概率论与数理统计（12）

清华大学

2020 年春季学期

例子：置信区间

- 各自从总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 中抽样，样本容量分别为 10 和 15, 样本方差分别为 56.5 和 52.4, 求 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信水平为 $1-\alpha$ 的置信区间，$\alpha \in (0,1)$。

例子：置信区间

- 各自从总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 中抽样，样本容量分别为 10 和 15, 样本方差分别为 56.5 和 52.4, 求 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信水平为 $1-\alpha$ 的置信区间，$\alpha \in (0,1)$。
- $m = 10$, $n = 15$, $(m-1)s_1^2/\sigma_1^2 \sim \chi^2(m-1)$, $(n-1)s_2^2/\sigma_2^2\chi^2(n-1)$,

例子：置信区间

- 各自从总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 中抽样，样本容量分别为 10 和 15, 样本方差分别为 56.5 和 52.4, 求 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信水平为 $1-\alpha$ 的置信区间，$\alpha \in (0,1)$。
- $m = 10$, $n = 15$, $(m-1)s_1^2/\sigma_1^2 \sim \chi^2(m-1)$, $(n-1)s_2^2/\sigma_2^2\chi^2(n-1)$,
- $F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F(m-1, n-1)$.

例子：置信区间

- 各自从总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 中抽样，样本容量分别为 10 和 15, 样本方差分别为 56.5 和 52.4, 求 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信水平为 $1-\alpha$ 的置信区间，$\alpha \in (0,1)$。
- $m = 10$, $n = 15$, $(m-1)s_1^2/\sigma_1^2 \sim \chi^2(m-1)$, $(n-1)s_2^2/\sigma_2^2\chi^2(n-1)$,
- $F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F(m-1, n-1)$.
- σ_1^2/σ_2^2 的 $1-\alpha$ 置信区间为

$$[\frac{s_1^2}{s_2^2}\frac{1}{F_{1-\alpha/2}(m-1,n-1)}, \frac{s_1^2}{s_2^2}\frac{1}{F_{\alpha/2}(m-1,n-1)}]$$
$$= [\frac{56.5}{52.4}\frac{1}{F_{1-\alpha/2}(9,15)}, \frac{56.5}{52.4}\frac{1}{F_{\alpha/2}(9,15)}].$$

假设检验

- 某公司生产的合金强度服从正态分布 $N(\theta, 16)$, 其中 θ 的设计值为不低于 110 帕。为保证质量，该公司每天都对生产情况做例行检查，以判断生产是否正常，即平均强度不低于 110 帕。某天进行了 25 块合金的检测，强度平均值为 108.2 帕，问当天生产是否正常？

假设检验

- 某公司生产的合金强度服从正态分布 $N(\theta, 16)$，其中 θ 的设计值为不低于 110 帕。为保证质量，该公司每天都对生产情况做例行检查，以判断生产是否正常，即平均强度不低于 110 帕。某天进行了 25 块合金的检测，强度平均值为 108.2 帕，问当天生产是否正常？
- 这不是个参数估计问题。

假设检验

- 某公司生产的合金强度服从正态分布 $N(\theta, 16)$，其中 θ 的设计值为不低于 110 帕。为保证质量，该公司每天都对生产情况做例行检查，以判断生产是否正常，即平均强度不低于 110 帕。某天进行了 25 块合金的检测，强度平均值为 108.2 帕，问当天生产是否正常？
- 这不是个参数估计问题。
- 要求的回答不是参数值 θ 是多少，而是否合格，即当天生产的合计的平均强度是否不小 110 帕。

假设检验

- 某公司生产的合金强度服从正态分布 $N(\theta, 16)$，其中 θ 的设计值为不低于 110 帕。为保证质量，该公司每天都对生产情况做例行检查，以判断生产是否正常，即平均强度不低于 110 帕。某天进行了 25 块合金的检测，强度平均值为 108.2 帕，问当天生产是否正常？
- 这不是个参数估计问题。
- 要求的回答不是参数值 θ 是多少，而是否合格，即当天生产的合计的平均强度是否不小 110 帕。
- 对于命题"生产的合金是合格的"的回答仅涉及参数 θ 的取值范围：

$$\Theta_0 = \{\theta \geqslant 110\}, \quad \Theta_1 = \{\theta < 110\}.$$

假设检验

- 某公司生产的合金强度服从正态分布 $N(\theta, 16)$，其中 θ 的设计值为不低于 110 帕。为保证质量，该公司每天都对生产情况做例行检查，以判断生产是否正常，即平均强度不低于 110 帕。某天进行了 25 块合金的检测，强度平均值为 108.2 帕，问当天生产是否正常？
- 这不是个参数估计问题。
- 要求的回答不是参数值 θ 是多少，而是否合格，即当天生产的合计的平均强度是否不小 110 帕。
- 对于命题“生产的合金是合格的”的回答仅涉及参数 θ 的取值范围：

$$\Theta_0 = \{\theta \geqslant 110\}, \quad \Theta_1 = \{\theta < 110\}.$$

这个两个非空不交的参数集合称为统计假设。

假设检验

- 某公司生产的合金强度服从正态分布 $N(\theta, 16)$, 其中 θ 的设计值为不低于 110 帕。为保证质量，该公司每天都对生产情况做例行检查，以判断生产是否正常，即平均强度不低于 110 帕。某天进行了 25 块合金的检测，强度平均值为 108.2 帕，问当天生产是否正常？
- 这不是个参数估计问题。
- 要求的回答不是参数值 θ 是多少，而是否合格，即当天生产的合计的平均强度是否不小 110 帕。
- 对于命题"生产的合金是合格的"的回答仅涉及参数 θ 的取值范围：

$$\Theta_0 = \{\theta \geqslant 110\}, \quad \Theta_1 = \{\theta < 110\}.$$

  这个两个非空不交的参数集合称为统计假设。
- 如何利用总体性质 $N(\theta, 16)$ 和样本信息 $\bar{x} = 108.2$ 来判断命题是否成立，为该假设的检验法则。

假设检验的基本步骤

- 建立假设：有两个不相交的参数集合 Θ_0, Θ_1，我们希望 $H_0 : \theta \in \Theta_0$ 成立，该假设为原假设或者零假设，对应的 $H_1 : \theta_1 \in \Theta_1$ 为对立假设或者备选假设。

假设检验的基本步骤

- 建立假设：有两个不相交的参数集合 Θ_0, Θ_1，我们希望 $H_0 : \theta \in \Theta_0$ 成立，该假设为原假设或者零假设，对应的 $H_1 : \theta_1 \in \Theta_1$ 为对立假设或者备选假设。
- 若 Θ_0 中只有一个值，则称其为简单原假设，否则称为复杂或者复合原假设：

假设检验的基本步骤

- 建立假设：有两个不相交的参数集合 Θ_0, Θ_1，我们希望 $H_0 : \theta \in \Theta_0$ 成立，该假设为原假设或者零假设，对应的 $H_1 : \theta_1 \in \Theta_1$ 为对立假设或者备选假设。
- 若 Θ_0 中只有一个值，则称其为简单原假设，否则称为复杂或者复合原假设：
  - $H_0 : \theta = \theta_0$, 为简单原假设，对立假设可以为：$\theta \neq \theta_0, \theta > \theta_0, \theta < \theta_0$.

假设检验的基本步骤

- 建立假设：有两个不相交的参数集合 Θ_0, Θ_1，我们希望 $H_0: \theta \in \Theta_0$ 成立，该假设为原假设或者零假设，对应的 $H_1: \theta_1 \in \Theta_1$ 为对立假设或者备选假设。
- 若 Θ_0 中只有一个值，则称其为简单原假设，否则称为复杂或者复合原假设：
  - $H_0: \theta = \theta_0$, 为简单原假设，对立假设可以为：$\theta \neq \theta_0, \theta > \theta_0, \theta < \theta_0$.
  - $H_0: \theta \geqslant 110$ 为复合原假设，对立假设可以为：$\theta < 110$.

假设检验的基本步骤

- 建立假设：有两个不相交的参数集合 Θ_0, Θ_1，我们希望 $H_0: \theta \in \Theta_0$ 成立，该假设为原假设或者零假设，对应的 $H_1: \theta_1 \in \Theta_1$ 为对立假设或者备选假设。
- 若 Θ_0 中只有一个值，则称其为简单原假设，否则称为复杂或者复合原假设：
  - $H_0: \theta = \theta_0$, 为简单原假设，对立假设可以为：$\theta \neq \theta_0, \theta > \theta_0, \theta < \theta_0$.
  - $H_0: \theta \geqslant 110$ 为复合原假设，对立假设可以为：$\theta < 110$.
- 选择检验统计量，给出拒绝区域的形式: 将样本空间划分为两个互不相交的部分，W 和 $\bar{W}$, 制定判断法则：
  - $(x_1, \ldots, x_n) \in W$, 则拒绝零假设 H_0,
  - $(x_1, \ldots, x_n) \in \bar{W}$, 不拒绝零假设，即接受它。

假设检验的基本步骤

- 建立假设：有两个不相交的参数集合 Θ_0, Θ_1，我们希望 $H_0: \theta \in \Theta_0$ 成立，该假设为原假设或者零假设，对应的 $H_1: \theta_1 \in \Theta_1$ 为对立假设或者备选假设。
- 若 Θ_0 中只有一个值，则称其为简单原假设，否则称为复杂或者复合原假设：
  - $H_0: \theta = \theta_0$, 为简单原假设，对立假设可以为：$\theta \neq \theta_0, \theta > \theta_0, \theta < \theta_0$.
  - $H_0: \theta \geqslant 110$ 为复合原假设，对立假设可以为：$\theta < 110$.
- 选择检验统计量，给出拒绝区域的形式：将样本空间划分为两个互不相交的部分，W 和 $\bar{W}$, 制定判断法则：
  - $(x_1, \ldots, x_n) \in W$, 则拒绝零假设 H_0,
  - $(x_1, \ldots, x_n) \in \bar{W}$, 不拒绝零假设，即接受它。
  - W 的划分一般和选择的统计量的有关：在前例中所选统计量为 $\bar{x}$, 可以选

$$W = \{(x_1, \ldots, x_n) : \bar{x} \leqslant c\}.$$

假设检验的基本步骤

- 选择显著水平：确定好拒绝区域后，假设检验可能发生的错误：

假设检验的基本步骤

- 选择显著水平：确定好拒绝区域后，假设检验可能发生的错误：
  - 第一类错误：$H_0 : \theta \in \Theta_0$ 为真，样本却落到拒绝区域，从而认为其为假：去真。

- 选择显著水平：确定好拒绝区域后，假设检验可能发生的错误：
  - 第一类错误：$H_0 : \theta \in \Theta_0$ 为真，样本却落到拒绝区域，从而认为其为假：去真。
  - 第二类错误：$H_0 : \theta \in \Theta_0$ 为假，样本没有落到拒绝区域，从而认为其为假：存伪。

假设检验的基本步骤

- 选择显著水平：确定好拒绝区域后，假设检验可能发生的错误：
  - 第一类错误：$H_0 : \theta \in \Theta_0$ 为真，样本却落到拒绝区域，从而认为其为假：去真。
  - 第二类错误：$H_0 : \theta \in \Theta_0$ 为假，样本没有落到拒绝区域，从而认为其为假：存伪。
  - 第一类错误发生的概率为 $\alpha = P(X \in \bar{W}|H_0)$，第二类错误发生的概率为 $\beta = P(X \in W|H_1)$.

