概率论与数理统计

第八章

韩潇
xhan011@ustc.edu.cn

第八章：　假设检验

大纲：

- 基本概念
- 正态总体参数检验
- 比例p的检验
- 似然比检验
- P值

8.4 似然比检验

- 记样本$\boldsymbol{X}$有联合密度函数或联合概率质量函数$f(\boldsymbol{x};\theta)$, 其为θ的连续函数.

- 考虑假设检验问题

$$H_0: \theta \in \Theta_0 \leftrightarrow H_1: \theta \in \Theta_1 = \Theta \backslash \Theta_0$$

其中Θ_0为参数空间Θ的非空真子集.

- 若 $f(\boldsymbol{x};\theta_1) < f(\boldsymbol{x};\theta_2)$, 则我们认为真参数为$\theta_2$的“似然性”较其为$\theta_1$的“似然性”大, 即观测的样本$\boldsymbol{X}$由$\theta_2$解释更好一些.

- 因此，我们考虑两个量 $L_{\Theta_0}(x) = \sup_{\theta \in \Theta_0} f(\boldsymbol{x};\theta), L_{\Theta_1}(x) = \sup_{\theta \in \Theta_1} f(\boldsymbol{x};\theta)$

- 若$L_{\Theta_1}(x)/ L_{\Theta_0}(x)$比较大，则倾向于拒绝原假设;反之倾向于不拒绝
- 令$L_{\Theta}(x) = \sup_{\theta\in\Theta} f(\boldsymbol{x};\theta)$, $LR(x) = L_{\Theta}(x)/ L_{\Theta_0}(x)$, 易得

$$LR(x) = \max\{1, L_{\Theta_1}(x)/ L_{\Theta_0}(x)\}$$

- 我们用$LR(x)$代替$L_{\Theta_1}(x)/ L_{\Theta_0}(x)$得到如下定义

定义 8.3 似然比检验

设样本 $\mathbf{X}$ 有联合密度函数或联合概率质量函数 $f(\mathbf{x};\theta), \theta \in \Theta \subseteq \mathbb{R}^k$, 则称统计量

$$LR(\mathbf{x}) = \sup_{\theta\in\Theta} f(\mathbf{x};\theta) \Big/ \sup_{\theta\in\Theta_0} f(\mathbf{x};\theta)$$

为检验问题(8.38)的似然比. 而由下式定义的检验

$$\phi: \quad \text{当 } LR(\mathbf{x}) > c \text{ 时, 拒绝原假设 } H_0, \text{不然不能拒绝 } H_0$$

称为检验 (8.38) 的一个似然比检验, 其中常数 c 可通过控制检验的水平来确定. ♣

8.4 似然比检验

例 8.21 设样本 $(X_1, \ldots, X_n)$ 来自总体 $N(\theta, \sigma^2)$, σ^2 未知. 求假设检验问题 $H_0 : \theta = \theta_0 \leftrightarrow H_1 : \theta \neq \theta_0$ 的一个水平 α 检验.

例 8.22 设 $\mathbf{X} = (X_1, \ldots, X_n)$ 为自均匀分布总体 $U(0, \theta)$ 中抽取的随机样本, 求

$$H_0 : \theta \leq \theta_0 \leftrightarrow H_1 : \theta > \theta_0 \tag{8.39}$$

水平 α 似然比检验. 此处 α 和 θ_0 给定.

8.4 似然比检验

例 8.23 设样本 $\mathbf{X}=(X_1,\ldots,X_n)$ 取自指数分布总体, 其密度函数为

$$f(x,\theta)=\frac{1}{2}\exp\{-(x-\theta)/2\},\ \ x\geq\theta,\quad -\infty<\theta<\infty$$

求检验问题

$$H_0:\ \theta=\theta_0\leftrightarrow H_1:\ \theta\neq\theta_0$$

水平 α 似然比检验. 此处 α 和 θ_0 给定.

8.4 似然比检验

- 若$X_1,\ldots,X_n$是简单随机样本, 在原假设成立之下, 当$n\to\infty$时, 似然比有一个简单的极限分布(Wilks, 1938).
- 极限分布的参数, 需要确定参数空间的维数. 例如 $\Theta=\{(\mu,\sigma^2),\mu\in R,\sigma^2>0\}$维数是2, 而 $\Theta_0=\{(\mu,\sigma^2),\mu=\mu_0,\sigma^2>0\}$维数是1.

定理 8.1 似然比的极限分布

设 Θ 的维数为 k, Θ_0 的维数为 s, 若 $k-s=t>0$, 则对检验问题 (8.38) 在原假设 H_0 成立之下, 当样本大小 $n\to\infty$ 时有

$$\mathbb{P}(2\ln LR(\mathbf{X})\le x)\ \to\ F_{\chi_t^2}(x),\quad \forall\, x\in\mathbb{R}.$$

记为 $2\ln LR(\mathbf{X})\xrightarrow{\mathscr{L}}\chi_t^2$.

