{1-\alpha}(n_2, n_1) = \frac{1}{F_\alpha(n_1, n_2)}$

7. 正态总体样本均值和方差的分佈

对象	对总体的要求	所用函数及其分布
均值 μ	一个正态总体, σ^2 已知	$U = \frac{\bar{X} - \mu}{\sigma} \sqrt{n} \sim N(0,1)$
均值 μ	一个正态总体, σ^2 未知	$T = \frac{\bar{X} - \mu}{S_n^*} \sqrt{n} \sim t(n-1)$
方差 σ^2	一个正态总体, μ 未知	$K = \frac{nS_n^2}{\sigma^2} = \frac{(n-1)S_n^{*2}}{\sigma^2} \sim \chi^2(n-1)$
方差 σ^2	一个正态总体, μ 已知	$K = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$
均值差 $\mu_1 - \mu_2$	两个正态总体, σ_1^2, σ_2^2 已知	$U = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$
均值差 $\mu_1 - \mu_2$	两个正态总体, $\sigma_1^2 = \sigma_2^2 = \sigma$ 未知	$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$ 其中 $S_w^2 = \frac{(n_1-1)S_1^{*2} + (n_2-1)S_2^{*2}}{n_1 + n_2 - 2}$
方差商 $\frac{\sigma_1^2}{\sigma_2^2}$	两个正态总体, μ_1, μ_2 未知	$F = \frac{S_1^{*2}/\sigma_1^2}{S_2^{*2}/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$
方差商 $\frac{\sigma_1^2}{\sigma_2^2}$	两个正态总体, μ_1, μ_2 已知	$F = \frac{\sigma_1^2 \sum_{i=1}^{n_2} (Y_i - \mu_2)^2 / n_2}{\sigma_2^2 \sum_{i=1}^{n_1} (X_i - \mu_1)^2 / n_1} \sim F(n_2, n_1)$

8. 最小次序统计量和最大次序统计量的分布密度

$$F(x) = P\{X \leq x\}$$

$$f_{X_{(1)}}(x) = n[1 - F(x)]^{n-1}f(x)$$

$$f_{X_{(n)}}(x) = n[F(x)]^{n-1}f(x)$$

$\bar{X}$ 与 S_n^2 (或 S_n^{*2}) 独立

9. 样本中位数

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

$$\tilde{X} = \begin{cases} X_{(\frac{n+1}{2})}, & n \text{ 为奇数} \\ \frac{1}{2} [X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}], & n \text{ 为偶数} \end{cases}$$

10. 样本极差

$$R = X_{(n)} - X_{(1)} = \max_{1 \leq i \leq n} X_i - \min_{1 \leq i \leq n} X_i$$

定义1.3 设 $X_1, X_2, \dots, X_n$ 是总体 X 的一个样本, $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 为总体 X 的样本 $(X_1, X_2, \dots, X_n)$ 的次序统计量.

$(x_{(1)}, x_{(2)}, \dots, x_{(n)})$ 为其观测值, 设 x 是任一实数, 称函数

$$F_n(x) = \begin{cases} 0, & x < x_{(1)}, \\ \frac{k}{n}, & x_{(k)} \leq x < x_{(k+1)}, \\ 1, & x \geq x_{(n)}. \end{cases}$$

为总体 X 的经验分布函数。

第2章 参数估计

1. 无偏估计

设 $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$ 是参数 θ 的估计量, 若

$$E\hat{\theta} = \theta$$

则称 $\hat{\theta}$ 是 θ 的无偏估计量。

$$MSE(\hat{\theta}, \theta) = E(\hat{\theta} - \theta)^2 = D(\hat{\theta}) + (E\hat{\theta} - \theta)^2$$

2. 均方误差准则

设 θ 为一个未知参数, $\hat{\theta}$ 为 θ 的一个估计量, $\hat{\theta}$ 的均方误差定义为

$$MSE(\hat{\theta}, \theta) = E(\hat{\theta} - \theta)^2$$

均方误差反映了估计量 $\hat{\theta}$ 与被估参数 θ 的平均误差。对于一个估计量, 它的均方误差越小就说明估计的效果越好。对于无偏估计, 均方误差越小越好的准则等价于估计量方差越小越好的准则。