假设检验的基本步骤

- 选择显著水平：确定好拒绝区域后，假设检验可能发生的错误：
  - 第一类错误：$H_0 : \theta \in \Theta_0$ 为真，样本却落到拒绝区域，从而认为其为假：去真。
  - 第二类错误：$H_0 : \theta \in \Theta_0$ 为假，样本没有落到拒绝区域，从而认为其为假：存伪。
  - 第一类错误发生的概率为 $\alpha = P(X \in \bar{W}|H_0)$, 第二类错误发生的概率为 $\beta = P(X \in W|H_1)$.只要有随机性，这两个概率一般情况下都不可能为零。有没有可能有找到能让两个概率同时都很小的方法呢?

假设检验的基本步骤

- 选择显著水平：确定好拒绝区域后，假设检验可能发生的错误：
  - 第一类错误：$H_0 : \theta \in \Theta_0$ 为真，样本却落到拒绝区域，从而认为其为假：去真。
  - 第二类错误：$H_0 : \theta \in \Theta_0$ 为假，样本没有落到拒绝区域，从而认为其为假：存伪。
  - 第一类错误发生的概率为 $\alpha = P(X \in \bar{W}|H_0)$，第二类错误发生的概率为 $\beta = P(X \in W|H_1)$.只要有随机性，这两个概率一般情况下都不可能为零。有没有可能有找到能让两个概率同时都很小的方法呢？一般也是不可能的。

- 选择显著水平：确定好拒绝区域后，假设检验可能发生的错误：
  - 第一类错误：$H_0 : \theta \in \Theta_0$ 为真，样本却落到拒绝区域，从而认为其为假：去真。
  - 第二类错误：$H_0 : \theta \in \Theta_0$ 为假，样本没有落到拒绝区域，从而认为其为假：存伪。
  - 第一类错误发生的概率为 $\alpha = P(X \in \bar{W}|H_0)$，第二类错误发生的概率为 $\beta = P(X \in W|H_1)$.只要有随机性，这两个概率一般情况下都不可能为零。有没有可能有找到能让两个概率同时都很小的方法呢？一般也是不可能的。
- 势函数或者功效函数：设假设检验问题：$H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1$ 的拒绝域为 W，则样本落在拒绝域的概率为该检验的势函数，记为

$$g(\theta) = P_\theta(X \in W), \quad \theta \in \Theta_0 \cup \Theta_1.$$

势函数

- 当 $\theta \in \Theta_0$ 时，$g(\theta) = \alpha(\theta)$，当 $\theta \in \Theta_1$ 时，$g(\theta) = 1 - \beta(\theta)$，也就有

$$\begin{cases} \alpha(\theta) = g(\theta), & \theta \in \Theta_0, \text{第一类错误的概率} \\ \beta(\theta) = 1 - g(\theta), & \theta \in \Theta_1 \text{ 第二类错误的概率.} \end{cases}$$

势函数

- 当 $\theta \in \Theta_0$ 时，$g(\theta) = \alpha(\theta)$，当 $\theta \in \Theta_1$ 时，$g(\theta) = 1 - \beta(\theta)$，也就有

$$\begin{cases} \alpha(\theta) = g(\theta), & \theta \in \Theta_0,\ \text{第一类错误的概率} \\ \beta(\theta) = 1 - g(\theta), & \theta \in \Theta_1\ \text{第二类错误的概率}. \end{cases}$$

- 在前面合金的例子中，拒绝域为 $W = \{\bar{x} \leqslant c\}$，则势函数为

$$g(\theta) = P_\theta(\bar{x} \leqslant c) = P_\theta\left(\frac{\bar{x} - \theta}{\sqrt{16}/\sqrt{25}} \leqslant \frac{c - \theta}{4/\sqrt{25}}\right) = \Phi\left(\frac{c - \theta}{4/5}\right).$$

也就是有

$$\begin{cases} \alpha(\theta) = \Phi(\frac{c-\theta}{4/5}), & \theta \in \Theta_0, \\ \beta(\theta) = 1 - \Phi(\frac{c-\theta}{4/5}), & \theta \in \Theta_1. \end{cases}$$

势函数

- 当 $\theta \in \Theta_0$ 时, $g(\theta) = \alpha(\theta)$, 当 $\theta \in \Theta_1$ 时, $g(\theta) = 1 - \beta(\theta)$, 也就有

$$\begin{cases} \alpha(\theta) = g(\theta), & \theta \in \Theta_0, \text{第一类错误的概率} \\ \beta(\theta) = 1 - g(\theta), & \theta \in \Theta_1 \text{ 第二类错误的概率}. \end{cases}$$

- 在前面合金的例子中，拒绝域为 $W = \{\bar{x} \leqslant c\}$, 则势函数为

$$g(\theta) = P_\theta(\bar{x} \leqslant c) = P_\theta\left(\frac{\bar{x} - \theta}{\sqrt{16}/\sqrt{25}} \leqslant \frac{c - \theta}{4/\sqrt{25}}\right) = \Phi\left(\frac{c - \theta}{4/5}\right).$$

也就是有

$$\begin{cases} \alpha(\theta) = \Phi(\frac{c-\theta}{4/5}), & \theta \in \Theta_0, \\ \beta(\theta) = 1 - \Phi(\frac{c-\theta}{4/5}), & \theta \in \Theta_1. \end{cases}$$

  - 减小 α, 就要降低 c, 降低 c 会增大 β.

势函数

- 当 $\theta \in \Theta_0$ 时，$g(\theta) = \alpha(\theta)$，当 $\theta \in \Theta_1$ 时，$g(\theta) = 1 - \beta(\theta)$，也就有

$$\begin{cases} \alpha(\theta) = g(\theta), & \theta \in \Theta_0，\text{第一类错误的概率} \\ \beta(\theta) = 1 - g(\theta), & \theta \in \Theta_1 \ \text{第二类错误的概率}. \end{cases}$$

- 在前面合金的例子中，拒绝域为 $W = \{\bar{x} \leqslant c\}$，则势函数为

$$g(\theta) = P_\theta(\bar{x} \leqslant c) = P_\theta\left(\frac{\bar{x} - \theta}{\sqrt{16}/\sqrt{25}} \leqslant \frac{c - \theta}{4/\sqrt{25}}\right) = \Phi\left(\frac{c - \theta}{4/5}\right).$$

也就是有

$$\begin{cases} \alpha(\theta) = \Phi(\frac{c-\theta}{4/5}), & \theta \in \Theta_0, \\ \beta(\theta) = 1 - \Phi(\frac{c-\theta}{4/5}), & \theta \in \Theta_1. \end{cases}$$

  - 减小 α，就要降低 c，降低 c 会增大 β.
  - 减小 β，要提升 c，从而增大 α.

- 当 $\theta \in \Theta_0$ 时, $g(\theta)=\alpha(\theta)$, 当 $\theta \in \Theta_1$ 时, $g(\theta)=1-\beta(\theta)$, 也就有

$$\begin{cases} \alpha(\theta)=g(\theta), & \theta \in \Theta_0, \text{第一类错误的概率} \\ \beta(\theta)=1-g(\theta), & \theta \in \Theta_1 \ \text{第二类错误的概率}. \end{cases}$$

- 在前面合金的例子中，拒绝域为 $W=\{\bar{x} \leqslant c\}$, 则势函数为

$$g(\theta)=P_\theta(\bar{x} \leqslant c)=P_\theta\left(\frac{\bar{x}-\theta}{\sqrt{16}/\sqrt{25}} \leqslant \frac{c-\theta}{4/\sqrt{25}}\right)=\Phi\left(\frac{c-\theta}{4/5}\right).$$

也就是有

$$\begin{cases} \alpha(\theta)=\Phi(\frac{c-\theta}{4/5}), & \theta \in \Theta_0, \\ \beta(\theta)=1-\Phi(\frac{c-\theta}{4/5}), & \theta \in \Theta_1. \end{cases}$$

  - 减小 α, 就要降低 c, 降低 c 会增大 β.
  - 减小 β, 要提升 c, 从而增大 α.要两类错误都很小，增大样本容量.

假设检验的基本步骤

- 显著水平：对于检验问题 $H_0: \theta \in \Theta_0$ vs $H_1: \theta_1 \in \Theta_1$，若对于所有 $\theta \in \Theta_0$ 均有
$$g(\theta) \leqslant \alpha,$$
则称该检验为显著水平为 α 的显著性检验，或者水平为 α 的检验。

假设检验的基本步骤

- 显著水平：对于检验问题 $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta_1 \in \Theta_1$，若对于所有 $\theta \in \Theta_0$ 均有
$$g(\theta) \leqslant \alpha,$$
则称该检验为显著水平为 α 的显著性检验，或者水平为 α 的检验。
- 显著性检验要限制第一类错误的概率，通常 $\alpha = 0.1$ 或 0.05.

