第八章: 假设检验

8.1 基本概念

8.2 正态总体均值的检验

8.3 正态总体方差的检验

8.1 基本概念

假设检验：引例

例 1.1

某工厂生产的一大批产品要卖给商店. 按规定次品率 p 不得超过 0.01, 今在其中抽取 *100* 件, 经检验有 *3* 件次品, 问这批产品可否出厂? 对这个问题存在两种可能性:

$$\text{甲}: 0 < p \leq 0.01 \sim \text{乙}: 0.01 < p < 1.$$

要通过这批产品抽样来决定甲、乙哪个更可能成立?

注 1.1

由于随机性的存在, 一次抽样估计次品率 p 存在误差. 进行 *1000* 次抽样, 每次抽 *100* 件都没有任何次品, 不会对“次品率小于 *0.01*“产生怀疑. 但进行一次抽样, 次品率为 $p = 0.03$, 这个样本提供的“证据”是否足以推翻 $p \leq 0.01$ 的假设?

假设检验：引例

要解决此问题需引入一个假设

$$H_0 : 0 < p \leq 0.01,$$

称为**零假设**(null hypothesis) 或**原假设**. 另一个可能是

$$H_1 : 0.01 < p < 1,$$

称为**对立假设**或**备择假设**(alternative hypothesis).

假设检验: 基本原理

例 1.2 (女士品茶)

一位女士 (*Muriel Bristol*) 坚称自己能够分辨一杯茶是先放奶 (*M*) 还是先放茶 (*T*). 为了检验这位女士是否真的有品鉴能力, 罗纳德·艾尔默·费歇尔 (*R. A. Fisher*) 首先引入了一个假设,

$$H_0: \text{该女士没有品鉴能力}.$$

如何利用 H_0 检验该女士是否真的有品鉴能力?

假设检验: 基本原理

解: **(1) 在 H_0 下, 构造一个小概率事件**.

- 先取 8 个完全相同的杯子随机排成一列, 4 杯先茶后奶 (TM), 4 杯先奶后茶 (MT). 让这位女士品鉴.
- 假设 H_0 成立, 即认为该女士没有品鉴能力.
- 设随机变量 X 表示该女士猜中 TM 的杯数, 则 X 服从超几何分布. 有

$$P(X=4)=\frac{\binom{4}{4}\binom{4}{0}}{\binom{8}{4}}=\frac{1}{70}\approx 0.0143,$$

可认为此事件为**小概率事件**.

(2) 如果该女士猜对了全部 TM 的 4 杯, 小概率事件发生 根据**反证法**推翻 H_0, 即可以认为此女士有品鉴能力. $\square$

假设检验: 基本原理

注 1.2

到底多小才算小概率呢? *Fisher* 当时给出了 *0.05* 这个值, 因为 $0.0143 < 0.05$, 所以 $X = 4$ 这个事件可认为是小概率事件. 但注意, *0.05* 不是黄金标准, 具体情况还需具体分析. 尤其是涉及人的生命安全时, *0.05* 就显得过大了.

定理 1.1 (假设检验的基本原理)

解决假设检验问题需要用到反证法.

- 设原假设 H_0 成立, 构造小概率事件, 进行一次试验.
- 如果小概率事件发生了, 则违背了"一次试验中小概率事件不发生"(小概率原理), 拒绝原假设.

例 1.3

设 $X_1, \cdots, X_n$ 为从正态总体 $N(\mu, \sigma^2)$ 中抽取的随机样本, 其中 σ^2 已知. 考虑检验问题

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

此处 μ_0 为给定的常数.

解: (1) 求 μ 的一个估计量. 一般地, $\overline{X}$ 是 μ 的一个优良估计.

- 若 $|\overline{X} - \mu_0|$ 较大, 就倾向于否定 H_0.
- 若 $|\overline{X} - \mu_0|$ 较小, 就认为抽样结果与 H_0 接近, 因而倾向于不能拒绝 H_0.

具体地, 确定一个数 A, 由 $X_1, \cdots, X_n$ 算出样本均值 $\overline{X}$.

- 当 $|\overline{X} - \mu_0| > A$ 时就否定 H_0.
- 当 $|\overline{X} - \mu_0| \leq A$ 时就接受 H_0.

(2) 称

$$D = \{\boldsymbol{X} = (X_1, \cdots, X_n) : |\overline{X} - \mu_0| > A\}$$

为**拒绝域**. 只要 A 确定, 检验否定域也就确定了. A 称为**临界值** (critical value). 要定下 A 的值需确定**检验统计量** $\overline{X}$ 的分布. 检验可视为如下法则

$$T : \begin{cases} 当\ |\overline{X} - \mu_0| > A\ 时, & 拒绝\ H_0; \\ 当\ |\overline{X} - \mu_0| \leq A\ 时, & 不能拒绝\ H_0. \end{cases}$$

假设检验: 两类错误

注 1.3

假设检验问题可看作一个二分类问题. 由于样本的随机性, 只能以一定的概率保证分类的可靠性. 在假设检验问题中常常存在两类错误.

定义 1.1 (第一类错误)

零假设 H_0 本来是对的, 由于样本的随机性, 观察值落入否定域 D, 错误地将 H_0 否定了, 即弃真错误, 称其为第一类错误, 记为 α.

定义 1.2 (第二类错误)

零假设 H_0 本来不对, 由于样本的随机性, 观察值却没有落入否定域 D, 没有拒绝 H_0, 即取伪错误, 称其为第二类错误, 记为 γ.

