Ch 8. 假设检验

钟友良

zhongyl@scut.edu.cn

Outline

8.0 前言

8.1 假设检验与两类错误

8.2 正态总体参数的假设检验

8.3 非正态总体均值的假设检验

8.4 非参数假设检验

Outline

8.0 前言

8.1 假设检验与两类错误

8.2 正态总体参数的假设检验

8.3 非正态总体均值的假设检验

8.4 非参数假设检验

Outline

8.0 前言

8.1 假设检验与两类错误

8.2 正态总体参数的假设检验

8.3 非正态总体均值的假设检验

8.4 非参数假设检验

什么是假设检验

recall: 最大似然估计: 发生 (观察到A) ⟶ 最大概率 (P(A)最大).

基本思想: 小概率事件在 一次试验 中几乎不会发生。如果在某 假设 下一个具有小概率的事件出现在试验结果中, 那么原假设不成立。

Fx=? ▶ 非参数假设检验 是指对总体分布提出假设, 再根据样本以及观测结果进行推断。

Fx已知 ▶ 参数假设检验 是指总体的分布类型已知的前提下对未知参数进行假设, 然后进行推断。

$F_X(\cdot;\theta)$ $\theta=?$

观察 A $\Rightarrow$ P[A] 比较大.

观察不到A $\Leftarrow$ P[A] 比较小

假设检验的基本元素

- 原假设: 对总体 参数 的基本假设，统计推断将基于这个基本假设。
- 备选假设: 与原假设不相容的假设，若一次观测值拒绝原假设，那么接受备选假设。
- 用于推断的统计量: 包含原假设中的已知参数以及样本的函数 Y，有确定 (与总体未知参数无关) 的分布。
- 拒绝域: 若一次观测值落在拒绝域中，那么拒绝原假设，接受备选假设。

假设检验的方法

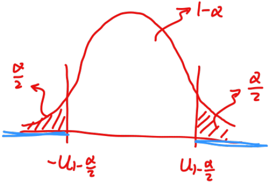
设总体 $X \sim N(\mu, \sigma_0^2)$，$\sigma_0^2$ 已知，$(X_1, X_2, \cdots, X_n)$ 是其样本，试对假设检验问题：

$$H_0(\text{原假设}) : \mu = \mu_0,\ H_1(\text{备选假设}) : \mu \neq \mu_0$$

做检验。
如果假设成立，即 $\mu = \mu_0$，则

$\overline{X} \sim N(\mu_0, \frac{1}{n}\sigma_0^2)$

$$\frac{\overline{X} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} \sim N(0, 1)$$



即

$$P\left(\left|\frac{\overline{X} - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}}\right| > u_{1-\frac{\alpha}{2}}\right) = \alpha,$$

假设检验的方法

当 α 很小的时候，事件

$$\left\{ \omega : \left| \frac{\overline{X} - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} \right| \geq u_{1-\frac{\alpha}{2}} \right\}$$

在一次试验中几乎不会发生。

若一次观测的结果 $\bar{x} = \frac{1}{n}\sum_i x_i$ 满足

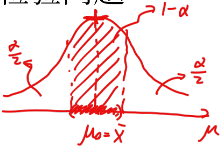
$$\left| \frac{\bar{x} - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} \right| > u_{1-\frac{\alpha}{2}},$$

我们就应该拒绝原假设 H_0，否则就可以接受原假设。

假设检验的方法

假设 $X \sim F_X(\cdot, \theta)$，参数的定义域为 Θ，假设检验问题

$$H_0 : \theta \in \Theta_0, \quad H_1 : \theta \in \Theta - \Theta_0$$



的统计推断遵循下面的几步：

- 假设 H_0 成立，我们将获得 θ 的经验值 θ_0。

  $X_i \sim F_X(\cdot\,; \theta_0)$　$Y(X_1,\dots,X_n)$

- 用样本 $X_1, X_2, \cdots, X_n$ 和 θ_0 构造一个新的随机变量 $Y(X_1, X_2, \cdots, X_n)$，Y 的分布已知。(Y 的分布参考表 7.1)

