

第八章 假设检验

 关键词:

假设检验

正态总体参数的假设检验

拟合优度检验



8.1 假设检验

统计推断的另一类重要问题是假设检验问题。它包括

- (1) 已知总体分布的形式，需对其中的未知参数给出假设检验。—参数检验
- (2) 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。—非参数检验

(一) 问题的提出

例1 设某种清漆的9个样品，其干燥时间（以小时计）分别为：

6.0 5.7 5.5 6.5 7.0 5.8 5.2 6.1 5.0

根据以往经验，干燥时间的总体服从正态分布 $N(6.0, 0.36)$ ，现根据样本检验均值是否与以往有显著差异？

例2 一种摄影药品被其制造商声称其贮藏寿命是均值**180**天、标准差不多于**10**天的正态分布。某位使用者担心标准差可能超过**10**天。他随机选取**12**个样品并测试，得到样本标准差为**14**天。根据样本有充分证据证明标准差大于**10**天吗？

例3 孟德尔遗传理论断言，当两个品种的豆杂交时，圆的和黄的、起皱的和黄的、圆的和绿的、起皱的和绿的豆的频数将以比例 $9:3:3:1$ 发生。在检验这个理论时，孟德尔分别得到频数 315 、 101 、 108 、 32 、这些数据提供充分证据拒绝该理论吗？

■ 假设:

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:

$$H_0: \theta \geq \theta_0, \quad H_1: \theta < \theta_0 \text{ (左边检验)}$$

$$H_0: \theta \leq \theta_0, \quad H_1: \theta > \theta_0 \text{ (右边检验)}$$

$$H_0: \theta = \theta_0, \quad H_1: \theta \neq \theta_0 \text{ (双边检验)}$$

(二) 检验统计量和拒绝域

■ 对例1的统计分析

设清漆的干燥时间为 X ，由已知 $X \sim N(\mu, \sigma^2)$ ，其中 $\sigma^2=0.36$ ，考虑有关参数 μ 的假设：

$$H_0: \mu = 6.0, H_1: \mu \neq 6.0 \text{ (双边检验)}$$

因样本均值 \bar{X} 是 μ 的无偏估计， \bar{X} 的取值大小反映了 μ 的取值大小，当原假设成立时， $|\bar{X} - 6.0|$ 取值应偏小。

检验规则：

当 $|\bar{X} - 6.0| \geq C$ 时，拒绝原假设 H_0 ；

当 $|\bar{X} - 6.0| < C$ 时，接受原假设 H_0 ，

其中 C 是待定的常数。

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**。

上述例子中，可取检验统计量为 \bar{X} (或 $\bar{X} - 6.0$)，拒绝域为

$$W = \{(X_1, \dots, X_n) : |\bar{X} - 6.0| \geq C\}$$

如何确定临界值 C ?

(三) 两类错误

- 由于样本的随机性，任一检验规则在应用时，都有可能发生错误的判断。

	原假设为真	原假设不真
根据样本拒绝原假设	第I类错误	正确
根据样本接受原假设	正确	第II类错误

第I类错误：拒绝真实的原假设(弃真)

第II类错误：接受错误的原假设(取伪)

犯第I类错误的概率

$$\begin{aligned}\alpha = P\{\text{第I类错误}\} &= P\{\text{拒绝}H_0 \mid H_0\text{是真的}\} \\ &= P_{H_0}\{\text{拒绝}H_0\}\end{aligned}$$

犯第II类错误的概率

$$\begin{aligned}\beta = P\{\text{第II类错误}\} &= P\{\text{接受}H_0 \mid H_0\text{是假的}\} \\ &= P_{H_1}\{\text{接受}H_0\}\end{aligned}$$

例1中，犯第I类错误的概率

$$\begin{aligned}\alpha(C) &= P\{\text{拒绝}H_0 \mid H_0\text{是真的}\} \\ &= P\{|\bar{X} - 6.0| \geq C \mid \mu = 6.0\} \\ &= P\left\{\frac{|\bar{X} - 6.0|}{\sigma / \sqrt{n}} \geq \frac{C}{\sigma / \sqrt{n}} \mid \mu = 6.0\right\}\end{aligned}$$