∴ 当 $E\hat{\theta} = \theta$, 比较 $D(\hat{\theta})$

3. 矩估计法

设总体 X 的分布函数 $F(x; \theta_1, \theta_2, \dots, \theta_m)$ 中有 m 个未知参数 $\theta_1, \theta_2, \dots, \theta_m$, 假定总体 X 的 m 阶矩存在, 记总体 X 的 k 阶原点矩为 α_k , 则

$$EX^k = \int_{-\infty}^{+\infty} x^k dF(x; \theta_1, \theta_2, \dots, \theta_m) \triangleq \alpha_k(\theta_1, \theta_2, \dots, \theta_m)$$

其中 $k = 1, 2, \dots, m$ 。现用样本的 k 阶原点矩作为总体 k 阶原点矩的估计, 即令

$$A_k = \frac{1}{n} \sum_{i=1}^n X_i^k = \alpha_k(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m), \quad k = 1, 2, \dots, m$$

$$\begin{cases} \text{总体 } k \text{ 阶原点矩} & \alpha_k = E(X^k) \\ \text{样本 } k \text{ 阶原点矩} & A_k = \frac{1}{n} \sum_{i=1}^n X_i^k \\ \text{总体 } k \text{ 阶中心矩} & \mu_k = E(X - E(X))^k \\ \text{样本 } k \text{ 阶中心矩} & B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k \end{cases}$$

解上述方程得 $\hat{\theta}_k = \hat{\theta}_k(X_1, X_2, \dots, X_n)$, $k = 1, 2, \dots, m$, 并以 $\hat{\theta}_k$ 作为参数 θ_k 的估计量, 则称 $\hat{\theta}_k$ 为未知参数 θ_k 的矩估计量, 这种求点估计量的方法称为矩估计法。若 $\hat{\theta}_k$ 为 θ_k 的矩估计量, $g(\theta)$ 为连续函数, 则也称 $g(\hat{\theta}_k)$ 为 $g(\theta_k)$ 的矩估计。

4. 似然函数

设总体 X 是连续型随机变量, 其分布密度为 $f(x, \theta)$, 其中 $\theta = (\theta_1, \theta_2, \dots, \theta_m)^T$ 是未知参数。若 $(X_1, X_2, \dots, X_n)^T$ 是来自总体 X 的一个样本, 则样本 $(X_1, X_2, \dots, X_n)^T$ 的联合分布密度为 $\prod_{i=1}^n f(x_i; \theta)$, 当给定样本值 $(x_1, x_2, \dots, x_n)^T$ 后, 它只是参数 $\theta = (\theta_1, \theta_2, \dots, \theta_m)^T$ 的函数, 记为 $L(\theta)$, 即

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta)$$

这个函数 L 称为似然函数, 即似然函数就是样本的联合分布密度。

若总体 X 是离散型随机变量, 其分布律为

$$P\{X = x\} = p(x; \theta), \quad x = x^{(1)}, x^{(2)}, \dots$$

其中 $\theta = (\theta_1, \theta_2, \dots, \theta_m)^T$ 是未知参数。 $(X_1, X_2, \dots, X_n)^T$ 是来自总体 X 的样本, 则样本 $(X_1, X_2, \dots, X_n)^T$ 的联合分布律 $\prod_{i=1}^n P\{X = x_i\}$ 称为似然函数, 记为 $L(\theta)$, 即

$$L(\theta) = \prod_{i=1}^n P\{X = x_i\} = \prod_{i=1}^n p(x_i; \theta)$$

5. 求最大似然估计量的一般步骤

(1) 写出似然函数 $L(\theta)$;

(2) 求出 $\ln L$ 及似然方程

$$\left. \frac{\partial \ln L}{\partial \theta_i} \right|_{\theta = \hat{\theta}} = 0, \quad i = 1, 2, \dots, m;$$

(3) 解似然方程得到最大似然估计 $\hat{\theta}_i(x_1, x_2, \dots, x_n)$ ($i = 1, 2, \dots, m$);

(4) 最后得到最大似然估计量 $\hat{\theta}_i(X_1, X_2, \dots, X_n)$ ($i = 1, 2, \dots, m$)。

三. 相合估计

1. 相合估计

对于 $\forall \theta \in \Theta$, 当 $n \rightarrow \infty$, $\hat{\theta}$ 依概率收敛于 θ

$$\lim_{n \rightarrow \infty} P\{|\hat{\theta}_n - \theta| < \varepsilon\} = 1$$

记为 $\hat{\theta}_n \xrightarrow{P} \theta$

2. 判别法则

定理2.1 设 $\hat{\theta}_n$ 是 θ 的一个估计量, 若

$$\lim_{n \rightarrow \infty} E(\hat{\theta}_n) = \theta, \quad \text{且} \quad \lim_{n \rightarrow \infty} D(\hat{\theta}_n) = 0$$

则 $\hat{\theta}_n$ 是 θ 的相合估计 (或一致估计).