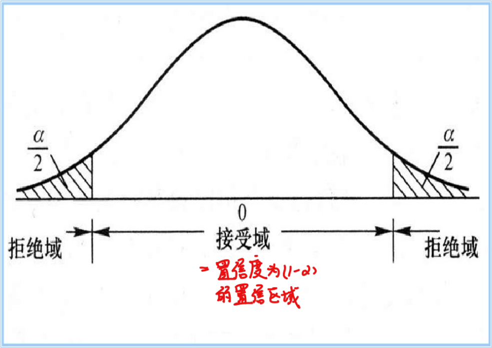
  ↳ 四类重要分布 $N(0,1)$, $\chi^2(n)$, $t(n)$, $F(m,n)$

- 确定 k_1, k_2 使得 $P(k_1 \le Y \le k_2) = 1 - \alpha$。确定 拒绝域 为 $\Omega_R = \{(x_1, x_2, \cdots, x_n) : Y(x_1, x_2, \cdots, x_n) \in (-\infty, k_1) \cup (k_1, \infty)\}$。

- 进行一次观测，得到样本的一组观测值 $(x_1, x_2, \cdots, x_n)$，若 $(x_1, x_2, \cdots, x_n) \in \Omega_R$ 则拒绝原假设 H_0，否则接受 H_0。

拒绝域和接受域

我们把不满足假设 H_0 中的点的集合称为该假设检验问题的 拒绝域， 把拒绝域的余集称为假设检验的 接受域 .



假设检验的两类错误

假设检验总是在原假设成立的前提下，构造 小概率事件 ，并且由一次的抽样观测值是否使得该小概率件发生来决定是否要拒绝原假设。
但判断总是有失误的:

假设检验的两类错误

- 第一类错误: 假设原假设 H_0 是正确的。观测值有 α 的概率落在拒绝域内，即我们拒绝 H_0 的概率是 α。这种错误被称为 第一类错误 ，或者 拒真错误 。

$$\alpha = P(\text{拒绝}H_0 \,\middle|\, H_0\text{真})$$

假设检验的两类错误

- 第二类错误: 假设原假设是错误的，即 H_1 正确。如果一次观测的结果落在接受域中，那么我们就接受 H_0 而否定 H_1。这是 第二类错误 ，或者 受伪错误 ，我们把这种错误的概率记为 β，即

$$\beta = P(\text{接受}H_0 \,|\, H_0\text{伪})$$

事实 \ 观察	样本落在 Ω_R $\Rightarrow$ 拒绝 H_0	…… $\Omega\setminus\Omega_R$ $\Rightarrow$ 接受 H_0
H_0真	$P(\Omega_R \| H_0) = \alpha$	✓
H_0伪	✓	$\beta = P[\Omega\setminus\Omega_R \| H_1] = 1 - P[\Omega_R \| H_1]$

两类错误的性质

- $\alpha + \beta \neq 1$，因为这两个条件概率的条件不一样。
- α 和 β 的变化趋势不一样，只能相反。
- 有一种检验方法可以控制犯第一类错误的概率，这种检验方法称为 显著性检验 。我们使用的就是显著性检验。显著性检验是在原假设条件下，若某个小概率事件发生了，则认为原假设显著的不真。第一类错误的概率 α 被称为 显著性水平 。

只关注 α

不关注 β.

两类错误图解

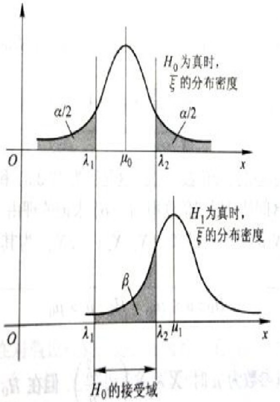


图 8.1　假设检验的两类错误示意图.

Outline

8.0 前言

8.1 假设检验与两类错误

8.2 正态总体参数的假设检验

8.3 非正态总体均值的假设检验

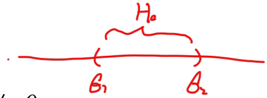
8.4 非参数假设检验

单个正态总体的双侧假设检验

假设的提法一般是

$$H_0:\theta\in\Theta_0,\quad H_1:\theta\in\Theta-\Theta_0$$

双侧假设: H_1 位于 H_0 的两侧。
判断下面的假设是否为双侧假设



$$\begin{aligned}
H_0:\theta=\theta_0,&\quad H_1:\theta\neq\theta_0\\
H_0:\theta=\theta_0,&\quad H_1:\theta<\theta_0\\
H_0:\theta=\theta_0,&\quad H_1:\theta>\theta_0\\
H_0:\theta\leq\theta_0,&\quad H_1:\theta>\theta_0
\end{aligned}$$