由于当 H_0 成立时，即 $\mu = 6.0$ 时， $\frac{\bar{X} - 6.0}{\sigma / \sqrt{n}} \sim N(0,1)$ ，因此

$$\alpha(C) = 2 - 2\Phi\left(\frac{C}{\sigma / \sqrt{n}}\right)$$

犯第II类错误的概率

$$\beta(C) = P\{\text{接受} H_0 \mid H_0 \text{是假的}\}$$

$$= P\{|\bar{X} - 6.0| < C \mid \mu \neq 6.0\}$$

$$= P\{6.0 - C < \bar{X} < 6.0 + C \mid \mu \neq 6.0\}$$

$$= P\left\{\frac{6.0 - C - \mu}{\sigma / \sqrt{n}} < \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} < \frac{6.0 + C - \mu}{\sigma / \sqrt{n}} \mid \mu \neq 6.0\right\}$$

$$= \Phi\left\{\frac{6.0 + C - \mu}{\sigma / \sqrt{n}}\right\} - \Phi\left\{\frac{6.0 - C - \mu}{\sigma / \sqrt{n}}\right\}, \quad \mu \neq 6.0$$

显然，犯第I类错误的概率 $\alpha(C)$ 关于 C 是单调减函数，而犯第II类错误的概率 $\beta(C)$ 关于 C 是单调增函数.

在给定的样本量下 n ，不可能找界值 C ，使得 $\alpha(C)$ 和 $\beta(C)$ 都尽可能小.

犯两类错误的概率相互制约！

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.

在例1中，若取显著水平 $\alpha = 0.05$ ，则有

$$\alpha(C) = 2 - 2\Phi\left(\frac{C}{\sigma / \sqrt{n}}\right) \leq 0.05$$

计算得

$$C \geq z_{0.025} \sigma / \sqrt{n} = 1.96 \times 0.6 / 3 = 0.392$$

由于犯第II类错误的概率 $\beta(C)$ 关于 C 单调增函数，根据Neyman-Pearson原则，应取

$$C = 0.392$$

因此拒绝域为

$$W = \{(X_1, \dots, X_9) : |\bar{X} - 6.0| \geq 0.392\}$$

根据实际样本资料，得 $\bar{x} = 6.4$, 有

$$|\bar{x} - 6.0| = 0.4 > 0.392,$$

样本落入拒绝域.

我们有95%的把握拒绝原假设 H_0 , 即认为油漆干燥时间与以往有显著差异.

根据上述检验规则，犯第I类错误的概率

$$\alpha(0.392) = 0.05 = \alpha$$

犯第II类错误的概率

$$\begin{aligned}\beta(0.392) &= \Phi\left\{\frac{6.0 + 0.392 - \mu}{0.6/\sqrt{9}}\right\} - \Phi\left\{\frac{6.0 - 0.392 - \mu}{0.6/\sqrt{9}}\right\} \\ &= \Phi\left\{\frac{6.392 - \mu}{0.2}\right\} - \Phi\left\{\frac{5.608 - \mu}{0.2}\right\}, \quad \mu \neq 6.0\end{aligned}$$

例如，当 $\mu=5.4$ 时，犯第II类错误的概率

$$\begin{aligned}\beta &= \Phi \left\{ \frac{6.392 - 5.4}{0.2} \right\} - \Phi \left\{ \frac{5.608 - 5.4}{0.2} \right\} \\ &= \Phi (4.96) - \Phi (1.04) \\ &\approx 1.00 - 0.85 \\ &= 0.15\end{aligned}$$

（四） P 值与统计显著性

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。

例1中 P 值的计算:

$$\begin{aligned} P_{-} &= P_{H_0} (|\bar{X} - 6.0| \geq |6.4 - 6.0|) = P_{H_0} \left(\left| \frac{\bar{X} - 6.0}{0.2} \right| \geq 2 \right) \\ &= 2 - 2\Phi(2) = 0.046 \end{aligned}$$

$|\bar{X} - 6.0| \geq 0.4$ 是小概率事件.