假设检验: 两类错误

例 1.4

假设 $X \sim N(\mu, 1)$ 服从正态分布, 其中 μ 未知, 抽出 16 个样本进行如下检验:

$$H_0 : \mu = 2 \sim H_1 : \mu \geq 2.$$

假定得到拒绝域为 $D = \{\overline{X} \geq 2.41\}$, 计算两类错误.

解: 由两类错误定义知,

$$\begin{aligned}\alpha &= P(X \in D | H_0 \text{ 为真}) = P(\overline{X} \geq 2.41 | \mu = 2) \\ &= P\left(\frac{\overline{X} - 2}{1/\sqrt{16}} \geq \frac{2.41 - 2}{1/\sqrt{16}}\right) = 0.05.\end{aligned}$$

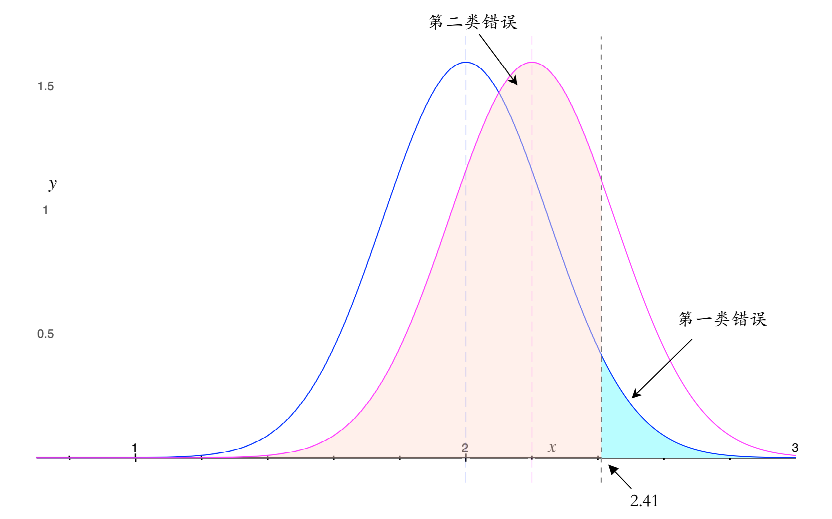
假设检验: 两类错误

第二类错误需在 H_1 成立条件下进行, 不妨假设 $\mu = 2.2$, 则

$$\begin{aligned}\gamma &= P(X \in \overline{D} | H_1 \text{ 为真}) \\ &= P(\overline{X} < 2.41 | \mu = 2.2) \\ &= P\left(\frac{\overline{X} - 2.2}{1/\sqrt{16}} \geq \frac{2.2 - 2}{1/\sqrt{16}}\right) \\ &\approx 0.80.\end{aligned}$$

下面将说明, 两类错误无法同时降低. 一般地, **在控制犯第一类错误概率的概率不超过指定数值 α ($0 < \alpha < 1$, 通常取较小的数) 的检验中, 寻找犯第二类错误尽可能小的检验**.

假设检验：两类错误



假设检验: 检验水平

定义 1.3 (检验水平)

设检验 $H_0 : \theta \in \Theta_0 \sim H_1 : \theta \in \Theta_1$. 令 $0 \le \alpha \le 1$. 如果检验犯第一类错误的概率总不超过 α, 则称 α 是该检验的显著性水平, 而检验称为显著性水平为 α 的检验.

注 1.4

习惯上把 α 取得比较小且标准化, 如 $\alpha = 0.01, 0.05$ 或 0.10. 构造 $\theta \in \Theta_0$ 时 $X \in D$ 的小概率事件. 假设 H_0 成立, 若样本落入拒绝域 D, 则小概率事件发生, 利用反证法推翻 H_0.

注 1.5

水平选得很低, 就需要更充足的理由去拒绝 H_0. 如果样本落入了否定域, 作出“拒绝 H_0”的结论就比较可靠.

假设检验: 检验水平

注 1.6

但是, 水平选得低时, 如果样本没有落入否定域, 我们只能得出"不能拒绝 H_0 "的结论, 只能表明: 在选定的水平下没有充分证据否定 H_0, 但决不意味着有充分证据说明 H_0 正确.

注 1.7

"不能拒绝 H_0"和"接受 H_0"是不同的概念. 以法庭为例, 在无罪推定 (H_0) 的条件下, 可找到充分证据推翻 H_0 判处被告人有罪, 但"没有充足证据证明被告有罪"(不能推翻 H_0), 并不意味着"被告真的无罪"(接受 H_0). 在现实中, 有罪却逃脱法律制裁的案例屡见不鲜, 如著名的辛普森案.

假设检验: 一般步骤

假设检验的一般步骤. 给定样本 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$.

(1) 根据问题的要求提出**零假设** H_0 和**备择假设** H_1;

(2) 导出**拒绝域**, 确定**检验统计量** $T(\boldsymbol{X})$, 其中**临界值** A 待定;

(3) 选择适当**显著性水平** α, 利用检验统计量的分布求出临界值 A;

(4) 由样本 $\boldsymbol{X}$ 计算出检验统计量 $T(\boldsymbol{X})$ 的具体值.

(5) 代入到拒绝域中, 与临界值相比, 作出是否拒绝 H_0 的结论.

假设检验: 发展历史

从理论上看, 假设检验是数理统计的一个分支. 它的系统发展始于 20 世纪初.