假设检验的基本步骤

- 显著水平：对于检验问题 $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta_1 \in \Theta_1$，若对于所有 $\theta \in \Theta_0$ 均有
$$g(\theta) \leqslant \alpha,$$
则称该检验为显著水平为 α 的显著性检验，或者水平为 α 的检验。
- 显著性检验要限制第一类错误的概率，通常 $\alpha = 0.1$ 或 0.05.
- 给出拒绝域: 在确定显著水平之后给出拒绝域：前合金例子中，要求
$$\theta \geqslant 110, \quad g(\theta) = \Phi(\frac{c-\theta}{4/5}) \leqslant \alpha.$$

- 显著水平：对于检验问题 $H_0:\theta\in\Theta_0$ vs $H_1:\theta_1\in\Theta_1$, 若对于所有 $\theta\in\Theta_0$ 均有
$$g(\theta)\leqslant\alpha,$$
则称该检验为显著水平为 α 的显著性检验，或者水平为 α 的检验。
- 显著性检验要限制第一类错误的概率，通常 $\alpha=0.1$ 或 0.05.
- 给出拒绝域: 在确定显著水平之后给出拒绝域：前合金例子中，要求
$$\theta\geqslant 110,\quad g(\theta)=\Phi(\frac{c-\theta}{4/5})\leqslant\alpha.$$
此处，$g(\theta)$ 关于 θ 递减，只要 $\theta=110$ 出成立即可: $c\leqslant 110+0.8u_\alpha$. 即拒绝域为
$$W=\{\bar{x}\leqslant 110+0.8u_\alpha\}=\{\frac{\bar{x}-110}{4/5}\leqslant u_\alpha\}.$$

- 做出判断:$u=\frac{108.2-110}{4/5}=-2.25.$

- 做出判断:$u = \frac{108.2-110}{4/5} = -2.25.$
  - $\alpha = 0.1$, $u_{0.1} = -1.282$, $u \leqslant u_\alpha$, 拒绝;

假设检验的基本步骤

- 做出判断:$u = \frac{108.2-110}{4/5} = -2.25.$
  - $\alpha = 0.1$, $u_{0.1} = -1.282$, $u \leqslant u_\alpha$, 拒绝;
  - $\alpha = 0.025$, $u_{0.025} = -1.96$, 拒绝;

假设检验的基本步骤

- 做出判断:$u = \frac{108.2-110}{4/5} = -2.25.$
  - $\alpha = 0.1$, $u_{0.1} = -1.282$, $u \leqslant u_\alpha$, 拒绝;
  - $\alpha = 0.025$, $u_{0.025} = -1.96$, 拒绝;
  - $\alpha = 0.01$, $u_{0.01} = -2.326$, 接受;

假设检验的基本步骤

- 做出判断:$u = \frac{108.2-110}{4/5} = -2.25.$
  - $\alpha = 0.1$, $u_{0.1} = -1.282$, $u \leqslant u_\alpha$, 拒绝;
  - $\alpha = 0.025$, $u_{0.025} = -1.96$, 拒绝;
  - $\alpha = 0.01$, $u_{0.01} = -2.326$, 接受;
  - $\alpha = 0.005$, $u_{0.005} = -2.576$, 接受。

假设检验的基本步骤

- 做出判断:$u = \frac{108.2-110}{4/5} = -2.25.$
  - $\alpha = 0.1$, $u_{0.1} = -1.282$, $u \leqslant u_\alpha$, 拒绝;
  - $\alpha = 0.025$, $u_{0.025} = -1.96$, 拒绝;
  - $\alpha = 0.01$, $u_{0.01} = -2.326$, 接受;
  - $\alpha = 0.005$, $u_{0.005} = -2.576$, 接受。
- 检验的 p 值：一个假设检验问题中，利用样本观测值能够做出拒绝原假设的最小显著水平称为检验的 p 值。

假设检验的基本步骤

- 做出判断:$u = \frac{108.2-110}{4/5} = -2.25.$
  - $\alpha = 0.1$, $u_{0.1} = -1.282$, $u \leqslant u_\alpha$, 拒绝;
  - $\alpha = 0.025$, $u_{0.025} = -1.96$, 拒绝;
  - $\alpha = 0.01$, $u_{0.01} = -2.326$, 接受;
  - $\alpha = 0.005$, $u_{0.005} = -2.576$, 接受。
- 检验的 p 值:一个假设检验问题中,利用样本观测值能够做出拒绝原假设的最小显著水平称为检验的 p 值。
- p 依赖于具体的样本数据。

- 做出判断:$u = \frac{108.2-110}{4/5} = -2.25$.
  - $\alpha = 0.1$, $u_{0.1} = -1.282$, $u \leqslant u_\alpha$, 拒绝;
  - $\alpha = 0.025$, $u_{0.025} = -1.96$, 拒绝;
  - $\alpha = 0.01$, $u_{0.01} = -2.326$, 接受;
  - $\alpha = 0.005$, $u_{0.005} = -2.576$, 接受。
- 检验的 p 值：一个假设检验问题中，利用样本观测值能够做出拒绝原假设的最小显著水平称为检验的 p 值。
- p 依赖于具体的样本数据。一般来说,
  - $\alpha \geqslant p$, 在显著水平 α 下拒绝 H_0;

假设检验的基本步骤

- 做出判断:$u = \frac{108.2-110}{4/5} = -2.25$.
  - $\alpha = 0.1$, $u_{0.1} = -1.282$, $u \leqslant u_\alpha$, 拒绝;
  - $\alpha = 0.025$, $u_{0.025} = -1.96$, 拒绝;
  - $\alpha = 0.01$, $u_{0.01} = -2.326$, 接受;
  - $\alpha = 0.005$, $u_{0.005} = -2.576$, 接受。
- 检验的 p 值：一个假设检验问题中，利用样本观测值能够做出拒绝原假设的最小显著水平称为检验的 p 值。
- p 依赖于具体的样本数据。一般来说,
  - $\alpha \geqslant p$, 在显著水平 α 下拒绝 H_0;
  - $\alpha < p$, 在显著水平 α 下接受 H_0.

假设检验的基本步骤

- 做出判断:$u = \frac{108.2-110}{4/5} = -2.25.$
  - $\alpha = 0.1$, $u_{0.1} = -1.282$, $u \leqslant u_\alpha$, 拒绝;
  - $\alpha = 0.025$, $u_{0.025} = -1.96$, 拒绝;
  - $\alpha = 0.01$, $u_{0.01} = -2.326$, 接受;
  - $\alpha = 0.005$, $u_{0.005} = -2.576$, 接受。
- 检验的 p 值:一个假设检验问题中,利用样本观测值能够做出拒绝原假设的最小显著水平称为检验的 p 值。
- p 依赖于具体的样本数据。一般来说,
  - $\alpha \geqslant p$, 在显著水平 α 下拒绝 H_0;
  - $\alpha < p$, 在显著水平 α 下接受 H_0.
  - p-值为在原假设成立的情况下,样本观测数值出现的概率!p-值越小,越应该拒绝原假设。
- 等价的判断方式:看样本是否在拒绝域,或者看检验的 p 值是否比显著水平小。哪个方便用哪个。

单个正态总体的均值的检验

- 总体为 $N(\mu, \sigma^2)$, $x_1, \ldots, x_n$ 为样本，考虑一下三种关于 μ 的检验问题：

$$\begin{cases} I: H_0: \ \mu \leqslant \mu_0 \ vs \ H_1: \mu > \mu_0; \\ II: H_0: \ \mu \geqslant \mu_0 \ vs \ H_1: \mu < \mu_0; \\ III: H_0: \ \mu = \mu_0 \ vs \ H_1: \mu \neq \mu_0. \end{cases}$$

单个正态总体的均值的检验

- 总体为 $N(\mu, \sigma^2)$, $x_1, \ldots, x_n$ 为样本，考虑一下三种关于 μ 的检验问题：

$$\begin{cases} I: H_0: \ \mu \leqslant \mu_0 \ vs \ H_1: \mu > \mu_0; \\ II: H_0: \ \mu \geqslant \mu_0 \ vs \ H_1: \mu < \mu_0; \\ III: H_0: \ \mu = \mu_0 \ vs \ H_1: \mu \neq \mu_0. \end{cases}$$

- 在 σ 已经知道的情况下，$\bar{x} \sim N(\mu, \sigma^2/n)$, 考虑统计量

$$u = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}.$$

单个正态总体的均值的检验

- 总体为 $N(\mu, \sigma^2)$, $x_1, \ldots, x_n$ 为样本，考虑一下三种关于 μ 的检验问题：

$$\begin{cases} I: H_0: \ \mu \leqslant \mu_0 \ vs \ H_1: \mu > \mu_0; \\ II: H_0: \ \mu \geqslant \mu_0 \ vs \ H_1: \mu < \mu_0; \\ III: H_0: \ \mu = \mu_0 \ vs \ H_1: \mu \neq \mu_0. \end{cases}$$

- 在 σ 已经知道的情况下，$\bar{x} \sim N(\mu, \sigma^2/n)$, 考虑统计量

$$u = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}.$$

- 对于检验 I，统计量太大的时候应该拒绝；对于检验 II，统计量太小的时候应该拒绝；对于检验 III，统计量的绝对值太大时应该拒绝。

单个正态总体均值的检验

- 检验问题 I： $H_0 : \mu \leqslant \mu_0$ vs $H_1 : \mu > \mu_0$.

单个正态总体均值的检验

- 检验问题 I： $H_0: \mu \leqslant \mu_0$ vs $H_1: \mu > \mu_0$.
- 拒绝域的形式为： $W_I = \{X : u \geqslant c\}$,

单个正态总体均值的检验

- 检验问题 I： $H_0 : \mu \leqslant \mu_0$ vs $H_1 : \mu > \mu_0$.
- 拒绝域的形式为： $W_I = \{X : u \geqslant c\}$,若要求显著水平为 α, 则可要求

$$P_{\mu_0}(u \geqslant c) = \alpha,$$

单个正态总体均值的检验

- 检验问题 I：$H_0 : \mu \leqslant \mu_0$ vs $H_1 : \mu > \mu_0$.
- 拒绝域的形式为：$W_I = \{X : u \geqslant c\}$,若要求显著水平为 α, 则可要求

$$P_{\mu_0}(u \geqslant c) = \alpha,$$

而当 $\mu = \mu_0$ 时，$u \sim N(0, 1)$, 所以 $c = u_{1-\alpha}$, 从而拒绝域为:

$$W_I = \{u \geqslant u_{1-\alpha}\}.$$

单个正态总体均值的检验

- 检验问题 I: $H_0: \mu \leqslant \mu_0$ vs $H_1: \mu > \mu_0$.
- 拒绝域的形式为: $W_I = \{X : u \geqslant c\}$,若要求显著水平为 α, 则可要求
$$P_{\mu_0}(u \geqslant c) = \alpha,$$
而当 $\mu = \mu_0$ 时, $u \sim N(0,1)$, 所以 $c = u_{1-\alpha}$, 从而拒绝域为:
$$W_I = \{u \geqslant u_{1-\alpha}\}.$$
- 该检验的势函数为
$$g(\mu) = P_\mu(X \in W_I) = P_\mu\left(\frac{\bar{x} - \mu + \mu - \mu_0}{\sigma/\sqrt{n}} \geqslant u_{1-\alpha}\right)$$
$$= 1 - \Phi(\sqrt{n}(\mu_0 - \mu)/\sigma + u_{1-\alpha}).$$

单个正态总体均值的检验

- 该检验的 p 值为

$$p_I = P(u \geqslant \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}) = 1 - \Phi(\frac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma}).$$

  - 当 $p_I > \alpha$, 或者 $u < u_{1-\alpha}$(不在拒绝域), 接受 H_0;

单个正态总体均值的检验

- 该检验的 p 值为

$$p_I = P(u \geqslant \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}) = 1 - \Phi(\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}).$$

  - 当 $p_I > \alpha$, 或者 $u < u_{1-\alpha}$(不在拒绝域），接受 H_0;
  - 当 $p_I < \alpha$, 或者 $u \geqslant u_{1-\alpha}$(在拒绝域), 拒绝 H_1.

