

数理统计笔记

强基数学 002 吴天阳

1 置信区间估计

定义 1.1 (置信区间). 令 X_1, \dots, X_n 是来自 $f(x; \theta)$ 的随机变量, $T_1 = T_1(X_1, \dots, X_n)$, $T_2 = T_2(X_1, \dots, X_n)$ 是两个关于样本的统计量, 且 $T_1 \leq T_2$, 满足 $P[T_1 < \tau(\theta) < T_2] = \gamma$, 其中 γ 与 θ 无关, 则称 $[T_1, T_2]$ 为随机区间, 也是 θ 的 100γ 置信度的置信区间, 其中 γ 称为置信系数, T_1 与 T_2 分别称为置信下限和置信上限.

注 1. 若 $\tau(\theta)$ 在随机区间中是严格单调的, 则 $P(\tau(T_1) < \tau(\theta) < \tau(T_2)) = P(T_1 < \theta < T_2) = \gamma$. 这告诉我们可以将许多问题转化为求解 θ 置信区间.

定义 1.2 (枢轴量). 设 X_1, \dots, X_n 是来自 $f(x; \theta)$ 的随机变量, 令 $Q = q(x_1, \dots, x_n; \theta)$, 若 Q 的分布与 θ 无关, 则称 Q 为枢轴量.

命题 1.3 (枢轴量方法). 给定置信系数 $0 < \gamma < 1$, 则 $\exists q_1, q_2$ 使得 $P(q_1 < Q < q_2)$, 则可通过 $q_1 < q(x_1, \dots, x_n; \theta) < q_2$ 反解出 $T_1(x_1, \dots, x_n) < \tau(\theta) < T_2(x_1, \dots, x_n)$, 于是 $P(T_1 < \tau(\theta) < T_2) = \gamma$. 于是得到 $\tau(\theta)$ 的置信区间为 (T_1, T_2) .

注 1. q_1, q_2 与 θ 无关.

注 2. 对于固定的 γ , q_1, q_2 的取值不唯一, 置信区间的期望长度 $\mathbb{E}(T_2 - T_1)$ 越小越好.

注 3. 若 Q 为 $\mu = 0$ 的对称分布, 如标准正态分布, t 分布, 取对称分布 $q_2 = -q_1 = z_{\frac{1+\gamma}{2}}$; 若为不对称分布, 通常取等尾分布 $q_2 = z_{\frac{1+\gamma}{2}}$, $q_1 = z_{\frac{1-\gamma}{2}}$, 其中 z_x 为 Q 的 x 分位数; 若为不对称分布, 还可通过转化为最优化极小化问题求解.

1.1 正态分布参数的置信区间

1.1.1 单参数估计

下面讨论来自正态分布的样本对其中未知参数的置信区间估计, 分为以下三种: (括号内为枢轴量满足的分布)

- (1) μ 未知, σ^2 已知. (标准正态分布)
- (2) μ 已知, σ^2 未知. ($\chi^2(n)$ 分布)
- (3) μ 未知, σ^2 未知. (μ 对应 $t(n)$ 分布, σ^2 对应 $\chi^2(n-1)$ 分布)

例 1.1. $X_1, \dots, X_n \stackrel{iid}{\sim} N(\theta, \sigma^2)$, 其中 σ^2 已知, 求 θ 的置信区间.

解答. 由于 $\bar{X} \sim N(\theta, \sigma^2/n)$, 则 $\frac{\bar{X} - \theta}{\sigma/\sqrt{n}} \sim N(0, 1)$, 设 $q_1, q_2 \in \mathbb{R}$, $q_1 < q_2$ 满足

$$\gamma = P\left(q_1 < \frac{\bar{X} - \theta}{\sigma/\sqrt{n}} < q_2\right) = P\left(\bar{X} - \frac{\sigma}{\sqrt{n}}q_2 < \theta < \bar{X} - \frac{\sigma}{\sqrt{n}}q_1\right)$$

由于 $N(0, 1)$ 为对称分布, 取 $q_2 = -q_1 = z_{\frac{1+\gamma}{2}}$, θ 的置信系数为 γ 的区间为

$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{1+\gamma}{2}} < \theta < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{1+\gamma}{2}} \right)$$

例 1.2. $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu, \sigma^2)$, 其中 μ 已知, 求 σ^2 的置信区间.