(1) K.Pearson 在 1900 年针对分类数据的检验问题, 提出了 χ^2 检验 (**拟合优度检验**);

(2) R.A.Fisher 在 20 世纪 20 年代的几篇论文创立了**显著性检验**方法. Fisher 建立了实用中常用的基于 t 分布和 F 分布的检验.

Pearson 和 Fisher 工作的不足之处是没有制定一些原则和标准对不同的检验方法进行比较和选择.

假设检验: 发展历史

(3) J.Neyman 和 E.S.Pearson 从 1928 年开始, 花了大约 10 年时间, 合作发表了一系列论文, 建立了假设检验的数学理论, **Neyman-Pearson (NP) 理论**. 该理论将统计问题的解转化为数学最优化问题.

(4) A.Wald 从 20 世纪 40 年代开始, 将 NP 理论的思想扩展到整个数理统计学领域, 在 1950 年建立了**统计决策理论**.

(5) 此外, 还需关注二战后贝叶斯学派的兴起. 贝叶斯方法的研究大致始于 20 世纪二三十年代, 到五六十年代引起人们广泛关注, 迅速崛起, 达到与频率学派分庭抗礼的程度. **贝叶斯方法已经渗透到数理统计的各个领域, 包括假设检验.**

8.2 正态总体均值的检验

单个正态总体均值的检验

设 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 为从正态总体 $N(\mu, \sigma^2)$ 中抽取的简单随机样本, 求下列三类检验问题:

$$
\begin{aligned}
&(1)\ \ H_0 : \mu = \mu_0 \sim H_1 : \mu \neq \mu_0, \\
&(2)\ \ H_0 : \mu = \mu_0 \sim H_1' : \mu > \mu_0, \\
&(3)\ \ H_0 : \mu = \mu_0 \sim H_1'' : \mu < \mu_0,
\end{aligned}
$$

其中 μ_0 和检验水平 α 给定.

- 检验问题 (1) 称为**双边检验** (two-side test).
- 检验问题 (2) 和 (3) 称为**单边检验** (one-side test).

单个正态总体均值的检验: 方差已知

对检验问题 (1), 考虑方差已知情况.

- 用直观方法构造拒绝域. 由于 $\overline{X} = \sum_{i=1}^{n} X_i/n$ 是 μ 的无偏估计. 直观上 $|\overline{X} - \mu_0|$ 越大, H_0 越不像成立.
- 拒绝域可取如下形式:

$$\{\boldsymbol{X} = (X_1, \cdots, X_n) : |\overline{X} - \mu_0| > A\},$$

  其中 A 待定.
- 当 σ^2 已知时, 在 H_0 成立时, $\overline{X} \sim N(\mu_0, \sigma^2/n)$, 将其标准化

$$U = \frac{\sqrt{n}(\overline{X} - \mu_0)}{\sigma} \sim N(0, 1).$$

单个正态总体均值的检验: 方差已知

- 将 U 作为检验统计量, 则拒绝域的等价形式可取为

  $$\{\boldsymbol{X} = (X_1, \cdots, X_n) : |U| > c\},$$

  其中 c 待定.
- 由检验水平 α,

  $$P(|U| > c | H_0) = P\left(\left|\frac{\sqrt{n}(\overline{X} - \mu_0)}{\sigma}\right| > c \middle| H_0\right) = \alpha.$$

  可得 $c = z_{\alpha/2}$.
- 因此拒绝域

  $$D = \{\boldsymbol{X} = (X_1, \cdots, X_n) : |U| > z_{\alpha/2}\}$$

  为检验问题 (1) 的水平为 α 的检验.

单个正态总体均值的检验: 方差已知

例 2.1

食品厂用自动装罐机装罐头食品, 每罐标准重量为 500g, 每天开工需检查机器的工作状况, 今抽得 *10* 罐, 测得其重量 *(*单位*: g)* 为 *495, 510, 505, 498, 503, 492, 502, 512, 497, 506.* 假定罐头重量服从正态分布 $N(\mu, \sigma^2)$, 已知 $\sigma = 6.5$, 问机器是否工作正常 *(*取检验水平 $\alpha = 0.05$*)?*

解: 检验问题为

$$H_0 : \mu = 500 \sim H_1 : \mu \neq 500.$$

由于 σ^2 已知, 故检验的拒绝域为

$$D = \left\{ (X_1, \cdots, X_n) : \left| \frac{\sqrt{n}(\overline{X} - \mu_0)}{\sigma} \right| > u_{\alpha/2} \right\},$$

单个正态总体均值的检验: 方差已知

其中,

$$n = 10, \alpha = 0.05, \sigma = 6.5, \mu_0 = 500,$$

查表得 $z_{0.025} = 1.96$, 由样本计算得 $\overline{X} = 502$, 因此

$$\begin{aligned} |U| &= \left| \frac{\sqrt{n}(\overline{X} - \mu_0)}{\sigma} \right| \\ &= \left| \frac{\sqrt{10}(502 - 500)}{6.5} \right| \\ &= 0.973 < 1.96, \end{aligned}$$

没有落在拒绝域, 无法拒绝 H_0. 所以在显著性水平 $\alpha = 0.05$ 下认为没有足够理由说明自动装罐机工作不正常. $\square$

单个正态总体均值的检验: 方差未知

对检验问题 (1), 考虑 σ^2 未知情况.

