

第八章 假设检验

参数检验：已知总体分布的形式，需对其中的未知参数给出假设检验

非参数检验：总体的分布形式完全未知的情况下，对总体的分布或数字特征进行假设检验

原假设(零假设) H_0 , 备择假设(对立假设) H_1 , 关于总体参数 θ 的假设:

$$\begin{cases} H_0: \theta \geq \theta_0; H_1: \theta < \theta_0 & \text{左边检验} \\ H_0: \theta \leq \theta_0; H_1: \theta > \theta_0 & \text{右边检验} \\ H_0: \theta = \theta_0; H_1: \theta \neq \theta_0 & \text{双边检验} \end{cases}$$

如果统计量 $T = T(X_1, \dots, X_n)$ 的取值大小和原假设 H_0 是否成立有密切联系, 可将之称为对应假设问题的 **检验统计量**, 对应于拒绝原假设 H_0 时, 样本值的范围称为 **拒绝域**, 记为 W , 其补集 $\bar{W}$ 称为 **接受域**.

第I类错误: 拒绝真实的原假设(弃真)

$$\alpha = P\{\text{第I类错误}\} = P_{H_0}\{\text{拒绝 } H_0\}$$

第II类错误: 接受错误的原假设(取伪)

$$\beta = P\{\text{第II类错误}\} = P_{H_1}\{\text{接受 } H_0\}$$

Neyman - Pearson 原则

首先控制犯第I类错误的概率不超过某个常数 $\alpha \in (0, 1)$, 再寻找检验, 使得犯第II类错误的概率尽可能小.

其中的常数 α 称为 **显著水平**.

常取 $\alpha = 0.01, 0.05, 0.1$ 等.

P-值: 当原假设成立时, 检验统计量取比观察到的

结果更为极端的数值的概率

P-值与显著水平 α 的关系:

若 $P \leq \alpha$, 等价于样本落在拒绝域内,

因此, 拒绝原假设,

此时称检验结果在水平 α 下是 **统计显著**的.

若 $P > \alpha$, 等价于样本不落在拒绝域内,

因此, 不拒绝(接受)原假设,

此时称检验结果在水平 α 下是 **统计不显著**.

处理假设检验问题的基本步骤

(1) 根据实际问题提出原假设和备择假设;

(2) 提出检验统计量和拒绝域的形式;

(3) 在给定的显著水平 α 下, 根据 Neyman-Pearson 原则求出拒绝域的临界值;

(3') 计算检验统计量的观测值与 P 值;

(4) 根据实际样本观测值作出判断.

(4') 根据给定的显著水平 α , 作出判断.

单个正态总体参数的假设检验

设样本 $x_1, x_2, \dots, x_n$ 来自正态总体 $N(\mu, \sigma^2)$, $\bar{x}$: 样本均值, S^2 : 方差, 显著性水平为 α

(一) 有关均值 μ 的检验

1) σ^2 已知时 —— Z 检验

双边假设问题: $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$, 取检验统计量为 $Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

在 H_0 为真时, $Z \sim N(0, 1)$

由 Neyman - Pearson 原则, 检验的拒绝域为 $W = \left\{ |Z| = \left| \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \right| \geq z_{\frac{\alpha}{2}} \right\}$

$$P_- = P_{H_0}\{|Z| \geq |z_0|\} = 2(1 - \phi(z_0))$$

当 P_- 小于 α 时, 拒绝原假设, 否则接受

左边假设问题: $H_0: \mu \geq \mu_0$, $H_1: \mu < \mu_0$, 取检验统计量为 $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

在 H_0 为真时, $Z \sim N(0,1)$

由 Neyman - Pearson 原则, 检验的拒绝域为 $W = \left\{ Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha \right\}$

$$P_- = \sup_{\mu \geq \mu_0} P_{H_0} \{ Z \leq z_0 \} = P_{H_0} \{ Z \leq z_0 \mid \mu = \mu_0 \} = \phi(z_0)$$