正态分布均值的双侧假设

设总体 $X \sim N(\mu, \sigma^2)$，$(X_1, X_2, \cdots, X_n)$ 为容量是 n 的样本。
总体 σ_0 已知，对总体均值 μ 的双侧假设检验

- 提出假设:

$$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0$$

- 定义统计量

$$Y = \frac{\overline{X} - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} \sim N(0, 1)$$

正态分布均值的双侧假设

- 确定 Y 的 $\frac{\alpha}{2}$ 和 $1-\frac{\alpha}{2}$ 分位点。对于正态分布，

$$P(|Y| \leq u_{1-\frac{\alpha}{2}}) = 1-\alpha$$

- 给出假设检验的拒绝域

$$\Omega_R = \{(x_1, x_2, \cdots, x_n) : \left|\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}\right| > u_{1-\frac{\alpha}{2}}\}$$

- 观测并判断：进行一次观测，观测值为 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$，若 $\boldsymbol{x} \in \Omega_R$，否定原假设 H_0，否则，接受 H_0。

正态分布均值的双侧检验

总体 σ 未知，总体均值的双侧假设检验问题

- 提出假设

$$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0$$

- 定义统计量

$$Y = \frac{\overline{X} - \mu_0}{\sqrt{\frac{S_n^{*2}}{n}}} \sim t(n-1)$$

$$\frac{\dfrac{\overline{X}-\mu_0}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)}{\sqrt{\dfrac{\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2 \sim \chi^2(n-1)}{n-1}}}$$

正态分布均值的双侧检验

- 确定 Y 的 $\frac{\alpha}{2}$ 和 $1-\frac{\alpha}{2}$ 分位点 $t_{1-\frac{\alpha}{2}}$。对于 t 分布，
$$P(|Y| \leq t_{1-\frac{\alpha}{2}}(n-1)) = 1-\alpha$$
- 给出假设检验的拒绝域:
$$\Omega_R = \left\{(x_1, x_2, \cdots, x_n) : \left|\frac{\bar{x}-\mu_0}{S_n^*/\sqrt{n}}\right| > t_{1-\frac{\alpha}{2}}(n-1)\right\}$$
- 观测并判断: 进行一次观测，观测值为 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$，若 $\boldsymbol{x} \in \Omega_R$，否定原假设 H_0，否则，接受 H_0。

正态总体方差的假设检验问题

设总体 $X \sim N(\mu, \sigma)$，$(X_1, X_2, \cdots, X_n)$ 为容量是 n 的样本。
总体均值 $\mu = \mu_0$ 已知，求总体 σ^2 的双侧假设检验

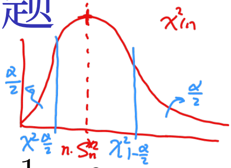
- 提出假设:

$$H_0 : \sigma^2 = \sigma_0^2,\quad H_1 : \sigma^2 \neq \sigma_0^2$$

- 定义统计量:

$$Y = \sum_{i=1}^{n} \frac{(X_i - \mu_0)^2}{\sigma_0^2} \sim \chi^2(n)$$

正态总体方差的假设检验问题



- 确定 Y 的 $\frac{\alpha}{2}$ 和 $1-\frac{\alpha}{2}$ 分位点:

$$P(\chi^2_{\frac{\alpha}{2}}(n) \le Y \le \chi^2_{1-\frac{\alpha}{2}}(n)) = 1-\alpha$$

- 给出假设检验的拒绝域:

$$\Omega_R = \{(x_1, x_2, \cdots, x_n) : \sum_{i=1}^{n} \frac{(x_i-\mu_0)^2}{\sigma_0^2} \in [0, \chi^2_{\alpha/2}(n)] \cup [\chi^2_{1-\alpha/2}(n), \infty)\}$$

- 观测并判断: 进行一次观测，获得观测值 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$，若 $\boldsymbol{x} \in \Omega_R$，否定原假设 H_0，否则，接受 H_0。