- 该检验的 p 值为

$$p_I = P(u \geqslant \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}) = 1 - \Phi(\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}).$$

  - 当 $p_I > \alpha$，或者 $u < u_{1-\alpha}$(不在拒绝域)，接受 H_0;
  - 当 $p_I < \alpha$，或者 $u \geqslant u_{1-\alpha}$(在拒绝域)，拒绝 H_1.
- 对于检验问题 II，$H_0: \mu \geqslant \mu_0 \ vs \ H_1: \mu < \mu_0$。当确定显著水平为 α 后，拒绝域为 $W_{II} = \{X: u \leqslant u_\alpha\}$;

- 该检验的 p 值为

$$p_I = P(u \geqslant \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}) = 1 - \Phi(\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}).$$

  - 当 $p_I > \alpha$，或者 $u < u_{1-\alpha}$(不在拒绝域），接受 H_0;
  - 当 $p_I < \alpha$，或者 $u \geqslant u_{1-\alpha}$(在拒绝域)，拒绝 H_1.
- 对于检验问题 II，$H_0: \mu \geqslant \mu_0 \ vs \ H_1: \mu < \mu_0$。当确定显著水平为 α 后，拒绝域为 $W_{II} = \{X: u \leqslant u_\alpha\}$;
  - 该检验的 p 值为：$p_{II} = P(u \leqslant \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}) = \Phi(\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma})$.

单个正态总体均值的检验

- 该检验的 p 值为

$$p_I = P(u \geqslant \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}) = 1 - \Phi(\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}).$$

  - 当 $p_I > \alpha$，或者 $u < u_{1-\alpha}$(不在拒绝域），接受 H_0;
  - 当 $p_I < \alpha$，或者 $u \geqslant u_{1-\alpha}$(在拒绝域)，拒绝 H_1.
- 对于检验问题 II，$H_0: \mu \geqslant \mu_0 \ vs \ H_1: \mu < \mu_0$。当确定显著水平为 α 后，拒绝域为 $W_{II} = \{X: u \leqslant u_\alpha\}$;
  - 该检验的 p 值为：$p_{II} = P(u \leqslant \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}) = \Phi(\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma})$.
  - 当 $p \leqslant \alpha$，或者 $u \leqslant u_\alpha$，拒绝 H_0；当 $p > \alpha$，或者 $u > u_\alpha$，保留 H_0.

- 该检验的 p 值为

$$p_I = P(u \geqslant \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}) = 1 - \Phi(\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}).$$

  - 当 $p_I > \alpha$，或者 $u < u_{1-\alpha}$(不在拒绝域)，接受 H_0;
  - 当 $p_I < \alpha$，或者 $u \geqslant u_{1-\alpha}$(在拒绝域)，拒绝 H_1.
- 对于检验问题 II，$H_0: \mu \geqslant \mu_0$ vs $H_1: \mu < \mu_0$。当确定显著水平为 α 后，拒绝域为 $W_{II} = \{X : u \leqslant u_\alpha\}$;
  - 该检验的 p 值为：$p_{II} = P(u \leqslant \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}) = \Phi(\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma})$.
  - 当 $p \leqslant \alpha$，或者 $u \leqslant u_\alpha$，拒绝 H_0; 当 $p > \alpha$，或者 $u > u_\alpha$，保留 H_0.
- 对于检验问题 III，$H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$。显著水平为 α 的拒绝域为 $W_{III} = \{|u| \geqslant u_{1-\alpha/2}\}$.
  - 该检验的 p 值为

$$p_{III} = 2(1 - \Phi(|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}|)).$$

例子

- 从甲地发送一个信号到乙地。设乙地接收到的信号值是一个服从正态分布 $N(\mu, 0.2^2)$ 的随机变量，其中 μ 为甲地发送的真实信号值。先甲地重复发送同一个信号 5 次，乙地收到的信号值为 8.05，8.15，8.2，8.1，8.25. 接受方有理由认为甲地发送的信号值为 8。问能否接受这一猜测。

例子

- 从甲地发送一个信号到乙地。设乙地接收到的信号值是一个服从正态分布 $N(\mu, 0.2^2)$ 的随机变量，其中 μ 为甲地发送的真实信号值。先甲地重复发送同一个信号 5 次，乙地收到的信号值为 8.05，8.15，8.2，8.1，8.25. 接受方有理由认为甲地发送的信号值为 8。问能否接受这一猜测。
- 总体为 $N(\mu, 0.2^2)$，检验问题为：$H_0 : \mu = 8$ vs $H_1 : \mu \neq 8$.

例子

- 从甲地发送一个信号到乙地。设乙地接收到的信号值是一个服从正态分布 $N(\mu, 0.2^2)$ 的随机变量，其中 μ 为甲地发送的真实信号值。先甲地重复发送同一个信号 5 次，乙地收到的信号值为 8.05，8.15，8.2，8.1，8.25. 接受方有理由认为甲地发送的信号值为 8。问能否接受这一猜测。
- 总体为 $N(\mu, 0.2^2)$，检验问题为：$H_0 : \mu = 8$ vs $H_1 : \mu \neq 8$.
- 给定显著水平 $\alpha = 0.05$, $u_{0.975} = 1.96$, $\bar{x} = 8.15$,
  $u = \frac{\bar{x}-8}{0.2/\sqrt{5}} = 1.68$.

例子

- 从甲地发送一个信号到乙地。设乙地接收到的信号值是一个服从正态分布 $N(\mu, 0.2^2)$ 的随机变量，其中 μ 为甲地发送的真实信号值。先甲地重复发送同一个信号 5 次，乙地收到的信号值为 8.05，8.15，8.2，8.1，8.25. 接受方有理由认为甲地发送的信号值为 8。问能否接受这一猜测。
- 总体为 $N(\mu, 0.2^2)$，检验问题为：$H_0 : \mu = 8$ vs $H_1 : \mu \neq 8$.
- 给定显著水平 $\alpha = 0.05$，$u_{0.975} = 1.96$，$\bar{x} = 8.15$，$u = \frac{\bar{x}-8}{0.2/\sqrt{5}} = 1.68$.不拒绝。

例子

- 从甲地发送一个信号到乙地。设乙地接收到的信号值是一个服从正态分布 $N(\mu, 0.2^2)$ 的随机变量，其中 μ 为甲地发送的真实信号值。先甲地重复发送同一个信号 5 次，乙地收到的信号值为 8.05，8.15，8.2，8.1，8.25. 接受方有理由认为甲地发送的信号值为 8。问能否接受这一猜测。
- 总体为 $N(\mu, 0.2^2)$，检验问题为：$H_0 : \mu = 8$ vs $H_1 : \mu \neq 8$.
- 给定显著水平 $\alpha = 0.05$，$u_{0.975} = 1.96$，$\bar{x} = 8.15$，$u = \frac{\bar{x}-8}{0.2/\sqrt{5}} = 1.68$.不拒绝。
- 该检验的 p 值为 $p = 2(1 - \Phi(1.68)) = 0.093$.

例子

- 从甲地发送一个信号到乙地。设乙地接收到的信号值是一个服从正态分布 $N(\mu, 0.2^2)$ 的随机变量，其中 μ 为甲地发送的真实信号值。先甲地重复发送同一个信号 5 次，乙地收到的信号值为 8.05，8.15，8.2，8.1，8.25. 接受方有理由认为甲地发送的信号值为 8。问能否接受这一猜测。
- 总体为 $N(\mu, 0.2^2)$，检验问题为：$H_0 : \mu = 8$ vs $H_1 : \mu \neq 8$.
- 给定显著水平 $\alpha = 0.05$，$u_{0.975} = 1.96$，$\bar{x} = 8.15$，$u = \frac{\bar{x}-8}{0.2/\sqrt{5}} = 1.68$.不拒绝。
- 该检验的 p 值为 $p = 2(1 - \Phi(1.68)) = 0.093$. 所有低于 0.093 的显著水平都不会被拒绝。若显著水平为 $\alpha = 0.1$，则该检验拒绝假设 H_0.

单个正态总体均值的检验

- 总体为 $N(\mu, \sigma)$，σ 为未知。

- 总体为 $N(\mu,\sigma)$，σ 为未知。考虑
$$t=\frac{\sqrt{n}(\bar{x}-\mu_0)}{s}.$$
当 $\mu=\mu_0$ 时，$t\sim t(n-1)$.
- 对于检验问题 I，$H_0:\ \mu\leqslant\mu_0\ vs\ H_1:\mu>\mu_0$, 拒绝域为 $W_I=\{t\geqslant t_{1-\alpha}\}$, p 值为 $p_I=P(t\geqslant t_0)$, t_0 为有样本数值计算出来的统计量的值。

- 总体为 $N(\mu,\sigma)$，σ 为未知。考虑

$$t=\frac{\sqrt{n}(\bar{x}-\mu_0)}{s}.$$

当 $\mu=\mu_0$ 时，$t\sim t(n-1)$.

- 对于检验问题 I，$H_0:\ \mu\leqslant\mu_0\ vs\ H_1:\mu>\mu_0$, 拒绝域为 $W_I=\{t\geqslant t_{1-\alpha}\}$, p 值为 $p_I=P(t\geqslant t_0)$, t_0 为有样本数值计算出来的统计量的值。
- 对于检验问题 II，$H_0:\ \mu\geqslant\mu_0\ vs\ H_1:\mu<\mu_0$, 拒绝域为 $W_{II}=\{t\leqslant t_\alpha\}$, p 值为 $p_{II}=P(t\leqslant t_0)$.

- 总体为 $N(\mu,\sigma)$，σ 为未知。考虑

$$t=\frac{\sqrt{n}(\bar{x}-\mu_0)}{s}.$$

当 $\mu=\mu_0$ 时，$t\sim t(n-1)$.