解答. 观察: 若直接使用上题的枢轴量, 由于 $q_1 < 0$, 通过计算发现, 只能获得 σ^2 的单边置信下限, 考虑构造值域非负的枢轴量, 例如 Γ, χ^2 .

由于 $\frac{X_i - \mu}{\sigma} \sim N(0, 1)$ 则 $\sum_{i=1}^N \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(n)$. 设 $q_1, q_2 \in [0, \infty)$, $q_1 < q_2$ 满足

$$\gamma = P \left(q_1 < \frac{\sum_{i=1}^N (X_i - \mu)^2}{\sigma^2} < q_2 \right) = P \left(\frac{\sum_{i=1}^N (X_i - \mu)^2}{q_2} < \sigma^2 < \frac{\sum_{i=1}^N (X_i - \mu)^2}{q_1} \right)$$

由于 χ^2 不是对称分布, 考虑等尾分布: $q_1 = \chi^2_{\frac{1-\gamma}{2}}(n)$, $q_2 = \chi^2_{\frac{1+\gamma}{2}}(n)$, 于是 σ^2 的置信度为 γ 的区间为

$$\left(\frac{\sum_{i=1}^N (X_i - \mu)^2}{\chi^2_{\frac{1+\gamma}{2}}(n)}, \frac{\sum_{i=1}^N (X_i - \mu)^2}{\chi^2_{\frac{1-\gamma}{2}}(n)} \right)$$

例 1.3. $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu, \sigma^2)$, 其中 μ, σ^2 均未知, 分别求 μ, σ^2 的置信区间.

解答. (i) 求解 μ 的置信区间.

观察: 考虑使用例 1.1 的做法, 由于分布处含有未知的 σ^2 , 会使得上下限中含有 σ^2 , 考虑利用随机变量相除将 σ^2 消去且满足某种分布, 于是我们引入了 t 分布: 若 $U \sim N(0, 1), V \sim \chi^2(n)$, 则 $\frac{U}{\sqrt{V/n}} \sim t(n)$.

由于 $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$, $\sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, 则

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \frac{\sigma\sqrt{n-1}}{\sqrt{(n-1)S^2}} = \sqrt{n} \frac{\bar{X} - \mu}{S} \sim t(n-1)$$

设 $q_1, q_2 \in \mathbb{R}$, $q_1 < q_2$ 满足

$$\gamma = P \left(q_1 < \sqrt{n} \frac{\bar{X} - \mu}{S} < q_2 \right) = P \left(\bar{X} - \frac{S}{\sqrt{n}} q_2 < \mu < \bar{X} - \frac{S}{\sqrt{n}} q_1 \right)$$

其中 $S^2 = \frac{1}{n-1} \sum_{i=1}^N (X_i - \bar{X})^2$. 由于 t 为对称分布, 取 $q_2 = -q_1 = t_{\frac{1+\gamma}{2}}(n-1)$, 则 μ 的置信度为 γ 的区间为

$$\left(\bar{X} - \frac{S}{\sqrt{n}} t_{\frac{1+\gamma}{2}}(n-1), \bar{X} + \frac{S}{\sqrt{n}} t_{\frac{1+\gamma}{2}}(n-1) \right)$$

(ii) 求 σ^2 的置信区间. 考虑使用例1.2的做法, 将 μ 用 \bar{X} 代替则

$$\sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

类似例1.2可得 σ^2 的置信度为 γ 的区间为

$$\left(\frac{(n-1)S^2}{\chi_{\frac{1+\gamma}{2}}^2(n-1)}, \frac{(n-1)S^2}{\chi_{\frac{1-\gamma}{2}}^2(n-1)} \right)$$