总体方差的假设检验

总体均值 μ 未知，求总体方差 σ^2 的假设检验问题

- 提出假设

$$H_0 : \sigma^2 = \sigma_0^2, \quad H_1 : \sigma^2 \neq \sigma_0^2$$

- 定义统计量

$$Y = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{\sigma_0^2} \sim \chi^2(n-1)$$

总体方差的假设检验

- 确定 Y 的 $\frac{\alpha}{2}$ 和 $1-\frac{\alpha}{2}$ 分位点

$$P(\chi^2_{\frac{\alpha}{2}}(n-1) \leq Y \leq \chi^2_{1-\frac{\alpha}{2}}(n-1)) = 1-\alpha$$

- 给出假设检验的拒绝域

$$\Omega_R = \{(x_1, x_2, \cdots, x_n): \\ \sum_{i=1}^{n} \frac{(x_i-\overline{x})^2}{\sigma_0^2} \in [0, \chi^2_{\alpha/2}(n-1)] \cup [\chi^2_{1-\alpha/2}(n-1),$$

- 观察并判断进行一次观测，获得观测值 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$，若 $\boldsymbol{x} \in \Omega_R$，否定原假设 H_0，否则，接受 H_0。

双侧假设检验问题的求解通法

- 认可原假设，定义分布已知的统计量 $Y(X_1, X_2, \cdots, X_n)$。 $\mu = \mu_0,\ \sigma^2 = \sigma_0^2$
- 给定显著性水平 α，找出 Y 的 $\frac{\alpha}{2}$ 和 $1 - \frac{\alpha}{2}$ 分位点 $\phi_{\alpha/2}$ 和 $\phi_{1-\frac{\alpha}{2}}$。
- 确定拒绝域

$$\Omega_R = \{(x_1, x_2, \cdots, x_n) : Y(x_1, x_2, \cdots, x_n) < \phi_{\alpha/2} \\ \text{or } Y(x_1, x_2, \cdots, x_n) > \phi_{1-\alpha/2}\}$$

- 进行一次观测，若观测点 $(x_1, x_2, \cdots, x_n)$ 落在 Ω_R 中，则拒绝原假设，否则接受原假设。

两独立正态总体参数双侧假设检验问题的拒绝域

两独立正态总体均值之差的假设检验：两个独立的正态总体 $X_1 \sim N(\mu_1, \sigma_1^2)$，$X_2 \sim N(\mu_2, \sigma_2^2)$，求 $\mu_1 - \mu_2$ 的假设检验问题

- 提出假设

$$H_0 : \mu_1 - \mu_2 = a, \quad H_1 : \mu_1 - \mu_2 \neq a$$

两独立正态总体参数双侧假设检验问题的拒绝域

- 定义统计量
  - σ_1^2, σ_2^2 已知,
  $$Y = \frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0, 1)$$
  - $\sigma_1^2 = \sigma_2^2$ 未知,
  $$Y = \frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{S_w\sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t(m + n - 2)$$

两独立正态总体参数双侧假设检验问题的拒绝域

- 确定拒绝域
  - σ_1^2, σ_2^2 已知,

$$\Omega_R = \left\{ (x_1, x_2, \cdots, x_m), (y_1, y_2, \cdots, y_n) : \left| \frac{(\bar{x} - \bar{y}) - a}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \right| \geq u_{1-\alpha/2} \right\}$$

  - $\sigma_1^2 = \sigma_2^2$ 未知,

$$\Omega_R = \left\{ (x_1, x_2, \cdots, x_m), (y_1, y_2, \cdots, y_n) : \left| \frac{(\bar{x} - \bar{y}) - a}{S_w \sqrt{\frac{1}{m} + \frac{1}{n}}} \right| \geq t_{1-\alpha/2}(m+n-2) \right\}$$

两独立正态总体参数双侧假设检验问题的拒绝域

- 观测并判断：进行一次观测，若观测点 $(x_1, x_2, \cdots, x_m)$，$(y_1, y_2, \cdots, y_n)$ 落在 Ω_R 中，则拒绝原假设，否则接受原假设。

Outline

8.0 前言

8.1 假设检验与两类错误

8.2 正态总体参数的假设检验

8.3 非正态总体均值的假设检验

8.4 非参数假设检验

Outline

8.0 前言

8.1 假设检验与两类错误

8.2 正态总体参数的假设检验

8.3 非正态总体均值的假设检验

8.4 非参数假设检验