- 对于检验问题 I，$H_0:\ \mu\leqslant\mu_0\ vs\ H_1:\mu>\mu_0$, 拒绝域为 $W_I=\{t\geqslant t_{1-\alpha}\}$, p 值为 $p_I=P(t\geqslant t_0)$, t_0 为有样本数值计算出来的统计量的值。
- 对于检验问题 II，$H_0:\ \mu\geqslant\mu_0\ vs\ H_1:\mu<\mu_0$, 拒绝域为 $W_{II}=\{t\leqslant t_\alpha\}$, p 值为 $p_{II}=P(t\leqslant t_0)$.
- 对于检验问题 III，$H_0:\ \mu=\mu_0\ vs\ H_1:\mu\neq\mu_0$, 拒绝域为 $W_{III}=\{|t|\geqslant t_{1-\alpha/2}\}$, p 值为 $p_{III}=P(|t|\geqslant|t_0|)=2P(t\geqslant t_0)$.

例子

- 某厂生产的某种铝材的长度服从正态分布，其均值设定为 240cm。先抽取 5 件产品，测得其长度为 239.7, 239.6, 239, 240, 239.2. 判断该厂此类铝材的长度是否满足设定要求。

例子

- 某厂生产的某种铝材的长度服从正态分布，其均值设定为 240cm。先抽取 5 件产品，测得其长度为 239.7, 239.6, 239, 240, 239.2. 判断该厂此类铝材的长度是否满足设定要求。
- $H_0 : \mu = 240$ vs $H_1 : \mu \neq 240$.

例子

- 某厂生产的某种铝材的长度服从正态分布，其均值设定为 240cm。先抽取 5 件产品，测得其长度为 239.7, 239.6, 239, 240, 239.2. 判断该厂此类铝材的长度是否满足设定要求。
- $H_0 : \mu = 240$ vs $H_1 : \mu \neq 240$.
- $t = \sqrt{5}|239.5 - 240|/0.4 = 2.795$.

例子

- 某厂生产的某种铝材的长度服从正态分布，其均值设定为 240cm。先抽取 5 件产品，测得其长度为 239.7, 239.6, 239, 240, 239.2. 判断该厂此类铝材的长度是否满足设定要求。
- $H_0 : \mu = 240$ vs $H_1 : \mu \neq 240$.
- $t = \sqrt{5}|239.5 - 240|/0.4 = 2.795$.
- 若显著水平 $\alpha = 0.05$, $t_{0.975} = 2.776$, 统计量数值进入拒绝域。所以拒绝 H_0.

例子

- 某厂生产的某种铝材的长度服从正态分布，其均值设定为 240cm。先抽取 5 件产品，测得其长度为 239.7, 239.6, 239, 240, 239.2. 判断该厂此类铝材的长度是否满足设定要求。
- $H_0 : \mu = 240$ vs $H_1 : \mu \neq 240$.
- $t = \sqrt{5}|239.5 - 240|/0.4 = 2.795$.
- 若显著水平 $\alpha = 0.05$, $t_{0.975} = 2.776$, 统计量数值进入拒绝域。所以拒绝 H_0.
- 该检验的 p 值为 $p = 2P(t \geqslant 2.795) = 0.0491$.

假设检验与置信区间的关系

- 设 $x_1, \ldots, x_n$ 是来自总体 $N(\mu, \sigma^2)$ 的样本。

假设检验与置信区间的关系

- 设 $x_1, \ldots, x_n$ 是来自总体 $N(\mu, \sigma^2)$ 的样本。
- 考虑双边检验问题 *III*, 显著水平为 α 的检验接受域为

$$\begin{aligned}\bar{W}_{III} &= \{|\bar{x} - \mu_0| \leqslant \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)\} \\ &= \{\mu_0 \in [\bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1), \bar{x} + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)]\}.\end{aligned}$$

假设检验与置信区间的关系

- 设 $x_1, \dots, x_n$ 是来自总体 $N(\mu, \sigma^2)$ 的样本。
- 考虑双边检验问题 *III*, 显著水平为 α 的检验接受域为

$$\bar{W}_{III} = \{|\bar{x} - \mu_0| \leqslant \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)\}$$
$$= \{\mu_0 \in [\bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1), \bar{x} + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)]\}.$$

- 也就是说，如果零假设里面的的 μ_0 值落到了 $1-\alpha$ 置信区间中，这接受零假设。

假设检验与置信区间的关系

- 设 $x_1, \ldots, x_n$ 是来自总体 $N(\mu, \sigma^2)$ 的样本。
- 考虑双边检验问题 *III*, 显著水平为 α 的检验接受域为

$$\begin{aligned}\bar{W}_{III} &= \{|\bar{x} - \mu_0| \leqslant \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)\} \\ &= \{\mu_0 \in [\bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1), \bar{x} + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)]\}.\end{aligned}$$

- 也就是说，如果零假设里面的的 μ_0 值落到了 $1-\alpha$ 置信区间中，这接受零假设。
- 对于单边假设检验问题 *I*, 显著性水平为 α 的接受域为

$$\bar{W}_I = \{\bar{x} - \mu_0 < \frac{s}{\sqrt{n}} t_{1-\alpha}\} = \{\bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha} < \mu_0\}.$$

假设检验与置信区间的关系

- 设 $x_1, \ldots, x_n$ 是来自总体 $N(\mu, \sigma^2)$ 的样本。
- 考虑双边检验问题 *III*, 显著水平为 α 的检验接受域为

$$\bar{W}_{III} = \{|\bar{x} - \mu_0| \leqslant \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)\}$$
$$= \{\mu_0 \in [\bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1), \bar{x} + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)]\}.$$

- 也就是说，如果零假设里面的的 μ_0 值落到了 $1-\alpha$ 置信区间中，这接受零假设。
- 对于单边假设检验问题 *I*, 显著性水平为 α 的接受域为

$$\bar{W}_I = \{\bar{x} - \mu_0 < \frac{s}{\sqrt{n}} t_{1-\alpha}\} = \{\bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha} < \mu_0\}.$$

- 若原假设与（单侧）置信区间相交非空，则接受之。

两个正态总体均值差的检验

- 设 $x_1, \ldots, x_m$ 是来自正态总体 $N(\mu_1, \sigma_1^2)$ 的样本，$y_1, \ldots, y_n$ 是来自另一个正态总体 $N(\mu_2, \sigma_2^2)$ 的样本，两个样本相互独立。考虑一下三类检验问题:
  - I: $H_0 : \mu_1 - \mu_2 \leqslant 0$ vs $H_1 : \mu_1 - \mu_2 > 0$.
  - II: $H_0 : \mu_1 - \mu_2 \geqslant 0$ vs $H_1 : \mu_1 - \mu_2 < 0$.
  - III: $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 \neq 0$.

- 设 $x_1, \ldots, x_m$ 是来自正态总体 $N(\mu_1, \sigma_1^2)$ 的样本，$y_1, \ldots, y_n$ 是来自另一个正态总体 $N(\mu_2, \sigma_2^2)$ 的样本，两个样本相互独立。考虑一下三类检验问题：
  - I: $H_0 : \mu_1 - \mu_2 \leqslant 0$ vs $H_1 : \mu_1 - \mu_2 > 0.$
  - II: $H_0 : \mu_1 - \mu_2 \geqslant 0$ vs $H_1 : \mu_1 - \mu_2 < 0.$
  - III: $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 \neq 0.$
- σ_1，σ_2 已知；
- σ_1，σ_2 未知。
  - $\sigma_1 = \sigma_2$ 具体数值未知；
  - $\sigma_1/\sigma_2 = c$，比例 c 已知。

- σ_1 和 σ_2 为已知。考虑统计量

$$u = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0,1)(\text{当 } \mu_1 - \mu_2 = 0 \text{ 时}).$$

两个正态总体均值差的检验

- σ_1 和 σ_2 为已知。考虑统计量

$$u = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0,1)(\text{当 } \mu_1 - \mu_2 = 0 \text{ 时}).$$

- 检验问题 I，$H_0 : \mu_1 - \mu_2 \leqslant 0$ vs $H_1 : \mu_1 - \mu_2 > 0$。拒绝域为 $W_I = \{u \geqslant u_{1-\alpha}\}$，$p$-值为 $p_I = 1 - \Phi(u_0)$，其中 u_0 为根据样本数据计算的统计量的值。

两个正态总体均值差的检验

- σ_1 和 σ_2 为已知。考虑统计量

$$u = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0, 1)(\text{当 } \mu_1 - \mu_2 = 0 \text{ 时}).$$

- 检验问题 I，$H_0 : \mu_1 - \mu_2 \leqslant 0$ vs $H_1 : \mu_1 - \mu_2 > 0$。拒绝域为 $W_I = \{u \geqslant u_{1-\alpha}\}$, p-值为 $p_I = 1 - \Phi(u_0)$，其中 u_0 为根据样本数据计算的统计量的值。
- 对于检验问题 II，$H_0 : \mu_1 - \mu_2 \geqslant 0$ vs $H_1 : \mu_1 - \mu_2 < 0$。拒绝域为 $W_{II} = \{u \leqslant u_\alpha\}$, p-值为 $p_{II} = \Phi(u_0)$.

- σ_1 和 σ_2 为已知。考虑统计量

$$u = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0,1)(\text{当 } \mu_1 - \mu_2 = 0 \text{ 时}).$$

- 检验问题 I，$H_0 : \mu_1 - \mu_2 \leqslant 0$ vs $H_1 : \mu_1 - \mu_2 > 0$。拒绝域为 $W_I = \{u \geqslant u_{1-\alpha}\}$，$p$-值为 $p_I = 1 - \Phi(u_0)$，其中 u_0 为根据样本数据计算的统计量的值。
- 对于检验问题 II，$H_0 : \mu_1 - \mu_2 \geqslant 0$ vs $H_1 : \mu_1 - \mu_2 < 0$。拒绝域为 $W_{II} = \{u \leqslant u_\alpha\}$，$p$-值为 $p_{II} = \Phi(u_0)$.
- 对于检验问题 III，$H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 \neq 0$。拒绝域为 $W_{III} = \{|u| \geqslant u_{1-\alpha/2}\}$，$p$-值为 $p_{III} = 2(1 - \Phi(|u_0|))$.