- 假定 $\mu = \mu_0$ 时, 令

$$T = \frac{\sqrt{n}(\overline{X} - \mu_0)}{S} \sim t_{n-1}.$$

  为检验统计量.

- 用与方差 σ^2 已知情形相同的方法得到检验问题 (1) 的检验水平为 α 的检验拒绝域为

$$D' = \{\boldsymbol{X} = (X_1, \cdots, X_n) : |T| > t_{n-1}(\alpha/2)\}.$$

对检验问题 (2), 考虑 σ^2 已知情况.

- 从直观上看检验问题 (2) 的拒绝域为
$$\{\sqrt{n}(\overline{X}-\mu_0)/\sigma > c\},$$
其中 c 待定. 取 $U=\sqrt{n}(\overline{X}-\mu_0)/\sigma$ 为检验统计量.
- 令
$$P(U>c|H_0^*)=P\left(\frac{\sqrt{n}(\overline{X}-\mu_0)}{\sigma}>c\middle|\mu=\mu_0\right)=\alpha,$$
由于 $U|\mu=\mu_0\sim N(0,1)$, 故取 $c=z_\alpha$.
- 得到拒绝域
$$D_2=\left\{(X_1,\cdots,X_n): U=\frac{\sqrt{n}(\overline{X}-\mu_0)}{\sigma}>z_\alpha\right\}.$$

单个正态总体均值的检验: 方差已知

对检验问题 (2), 考虑 σ^2 未知情况.

- 假定 $\mu = \mu_0$ 时, 令

$$T = \frac{\sqrt{n}(\overline{X} - \mu_0)}{S} \sim t_{n-1}.$$

为检验统计量.

- 用与方差 σ^2 已知情形相同的方法得到检验问题 (2) 的检验水平为 α 的检验拒绝域为

$$D_2' = \{\boldsymbol{X} = (X_1, \cdots, X_n) : |T| > t_{n-1}(\alpha)\}.$$

对检验问题 (3), 情况与问题 (2) 完全类似.

单个正态总体均值的检验: 方差未知

例 2.2

某砖厂所生产的地砖的抗断强度 X 服从正态分布 $\mathcal{N}(\mu, \sigma^2)$, 今从该厂生产的地砖中随机抽取 6 块测得抗断强度 (单位: kg/cm^2) 如下: 32.56, 29.66, 31.64, 30.00, 31.87, 31.03. 问这一批地砖的平均抗断强度可否认为不低于 $32.50kg/cm^2$ (取检验水平 $\alpha = 0.05$)?

解: 由样本计算 $n = 6, \overline{X} = 31.13, S = 1.123$. 要验证平均抗断强度低于 $32.50kg/cm^2$, 需推翻 $H_0 : \mu \geq 32.50 (\mu = 32.50)$. 检验问题为

$$H_0 : \mu \geq 32.50 \sim H_1 : \mu < 32.50.$$

单个正态总体均值的检验: 方差未知

其中, σ^2 未知, 采用 t 检验法, 拒绝域为

$$D = \left\{ (X_1, \cdots, X_n) : \frac{\sqrt{n}(\overline{X} - \mu_0)}{S} < -t_{n-1}(\alpha) \right\},$$

因为 $\alpha = 0.05$, $t_5(0.05) = 2.015$, 因此有

$$T = \frac{\sqrt{n}(\overline{X} - \mu_0)}{S} = \frac{\sqrt{6}(31.13 - 32.50)}{1.123} = -2.99 < -2.015,$$

落在拒绝域, 拒绝 H_0, 即认为地砖强度达不到 $32.50kg/cm^2$. $\square$

单个正态总体均值的检验: 总结

	H_0	H_1	检验量及其分布	拒绝域
σ^2 已知	$\mu = \mu_0$	$\mu \neq \mu_0$	$U = \sqrt{n}(\overline{X} - \mu_0)/\sigma$	$\lvert U \rvert > z_{\alpha/2}$
	$\mu = \mu_0$	$\mu > \mu_0$	$U \vert \mu = \mu_0 \sim N(0,1)$	$U > z_\alpha$
	$\mu = \mu_0$	$\mu < \mu_0$		$U < -z_\alpha$
σ^2 未知	$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \sqrt{n}(\overline{X} - \mu_0)/S$	$\lvert T \rvert > t_{n-1}(\alpha/2)$
	$\mu = \mu_0$	$\mu > \mu_0$	$T \vert \mu = \mu_0 \sim t_{n-1}$	$T > t_{n-1}(\alpha)$
	$\mu = \mu_0$	$\mu < \mu_0$		$T < -t_{n-1}(\alpha)$

两个正态总体均值差的检验

- 设 $\boldsymbol{X} = (X_1, \cdots, X_m)$ 为正态总体 $N(\mu_1, \sigma_1^2)$ 中抽取的简单随机样本.
- 设 $\boldsymbol{Y} = (Y_1, \cdots, Y_n)$ 为自正态总体 $N(\mu_2, \sigma_2^2)$ 中抽取的简单随机样本.
- 且样本 $X_1, \cdots, X_m$ 和 $Y_1, \cdots, Y_m$ 相互独立.

讨论检验问题 (4):

$$H_0 : \mu_2 - \mu_1 = 0 \sim \mu_2 - \mu_1 \neq 0,$$

其中检验水平 α 给定.

两个正态总体均值差的检验: 方差已知

对检验问题 (4), 考虑 σ_1^2 和 σ_2^2 已知.