1.1.2 双参数估计

我们只关心两个独立的正态分布的均值之差与方差之比的区间估计, 令 X_1, \dots, X_n 是来自 $N(\mu_1, \sigma_1^2)$ 的随机样本, Y_1, \dots, Y_m 是来自 $N(\mu_2, \sigma_2^2)$ 的随机样本, 考虑 $\mu_2 - \mu_1$ 的置信区间, 主要分为以下三种情况 (括号内为枢轴量满足的分布)

- (1) σ_1^2, σ_2^2 已知. (标准正态分布)
- (2) $\sigma_1^2 = \sigma_2^2 = \sigma^2$, σ 未知. ($t(n+m-2)$ 分布)
- (3) σ_1^2, σ_2^2 未知. (样本量大利用大数定律构造标准正态分布, 样本量小记结论构造 $t(n+m-2)$ 分布)

不难发现上述构造的枢轴量都是对称分布, 所以最终求出来的置信区间也都是对称的.

例 1.4. 设 $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu_1, \sigma_1^2), Y_1, \dots, Y_m \stackrel{iid}{\sim} N(\mu_2, \sigma_2^2)$, σ_1^2, σ_2^2 已知, 求 $\mu_2 - \mu_1$ 的区间估计.

解答. 由于 $\bar{Y} - \bar{X} \sim N\left(\mu_2 - \mu_1, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)$, 则 $\frac{\bar{Y} - \bar{X} - (\mu_2 - \mu_1)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim N(0, 1)$, 取对称分布点 $q_1 = -q_2 = z_{\frac{1+\gamma}{2}}$, 则 $\mu_2 - \mu_1$ 的置信度为 μ 的区间为

$$(Q_1, Q_2), \quad Q_1, Q_2 = \bar{Y} - \bar{X} \pm z_{\frac{1+\gamma}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}$$

例 1.5. 设 $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu_1, \sigma_1^2), Y_1, \dots, Y_m \stackrel{iid}{\sim} N(\mu_2, \sigma_2^2)$, $\sigma_1 = \sigma_2 = \sigma$, σ 未知, 求 $\mu_2 - \mu_1$ 的区间估计.

解答. 观察: 由于分母中有未知参数 σ , 类似例1.2的方法, 通过除以 χ^2 分布消去 σ 即可得到枢轴量.

由于 $U = \frac{\bar{Y} - \bar{X} - (\mu_2 - \mu_1)}{\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{m}}} \sim N(0, 1)$, $\frac{(n-1)S_1^2}{\sigma^2} \sim \chi^2(n-1)$, $\frac{(m-1)S_2^2}{\sigma^2} \sim \chi^2(m-1)$, 于是 $V = \frac{(n-1)S_1^2 + (m-1)S_2^2}{\sigma^2} \sim \chi^2(n+m-2)$, 则

$$\frac{U}{\sqrt{V/(n+m-2)}} = \frac{\bar{Y} - \bar{X} - (\mu_2 - \mu_1)}{S_p \sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t(n+m-2)$$

其中 $S_p^2 = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$. 取对称分布点 $q_1 = -q_2 = t_{\frac{1+\gamma}{2}}(n+m-2)$ 可得 $\mu_2 - \mu_1$ 置信度为 γ 的置信区间为

$$(Q_1, Q_2), \quad Q_1, Q_2 = \bar{Y} - \bar{X} \pm t_{\frac{1+\gamma}{2}}(n+m-2)S_p \sqrt{\frac{1}{n} + \frac{1}{m}}$$

注 1. 当 $\sigma_1 = \lambda\sigma_2$ 时, 也可转化为上述问题, 最终得到的区间估计为

$$(Q_1, Q_2), \quad Q_1, Q_2 = \bar{Y} - \bar{X} \pm t_{\frac{1+\gamma}{2}}(m+n-2)S \sqrt{\frac{1}{n} + \frac{\lambda}{m}}, \text{ 其中 } S^2 = \frac{(n-1)S_1^2 + \frac{m-1}{\lambda}S_2^2}{m+n-2}$$

例 1.6. 设 $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu_1, \sigma_1^2), Y_1, \dots, Y_m \stackrel{iid}{\sim} N(\mu_2, \sigma_2^2)$, $\sigma_1 \neq \sigma_2$, σ_1, σ_2 均未知, 求 $\mu_2 - \mu_1$ 的区间估计.