两个正态总体均值差的检验

- $\sigma_1 = \sigma_2 = \sigma$, σ 未知。考虑统计量

$$t = \frac{\bar{x} - \bar{y}}{s_w \sqrt{\frac{1}{m} + \frac{1}{n}}}, \quad s_w^2 = \frac{\sum_{i=1}^{m}(x_i - \bar{x})^2 + \sum_{j=1}^{n}(y_i - \bar{y})^2}{m + n - 2}.$$

当 $\mu_1 - \mu_2 = 0$ 时，$t \sim t(m + n - 2)$

- $\sigma_1=\sigma_2=\sigma$, σ 未知。考虑统计量

$$t=\frac{\bar{x}-\bar{y}}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}},\quad s_w^2=\frac{\sum_{i=1}^m(x_i-\bar{x})^2+\sum_{j=1}^n(y_i-\bar{y})^2}{m+n-2}.$$

当 $\mu_1-\mu_2=0$ 时，$t\sim t(m+n-2)$

- 检验问题 I，$H_0:\mu_1-\mu_2\leqslant 0$ vs $H_1:\mu_1-\mu_2>0$，
$W_I=\{t\geqslant t_{1-\alpha}(m+n-2)\}$, $p_I=P(t\geqslant t_0)$;

- $\sigma_1 = \sigma_2 = \sigma$, σ 未知。考虑统计量

$$t = \frac{\bar{x} - \bar{y}}{s_w\sqrt{\frac{1}{m} + \frac{1}{n}}}, \quad s_w^2 = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 + \sum_{j=1}^n (y_i - \bar{y})^2}{m + n - 2}.$$

当 $\mu_1 - \mu_2 = 0$ 时，$t \sim t(m + n - 2)$

- 检验问题 I，$H_0 : \mu_1 - \mu_2 \leqslant 0$ vs $H_1 : \mu_1 - \mu_2 > 0$，
$W_I = \{t \geqslant t_{1-\alpha}(m + n - 2)\}$, $p_I = P(t \geqslant t_0)$;
- 检验问题 II, $H_0 : \mu_1 - \mu_2 \geqslant 0$ vs $H_1 : \mu_1 - \mu_2 < 0$,
$W_{II} = \{t \leqslant t_\alpha\}$, $p_{II} = P(t \leqslant t_0)$;

两个正态总体均值差的检验

- $\sigma_1 = \sigma_2 = \sigma$, σ 未知。考虑统计量

$$t = \frac{\bar{x} - \bar{y}}{s_w\sqrt{\frac{1}{m} + \frac{1}{n}}}, \quad s_w^2 = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 + \sum_{j=1}^n (y_i - \bar{y})^2}{m+n-2}.$$

当 $\mu_1 - \mu_2 = 0$ 时，$t \sim t(m+n-2)$

- 检验问题 I，$H_0 : \mu_1 - \mu_2 \leqslant 0$ vs $H_1 : \mu_1 - \mu_2 > 0$，
$W_I = \{t \geqslant t_{1-\alpha}(m+n-2)\}$, $p_I = P(t \geqslant t_0)$;
- 检验问题 II, $H_0 : \mu_1 - \mu_2 \geqslant 0$ vs $H_1 : \mu_1 - \mu_2 < 0$,
$W_{II} = \{t \leqslant t_\alpha\}$, $p_{II} = P(t \leqslant t_0)$;
- 检验问题 III, $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 \neq 0$,
$W_{III} = \{|t| \geqslant t_{1-\alpha/2}\}$, $p_{III} = P(|t| \geqslant t_0)$.

两个正态总体均值差的检验

- $\sigma_1/\sigma_2 = c$，比例已知。

两个正态总体均值差的检验

- $\sigma_1/\sigma_2 = c$, 比例已知。考虑统计量

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{(m-1)s_x^2 + (n-1)s_y^2/c}} \sqrt{\frac{mn(m+n-2)}{mc+n}}.$$

- $\sigma_1/\sigma_2 = c$, 比例已知。考虑统计量

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{(m-1)s_x^2 + (n-1)s_y^2/c}} \sqrt{\frac{mn(m+n-2)}{mc+n}}.$$

当 $\mu_1 - \mu_2 = 0$ 时，$t \sim t(m+n-2)$.

- $\sigma_1/\sigma_2 = c$，比例已知。考虑统计量

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{(m-1)s_x^2+(n-1)s_y^2/c}}\sqrt{\frac{mn(m+n-2)}{mc+n}}.$$

当 $\mu_1-\mu_2=0$ 时，$t \sim t(m+n-2)$.

- 检验问题 I，$H_0:\mu_1-\mu_2 \leqslant 0$ vs $H_1:\mu_1-\mu_2>0$，
$W_I=\{t \geqslant t_{1-\alpha}(m+n-2)\}$，$p_I = P(t \geqslant t_0)$；

两个正态总体均值差的检验

- $\sigma_1/\sigma_2 = c$, 比例已知。考虑统计量

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{(m-1)s_x^2+(n-1)s_y^2/c}}\sqrt{\frac{mn(m+n-2)}{mc+n}}.$$

  当 $\mu_1-\mu_2=0$ 时, $t\sim t(m+n-2)$.
- 检验问题 I, $H_0:\mu_1-\mu_2\leqslant 0$ vs $H_1:\mu_1-\mu_2>0$,
  $W_I=\{t\geqslant t_{1-\alpha}(m+n-2)\}$, $p_I=P(t\geqslant t_0)$;
- 检验问题 II, $H_0:\mu_1-\mu_2\geqslant 0$ vs $H_1:\mu_1-\mu_2<0$,
  $W_{II}=\{t\leqslant t_\alpha\}$, $p_{II}=P(t\leqslant t_0)$;

- $\sigma_1/\sigma_2 = c$, 比例已知。考虑统计量

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{(m-1)s_x^2+(n-1)s_y^2/c}}\sqrt{\frac{mn(m+n-2)}{mc+n}}.$$

当 $\mu_1-\mu_2=0$ 时, $t\sim t(m+n-2)$.

- 检验问题 I, $H_0:\mu_1-\mu_2\leqslant 0$ vs $H_1:\mu_1-\mu_2>0$,
  $W_I=\{t\geqslant t_{1-\alpha}(m+n-2)\}$, $p_I=P(t\geqslant t_0)$;
- 检验问题 II, $H_0:\mu_1-\mu_2\geqslant 0$ vs $H_1:\mu_1-\mu_2<0$,
  $W_{II}=\{t\leqslant t_\alpha\}$, $p_{II}=P(t\leqslant t_0)$;
- 检验问题 III, $H_0:\mu_1-\mu_2=0$ vs $H_1:\mu_1-\mu_2\neq 0$,
  $W_{III}=\{|t|\geqslant t_{1-\alpha/2}\}$, $p_{III}=P(|t|\geqslant t_0)$.

例子

- 为提高钢材的硬度，某厂采用了新的锻造工艺。现有数据：
  - 新，样本容量 8，硬度均值 $\bar{x} = 73.39$，样本方差 $s_x^2 = 191.7958/8$；
  - 旧，样本容量 9，硬度均值 $\bar{y} = 68.2756$，样本方差 $s_y^2 = 91.1548/9$.

例子

- 为提高钢材的硬度，某厂采用了新的锻造工艺。现有数据：
  - 新，样本容量 8，硬度均值 $\bar{x} = 73.39$，样本方差 $s_x^2 = 191.7958/8$；
  - 旧，样本容量 9，硬度均值 $\bar{y} = 68.2756$，样本方差 $s_y^2 = 91.1548/9$.
- 假设新工艺对钢材硬度的方差不产生影响。新工艺有用吗？

例子

- 为提高钢材的硬度，某厂采用了新的锻造工艺。现有数据：
  - 新，样本容量 8，硬度均值 $\bar{x} = 73.39$，样本方差 $s_x^2 = 191.7958/8$;
  - 旧，样本容量 9，硬度均值 $\bar{y} = 68.2756$，样本方差 $s_y^2 = 91.1548/9$.
- 假设新工艺对钢材硬度的方差不产生影响。新工艺有用吗？
- $H_0: \ x = y \quad vs \quad H_1: \ x > y$.

例子

- 为提高钢材的硬度，某厂采用了新的锻造工艺。现有数据：
  - 新，样本容量 8，硬度均值 $\bar{x} = 73.39$，样本方差 $s_x^2 = 191.7958/8$；
  - 旧，样本容量 9，硬度均值 $\bar{y} = 68.2756$，样本方差 $s_y^2 = 91.1548/9$.
- 假设新工艺对钢材硬度的方差不产生影响。新工艺有用吗？
- $H_0: \ x = y \quad vs \quad H_1: \ x > y$.
- $s_w = \sqrt{\frac{1}{8+9-2}(191.7958 + 91.1548)} = 4.3432$,

$$t = \frac{\bar{x} - \bar{y}}{s_w\sqrt{\frac{1}{8} + \frac{1}{9}}} = 2.4233.$$

例子

- 为提高钢材的硬度，某厂采用了新的锻造工艺。现有数据：
  - 新，样本容量 8，硬度均值 $\bar{x}=73.39$，样本方差 $s_x^2=191.7958/8$；
  - 旧，样本容量 9，硬度均值 $\bar{y}=68.2756$，样本方差 $s_y^2=91.1548/9$.
- 假设新工艺对钢材硬度的方差不产生影响。新工艺有用吗？
- $H_0: \ x=y \quad vs \quad H_1: \ x>y$.
- $s_w=\sqrt{\frac{1}{8+9-2}(191.7958+91.1548)}=4.3432$,

$$t=\frac{\bar{x}-\bar{y}}{s_w\sqrt{\frac{1}{8}+\frac{1}{9}}}=2.4233.$$

- $\alpha=0.05$, $t_{0.95}(15)=1.7531$, $t>t_{0.95}$，所以拒绝原假设。即，新工艺提高硬度。

例子

- 为提高钢材的硬度，某厂采用了新的锻造工艺。现有数据：
  - 新，样本容量 8，硬度均值 $\bar{x} = 73.39$，样本方差 $s_x^2 = 191.7958/8$;
  - 旧，样本容量 9，硬度均值 $\bar{y} = 68.2756$，样本方差 $s_y^2 = 91.1548/9$.
- 假设新工艺对钢材硬度的方差不产生影响。新工艺有用吗？
- $H_0: x = y \quad vs \quad H_1: x > y$.
- $s_w = \sqrt{\frac{1}{8+9-2}(191.7958 + 91.1548)} = 4.3432$,

$$t = \frac{\bar{x} - \bar{y}}{s_w\sqrt{\frac{1}{8} + \frac{1}{9}}} = 2.4233.$$

- $\alpha = 0.05$, $t_{0.95}(15) = 1.7531$, $t > t_{0.95}$，所以拒绝原假设。即，新工艺提高硬度。
- p-值为 $p = P(t \geqslant 2.433) = 0.0142$.