作出拒绝原假设的判断.

用 P 值计算, 不仅简单, 而且能知道概率大小。

两种检验方式比较：

根据拒绝域的形式： $|\bar{X} - 6.0| \geq C$,

方式一：由犯第I类错误的概率

$$\alpha(C) = P\{|\bar{X} - 6.0| \geq C | \mu = 6.0\} \leq 0.05$$

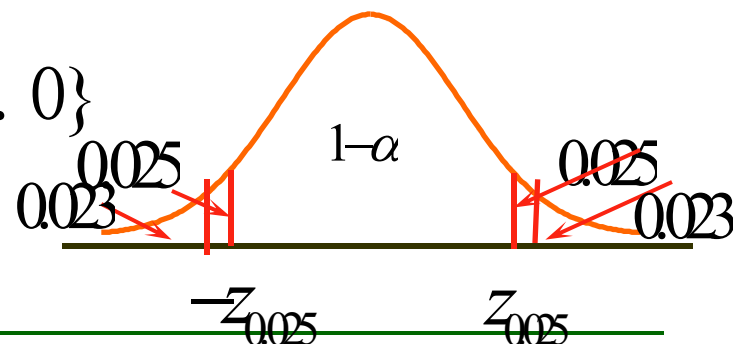
确定 $C = 0.396$ ，由样本值计算得，

$|6.4 - 6.0| \geq 0.396$ ，落在拒绝域内，拒绝原假设。

方式二：计算：

$$P_- = P\{|\bar{X} - 6.0| \geq |6.4 - 6.0| | \mu = 6.0\}$$

$$= 2 - 2\Phi(2) = 0.046 < 0.05$$



拒绝原假设。

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。
- (4) 根据实际样本观测值作出判断。

(3') 计算检验统计量的观测值与 P 值;

(4') 根据给定的显著水平 α , 作出判断.

8.2 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} 和 S^2 分别为样本均值和方差,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

$H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$, 其中 μ_0 是已知的常数

由前一节讨论知, 可取检验统计量为

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$$

在 H_0 为真时, $Z \sim N(0,1)$.

根据Neyman-Pearson原则, 检验的拒绝域为

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

P_- 值的计算

对给定的样本观察值 x_1, \dots, x_n , 记检验统计量 Z 的取值为

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}, \text{ 则有}$$

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

当 P_- 小于显著水平 α 时, 拒绝原假设,
否则, 接受原假设.

左边假设问题

$H_0 : \mu \geq \mu_0, H_1 : \mu < \mu_0$, 其中 μ_0 是已知的常数

检验统计量仍取为

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$$

拒绝域形式为

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq C.$$

根据Neyman-Pearson原则，可得检验的拒绝域为

$$W = \left\{ Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha \right\}$$

事实上，犯第I类错误的概率

$$\begin{aligned} P\{\text{拒绝}H_0 | H_0 \text{是真的}\} &= P\left\{ \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq C \mid \mu \geq \mu_0 \right\} \\ &= P\left\{ \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq C + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}} \mid \mu \geq \mu_0 \right\} = \Phi\left(C + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right) \leq \alpha \\ &\Rightarrow C = -z_\alpha \end{aligned}$$

P_- 值的计算

$$\begin{aligned}P_- &= \sup_{\mu \geq \mu_0} P_{H_0} \{Z \leq z_0\} = \sup_{\mu \geq \mu_0} P\left\{\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \mid \mu \geq \mu_0\right\} \\&= \sup_{\mu \geq \mu_0} P\left\{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}} \mid \mu \geq \mu_0\right\} \\&= P\left\{\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \mid \mu = \mu_0\right\} = P_{H_0} \{Z \leq z_0 \mid \mu = \mu_0\} \\&= \Phi(z_0).\end{aligned}$$

对给定的样本观察值 x_1, \dots, x_n , $z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$.