当 P_- 小于 α 时, 拒绝原假设, 否则接受

右边假设问题: $H_0: \mu \leq \mu_0$, $H_1: \mu > \mu_0$, 取检验统计量为 $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

在 H_0 为真时, $Z \sim N(0,1)$

由 Neyman - Pearson 原则, 检验的拒绝域为 $W = \left\{ Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha \right\}$

$$P_- = \sup_{\mu \leq \mu_0} P_{H_0} \{ Z \geq z_0 \} = P_{H_0} \{ Z \geq z_0 \mid \mu = \mu_0 \} = 1 - \phi(z_0)$$

当 P_- 小于 α 时, 拒绝原假设, 否则接受

(2) σ^2 未知时 — t 检验

双边假设问题: $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$, 用 σ 的估计量 S 代替 σ , 采用 $T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$ 作检验统计量

当原假设成立时, $\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$

由 Neyman - Pearson 原则, 可得拒绝域 $T = \frac{|\bar{X} - \mu_0|}{S/\sqrt{n}} \geq t_{\frac{\alpha}{2}}(n-1)$

$$\text{记 } t_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}, \quad P_- = P_{H_0} \{ |T| \geq |t_0| \} = 2P \{ t(n-1) \geq |t_0| \}$$

当 $P_- \leq \alpha$ 时, 拒绝原假设, 否则接受

左边假设问题: $H_0: \mu \geq \mu_0$, $H_1: \mu < \mu_0$

拒绝域: $W = \left\{ T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq -t_\alpha(n-1) \right\}$

$$P_- = \sup_{\mu \geq \mu_0} \{ T \leq t_0 \} = P \{ t(n-1) \leq t_0 \}$$

右边假设问题: $H_0: \mu \leq \mu_0$, $H_1: \mu > \mu_0$

拒绝域: $W = \left\{ T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq t_\alpha(n-1) \right\}$

$$P_- = \sup_{\mu \leq \mu_0} \{ T \geq t_0 \} = P \{ t(n-1) \geq t_0 \}$$

(二) 成对数据的 t 检验

成对样本 $(x_1, y_1), \dots, (x_n, y_n)$,

设差值 $D_i = x_i - y_i, i = 1, 2, \dots, n$

可视作来自正态总体 $N(\mu_d, \sigma_d^2)$ 的样本

为比较两总体均值是否有显著差异,

可考虑假设问题

$$H_0: \mu_d = 0 \leftrightarrow \mu_d \neq 0$$

$$\text{记 } \bar{D} = \frac{1}{n} \sum_{i=1}^n D_i, S_D = \frac{1}{n-1} \sum_{i=1}^n (D_i - \bar{D})^2$$

$$\text{则检验统计量为 } T = \frac{\sqrt{n}\bar{D}}{S_D},$$

$$\text{检验的拒绝域为 } W = \{ |T| \geq t_{\alpha/2}(n-1) \},$$

$$P\text{-值为 } \text{观察值为 } t_0 = \frac{\sqrt{n}\bar{d}}{s_d}.$$

转化为单个正态总体的均值的假设检验。

$$P_- = P_{H_0} \{ |T| \geq |t_0| \} = 2P \{ t(n-1) \geq |t_0| \}$$

(三) 有关参数 σ^2 的假设检验 (设 μ 未知)

双边检验 $H_0: \sigma^2 = \sigma_0^2, H_1: \sigma^2 \neq \sigma_0^2$

此时 σ^2 的无偏估计量为样本方差 S^2 , 且有 $\frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$