成对数据检验

- $(x_1, y_1), \ldots, (x_n, y_n)$，为 (X, Y) 的样本，$X \sim N(\mu_1, \sigma^2)$，$Y \sim N(\mu_2, \sigma^2)$，$X$ 与 Y 相互独立。考虑检验问题

$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2.$$

成对数据检验

- $(x_1, y_1), \ldots, (x_n, y_n)$，为 (X, Y) 的样本，$X \sim N(\mu_1, \sigma^2)$，$Y \sim N(\mu_2, \sigma^2)$，$X$ 与 Y 相互独立。考虑检验问题
$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2.$$

- 有两种处理方式，

成对数据检验

- $(x_1, y_1), \ldots, (x_n, y_n)$，为 (X, Y) 的样本，$X \sim N(\mu_1, \sigma^2)$，$Y \sim N(\mu_2, \sigma^2)$，$X$ 与 Y 相互独立。考虑检验问题
$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2.$$

- 有两种处理方式，一是将其看成分别来自两个正态总体的样本，考虑
$$t_1 = \frac{\bar{x} - \bar{y}}{s_w / \sqrt{1/n + 1/n}}, s_w^2 = \frac{s_x^2 + s_y^2}{2}.$$

成对数据检验

- $(x_1,y_1),\ldots,(x_n,y_n)$，为 (X,Y) 的样本，$X\sim N(\mu_1,\sigma^2)$，$Y\sim N(\mu_2,\sigma^2)$，X 与 Y 相互独立。考虑检验问题

$$H_0:\mu_1=\mu_2 \quad \text{vs} \quad H_1:\mu_1\neq\mu_2.$$

- 有两种处理方式，一是将其看成分别来自两个正态总体的样本，考虑

$$t_1=\frac{\bar{x}-\bar{y}}{s_w/\sqrt{1/n+1/n}},\ s_w^2=\frac{s_x^2+s_y^2}{2}.$$

- 另一种是，考虑 $d_i=x_i-y_i$，$d=X-Y\sim N(d,2\sigma^2)$，$d=\mu_1-\mu_2$，则原检验问题就变成

$$H_0:d=0 \quad \text{vs} \quad H_1:d\neq 0.$$

再用单个正态总体的检验，考虑统计量 $t_2=\bar{d}/(s_d/\sqrt{n})$。

成对数据检验

- $(x_1, y_1), \ldots, (x_n, y_n)$，为 (X, Y) 的样本，$X \sim N(\mu_1, \sigma^2)$，$Y \sim N(\mu_2, \sigma^2)$，$X$ 与 Y 相互独立。考虑检验问题

$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2.$$

- 有两种处理方式，一是将其看成分别来自两个正态总体的样本，考虑

$$t_1 = \frac{\bar{x} - \bar{y}}{s_w/\sqrt{1/n + 1/n}}, s_w^2 = \frac{s_x^2 + s_y^2}{2}.$$

- 另一种是，考虑 $d_i = x_i - y_i$，
  $d = X - Y \sim N(d, 2\sigma^2), d = \mu_1 - \mu_2$，则原检验问题就变成

$$H_0 : d = 0 \quad \text{vs} \quad H_1 : d \neq 0.$$

  再用单个正态总体的检验，考虑统计量 $t_2 = \bar{d}/(s_d/\sqrt{n})$。
- 有时候会得到不一样的结论。

成对数据检验

- $(x_1,y_1),\ldots,(x_n,y_n)$，为 (X,Y) 的样本，$X\sim N(\mu_1,\sigma^2)$，$Y\sim N(\mu_2,\sigma^2)$，X 与 Y 相互独立。考虑检验问题
$$H_0:\mu_1=\mu_2 \quad \text{vs} \quad H_1:\mu_1\neq\mu_2.$$
- 有两种处理方式，一是将其看成分别来自两个正态总体的样本，考虑
$$t_1=\frac{\bar{x}-\bar{y}}{s_w/\sqrt{1/n+1/n}},\ s_w^2=\frac{s_x^2+s_y^2}{2}.$$
- 另一种是，考虑 $d_i=x_i-y_i$，$d=X-Y\sim N(d,2\sigma^2), d=\mu_1-\mu_2$，则原检验问题就变成
$$H_0:d=0 \quad \text{vs} \quad H_1:d\neq 0.$$
再用单个正态总体的检验，考虑统计量 $t_2=\bar{d}/(s_d/\sqrt{n})$。
- 有时候会得到不一样的结论。倾向于第二种方式，利用了数据成对出现的特点

两个正态总体均值的检验

- $x_1, \ldots, x_m$ 是总体 $N(\mu_1, \sigma^2)$ 的样本，$y_1, \ldots, y_n$ 是总体 $N(\mu_2, \sigma^2)$ 的样本，两个样本相互独立。考虑一下假设检验问题
$$H_0 : c\mu_1 + d\mu_2 = \delta \quad \text{vs} \quad H_1 : c\mu_1 + d\mu_2 \neq \delta,$$
其中 $c \neq 0$, $d \neq 0$, δ 均为已知常数。

两个正态总体均值的检验

- $x_1, \ldots, x_m$ 是总体 $N(\mu_1, \sigma^2)$ 的样本，$y_1, \ldots, y_n$ 是总体 $N(\mu_2, \sigma^2)$ 的样本，两个样本相互独立。考虑一下假设检验问题

$$H_0 : c\mu_1 + d\mu_2 = \delta \quad \text{vs} \quad H_1 : c\mu_1 + d\mu_2 \neq \delta,$$

  其中 $c \neq 0$, $d \neq 0$, δ 均为已知常数。
- $c\bar{x} + d\bar{y} \sim N(c\mu_1 + d\mu_2, (\frac{c^2}{m} + \frac{d^2}{n})\sigma^2)$.

- $x_1,\ldots,x_m$ 是总体 $N(\mu_1,\sigma^2)$ 的样本，$y_1,\ldots,y_n$ 是总体 $N(\mu_2,\sigma^2)$ 的样本，两个样本相互独立。考虑一下假设检验问题
$$H_0: c\mu_1+d\mu_2=\delta \quad \text{vs} \quad H_1: c\mu_1+d\mu_2\neq\delta,$$
其中 $c\neq 0$, $d\neq 0$, δ 均为已知常数。
- $c\bar{x}+d\bar{y}\sim N(c\mu_1+d\mu_2,(\frac{c^2}{m}+\frac{d^2}{n})\sigma^2)$.
- 如果原假设成立，则
$$u=\frac{c\bar{x}+d\bar{y}-\delta}{\sigma\sqrt{\frac{c^2}{m}+\frac{d^2}{n}}}\sim N(0,1).$$
进而
$$t=\frac{c\bar{x}+d\bar{y}-\delta}{s_w\sqrt{\frac{c^2}{m}+\frac{d^2}{n}}}\sim t(m+n-2).$$
- 显著水平为 α 的拒绝域为 $W=\{|t|\geqslant t_{1-\alpha/2}(m+n-2)\}$.

正态总体方差的检验

- 单个正态总体方差的 χ^2 检验。

- 单个正态总体方差的 χ^2 检验。$x_1, \ldots, x_n$ 是来自 $N(\mu, \sigma^2)$ 的样本，对方差可考虑一下三个检验问题：
  - $I: \quad H_0: \sigma^2 \leqslant \sigma_0^2$ vs $H_1: \sigma^2 > \sigma_0^2$.
  - $II: \quad H_0: \sigma^2 \geqslant \sigma_0^2$ vs $H_1: \sigma^2 < \sigma_0^2$.
  - $III: \quad H_0: \sigma^2 = \sigma_0^2$ vs $H_1: \sigma^2 \neq \sigma_0^2$.

- 单个正态总体方差的 χ^2 检验。$x_1, \ldots, x_n$ 是来自 $N(\mu, \sigma^2)$ 的样本，对方差可考虑一下三个检验问题：
  - I : $H_0 : \sigma^2 \leqslant \sigma_0^2$ vs $H_1 : \sigma^2 > \sigma_0^2$.
  - II : $H_0 : \sigma^2 \geqslant \sigma_0^2$ vs $H_1 : \sigma^2 < \sigma_0^2$.
  - III : $H_0 : \sigma^2 = \sigma_0^2$ vs $H_1 : \sigma^2 \neq \sigma_0^2$.
- 考虑统计统计量 $\chi^2 = (n-1)s^2/\sigma_0^2 \sim \chi^2(n-1)$(当 $\sigma = \sigma_0$ 时).

正态总体方差的检验

- 单个正态总体方差的 χ^2 检验。$x_1, \ldots, x_n$ 是来自 $N(\mu, \sigma^2)$ 的样本，对方差可考虑一下三个检验问题：
  - $I: \quad H_0: \sigma^2 \leqslant \sigma_0^2 \quad$ vs $\quad H_1: \sigma^2 > \sigma_0^2$.
  - $II: \quad H_0: \sigma^2 \geqslant \sigma_0^2 \quad$ vs $\quad H_1: \sigma^2 < \sigma_0^2$.
  - $III: \quad H_0: \sigma^2 = \sigma_0^2 \quad$ vs $\quad H_1: \sigma^2 \neq \sigma_0^2$.
- 考虑统计统计量 $\chi^2 = (n-1)s^2/\sigma_0^2 \sim \chi^2(n-1)$(当 $\sigma = \sigma_0$ 时).
  - 对于检验问题 I，$W_I = \{\chi^2 \geqslant \chi^2_{1-\alpha}(n-1)\}$, $p_I = P(\chi^2 \geqslant \chi_0^2)$.
  - 对于检验问题 II，$W_{II} = \{\chi^2 \leqslant \chi^2_{\alpha}(n-1)\}$, $p_{II} = p(\chi^2 \leqslant \chi_0^2)$.
  - 对于检验问题 III，
    $W_{III} = \{\chi^2 \leqslant \chi^2_{\alpha/2}(n-1)\} \cup \{\chi^2 \geqslant \chi^2_{1-\alpha/2}(n-1)\}$.
    $p_{III} = 2\min\{P(\chi^2 \geqslant \chi_0^2), P(\chi^2 \leqslant \chi_0^2)\}$.