思考题：比较 双边假设问题

$H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$ ，其中 μ_0 是已知的常数

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

与

左边假设问题

$H_0 : \mu \geq \mu_0, H_1 : \mu < \mu_0$ ，其中 μ_0 是已知的常数

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

你能写出右边假设问题检验的拒绝域吗？

右边假设问题

$H_0 : \mu \leq \mu_0, H_1 : \mu > \mu_0$, 其中 μ_0 是已知的常数

检验统计量仍取为

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}.$$

根据Neyman-Pearson原则，可得拒绝域为

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

P_- 值的计算

$$\begin{aligned} P_- &= \sup_{\mu \leq \mu_0} P_{H_0} \{Z \geq z_0\} \\ &= P\{Z \geq z_0 \mid \mu = \mu_0\} \\ &= 1 - \Phi(z_0). \end{aligned}$$

对给定的样本观察值 x_1, \dots, x_n , $z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$.

(2) σ^2 未知时—— t 检验

双边假设问题

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

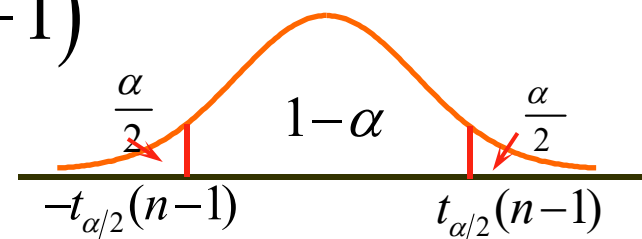
由于 σ^2 未知，故不能用 $\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ 来确定拒绝域。

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。

即检验拒绝域的形式为 $\frac{|\bar{X} - \mu_0|}{S/\sqrt{n}} \geq k$.

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



根据Neyman-Pearson原则, 可得拒绝域为

$$T = \frac{|\bar{X} - \mu_0|}{S/\sqrt{n}} \geq t_{\alpha/2}(n-1)$$

P_- 值的计算

对给定的样本观察值 x_1, \dots, x_n , 记检验统计量 T 的取值为

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \text{ 则有}$$

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

当 $P_- \leq \alpha$ 时, 拒绝原假设,
否则, 接受原假设.

左边假设问题

$H_0 : \mu \geq \mu_0, H_1 : \mu < \mu_0$, 其中 μ_0 是已知的常数

拒绝域为

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

P_- 值为

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

右边假设问题

$H_0 : \mu \leq \mu_0, H_1 : \mu > \mu_0$, 其中 μ_0 是已知的常数

拒绝域为

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

P_- 值为

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

例 某种元件的寿命 X （以小时记）服从正态分布 $N(\mu, \sigma^2)$, μ, σ 均未知。现测得16只元件的寿命如下：

159 280 101 212 224 379 179 264
222 362 168 250 149 260 485 170

问是否有理由认为元件的平均寿命大于225（小时）？
（取显著性水平为0.05）

解：按题意需检验

$$H_0 : \mu \leq \mu_0 = 225, \quad H_1 : \mu > 225.$$

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

$$n = 16, t_{0.05}(15) = 1.7531. \quad \bar{x} = 241.5, s = 98.7259$$

计算得： $t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = 0.6685 < 1.7531 = t_{0.05}(15).$

没有落在拒绝域内，故不能拒绝原假设，
认为元件的平均寿命不大于**225**小时。

由 $Excel$ 可计算 P_- 值为

$$P_- = P_{H_0} \{T \geq t_0\} = P \{t(15) \geq 0.6685\} \approx 0.257 > 0.05$$

因此接受原假设，即认为元件的平均寿命不大于225小时。

判断结果与前面一致！

- 问：若将原假设和备择假设互换，即考虑左边检验

$$H_0 : \mu \geq \mu_0 = 225, \quad H_1 : \mu < 225.$$

- 检验结果怎么样？请给出合理的解释。

- 一般地，在有关参数的假设检验中，备择假设是我们根据样本资料希望得到支持的假设。

例3 要求某种元件的平均使用寿命不得低于1000小时，生产者从一批这种元件中随机抽取25件，测得其平均寿命为950小时，标准差为100小时。已知这批元件的寿命服从正态分布。试在显著性水平0.05下确定这批元件是否合格？