* 6. 最小方差无偏估计 (MVUE) 的求解方法

- (1) 寻找 θ 的一个充分完备统计量 T ;
- (2) 寻找 θ 的一个无偏估计 $\hat{\theta}$;
- (3) 求 $\hat{\theta}^* = E(\hat{\theta}|T)$.

7. Fisher 信息量

$$I(\theta) = E \left[\frac{\partial}{\partial \theta} \ln f(X; \theta) \right]^2$$

$$\left. \begin{aligned} I(\theta) &= E \left(\frac{\partial \ln f(X; \theta)}{\partial \theta} \right)^2 \\ I(\theta) &= -E \left[\frac{\partial^2 \ln f(X; \theta)}{\partial \theta^2} \right] \end{aligned} \right\} \text{为 Fisher 信息量.}$$

8. 罗-克拉默不等式

 $T(X) = T(X_1, X_2, \dots, X_n)$ 是 $g(\theta)$ 的一个无偏估计量, 对一切 $\theta \in \Theta$, 有

$$D[T(X)] \geq \frac{[g'(\theta)]^2}{nI(\theta)}$$

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

特别当 $g(\theta) = \theta$ 时, 上式成为

$$D[T(X)] \geq \frac{1}{nI(\theta)}$$

以上两式称为信息不等式, 或称为罗-克拉默不等式。不等式的右端项称为估计量 $T(X)$ 方差的罗-克拉默下界。

9. 有效估计

若 θ (或 $g(\theta)$) 的一个无偏估计量 $\hat{\theta}$ (或 $T(X)$) 的方差达到罗-克拉默下界, 即

$$D\hat{\theta} = \frac{1}{nI(\theta)} \quad (\text{或} \quad D[T(X)] = \frac{[g'(\theta)]^2}{nI(\theta)})$$

* 则称 $\hat{\theta}$ (或 $T(X)$) 为 θ (或 $g(\theta)$) 的有效估计量。

10. 估计量的效率

设 $\hat{\theta}$ 是 θ 的任一无偏估计, 称

$$e(\hat{\theta}) = \frac{1}{nI(\theta)} / D\hat{\theta}$$

为 $\hat{\theta}$ 的效率。由罗-克拉默不等式知, 对于任意一个无偏估计 $\hat{\theta}$, 其效率满足 $0 < e(\hat{\theta}) \leq 1$ 。如果 $e(\hat{\theta}) = 1$, 则 $\hat{\theta}$ 是 θ 的有效估计。

11. 正态总体均值和方差的置信区间

估计对象	对总体的要求	所用函数及其分布	置信区间
均值 μ	一个正态总体, σ^2 已知	$U = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\left(\bar{X} - u_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + u_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right)$
均值 μ	一个正态总体, σ^2 未知	$T = \frac{\bar{X} - \mu}{S_n^*/\sqrt{n}} \sim t(n-1)$	$\left(\bar{X} - t_{\frac{\alpha}{2}}(n-1) \frac{S_n^*}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2}}(n-1) \frac{S_n^*}{\sqrt{n}} \right)$
方差 σ^2	一个正态总体, μ 未知	$K = \frac{nS_n^2}{\sigma^2} = \frac{(n-1)S_n^{*2}}{\sigma^2} \sim \chi^2(n-1)$	$\left(\frac{(n-1)S_n^{*2}}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1)S_n^{*2}}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \right)$