解答. (i) 当样本量 n, m 非常大 (一般为 ≥ 30) 时, 可以认为 $S_1^2 \xrightarrow{p} \sigma_1^2, S_2^2 \xrightarrow{p} \sigma_2^2$, 则 $\frac{\bar{Y} - \bar{X} - (\mu_2 - \mu_1)}{\sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{m}}}$ 近似服从 $N(0, 1)$, 则 $\mu_2 - \mu_1$ 置信度为 γ 的区间为

$$(Q_1, Q_2), \quad Q_1, Q_2 = \bar{Y} - \bar{X} \pm z_{\frac{1+\gamma}{2}} \sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{m}}$$

(ii) 令 $S_k^2 = \frac{S_1^2}{n} + \frac{S_2^2}{m}$ 则 $T = \frac{\bar{Y} - \bar{X} - (\mu_2 - \mu_1)}{S_k}$ 近似服从 $t(r)$, 其中 $r = \frac{S_k^4}{\frac{S_1^4}{n^2(n-1)} + \frac{S_2^4}{m^2(m-1)}}$,

则 $\mu_2 - \mu_1$ 置信度为 γ 的区间为

$$(Q_1, Q_2), \quad Q_1, Q_2 = \bar{Y} - \bar{X} \pm t_{\frac{1+\gamma}{2}}(r)S_k$$

例 1.7 (联合正态分布). 令 $(X_1, Y_1), \dots, (X_n, Y_n) \stackrel{iid}{\sim} N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right)$ 服从二维正态分布, 其中 $\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$, 求 $\mu_2 - \mu_1$ 的区间估计.

解答. 令 $D_i = Y_i - X_i$, 则 $D_i \sim N(\mu_2 - \mu_1, \sigma_D^2)$, $\sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$, 转化为 μ, σ 都未知, 求 μ 的单参数区间估计问题 (转化为 t 分布), 由例 1.3(i) 可得, $\mu_2 - \mu_1$ 的置信度为 γ 的区间为

$$(Q_1, Q_2), \quad Q_1, Q_2 = \bar{D} \pm \frac{S_D}{\sqrt{n}} t_{\frac{1+\gamma}{2}}(n-1)$$

例 1.8 (方差之比). 设 $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu_1, \sigma_1^2), Y_1, \dots, Y_m \stackrel{iid}{\sim} N(\mu_2, \sigma_2^2)$, $\sigma_1 \neq \sigma_2$, μ_1, μ_2 均未知, 求 σ_2^2/σ_1^2 的区间估计.

解答. 由于 $T_1 = \frac{\sum_{i=1}^n (X_i - \mu_1)^2}{\sigma_1^2} \sim \chi^2(n)$, $T_2 = \frac{\sum_{j=1}^m (Y_j - \mu_2)^2}{\sigma_2^2} \sim \chi^2(m)$, 则

$$F = \frac{T_1/n}{T_2/m} = \frac{\sigma_2^2 \sum_{i=1}^n (X_i - \mu_1)^2}{\sigma_1^2 \sum_{j=1}^m (Y_j - \mu_2)^2} \sim F(n, m)$$

由于 F 分布不是对称的, 取等尾分布 $q_1 = F_{\frac{1-\gamma}{2}}(n, m), q_2 = F_{\frac{1+\gamma}{2}}(n, m)$ 则 σ_2^2/σ_1^2 置信度为 γ 的区间为

$$\left(\frac{\sum_{j=1}^m (Y_j - \mu_2)^2/m}{\sum_{i=1}^n (X_i - \mu_1)^2/n} F_{\frac{1-\gamma}{2}}(n, m), \frac{\sum_{j=1}^m (Y_j - \mu_2)^2/m}{\sum_{i=1}^n (X_i - \mu_1)^2/n} F_{\frac{1+\gamma}{2}}(n, m) \right)$$