- 由于 $\bar{Y}-\bar{X}$ 是 $\mu_2-\mu_1$ 的一个良好估计, 直观上 $|\bar{Y}-\bar{X}|$ 越大, H_0 越倾向于拒绝, 故拒绝域有如下形式:

$$\{(\boldsymbol{X},\boldsymbol{Y}):|\bar{Y}-\bar{X}-\mu_0|>A\},$$

  其中 A 待定.

- 当 $\mu_2-\mu_1=\mu_0$ 时,

$$U=\frac{\bar{Y}-\bar{X}-\mu_0}{\sqrt{\sigma_1^2/m+\sigma_2^2/n}}\sim N(0,1).$$

- 故可取检验统计量 $U=(\bar{Y}-\bar{X}-\mu_0)/\sqrt{\sigma_1^2/m+\sigma_2^2/n}$.

两个正态总体均值差的检验: 方差已知

- 其拒绝域形式为

$$\{(X_1, \cdots, X_m; Y_1, \cdots, Y_n) : |U| > c\}.$$

  其中 c 待定.
- 记 $\theta = \mu_2 - \mu_1$, 为确定 c, 令

$$\begin{aligned}\alpha &= P_\theta(|U| > c|H_0) \\ &= P\left(\left|\frac{\bar{Y} - \bar{X} - \mu_0}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}}\right| > cH_0\right) = 2 - 2\Phi(c),\end{aligned}$$

  由此确定临界值 $c = z_{\alpha/2}$.
- 因此检验问题水平为 α 的拒绝域为

$$D = \{(X_1, \cdots, X_m; Y_1, \cdots, Y_n) : |U| > z_{\alpha/2}\}.$$

两个正态总体均值差的检验: 方差未知

对检验问题 (4), 考虑 $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知.

- 若 $\sigma_1^2=\sigma_2^2=\sigma^2$ 已知时, 利用两样本 U 检验的统计量

$$U=\frac{\bar{Y}-\bar{X}}{\sqrt{\sigma_1^2/m+\sigma_2^2/n}}=\frac{\bar{Y}-\bar{X}}{\sigma}\sqrt{\frac{mn}{m+n}}.$$

- 若 $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知时, 上述表达式中 σ^2 用 S_ω^2 估计,

$$\begin{aligned}S_\omega^2&=\frac{1}{n+m-2}[(m-1)S_1^2+(n-1)S_2^2]\\&=\frac{1}{n+m-2}\left[\sum_{i=1}^{m}(X_i-\bar{X})^2+\sum_{j=1}^{n}(Y_j-\bar{Y})^2\right]\end{aligned}$$

两个正态总体均值差的检验: 方差未知

- 其中

$$S_1^2 = \frac{1}{m-1}\sum_{i=1}^{m}(X_i - \bar{X})^2, \quad S_2^2 = \frac{1}{n-1}\sum_{j=1}^{n}(Y_j - \bar{Y})^2$$

分别为样本 $X_1, \cdots, X_m$ 和 $Y_1, \cdots, Y_n$ 的样本方差.

- 利用 S_ω^2 代替 σ^2 可知

$$T_\omega = \frac{\bar{Y} - \bar{X} - \mu_0}{S_\omega}\sqrt{\frac{mn}{m+n}} \sim t_{n+m-2}.$$

故取 T_ω 作为检验统计量.

两个正态总体均值差的检验: 方差未知

例 2.3

为研究正常成年男女血液红细胞平均数的差别, 检验某地正常成年男子 156 人, 女子 74 人, 计算男女红细胞的平均数和标准差分别为

$$\text{男}:\ \overline{X} = 465.13, \quad S_1 = 54.80,$$
$$\text{女}:\ \overline{Y} = 422.16, \quad S_2 = 49.20.$$

假定正常成年男女红细胞分别服从正态分布, 且方差相同. 检验正常成年人红细胞数是否与性别有关 ($\alpha = 0.01$).

解: 设 $X_1, \cdots, X_m \sim N(\mu_1, \sigma^2)$; $Y_1, \cdots, Y_n \sim N(\mu_2, \sigma^2)$, 且假定两组样本独立.

两个正态总体均值差的检验: 方差未知

检验问题为

$$H_0 : \mu_2 - \mu_1 = 0 \sim H_1 : \mu_2 - \mu_1 \neq 0.$$

检验的拒绝域为

$$D = \left\{(X_1, \cdots, X_m; Y_1, \cdots, Y_n) : \left|\frac{\bar{Y} - \bar{X}}{S_\omega}\sqrt{\frac{mn}{m+n}}\right| > t_{n+m-2}(\alpha/2)\right\},$$

其中

$$m = 156,\ n = 74,\ \bar{X} = 465.13,\ \bar{Y} = 422.16,\ S_1 = 54.80,\ S_2 = 49.20,$$

$$S_\omega^2 = \frac{1}{n+m-2}[(m-1)S_1^2 + (n-1)S_2^2] = 2816.6,\ S_\omega = 53.07.$$

查表得 $t_{228}(0.005) = 2.576$. 因此

$$
\begin{aligned}
|T_\omega| &= \left| \frac{\bar{Y} - \bar{X}}{S_\omega} \sqrt{\frac{mn}{m+n}} \right| \\
&= \left| \frac{422.16 - 465.13}{53.07} \sqrt{\frac{156 \times 74}{156 + 74}} \right| \\
&= 5.74 > 2.575,
\end{aligned}
$$