方差 σ^2	一个正态总体, μ 已知	$K = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$	$\left(\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{\frac{\alpha}{2}}^2(n)}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{1-\frac{\alpha}{2}}^2(n)} \right)$
均值 差 $\mu_1 - \mu_2$	两个正态总体, σ_1^2, σ_2^2 已知	$U = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$	$\left(\bar{X} - \bar{Y} - u_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \bar{X} - \bar{Y} + u_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right)$
均值 差 $\mu_1 - \mu_2$	两个正态总体, $\sigma_1^2 = \sigma_2^2 = \sigma$ 未知	$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$ 其中 $S_w^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}$	$\left(\bar{X} - \bar{Y} - S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{\frac{\alpha}{2}}(n_1 + n_2 - 2), \bar{X} - \bar{Y} + S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{\frac{\alpha}{2}}(n_1 + n_2 - 2) \right)$
方差 商 $\frac{\sigma_1^2}{\sigma_2^2}$	两个正态总体, μ_1, μ_2 未知	$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$	$\left(F_{1-\frac{\alpha}{2}}(n_2 - 1, n_1 - 1) \frac{S_1^{*2}}{S_2^{*2}}, F_{\frac{\alpha}{2}}(n_2 - 1, n_1 - 1) \frac{S_1^{*2}}{S_2^{*2}} \right)$
方差 商 $\frac{\sigma_1^2}{\sigma_2^2}$	两个正态总体, μ_1, μ_2 已知	$F = \frac{\sigma_1^2 \sum_{i=1}^{n_2} (Y_i - \mu_2)^2 / n_2}{\sigma_2^2 \sum_{i=1}^{n_1} (X_i - \mu_1)^2 / n_1} \sim F(n_2, n_1)$	$\left(F_{1-\frac{\alpha}{2}}(n_2, n_1) \frac{\sum_{i=1}^{n_1} (X_i - \mu_1)^2 / n_1}{\sum_{i=1}^{n_2} (Y_i - \mu_2)^2 / n_2}, F_{\frac{\alpha}{2}}(n_2, n_1) \frac{\sum_{i=1}^{n_1} (X_i - \mu_1)^2 / n_1}{\sum_{i=1}^{n_2} (Y_i - \mu_2)^2 / n_2} \right)$

第3章 统计决策与贝叶斯估计

1. 贝叶斯风险（平方损失函数情况）

(1) 损失函数

$$L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$$

(2) 风险函数

$$R(\hat{\theta}, \theta) = EL(\hat{\theta}, \theta) = E(\hat{\theta} - \theta)^2$$

(3) 贝叶斯风险

$$R_B(\hat{\theta}) = ER(\hat{\theta}, \theta) = \int_{\Theta} \pi(\theta) R(\hat{\theta}, \theta) d\theta$$

2. 贝叶斯估计

(1) 样本关于 θ 的条件分布

$$q(x|\theta) = \prod_{i=1}^n p(x_i|\theta)$$

(2) 样本和 θ 的联合分布

$$f(x, \theta) = \pi(\theta) q(x|\theta)$$

(3) 样本的边缘分布

$$m(x) = \int_{\Theta} f(x, \theta) d\theta$$

(4) θ 关于样本的后验分布

$$h(\theta|x) = \frac{f(x, \theta)}{m(x)}$$

(5) 贝叶斯估计

① $L(\theta, d) = [\theta - d(x)]^2$ (平方损失函数)

$$\hat{\theta}^* = E(\theta|x) = \int_{\Theta} \theta \cdot h(\theta|x) d\theta$$

② $L(\theta, d) = \lambda(\theta)[\theta - d(x)]^2$ (加权平方损失函数)

$$\hat{\theta}^* = \frac{E(\theta\lambda(\theta)|x)}{E(\lambda(\theta)|x)} = \frac{\int_{\Theta} \theta\lambda(\theta) \cdot h(\theta|x) d\theta}{\int_{\Theta} \lambda(\theta) \cdot h(\theta|x) d\theta}$$

3. 三种重要积分

(1)

$Gamma(\alpha, \beta)$ $\begin{cases} EX = \frac{\alpha}{\beta} \\ DX = \frac{\alpha}{\beta^2} \end{cases}$
伽玛分布

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx \quad (\alpha > 0)$$

$$\int u dv = uv - \int v du$$

性质: $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$; $\Gamma(1) = 1$; $\Gamma(n + 1) = n!$; $\Gamma(1/2) = \sqrt{\pi}$ 。

(2)

$$\Gamma(2) = 1$$

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx \quad (a > 0, b > 0)$$

性质: $B(a, b) = B(b, a)$; $B(a + 1, b) = \frac{a}{a+b} B(a, b)$; $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ 。

(3)

$$\int_a^b \frac{1}{x^{\alpha+1}} dx = \frac{1}{\alpha} \frac{1}{x^{\alpha}} \Big|_a^b$$

第4章 假设检验

1. 势函数

对于检验 $\delta(x)$, 可以定义一个函数

$$\beta(\theta) = E_{\theta}[\delta(X)] = P_{\theta}\{X \in W\}$$

称 $\beta(\theta)$ 为这个检验的势函数, 又称为功率函数。

置信区间: $|U| \leq u_{\frac{\alpha}{2}}$

↑ (相反)