注 1. μ_1, μ_2 未知, 则问题转化为类似 μ, σ 都未知求 σ 分布, 见例 1.3(ii), 利用 \bar{X}, \bar{Y} 分别替换 μ_1, μ_2 , 将自由度 -1 即可, σ_2^2/σ_1^2 的置信度为 γ 的区间为

$$\left(\frac{S_2^2}{S_1^2} F_{\frac{1-\gamma}{2}}(n-1, m-1), \frac{S_2^2}{S_1^2} F_{\frac{1+\gamma}{2}}(n-1, m-1) \right)$$

1.2 一般的枢轴量方法

命题 1.4 (分布函数的枢轴量方法). 设 X_1, \dots, X_n 是来自 $f(x; \theta)$ 的随机样本, 将 $F(X; \theta)$ 视为随机变量, 则 $F(X; \theta) \sim I(0, 1)$, 于是 $-\log F(X; \theta) \sim \Gamma(1, 1)$, 则 $-\sum_{i=1}^n \log F(X_i; \theta) \sim \Gamma(n, 1)$ 为枢轴量.

注 1. 上述方法在分布函数便于计算时可以使用.

注 2. 一般我们期望转化为关于 χ^2 分布的枢轴量 (便于查表), 于是

$$-2 \sum_{i=1}^n \log F(X_i; \theta) \sim \Gamma(n, 1/2) = \chi^2(2n)$$

再利用 $\chi^2(2n)$ 等尾概率求出置信区间.

注 3. $F(x; \theta) \sim I(0, 1)$ 是因为 $P(F(x; \theta) \leq y) = P(x \leq F^{-1}(y)) = F(F^{-1}(y)) = y$.

注 4. 令 $z = -\log Y$ 且 $Y \sim I(0, 1)$, 则 $Z \in (0, \infty)$ 且 $f_Z(z) = f_Y(e^{-z})e^{-z} = e^{-z} \sim \Gamma(1, 1)$.

2 假设检验

2.1 基本概念

定义 2.1 (假设 Hypothesis). 对参数取值的一个断言, 将假设记为 H , H 后接一个冒号, 冒号后的描述该断言内容.

若可通过该断言确定样本服从的分布函数, 则称为**简单假设**; 否则称为**复合假设**.

注 1. 上述的断言就是对未知的参数 θ 进行的一种前提假设, 我们期望通过样本的相关信息判断该断言是否正确, 这就是假设检验.

注 2. 设 $X_1, \dots, X_n \stackrel{iid}{\sim} f(x; \theta)$, $\theta_0 \in \Theta, \Theta_0 \subset \Theta$, 其中 Θ 为参数空间. 则 $H: \theta = \theta_0$ 是简单断言; $H: \theta \in \Theta_0$ 是复合断言. (简单断言就是将参数取定, 而符合断言中参数仍然不定)

定义 2.2 (两个对立假设). 根据实际问题, 提出**两个对立假设**: H_0 称为**原假设 (The null hypothesis)**, H_1 称为**备择假设 (The alternative hypothesis)**.

注 1. 对立假设的含义是：若拒绝 H_0 ，则一定接受 H_1 ；反之若接受 H_0 ，则一定拒绝 H_1 。

注 2. 实际含义上：原假设是实际问题中期望得到的结论（如：新产品的收益率比旧产品更高，女士不具有分辨茶和奶加入顺序）。

定义 2.3 (检验 Test). 根据样本结果，判断是否**拒绝**原假设 H_0 ，将这个判断规则称为**检验**记为 T 。判断规则一般通过**检验统计量** T^* 是否在某个集合内来判断。