落入拒绝域, 故拒绝 H_0, 即认为正常成年人的红细胞数与性别有关. □

两个正态总体均值差的检验: 成对样本检验

前面的讨论假定了两个正态总体的样本是相互独立的. 但在实际问题中, **可能两个正态总体的样本是来自同一个总体上的重复观察, 它们是成对出现的, 而且是相关的.**

- 例如, 为了考察一种安眠药的效果, 记录了 n 个失眠患者服药前每晚睡眠时间 $X_1, \cdots, X_n$ 和服用此安眠药后每晚的睡眠时间. 它们是相关的.
- 另一方面, $X_1, \cdots, X_n$ 是 n 个不同失眠患者的睡眠时间, 由于个人体质诸方面的条件不同, 这 n 个观察值不能认为是来自同一个正态总体的样本.

这样的数据称为**成对数据**.

两个正态总体均值差的检验：成对样本检验

这样的数据模型利用两样本 t 检验就不合适.

- 因为 X_i 和 Y_i 是同在第 i 个患者身上观察到的夜晚睡眠时间，所以 $Z_i = Y_i - X_i$ 就消除了人的体质诸方面的差异，仅剩下安眠药的效果.
- 若安眠药无效，Z_i 的差异仅由随机误差引起，且可认为服从正态分布 $N(0, \sigma^2)$. 假定 $Z_1, \cdots, Z_n$ 来自 $N(\mu, \sigma^2)$ 中抽取的简单随机样本，μ 是安眠药的平均效果.

两个正态总体均值差的检验: 成对样本检验

- 安眠药是否有效, 就归结为如下假设:

$$H_0: \mu = 0 \sim H_1: \mu \neq 0.$$

- 利用单个正态总体均值的 t 检验方法, 其拒绝域为

$$D = \{(Z_1, \cdots, Z_n) : |T_z| > t_{n-1}(\alpha/2)\},$$

此处 α 为检验水平, $T_z = \sqrt{n}\bar{Z}/S_z$ 为检验统计量, 其中 $\bar{Z}$ 和 S_z^2 分别为 $Z_1, \cdots, Z_n$ 的样本均值和样本方差.

两个正态总体均值差的检验: 成对样本检验

例 2.4

今有两台测量材料中某种金属含量的光谱仪 A 和 B, 为鉴定它们的质量有无显著差异, 对金属含量不同的 9 件材料样品进行测量, 得到 9 对观察值为

$$A: 0.20,\ 0.30,\ 0.40,\ 0.50,\ 0.60,\ 0.70,\ 0.80,\ 0.90,\ 1.00.$$

$$B: 0.10,\ 0.21,\ 0.52,\ 0.32,\ 0.78,\ 0.59,\ 0.68,\ 0.77,\ 0.89.$$

问根据实验结果, 在 $\alpha = 0.01$ 下, 能否判断这两台光谱仪的质量有无显著性差异?

解: 将光谱仪 A 和 B 对 9 件样品的测定值记为 $X_1, \cdots, X_9$, 和 $Y_1, \cdots, Y_9$. 由于 9 件样品金属含量不同, 所以 $X_1, \cdots, X_9$ 不能看成来自同一个总体.

两个正态总体均值差的检验: 成对样本检验

每个对子 (X_i, Y_i) 中 X_i 与 Y_i 不独立. 需利用成对比较. 记

$$Z_i = Y_i - X_i, \quad i = 1, 2, \cdots, 9.$$

若这两台光谱仪质量一样, 测量得到的每对数据的差异仅由随机误差引起, 且服从正态分布 $N(0, \sigma^2)$. 要检验

$$H_0 : \mu = 0 \sim H_1 : \mu \neq 0.$$

其拒绝域为

$$\{(Z_1, \cdots, Z_n) : |T_z| > t_{n-1}(\alpha/2)\},$$

两个正态总体均值差的检验: 成对样本检验

计算得

$$\bar{Z} = \frac{1}{9}\sum_{i=1}^{9} Z_i = 0.06, \quad S_z^2 = \frac{1}{8}\sum_{i=1}^{9}(Z_i - \bar{Z})^2 = 0.01505,$$
$$S_z = 0.12268,$$

查表得 $t_{n-1}(\alpha/2) = t_8(0.005) = 3.3554$. 由于

$$|T_z| = \left|\frac{\sqrt{n}\bar{Z}}{S_z}\right| = \frac{3 \times 0.06}{0.12268} = 1.47 < 3.3554,$$

没有落在拒绝域, 无法拒绝 H_0, 即没有足够证据显示两台仪器有显著差异. □

8.3 正态总体方差的检验

单个正态总体方差的检验

设 $X_1, \cdots, X_n$ 为正态总体 $\mathcal{N}(\mu, \sigma^2)$ 中抽取的简单随机样本,讨论三类检验问题:

$$
\begin{aligned}
&(5)\ \ H_0 : \sigma^2 = \sigma_0^2 \sim H_1 : \sigma^2 \neq \sigma_0^2, \\
&(6)\ \ H_0 : \sigma^2 \leq \sigma_0^2 \sim H_1 : \sigma^2 > \sigma_0^2, \\
&(7)\ \ H_0 : \sigma^2 \geq \sigma_0^2 \sim H_1 : \sigma^2 < \sigma_0^2,
\end{aligned}
$$

其中 σ_0^2 的检验水平 α 给定. **在实际情况中, 均值 μ 常常是未知的.**

单个正态总体方差的检验: 均值未知

对检验问题 (5), 考虑均值 μ 未知.