由 χ^2 检验法, 拒绝域: $\frac{(n-1)S^2}{\sigma_0^2} \leq \chi_{1-\frac{\alpha}{2}}^2(n-1)$ 或 $\frac{(n-1)S^2}{\sigma_0^2} \geq \chi_{\frac{\alpha}{2}}^2(n-1)$

$$P_- = 2 \min \{ P, 1-P \}, \quad P\text{-小接大拒}$$

左边检验 $H_0: \sigma^2 \geq \sigma_0^2, H_1: \sigma^2 < \sigma_0^2$

拒绝域 $\frac{(n-1)S^2}{\sigma_0^2} \leq \chi_{1-\alpha}^2(n-1)$

$$P_- = \sup P_{H_0} \left\{ \frac{(n-1)S^2}{\sigma_0^2} \leq \frac{(n-1)S^2}{\sigma_0^2} \right\} = P \{ \chi^2(n-1) \leq \chi_0^2 \}$$

右边检验 $H_0: \sigma^2 \leq \sigma_0^2, H_1: \sigma^2 > \sigma_0^2$

拒绝域 $\frac{(n-1)S^2}{\sigma_0^2} \geq \chi_{\alpha}^2(n-1)$

$$P_- = \sup P_{H_0} \left\{ \frac{(n-1)S^2}{\sigma_0^2} \geq \frac{(n-1)S^2}{\sigma_0^2} \right\} = P \{ \chi^2(n-1) \geq \chi_0^2 \}$$

两个正态总体参数的假设检验

设 $X_1, X_2, \dots, X_{n_1}$ 是来自 $N(\mu_1, \sigma_1^2)$ 的样本,

$Y_1, Y_2, \dots, Y_{n_2}$ 是来自 $N(\mu_2, \sigma_2^2)$ 的样本,

两样本相互独立, 并记

$\bar{X}, \bar{Y}, S_1^2, S_2^2$ 分别为两样本的均值和方差。

(一) 比较两个正态总体均值的检验

双边检验 $H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$

取检验统计量为 $\bar{X} - \bar{Y}$, 此时拒绝域的形式为 $|\bar{X} - \bar{Y}| \geq c$

1. 当 σ_1^2 和 σ_2^2 已知时

$$\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

当 H_0 成立时, $\bar{X} - \bar{Y} \sim N(0, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$

$$\text{记 } Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

则检验拒绝域为 $|Z| \geq z_{\frac{\alpha}{2}}$

$$P_- = P_{H_0} \{ |Z| \geq |z_0| \} = 2(1 - \phi(|z_0|))$$

$$\text{其中 } z_0 = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

2. 当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 但未知时

参数 σ^2 的无偏估计量

$$S_w^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$\text{可取检验统计量 } T = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

在原假设成立时, $T \sim t(n_1 + n_2 - 2)$

$$\text{拒绝域: } T = \frac{|\bar{x} - \bar{y}|}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \geq t_{\frac{\alpha}{2}}(n_1 + n_2 - 2)$$

$$P_- = P_{H_0} \{ |T| \geq |t_0| \} = 2P \{ t(n_1 + n_2 - 2) \geq |t_0| \}$$

$$\text{其中 } t_0 = \frac{\bar{x} - \bar{y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

例如: 当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (未知) 时

左边检验 $H_0: \mu_1 \geq \mu_2, H_1: \mu_1 < \mu_2$

$$\text{的检验拒绝域为: } T = \frac{(\bar{X} - \bar{Y})}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \leq -t_{\alpha}(n_1 + n_2 - 2)$$

$$P_- \text{ 值为 } P_- = \sup P_{H_0} \{ T \leq t_0 \} = P \{ t(n_1 + n_2 - 2) \leq t_0 \}$$

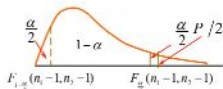
$$\text{其中 } t_0 = \frac{\bar{x} - \bar{y}}{s_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

(二) 比较两个正态总体方差的检验

(设 μ_1, μ_2 未知)

$$H_0: \sigma_1^2 = \sigma_2^2, H_1: \sigma_1^2 \neq \sigma_2^2$$

可取检验统计量为 $F = \frac{S_1^2}{S_2^2}$



在原假设成立时, $F \sim F(n_1 - 1, n_2 - 1)$

检验拒绝域为:

$$F \leq F_{1-\alpha/2}(n_1 - 1, n_2 - 1), \text{ 或 } F \geq F_{\alpha/2}(n_1 - 1, n_2 - 1)$$

$$P_- = 2 \min \{ P(F \geq f_0), P(F \leq f_0) \}$$

$P_- \leq \alpha$, 拒绝原假设, $P_- > \alpha$, 接受原假设. 其中 $f_0 = s_1^2 / s_2^2$.