2. 正态总体均值和方差的假设检验

H_0	适用范围	统计量	拒绝域
$\mu = \mu_0$	一个正态总体, σ_0^2 已知	$U = \frac{\bar{X} - \mu_0}{\sigma_0} \sqrt{n}$	$ U \geq u_{\frac{\alpha}{2}}$
$\mu = \mu_0$	一个正态总体, σ^2 未知	$T = \frac{\bar{X} - \mu_0}{S_n^*} \sqrt{n}$	$ T \geq t_{\frac{\alpha}{2}}(n-1)$
$\sigma^2 = \sigma_0^2$	一个正态总体, μ 未知	$K = \frac{nS_n^2}{\sigma_0^2} = \frac{(n-1)S_n^{*2}}{\sigma_0^2}$	$K \geq \chi_{\frac{\alpha}{2}}^2(n-1)$ 或 $K \leq \chi_{1-\frac{\alpha}{2}}^2(n-1)$
$\sigma^2 = \sigma_0^2$	一个正态总体, μ_0 已知	$K = \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma_0^2}$	$K \geq \chi_{\frac{\alpha}{2}}^2(n)$ 或 $K \leq \chi_{1-\frac{\alpha}{2}}^2(n)$
$\mu_1 = \mu_2$	两个正态总体, σ_1^2, σ_2^2 已知	$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$ U \geq u_{\frac{\alpha}{2}}$
$\mu_1 = \mu_2$	两个正态总体, $\sigma_1^2 = \sigma_2^2 = \sigma$ 未知	$T = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$ T \geq t_{\frac{\alpha}{2}}(n_1 + n_2 - 2)$

		其中 $S_w^2 = \frac{(n_1-1)S_1^{*2} + (n_2-1)S_2^{*2}}{n_1+n_2-2}$	
$\sigma_1^2 = \sigma_2^2$	两个正态总体, μ_1, μ_2 未知	$F = \frac{S_1^{*2}}{S_2^{*2}}$	$F \geq F_{\frac{\alpha}{2}}(n_1-1, n_2-1)$ 或 $F \leq F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)$
$\sigma_1^2 = \sigma_2^2$	两个正态总体, μ_1, μ_2 已知	$F = \frac{\sum_{i=1}^{n_2} (Y_i - \mu_2)^2 / n_2}{\sum_{i=1}^{n_1} (X_i - \mu_1)^2 / n_1}$	$F \geq F_{\frac{\alpha}{2}}(n_2, n_1)$ 或 $F \leq F_{1-\frac{\alpha}{2}}(n_2, n_1)$

3. 多项分布的 χ^2 检验法

检验假设 $H_0: p_i = p_{i0} \leftrightarrow H_1: p_i \neq p_{i0} \ (i = 1, 2, \dots, m)$, 其中 p_{i0} 是已知数。可使用皮尔逊统计量

$$\chi_n^2 = \sum_{i=1}^m \frac{(N_i - np_{i0})^2}{np_{i0}}$$

来衡量 $\frac{N_i}{n}$ 与 p_{i0} 之间的差异程度。如果 $\hat{\chi}_n^2 \geq \chi_{\alpha}^2(m-1)$, 则拒绝假设 H_0 , 即认为总体的分布与假设 H_0 中的分布有显著差异; 若 $\hat{\chi}_n^2 < \chi_{\alpha}^2(m-1)$, 则接受假设 H_0 , 即认为总体的分布与假设 H_0 中的分布无显著差异。

4. 科尔莫戈罗夫检验

设总体 X 的分布函数为 $F(x)$, 而且假定 $F(x)$ 是 x 的连续函数。设 $(X_1, X_2, \dots, X_n)^T$ 是来自 X 的一个容量为 n 的样本, 根据此样本作经验分布函数 $F_n(x)$ 。由格列文科定理断定

$$P \left\{ \lim_{n \rightarrow \infty} \sup_{-\infty < x < +\infty} |F_n(x) - F(x)| = 0 \right\} = 1$$

这个定理表示随机变量 $D_n = \sup_{-\infty < x < +\infty} |F_n(x) - F(x)|$ 以概率 1 是无穷小的。如果 $\hat{D}_n \geq D_{n,\alpha}$, 则拒绝假设 H_0 ($F(x) = F_0(x)$), 否则接受 H_0 。

5. 斯米尔诺夫检验

设 $(X_1, X_2, \dots, X_{n_1})^T$ 是来自具有连续分布函数 $F(x)$ 的总体 X 的样本, $(Y_1, Y_2, \dots, Y_{n_2})^T$ 是来自具有连续分布函数 $G(x)$ 的总体 Y 的样本, 且假定两个样本相互独立。欲检验假设:

$$H_0: F(x) = G(x) \leftrightarrow H_1: F(x) \neq G(x), \quad -\infty < x < +\infty$$