- 均值 μ 未知时 $S^2 = \sum_{i=1}^n (X_i - \overline{X})^2/(n-1)$ 是 σ^2 的一个无偏估计, 直观上看当 S^2/σ_0^2 太小或太大时, H_0 倾向于不成立.
- 检验的拒绝域可取如下形式:

$$D = \{(X_1, \cdots, X_n) : S^2/\sigma_0^2 < A_1 \text{ 或 } S^2/\sigma_0^2 > A_2\},$$

  其中 A_1, A_2 待定.
- 在给定 $\sigma^2 = \sigma_0^2$ 的条件下,

$$\frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}.$$

  故取检验统计量 $\chi^2 = (n-1)S^2/\sigma_0^2$.

单个正态总体方差的检验: 均值未知

- 拒绝域的等价形式可取为

$$D = \left\{ (X_1, \cdots, X_n) : \frac{(n-1)S^2}{\sigma_0^2} < c_1 \text{ 或 } \frac{(n-1)S^2}{\sigma_0^2} > c_2 | H_0 \right\},$$

- 记 $\theta = (\mu, \sigma^2)$, 为了确定 c_1, c_2, 令

$$\alpha = P_\theta \left(\frac{(n-1)S^2}{\sigma_0^2} < c_1 \text{ 或 } \frac{(n-1)S^2}{\sigma_0^2} > c_2 | H_0 \right),$$

- 简单地, 令

$$P_\theta \left(\frac{(n-1)S^2}{\sigma_0^2} < c_1 \middle| H_0 \right) = \frac{\alpha}{2},$$
$$P_\theta \left(\frac{(n-1)S^2}{\sigma_0^2} > c_2 \middle| H_0 \right) = \frac{\alpha}{2}.$$

单个正态总体方差的检验: 均值未知

- 由上述两式易知临界值

$$c_1 = \chi^2_{n-1}(1-\alpha/2), c_2 = \chi^2_{n-1}(\alpha/2).$$

- 故检验问题 (5) 的水平为 α 的拒绝域为

$$\begin{aligned} D_5 = \Bigg\{(X_1, \cdots, X_n) : & \frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{n-1}\left(1 - \frac{\alpha}{2}\right) \\ & \text{或 } \frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{n-1}\left(\frac{\alpha}{2}\right)\Bigg\}. \end{aligned}$$

单个正态总体方差的检验: 均值未知

利用求检验问题 (2) 和 (3) 的方法可求得检验问题 (6) 和 (7) 的水平为 α 的拒绝域如下:

$$D_6 = \left\{(X_1, \cdots, X_n) : \frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{n-1}(\alpha)\right\},$$

$$D_7 = \left\{(X_1, \cdots, X_n) : \frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{n-1}(1-\alpha)\right\}.$$

单个正态总体方差的检验: 均值未知

例 3.1

某工厂生产的一种细纱支数服从正态分布, 其标准差为 *1.2*. 现从某日生产的一批产品中抽取 *16* 缕进行支数测量, 测得样本标准差为 *2.1*, 问纱的均匀度是否改变 $(\alpha = 0.05)$?

解: 由数据计算得 $S^2 = 4.41$. 检验问题为

$$H_0 : \sigma^2 = 1.44 \sim H_1 : \sigma^2 \neq 1.44.$$

检验的拒绝域为

$$\left\{(X_1, \cdots, X_n) : \frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{n-1}\left(1-\frac{\alpha}{2}\right) \text{ 或 } \frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{n-1}\left(\frac{\alpha}{2}\right)\right\}.$$

单个正态总体方差的检验: 均值未知

其中

$$n = 16, \alpha = 0.05,$$
$$\chi^2_{15}(0.975) = 6.262, \chi^2_{15}(0.025) = 27.488.$$

故有

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{15 \times 4.41}{1.44} = 45.94 > 27.488.$$

落入拒绝域, 可推翻 H_0, 即认为棉纱的均匀度发生改变. $\square$

两个正态总体方差比的检验

- 设 $\boldsymbol{X} = (X_1, \cdots, X_m)$ 为正态总体 $N(\mu_1, \sigma_1^2)$ 中抽取的简单随机样本.
- 设 $\boldsymbol{Y} = (Y_1, \cdots, Y_n)$ 为自正态总体 $N(\mu_2, \sigma_2^2)$ 中抽取的简单随机样本.
- 且样本 $X_1, \cdots, X_m$ 和 $Y_1, \cdots, Y_m$ 相互独立.

讨论三类检验问题:

$$
\begin{aligned}
&(8)\quad H_0 : \sigma_2^2/\sigma_1^2 = 1 \sim \sigma_2^2/\sigma_1^2 \neq 1,\\
&(9)\quad H_0 : \sigma_2^2/\sigma_1^2 \leq 1 \sim \sigma_2^2/\sigma_1^2 > 1,\\
&(10)\quad H_0 : \sigma_2^2/\sigma_1^2 \geq 1 \sim \sigma_2^2/\sigma_1^2 < 1,
\end{aligned}
$$

其中检验水平 α 给定. 记 $\bar{X}$ 和 S_1^2 为 $X_1, \cdots, X_m$ 的样本均值和样本方差; $\bar{Y}$ 和 S_2^2 为 $Y_1, \cdots, Y_n$ 的样本均值和样本方差.