8.4 似然比检验

例 8.24 设样本 $X_{i1}, \ldots, X_{in_i}$, i.i.d. $\sim N(\mu_i, \sigma_i^2)$, $1 \leq i \leq m$, 且全部样本相互独立. 求假设

$$H_0: \ \sigma_1^2 = \cdots = \sigma_m^2 \leftrightarrow H_1: \ \sigma_1^2 = \cdots = \sigma_m^2 \text{不完全相同.}$$

8.5 P值

- 如何衡量假设检验中拒绝原假设的证据的强度？- p值

p值=P(得到当前样本下检验统计量的值或更极端值|原假设下)

- 例：正态总体均值假设检验 $H_0: \mu = 2 \leftrightarrow H_1: \mu > 2$, 方差$\sigma^2 = 1$，则

$$p\text{值}=P(Z \geq z_{obs}|H_0),\ \text{其中}Z = \sqrt{n}(\bar{X} - 2)/\sigma$$

若观测到的第一组样本值$\bar{x} = 3$，样本量=4, 则$z_{obs} = \frac{\sqrt{n}(\bar{x}-2)}{\sigma} = 2.$

类似地，假设观测到的第二组(样本量相同)样本值得到的$\bar{x} =$

3.75, 则$z_{obs} = 3.5$

- 第二组样本的p值显著小于第一组样本的p值

8.5 P值

- p值越小，否定原假设的程度越强烈
- 取检验的水平为α, 当一个检验法则的p值不超过α时, 检验统计量T的值落在了拒绝域内, 我们即拒绝原假设; 反之则没有足够的证据拒绝原假设. 这样即得到一个水平α检验法则:

  p值=P(得到当前样本下检验统计量的值或更极端值|原假设下)

  ϕ:当p值$<\alpha$时，拒绝原假设H_0
- p值表示了在当前样本值下观测到的显著性水平.

8.5 P值

- 检验统计量T在原假设下的分布难以得到时候，应用中常用自助法计算p值，即在原假设下使用自助法得到统计量T的B个自助版本$T_1^*,\dots,T_B^*$值，若 p值=P(T $\geq \boldsymbol{t_{obs}}|\boldsymbol{H_0}$)可以近似为

$$p\text{值}\approx \frac{1}{B}\sum_{b=1}^{B} I(T_b^* \geq \boldsymbol{t_{obs}})$$

8.5 P值

例 8.25 废弃电池破损后会释放出金属成分, 进而随着雨水对周围环境造成重金属污染. 一项研究随机调查了具有丢弃破损 AAA 电池的 $n=51$ 个地点, 测得其中土壤中锌含量 (单位: g) 分别为

$$
\begin{array}{llllllllllllll}
1.94 & 2.06 & 1.88 & 1.96 & 2.25 & 2.26 & 2.34 & 1.96 & 2.22 & 2.15 & 2.03 & 2.00 & 1.91 & 1.82 \\
1.96 & 2.13 & 2.20 & 2.01 & 2.12 & 1.98 & 2.31 & 1.97 & 1.95 & 2.09 & 1.97 & 2.02 & 2.11 & 1.91 \\
1.97 & 1.91 & 2.29 & 2.22 & 2.01 & 2.23 & 2.18 & 1.95 & 2.01 & 2.04 & 2.22 & 2.27 & 2.14 & 2.18 \\
1.98 & 1.89 & 1.84 & 2.19 & 2.07 & 2.09 & 2.07 & 2.14 & 1.70
\end{array}
$$

试问可否得出总体平均锌含量超过 2.0 g 的结论?

8.5.1置信区间和假设检验之间的关系

- 置信区间和双边检验之间的关系：
  设$X_1,\ldots,X_n$为从总体$F(x;\theta),\theta\in\Theta$中抽取的样本，参数$\theta$的$1-\alpha$置信区间为$[\underline{\theta},\overline{\theta}]$, 即

$$P_\theta\left(\underline{\theta}\le\theta\le\overline{\theta}\right)\ge 1-\alpha,\theta\in\Theta$$

P_θ表示在θ对应分布下计算的概率. 考虑假设检验

$$H_0:\theta=\theta_0\leftrightarrow H_1:\theta\ne\theta_0$$

在原假设下有

$$P_{\theta_0}\left(\underline{\theta}\le\theta_0\le\overline{\theta}\right)\ge 1-\alpha$$

即

$$P_{\theta_0}\left(\theta_0>\overline{\theta}\right)+P_{\theta_0}\left(\theta_0<\underline{\theta}\right)\le\alpha$$

8.5.1置信区间和假设检验之间的关系

- 按照显著性定义，我们有如下检验

$$\phi: \theta_0 > \overline{\theta}\text{或}\theta_0 < \underline{\theta}\text{时拒绝}H_0,\text{否则不能拒绝}H_0$$

反之，如果假设$H_0: \theta = \theta_0 \leftrightarrow H_1: \theta \neq \theta_0 (*)$的接受域为

$$\underline{\theta}(x_1, \dots, x_n) \leq \theta_0 \leq \overline{\theta}\ (x_1, \dots, x_n),$$

即

$$P_{\theta_0}\left(\underline{\theta} \leq \theta_0 \leq \overline{\theta}\right) \geq 1 - \alpha$$

由θ_0的任意性，对任意的θ有

$$P_{\theta}\left(\underline{\theta} \leq \theta \leq \overline{\theta}\right) \geq 1 - \alpha$$

- 因此, 为求参数θ的置信区间，只需找出检验$(*)$对应的接受域，反之，参数θ的置信区间的补集就是$(*)$的拒绝域

8.5.1置信区间和假设检验之间的关系

- 类似地， 置信水平为$1-\alpha$的单侧置信区间$[\underline{\theta},\infty)$（或者$(-\infty,\ \overline{\theta}]$）与显著性水平为$\alpha$的右（或者左）边检验问题$H_0:\theta \le \theta_0 \leftrightarrow H_1:\theta > \theta_0$(或者$H_0:\theta \ge \theta_0 \leftrightarrow H_1:\theta < \theta_0$)也有类似的对应关系.

Thank you!