设 $F_{n_1}(x)$ 和 $G_{n_2}(x)$ 分别是这两个样本所对应的经验分布函数, 作统计量

$$D_{n_1, n_2} = \sup_{-\infty < x < +\infty} |F_{n_1}(x) - G_{n_2}(x)|$$

如果 $\hat{D}_{n_1, n_2} \geq D_{n, \alpha}$, 则拒绝假设 H_0 , 否则接受 H_0 。

第 5 章 方差分析与试验设计

类型 1: 单因素方差分析

(1) 总离差平方和分解

$$Q_T = Q_E + Q_A$$

其中

$$Q_E = \sum_{i=1}^r \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

(1) 总离差平方和分解

其中

$$Q_T = Q_E + Q_A$$

资料请关注微信公众号“知识贩卖机”

$$Q_E = \sum_{i=1}^r \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$
$$Q_A = \sum_{i=1}^r n_i (\bar{X}_i - \bar{X})^2$$

分别称 Q_E 与 Q_A 为组内离差平方和与组间离差平方和。 Q_E 表示 X_{ij} 与其组内平均 $\bar{X}_i$ 的离差平方和， Q_A 是组内平均与总平均的离差平方和。

r : 行数(有几家企业)
 n_i : 不同企业样本容量

(2) 显著性检验

构造检验统计量

$$F = \frac{\frac{Q_A}{r-1}}{\frac{Q_E}{n-r}} = \frac{\bar{Q}_A}{\bar{Q}_E}$$

服从自由度为 $(r-1, n-r)$ 的F分布。

一次抽样后由样本值计算得 F 的数值，若

$$F \geq F_{\alpha}(r-1, n-r)$$

则拒绝假设 H_0 ，即认为在显著性水平 α 下，因素的不同水平对试验结果有显著影响；若

$$F < F_{\alpha}(r-1, n-r)$$

则接受假设 H_0 ，即认为在显著性水平 α 下，因素的不同水平对试验结果无显著影响。

(3) 方差分析表

方差来源	离差平方和	自由度	平均离差平方和	F 值	显著性
组间	$Q_A = \sum_{i=1}^r n_i (\bar{X}_i - \bar{X})^2$	$r-1$	$\bar{Q}_A = \frac{Q_A}{r-1}$	$F = \frac{\bar{Q}_A}{\bar{Q}_E}$	
组内	$Q_E = \sum_{i=1}^r \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$	$n-r$	$\bar{Q}_E = \frac{Q_E}{n-r}$		
总和	$Q_T = \sum_{i=1}^r \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2$	$n-1$			

类型 2: 两因素非重复试验的方差分析

(1) 总离差平方和分解公式

记

$$\bar{X}_{i\cdot} = \frac{1}{s} \sum_{j=1}^s X_{ij} \quad (i = 1, 2, \dots, r)$$

$$\bar{X}_{\cdot j} = \frac{1}{r} \sum_{i=1}^r X_{ij} \quad (j = 1, 2, \dots, s)$$

$$\bar{X} = \frac{1}{rs} \sum_{i=1}^r \sum_{j=1}^s X_{ij} = \frac{1}{r} \sum_{i=1}^r \bar{X}_{i\cdot} = \frac{1}{s} \sum_{j=1}^s \bar{X}_{\cdot j}$$

总离差平方和

$$Q_T = Q_A + Q_B + Q_E$$

其中

总离差平方和

$$\begin{aligned}\bar{X}_{i\cdot} &= \frac{1}{s} \sum_{j=1}^s X_{ij} \quad (i = 1, 2, \dots, r) \\ \bar{X}_{\cdot j} &= \frac{1}{r} \sum_{i=1}^r X_{ij} \quad (j = 1, 2, \dots, s) \\ \bar{X} &= \frac{1}{rs} \sum_{i=1}^r \sum_{j=1}^s X_{ij} = \frac{1}{r} \sum_{i=1}^r \bar{X}_{i\cdot} = \frac{1}{s} \sum_{j=1}^s \bar{X}_{\cdot j} \\ Q_T &= Q_A + Q_B + Q_E\end{aligned}$$

$$Q_A = s \sum_{i=1}^r (\bar{X}_{i\cdot} - \bar{X})^2$$

$$Q_B = r \sum_{j=1}^s (\bar{X}_{\cdot j} - \bar{X})^2$$

$$Q_E = \sum_{i=1}^r \sum_{j=1}^s (X_{ij} - \bar{X}_{i\cdot} - \bar{X}_{\cdot j} + \bar{X})^2$$