正态总体方的检验

- 两个正态总体方差比的 F 检验:

正态总体方的检验

- 两个正态总体方差比的 F 检验：$x_1, \ldots, x_m$ 是来自 $N(\mu_1, \sigma_1^2)$ 的样本，$y_1, \ldots, y_n$ 是来自 $N(\mu_1, \sigma_2^2)$ 的样本。

- 两个正态总体方差比的 F 检验：$x_1, \ldots, x_m$ 是来自 $N(\mu_1, \sigma_1^2)$ 的样本，$y_1, \ldots, y_n$ 是来自 $N(\mu_1, \sigma_2^2)$ 的样本。考虑以下三个假设检验问题：
  - $I: \quad H_0: \sigma_1^2 \leqslant \sigma_2^2 \quad \text{vs} \quad H_1: \sigma_1^2 > \sigma_2^2$;
  - $II: \quad H_0: \sigma_1^2 \geqslant \sigma_2^2 \quad \text{vs} \quad H_1: \sigma_1^2 < \sigma_2^2$;
  - $III: \quad H_0: \sigma_1^2 = \sigma_2^2 \quad \text{vs} \quad H_1: \sigma_1^2 \neq \sigma_2^2$.
- 考虑统计量 $F = s_x^2/s_y^2$，当 $\sigma_1 = \sigma_2$ 时，$F \sim F(m-1, n-1)$.
  - 对于检验问题 I，$W_I = \{F \geqslant F_{1-\alpha}(m-1, n-1)\}$, $p_I = P(F \geqslant F_0)$.
  - 对于检验问题 II，$W_{II} = \{F \leqslant F_{\alpha}(m-1, n-1)\}$, $p_{II} = P(F \leqslant F_0)$.
  - 对于检验问题 III，
    $W_{III} = \{F \leqslant F_{\alpha/2}(m-1, n-1), F \geqslant F_{1-\alpha/2}(m-1, n-1)\}$,
    $p_{III} = 2\min\{P(F \geqslant F_0), P(F \leqslant F_0)\}$.

例子

- 为比较正常成年男女所含红血球的差异，对某地区的男女进行了测量：
  - 男，156 名，样本均值 465.13，样本方差 54.80^2;
  - 女，74 名，样本均值 422.16，样本方差 49.20^2.
  - 假设红血球含量服从正态分布。男女所含的红血球量的均值有差异吗？

例子

- 为比较正常成年男女所含红血球的差异，对某地区的男女进行了测量：
  - 男，156 名，样本均值 465.13，样本方差 54.80^2;
  - 女，74 名，样本均值 422.16，样本方差 49.20^2.
  - 假设红血球含量服从正态分布。男女所含的红血球量的均值有差异吗？
- 现检验正态分布的方差是和否相等。
  $H_0: \sigma_1 = \sigma_2$, $vs\ H_1: \sigma_1 \neq \sigma_2$.
- $\alpha = 0.05$, $F = \frac{s_x^2}{s_y^2} = \frac{54.80^2}{49.20^2} = 1.24$, $F_{0.975}(155, 73) = 1.5$, $F_{0.025}(155, 73) = 0.68$. 不拒绝原假设。

例子

- 为比较正常成年男女所含红血球的差异，对某地区的男女进行了测量：
  - 男，156 名，样本均值 465.13，样本方差 54.80^2;
  - 女，74 名，样本均值 422.16，样本方差 49.20^2.
  - 假设红血球含量服从正态分布。男女所含的红血球量的均值有差异吗？
- 现检验正态分布的方差是和否相等。
  $H_0: \sigma_1 = \sigma_2$, vs $H_1: \sigma_1 \neq \sigma_2$.
- $\alpha = 0.05$, $F = \frac{s_x^2}{s_y^2} = \frac{54.80^2}{49.20^2} = 1.24$, $F_{0.975}(155, 73) = 1.5$, $F_{0.025}(155, 73) = 0.68$. 不拒绝原假设。
- 在方差相等的假设下，进行检验：$H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$.

例子

- 为比较正常成年男女所含红血球的差异，对某地区的男女进行了测量：
  - 男，156 名，样本均值 465.13，样本方差 54.80^2;
  - 女，74 名，样本均值 422.16，样本方差 49.20^2.
  - 假设红血球含量服从正态分布。男女所含的红血球量的均值有差异吗？
- 现检验正态分布的方差是和否相等。
  $H_0: \sigma_1 = \sigma_2$, vs $H_1: \sigma_1 \neq \sigma_2$.
- $\alpha = 0.05$, $F = \frac{s_x^2}{s_y^2} = \frac{54.80^2}{49.20^2} = 1.24$, $F_{0.975}(155, 73) = 1.5$, $F_{0.025}(155, 73) = 0.68$. 不拒绝原假设。
- 在方差相等的假设下，进行检验：$H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$.
- $t = 5.96$, $t_{0.975}(228) = 1.96 < 5.96$, 在拒绝域，拒绝 H_0. 即可认为有差异。

指数分布参数的假设检验

- $x_1, \ldots, x_n$ 是来自总体 $Exp(\frac{1}{\theta})$ 的样本，$p(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$，$x > 0$.

指数分布参数的假设检验

- $x_1, \ldots, x_n$ 是来自总体 $Exp(\frac{1}{\theta})$ 的样本，$p(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$，$x > 0$.考虑关于 θ 的检验问题:
    - $I:\quad H_0 : \theta \leqslant \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0.$
    - $II:\quad H_0 : \theta \geqslant \theta_0 \quad \text{vs} \quad H_1 : \theta < \theta_0.$
    - $III:\quad H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta \neq \theta_0.$

指数分布参数的假设检验

- $x_1, \ldots, x_n$ 是来自总体 $Exp(\frac{1}{\theta})$ 的样本，$p(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$, $x > 0$.考虑关于 θ 的检验问题:
  - I : $H_0 : \theta \leqslant \theta_0$ vs $H_1 : \theta > \theta_0$.
  - II : $H_0 : \theta \geqslant \theta_0$ vs $H_1 : \theta < \theta_0$.
  - III : $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$.
- 考虑统计量 $n\bar{x} = \sum_{i=1}^n x_i \sim Ga(n, \frac{1}{\theta_0})$(当 $\theta = \theta_0$ 时).

指数分布参数的假设检验

- $x_1, \dots, x_n$ 是来自总体 $Exp(\frac{1}{\theta})$ 的样本，$p(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$，$x > 0$.考虑关于 θ 的检验问题：
  - $I: \quad H_0 : \theta \leqslant \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0.$
  - $II: \quad H_0 : \theta \geqslant \theta_0 \quad \text{vs} \quad H_1 : \theta < \theta_0.$
  - $III: \quad H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta \neq \theta_0.$
- 考虑统计量 $n\bar{x} = \sum_{i=1}^n x_i \sim Ga(n, \frac{1}{\theta_0})$(当 $\theta = \theta_0$ 时).即有

$$\chi^2 = \frac{2n\bar{x}}{\theta_0} \sim \chi^2(2n) = Ga(\frac{2n}{2}, \frac{1}{2}).$$

  - 对于检验问题 I，$W_I = \{\chi^2 \geqslant \chi^2_{1-\alpha}(2n)\}$, $p_I = P(\chi^2 \geqslant \chi^2_0)$.
  - 对于检验问题 II，$W_{II} = \{\chi^2 \leqslant \chi^2_{\alpha}(2n)\}$, $p_{II} = p(\chi^2 \leqslant \chi^2_0)$.
  - 对于检验问题 III，$W_{III} = \{\chi^2 \leqslant \chi^2_{\alpha/2}(2n)\} \cup \{\chi^2 \geqslant \chi^2_{1-\alpha/2}(2n)\}$.
    $p_{III} = 2\min\{P(\chi^2 \geqslant \chi^2_0), P(\chi^2 \leqslant \chi^2_0)\}$.

大样本检验

- $x_1, \ldots, x_n$ 来自某总体 $F(x;\theta)$ 的样本，设总体均值为 θ，方差为 θ 的函数 $\sigma^2(\theta)$，

- $x_1, \ldots, x_n$ 来自某总体 $F(x;\theta)$ 的样本，设总体均值为 θ，方差为 θ 的函数 $\sigma^2(\theta)$，可考虑一下假设检验问题：
  - $I:\quad H_0: \theta \leqslant \theta_0 \quad$ vs $\quad H_1: \theta > \theta_0$.
  - $II:\quad H_0: \theta \geqslant \theta_0 \quad$ vs $\quad H_1: \theta < \theta_0$.
  - $III:\quad H_0: \theta = \theta_0 \quad$ vs $\quad H_1: \theta \neq \theta_0$.
- 考虑统计量

$$u = \frac{\sqrt{n}(\bar{x} - \theta_0)}{\sqrt{\sigma^2(\hat{\theta})}} \dot{\sim} N(0,1)(\text{当 } \theta = \theta_0 \text{ 时}),$$

其中 $\hat{\theta}$ 是 θ 的最大似然估计。
  - $W_I = \{u \geqslant u_{1-\alpha}\}$.
  - $W_{II} = \{u \leqslant u_\alpha\}$.
  - $W_{III} = \{|u| \geqslant u_{1-\alpha/2}\}$.

例子：抛硬币

- 给你一块硬币，如何决定改硬币是“公平”的？

例子：抛硬币

- 给你一块硬币，如何决定改硬币是“公平”的？
- $H_0:\ p=\frac{1}{2}\ vs\ H_1:\ p\neq\frac{1}{2}$.

例子：抛硬币

- 给你一块硬币，如何决定改硬币是“公平”的？
- $H_0: \ p = \frac{1}{2} \ vs \ H_1: \ p \neq \frac{1}{2}$.
- 统计量：$u = \frac{\bar{x} - \frac{1}{2}}{\sqrt{\bar{x}(1-\bar{x})/n}}$.

例子：抛硬币

- 给你一块硬币，如何决定改硬币是“公平”的？
- $H_0: \ p = \frac{1}{2}$ vs $H_1: \ p \neq \frac{1}{2}$.
- 统计量：$u = \frac{\bar{x} - \frac{1}{2}}{\sqrt{\bar{x}(1-\bar{x})/n}}$.
- 给定显著水平 α, 拒绝域：$W_{III} = \{|u| \geqslant u_{1-\alpha/2}\}$.

例子：抛硬币

- 给你一块硬币，如何决定改硬币是“公平”的？
- $H_0: \ p = \frac{1}{2}$ vs $H_1: \ p \neq \frac{1}{2}$.
- 统计量：$u = \frac{\bar{x}-\frac{1}{2}}{\sqrt{\bar{x}(1-\bar{x})/n}}$.
- 给定显著水平 α, 拒绝域：$W_{III} = \{|u| \geqslant u_{1-\alpha/2}\}$.
- 近似 p-值?

例子：抛硬币

- 给你一块硬币，如何决定改硬币是“公平”的？
- $H_0: \ p = \frac{1}{2}$ vs $H_1: \ p \neq \frac{1}{2}$.
- 统计量：$u = \frac{\bar{x} - \frac{1}{2}}{\sqrt{\bar{x}(1-\bar{x})/n}}$.
- 给定显著水平 α，拒绝域：$W_{III} = \{|u| \geqslant u_{1-\alpha/2}\}$.
- 近似 p-值？$P(|u| \geqslant u_0)$.

- 如果 $x_1, \ldots, x_n$ 来自某总体 $p(x;\theta)$, θ 不是总体的均值。有什么办法进行假设检验?