如果 T^* 属于某个集合内 Ω 时就拒绝 H_0 ，则称该检验 T 为**非随机化检验**。

如果以一定概率拒绝 H_0 ，则称该检验 T 为**随机化检验**。若样本落在某个集合 C_T 内时，拒绝原假设，则称 C_T 为**拒绝域**，无歧义时简记为 C 。

将拒绝原假设概率记为**检验函数** $\psi_T = \psi_T(x_1, \dots, x_n) = \psi_T(\text{reject } H_0 | x_1, \dots, x_n)$ 。则非随机化检验的 ψ_T 只有 0 和 1 的取值；而随机化检验中 ψ_T 有介于 0, 1 之间的取值。

注 1. 我们一般只讨论**拒绝**原假设的情况（因为一般我们不认为原假设为真，例：不会轻易接受新产品）

注 2. 设拒绝域 C 为开集，则若 $\psi_T = I_C = \begin{cases} 0, & (x_1, \dots, x_n) \in C, \\ 1, & (x_1, \dots, x_n) \notin C. \end{cases}$ 则检验 T 是一个

非随机化检验，

若 $\psi_T = \begin{cases} 0, & (x_1, \dots, x_n) \in C^\circ, \\ 1/2, & (x_1, \dots, x_n) \in \partial C, \\ 1, & (x_1, \dots, x_n) \notin \bar{C}. \end{cases}$ 则检验 T 是一个随机化检验。

注 3. 随机化检验一般是带有主观色彩得，因为它会以一定的概率去拒绝原假设，并不果断。（一般在离散随机变量时用到）

注 4. 一般难以直接找到拒绝域 C ，通过转化为检验统计量找等价条件，即

$$(x_1, \dots, x_n) \in C \iff T^*(x_1, \dots, x_n) \in C^*.$$

定义 2.4 (两类错误，势函数 Power function). 设 T 为 H_0 vs. H_1 的检验

第一类错误（弃真）： H_0 正确但是拒绝了 H_0 。记为 $\alpha = \alpha(\theta) = P_{H_0}(\text{reject } H_0)$

第二类错误（取伪）： H_0 错误但是接受了 H_0 。记为 $\beta = \beta(\theta) = P_{H_1}(\text{reject } H_1)$

T 的势函数： $\pi_T(\theta) = P_\theta(\text{reject } H_0)$ ，即参数取值为 θ 时，拒绝 H_0 的概率。

势函数与两类错误的关系：

1. 若 H_0, H_1 均为简单假设时，设检验 T 为 $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$ ，则 $\alpha = P_{\theta_0}(\text{reject } H_0) = \pi_T(\theta_0)$ ， $\beta = P_{\theta_1}(\text{reject } H_1) = 1 - \pi_T(\theta_1)$ 。
2. 若 H_0, H_1 均为复合假设时，设检验 T 为 $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1$ ，则

$$\begin{cases} \alpha(\theta) = \pi_T(\theta), & \theta \in \Theta_0, \\ \beta(\theta) = 1 - \pi_T(\theta), & \theta \in \Theta_1. \end{cases}$$
 此处的 α, β 不是常数，与简单假设时不同。

注 1. 在复合假设下， α, β 都是与 θ 相关的函数，即在不同的参数下，犯错误的概率不同。我们期望在任何情况下，我们犯错误能达到最小，于是有了检验水平的定义。

注 2. 脚标 P_{H_0} 表示 H_0 对应的断言成立前提下，计算出的概率。

注 3. 当 T 为非随机化检验时 $\pi_T(\theta) = P_\theta[(x_1, \dots, x_n) \in C]$ 。

当 T 为随机化检验时 (该结果可推出非随机化检验情况, 取 $\psi_T = I_C$ 即可)

$$\begin{aligned}\pi_T(\theta) &= P_\theta[\text{reject } H_0] = \int \cdots \int P(\text{reject } H_0 | x_1, \cdots, x_n) f_{X_1, \dots, X_n}(x_1, \cdots, x_n; \theta) \prod_{i=1}^n dx_i \\ &= \int \cdots \int \psi_T(x_1, \cdots, x_n) f_{X_1, \dots, X_n}(x_1, \cdots, x_n) \prod_{i=1}^n dx_i = \mathbb{E}_\theta[\psi_T(x_1, \cdots, x_n)]\end{aligned}$$