其中 Q_A 为因素 A 引起的离差平方和， Q_B 为因素 B 引起的离差平方和， Q_E 称为随机误差平方和。

(2) 显著性检验

构造检验统计量

$$F_A = \frac{\frac{Q_A}{r-1}}{\frac{Q_E}{(r-1)(s-1)}} = \frac{\bar{Q}_A}{\bar{Q}_E}$$

$$F_B = \frac{\frac{Q_B}{s-1}}{\frac{Q_E}{(r-1)(s-1)}} = \frac{\bar{Q}_B}{\bar{Q}_E}$$

$F_A \sim F(r-1, (r-1)(s-1))$ ； $F_B \sim F(s-1, (r-1)(s-1))$ 。

一次抽样后由样本值计算得 F_A 的数值，若

$$F_A \geq F_{\alpha}(r-1, (r-1)(s-1))$$

则拒绝 H_{01} ，即认为因素 A 对试验结果有显著影响；若

$$F_A < F_{\alpha}(r-1, (r-1)(s-1))$$

则接受 H_{01} ，即认为因素 A 对试验结果无显著影响。

同样，根据一次抽样后由样本值计算得 F_B 的数值，若

$$F_B \geq F_{\alpha}(s-1, (r-1)(s-1))$$

则拒绝 H_{02} ，即认为因素 B 对试验结果有显著影响；若

$$F_B < F_{\alpha}(s-1, (r-1)(s-1))$$

则接受 H_{02} ，即认为因素 B 对试验结果无显著影响。

(3) 方差分析表

方差来源	离差平方和	自由度	均方误差	F 值	显著性
因素 A	$Q_A = s \sum_{i=1}^r (\bar{X}_{i\cdot} - \bar{X})^2$	$r - 1$	$\bar{Q}_A = \frac{Q_A}{r - 1}$	$F_A = \frac{\bar{Q}_A}{\bar{Q}_E}$	
因素 B	$Q_B = r \sum_{j=1}^s (\bar{X}_{\cdot j} - \bar{X})^2$	$s - 1$	$\bar{Q}_B = \frac{Q_B}{s - 1}$	$F_B = \frac{\bar{Q}_B}{\bar{Q}_E}$	
误差	$Q_E = \sum_{i=1}^r \sum_{j=1}^s (X_{ij} - \bar{X}_{i\cdot} - \bar{X}_{\cdot j} + \bar{X})^2$	$(r - 1)(s - 1)$	$\bar{Q}_E = \frac{Q_E}{(r - 1)(s - 1)}$		
总和	$Q_T = \sum_{i=1}^r \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2$	$rs - 1$			

第6章 回归分析

类型 1: 一元线性回归分析

(1) 一元线性回归模型

$$Y_i = \alpha + \beta x_i + \varepsilon_i, \quad i = 1, \dots, n$$

假设 ε_i 相互独立且 $\varepsilon_i \sim N(0, \sigma^2)$ ($i = 1, \dots, n$)。

(2) 参数的最小二乘估计量

$$\begin{cases} \hat{\alpha} = \bar{Y} - \hat{\beta} \bar{x} \\ \hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{cases}$$

(3) 方差的无偏估计量

$$\hat{\sigma}^{*2} = \frac{1}{n-2} \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta} x_i)^2$$

(4) 显著性检验

构造统计量

$$T = \frac{\hat{\beta}}{\hat{\sigma}^*} \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sim t(n-2)$$

当 $|T| \geq t_{\frac{\alpha}{2}}(n-2)$ 时, 拒绝 H_0 。

如果 $(x_{i1}, x_{i2}, \dots, x_{im}, Y_i) (i = 1, 2, \dots, n)$ 是 $(x_1, x_2, \dots, x_m, Y)$ 的 n 个观测值,

类型 2: 多元线性回归分析

(1) 多元线性回归模型

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im} + \varepsilon_i, \quad i = 1, \dots, n$$

假设 ε_i 相互独立且 $\varepsilon_i \sim N(0, \sigma^2)$ ($i = 1, \dots, n$)。

令

$$Y = (Y_1, Y_2, \dots, Y_n)^T, \quad \beta = (\beta_0, \beta_1, \dots, \beta_m)^T, \quad \varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$$

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1m} \\ 1 & x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nm} \end{pmatrix}$$

可得矩阵表达式为

$$Y = X\beta + \varepsilon$$

$$EY = X\beta, \quad \text{Cov}(Y, Y) = \sigma^2 I_n$$

参数估计

(2) 参数的最小二乘估计量

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

(3) 方差的无偏估计量

$$\hat{\sigma}^{*2} = \frac{1}{n-m-1} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (\hat{Y}_i = \sum_{j=0}^m \hat{\beta}_j x_{ij})$$