两个正态总体方差比的检验

对检验问题 (8), 考虑 μ_1 和 μ_2 未知.

- 由于 S_1^2 和 S_2^2 分别是 σ_1^2 和 σ_2^2 的无偏估计, 直观上, S_2^2/S_1^2 太小或太大时, H_0 越不成立.
- 拒绝域形式为
$$\left\{(X_1,\cdots,X_m;Y_1,\cdots,Y_n):\frac{S_2^2}{S_1^2}<c_1 \text{ 或 } \frac{S_2^2}{S_1^2}>c_2\right\},$$
其中 c_1 和 c_2 待定.
- H_0 成立的条件下,
$$F=\frac{S_2^2}{S_1^2}\sim F_{n-1,m-1}.$$
故取检验统计量 $F=S_2^2/S_1^2$.

两个正态总体方差比的检验

- 记 $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$, 为确定拒绝域的临界值 c_1, c_2, 令

$$P_\theta\left(\frac{S_2^2}{S_1^2} < c_1 \text{ 或 } \frac{S_2^2}{S_1^2} > c_2 | H_0\right) = \alpha.$$

- 简单地, 令

$$P_\theta\left(\frac{S_2^2}{S_1^2} < c_1\right) = \frac{\alpha}{2}, \quad P_\theta\left(\frac{S_2^2}{S_1^2} > c_2\right) = \frac{\alpha}{2}.$$

解得 $c_1 = F_{n-1,m-1}(1-\alpha/2), c_2 = F_{n-1,m-1}(\alpha/2)$.

两个正态总体方差比的检验

检验问题 (8) 水平为 α 的拒绝域为

$$D_8 = \left\{(\boldsymbol{X};\boldsymbol{Y}) : \frac{S_2^2}{S_1^2} < F_{n-1,m-1}(1-\alpha/2)\right. \\ \left. \text{或 } \frac{S_2^2}{S_1^2} > F_{n-1,m-1}(\alpha/2)\right\}.$$

类似地, 检验问题 (9) 和 (10) 的水平为 α 的拒绝域如下:

$$D_9 = \left\{(X_1,\cdots,X_m;Y_1,\cdots,Y_n) : \frac{S_2^2}{S_1^2} > F_{n-1,m-1}(\alpha)\right\},$$

$$D_{10} = \left\{(X_1,\cdots,X_m;Y_1,\cdots,Y_n) : \frac{S_2^2}{S_1^2} < F_{n-1,m-1}(1-\alpha)\right\}.$$

两个正态总体方差比的检验

例 3.2

测得样本大小为 6 的器材电阻均值 $\overline{X} = 0.14, \overline{Y} = 0.139$, 样本标准差为 $S_1 = 0.0026, S_2 = 0.0024$, 假设这两批器材的电阻分别服从 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$, 均值方差皆未知且两组样本独立, 问这两批电子器材的电阻是否相同 $(\alpha = 0.05)$?

解: 要解决两个正态总体均值差的检验, 首先要求两个正态总体方差是否相同的检验. (1) 首先考虑检验

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

$\alpha = 0.05$. 此检验的拒绝域为

$$\left\{(\boldsymbol{X}; \boldsymbol{Y}) : \frac{S_2^2}{S_1^2} < F_{n-1,m-1}(1-\alpha/2) \text{ 或 } \frac{S_2^2}{S_1^2} > F_{n-1,m-1}(\alpha/2)\right\}.$$

两个正态总体方差比的检验

其中

$$m = n = 6, S_1^2/S_2^2 = 0.0026^2/0.0024^2 = 1.17.$$

由 $\alpha = 0.05$ 可得

$$F_{5,5}(0.025) = 7.15, F_{5,5}(0.975) = 1/7.15.$$

因为

$$\frac{1}{7.15} < F = \frac{S_2^2}{S_1^2} < 7.15,$$

没有落在拒绝域, 因此无法拒绝 H_0, 即无法认为 σ_1^2 和 σ_2^2 不等.

两个正态总体方差比的检验

(2) 假定 $\sigma_1^2 = \sigma_2^2$, 考虑如下检验问题:

$$H_0' : \mu_1 = \mu_2 \sim H_1' : \mu_1 \neq \mu_2.$$

该检验的拒绝域为

$$\{(\boldsymbol{X}, \boldsymbol{Y}) : |T_\omega| > t_{n+m-2}(\alpha/2)\},$$

$m = n = 6, \bar{X} = 0.14, \bar{Y} = 0.139, S_1 = 0.0026, S_2 = 0.0024$, 因此有

$$S_\omega^2 = \frac{1}{10}(5 \times 0.0026^2 + 5 \times 0.0024^2) = 6.26 \times 10^{-6},$$

且 $S_\omega = 0.0025$.

由 $\alpha = 0.05$ 得, $t_{10}(0.0025) = 2.228$. 由于

$$\begin{aligned}|T_\omega| &= \sqrt{\frac{mn}{m+n}}\left|\frac{\bar{Y}-\bar{X}}{S_\omega}\right| \\ &= \sqrt{3}\left|\frac{0.14-0.139}{0.0025}\right| \\ &= 0.6928 \\ &< 2.228,\end{aligned}$$

没有落在拒绝域, 故无法拒绝 H_0', 即没有充足的理由拒绝两批电子器件的电阻值相同. □