定义 2.5 (检验水平 Size of test). 令原假设 $H_0: \theta \in \Theta_0$, 则检验 T 关于 H_0 的检验水平为 $\alpha = \sup_{\theta \in \Theta_0} \pi_T(\theta)$, 也称显著水平 (Significance level)

注 1. 若 $\Theta_0 = \{\theta_0\}$ 即为简单假设时, 检验水平 α 就是犯第一类错误的概率 $\pi_T(\theta_0)$.

解决假设检验的一般思路:

1. 根据实际问题提出原假设和备择假设 H_0, H_1 .
2. 通过样本的分布和假设, 确定检验模型 T (简单似然比检验、广义似然比检验).
3. 根据检验水平 α 确定检验的拒绝域 C_T . (此处拒绝域一般找某个充分统计量下的等价拒绝域)
4. 根据样本实际取值做出判断, 是否要拒绝 H_0 .

例 2.1. 设 $X \sim \theta x^{\theta-1} I_{(0,1)}(x)$, $\theta > 0$, 求关于 $H_0: \theta \geq 2$ vs. $H_1: \theta < 2$ 检验水平为 α 的 GLR 检验.

解答. 由于 $\frac{dL(\theta)}{d\theta} = x^{\theta-1}(1 + \theta \log x)$, 则 $\theta = -\frac{1}{\log x}$ 处有最大值, 于是

$$\sup_{\theta > 0} L(\theta) = L\left(-\frac{1}{\log x}\right) = -\frac{1}{\log x} x^{-\frac{1}{\log x} - 1}$$

下讨论 $\theta \geq 2$ 时, $L(\theta)$ 的最大值取值, 当 $2 \leq -\frac{1}{\log x} \Rightarrow e^{-\frac{1}{2}} < x < 1$ 时, 最大值在 $-\frac{1}{\log x}$ 处取到;

当 $-\frac{1}{\log x} < 2 \Rightarrow 0 < x < e^{-\frac{1}{2}}$ 时, 最大值在 2 处取到, $L(2) = 2x$. 于是

$$\lambda = \frac{\sup_{\theta \geq 2} L(\theta)}{\sup_{\theta > 0} L(\theta)} = \begin{cases} 1, & e^{-\frac{1}{2}} < x < 1, \\ -2x^{\frac{1}{\log x} + 2} \log x, & 0 < x < e^{-\frac{1}{2}}. \end{cases}$$

则拒绝 $H \iff -2x^{\frac{1}{\log x} + 2} \log x < \lambda_0$ 且 $0 < x < e^{-\frac{1}{2}}$, 令 $y = -\log x \Rightarrow x = e^{-y}$, ($y > 1/2$), 于是

$$-2x^{\frac{1}{\log x} + 2} \log x = -2(e^{-y})^{-\frac{1}{y} + 2}(-y) = 2ye^{1-2y} < \lambda_0$$

令 $g(y) = 2ye^{1-2y}$, 则 $g'(y) = (2 - 4y)e^{1-2y} < 0$, 则在 $y > 1/2$ 上单调递减, 于是 $g(y) < \lambda_0 \iff y > k$.

于是检验水平

$$\alpha = \sup_{\theta \geq 2} P_\theta[y > k] = \sup_{\theta \geq 2} \int_k^\infty \theta e^{-\theta y} dy = \sup_{\theta \geq 2} e^{-\theta k} = e^{-2k} \Rightarrow k = -\frac{1}{2} \log \alpha$$

GLR 检验的拒绝域为 $C = \left\{x : -\log x > -\frac{1}{2} \log \alpha\right\} = \{x : x < \sqrt{\alpha}\}$.