(4) 显著性检验

检验 $H_0: \alpha^T \beta = 0 \leftrightarrow H_1: \alpha^T \beta \neq 0$ 。令 $L = \alpha^T \hat{\beta}$, 由于 $\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$, 则

$$EL = \alpha^T E\hat{\beta} = \alpha^T \beta$$

$$K = \frac{(n-m-1)\hat{\sigma}^{*2}}{\sigma^2} \sim \chi^2(n-m-1)$$

【相关知识点】估计量的分布

$\hat{\beta}$ 的任一分量均是独立正态随机变量 $(Y_1, \dots, Y_n)^T$ 的线性组合, 随机向量 $\hat{\beta}$ 服从均值为 β , 协方差阵为 $\sigma^2 (X^T X)^{-1}$ 的正态分布。

所以, $L \sim N(\alpha^T \hat{\beta}, \sigma^2 \alpha^T (X^T X)^{-1} \alpha)$ 。

令

$$DL = \alpha^T \text{Cov}(\hat{\beta}, \hat{\beta}) \alpha = \sigma^2 \alpha^T (X^T X)^{-1} \alpha$$

$$U = \frac{\alpha^T \hat{\beta} - \alpha^T \beta}{\sigma \sqrt{\alpha^T (X^T X)^{-1} \alpha}} \sim N(0, 1)$$

当 H_0 成立时,

$$U = \frac{\alpha^T \hat{\beta}}{\sigma \sqrt{\alpha^T (X^T X)^{-1} \alpha}} \sim N(0, 1)$$

而

$$K = \frac{(n - m - 1) \hat{\sigma}^{*2}}{\sigma^2} \sim \chi^2(n - m - 1)$$

K 与 U 相互独立, 则

$$T = \frac{U}{\sqrt{K/(n - m - 1)}} = \frac{\alpha^T \hat{\beta}}{\hat{\sigma}^* \sqrt{\alpha^T (X^T X)^{-1} \alpha}} \sim t(n - m - 1)$$

当 $|T| \geq t_{\frac{\alpha}{2}}(n - m - 1)$ 时, 拒绝 H_0 。

对于 $\hat{\beta} = (X^T X)^{-1} X^T Y$,

经常需要计算 $(X^T X)^{-1}$ 。

简单的2阶矩阵的逆矩阵的计算口诀:

主对调, 次换号, 除以行列式

推导: 假设 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $a, b, c, d \in R$, 且 A 可逆, 那么根据知识储备 1.2 $A^* = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

所以呢, $A^{-1} = \frac{A^*}{|A|} = \frac{\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} \quad X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad \hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$$

$$Y = X\beta + \varepsilon$$

$$\text{估计量} \begin{cases} \text{最小二乘} & \hat{\beta} = (X^T X)^{-1} X^T Y \\ \text{无偏} & \sigma^2 = \frac{1}{n-m-1} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (\hat{Y}_i = \sum_{j=0}^m \hat{\beta}_j x_{ij}) \end{cases}$$

$$H_0: \alpha^T \beta = 0$$

$$\star \text{构造 } \hat{Y} = \alpha^T \hat{\beta}$$

$$E\hat{Y} = \alpha^T E\hat{\beta} = \alpha^T \beta$$

$$D\hat{Y} = D(\alpha^T \hat{\beta}) = \alpha^T D(\hat{\beta}) \alpha = \sigma^2 \alpha^T (X^T X)^{-1} \alpha$$

$$\therefore \hat{Y} \sim N(\alpha^T \beta, \sigma^2 \alpha^T (X^T X)^{-1} \alpha)$$

$$\text{令 } U = \frac{\alpha^T \hat{\beta} - \alpha^T \beta}{\sigma \sqrt{\alpha^T (X^T X)^{-1} \alpha}} \sim N(0, 1)$$

$$\text{而 } K = \frac{(n-m-1) \hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-m-1)$$

在 H_0 成立时 ($\alpha^T \beta = 0$)

$$T = \frac{U}{\sqrt{K/(n-m-1)}} = \frac{\alpha^T \hat{\beta}}{\hat{\sigma}^* \sqrt{\alpha^T (X^T X)^{-1} \alpha}} \sim t(n-m-1)$$

当 $|T| \geq t_{\frac{\alpha}{2}}(n-m-1)$ 拒绝 H_0