左边检验 $H_0: \sigma_1^2 \geq \sigma_2^2, H_1: \sigma_1^2 < \sigma_2^2$

的检验拒绝域为: $F \leq F_{1-\alpha}(n_1 - 1, n_2 - 1)$

$$P_- = P(F \leq f_0), \text{ 其中 } f_0 = s_1^2 / s_2^2.$$

$P_- \leq \alpha$, 拒绝原假设, $P_- > \alpha$, 接受原假设.

右边检验 $H_0: \sigma_1^2 \leq \sigma_2^2, H_1: \sigma_1^2 > \sigma_2^2$

的检验拒绝域为: $F \geq F_{\alpha}(n_1 - 1, n_2 - 1)$

$$P_- = P(F \geq f_0), \text{ 其中 } f_0 = s_1^2 / s_2^2.$$

$P_- \leq \alpha$, 拒绝原假设, $P_- > \alpha$, 接受原假设.

假设检验与区间估计

作区间估计时，对参数没有先验的认识，但确定参数是固定不变的，只是未知，所以区间估计的目的是：**根据样本对参数进行估计**；

作假设检验时，对参数有一个先验的认识（例如 $\mu=\mu_0$ ），但由于某种情形的出现（如工艺改良等），猜测真实参数值可能发生了变化，所以假设检验的目的是：**根据样本确认参数是否真的发生了改变**。

但置信区间与假设检验的拒绝域之间又有密切的关系。

假设检验问题 $H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$,

显著性水平为 α 的检验拒绝域为

$$W = \left\{ \frac{|\bar{X} - \mu_0|}{\sigma / \sqrt{n}} \geq z_{\alpha/2} \right\},$$

接受域为

$$\begin{aligned} \bar{W} &= \left\{ \frac{|\bar{X} - \mu_0|}{\sigma / \sqrt{n}} < z_{\alpha/2} \right\} \\ &= \left\{ \bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} < \mu_0 < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right\} \end{aligned}$$

一般地，若假设检验问题 $H_0: \theta = \theta_0$ $H_1: \theta \neq \theta_0$ 的显著水平为 α 的接受域能等价地写成

$$\hat{\theta}_L < \theta_0 < \hat{\theta}_U$$

那么 $(\hat{\theta}_L, \hat{\theta}_U)$ 是参数 θ 的置信水平为 $1-\alpha$ 的置信区间。

考虑单个正态总体方差已知时有关均值的统计推断。

设 $X_1, X_2, \dots, X_n$ 是来自总体 $N(\mu, \sigma^2)$ 样本， σ^2 已知。

μ 的置信水平为 $1-\alpha$ 的置信区间为

$$\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$$

将接受域中的 μ_0 改写成 μ 时，所得结果正好是参数 μ 置信水平为 $1-\alpha$ 的置信区间。

反之，若 $(\hat{\theta}_L, \hat{\theta}_U)$ 是 θ 的置信水平为 $1-\alpha$ 的置信区间，则当 $\theta_0 \in (\hat{\theta}_L, \hat{\theta}_U)$ 时，接受双边检验 $H_0: \theta = \theta_0$, $H_1: \theta \neq \theta_0$ 中的原假设 H_0 ，且检验的拒绝域为 $\theta_0 \leq \hat{\theta}_L$ 或 $\theta_0 \geq \hat{\theta}_U$ 。