解：按题意需检验

$$H_0 : \mu \geq \mu_0 = 1000, \quad H_1 : \mu < 1000.$$

$$\text{拒绝域为: } t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} < -t_{\alpha}(n-1).$$

$$n = 25, t_{0.05}(24) = 1.7109. \quad \bar{x} = 950, s = 100$$

$$\text{计算得: } t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = -2.5 < -1.7109 = -t_{0.05}(24).$$

t 落在拒绝域内，故拒绝原假设，

认为这批元件的平均寿命小于**1000**小时，不合格。

P_- 值为

$$P_- = P_{H_0} \{T \leq t_0\} = P\{t(24) \leq -2.5\} \approx 0.000866 < 0.05$$

因此拒绝原假设，判断结果与前面一致！

(二) 成对数据的 t 检验

成对数据问题在7.4节中已作过介绍.

成对样本 $(X_1, Y_1), \dots, (X_n, Y_n),$

设差值 $D_i = X_i - Y_i, i = 1, \dots, n.$

可以看成来自正态总体 $N(\mu_d, \sigma_d^2)$ 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

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

转化为单个正态总体的均值的假设检验。

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

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

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

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

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

例4: 为了试验两种不同谷物种子的优劣, 选取了十块土质不同的土地, 并将每块土地分为面积相同的两部分, 分别种植这两种种子。设在每块土地的两部分人工管理等条件完全一样。下面给出各块土地上的产量。

土地	1	2	3	4	5	6	7	8	9	10
种子A(x_i)	23	35	29	42	39	29	37	34	35	28
种子B(y_i)	26	39	35	40	38	24	36	27	41	27
$d_i = x_i - y_i$	-3	-4	-6	2	1	5	1	7	-6	1

问: 以这两种种子种植的谷物产量是否有显著的差异 (取显著性水平为0.05)?

解：检验假设 $H_0 : \mu_D = 0, H_1 : \mu_D \neq 0$

分别将 D_1, D_2, \dots, D_n 的样本均值和样本方差记为 \bar{D}, S_D^2 ,

$$\text{拒绝域为: } \frac{|\bar{D}|}{S_D / \sqrt{n}} \geq t_{\alpha/2}(n-1),$$

$n = 10$, 查表得: $t_{0.025}(9) = 2.2622, \bar{d} = -0.2, s_d = 4.442,$

$$\text{计算得: } \frac{|\bar{d}|}{s_d / \sqrt{n}} = 0.142 < 2.2622$$

不拒绝原假设 H_0 ,

认为两种种子的产量没有显著差异。

$$P_- = P_{H_0} \{ |T| \geq |t_0| \} = 2P \{ t(n-1) \geq 0.142 \} = 0.89.$$

■ 在Excel中的实现-----TTEST函数

本例的分析步骤如下：

(1) 将两品种种子的产量数据输入Excel表中，设数据区域分别为A1:A10和B1:B10；

(2) 下拉菜单“插入”选项卡=>单击“函数”=>
在类别的下拉式菜单中选择“统计”=>选“TTEST”；

(3) 在“**Array1**”文本框中输入“**A1:A10**”，在“**Array2**”文本框中输入“**B1:B10**”，“**Tails**”文本框中输入“**2**”

(“**1**”代表单尾概率，“**2**”代表双尾概率)， “**Type**”文本框中输入“**1**”（ “**1**”代表成对数据的**t**检验，“**2**”代表方差齐性的两样本**t**检验，“**3**”代表异方差的两样本**t**检验）；

(4) 点击**Enter**键，即显示**P_**值为“**0.889921**”，因此认为两品种种子产量没有显著差异。

（三）有关参数 σ^2 的假设检验

（不妨设 μ 未知）

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

其中 σ_0^2 是已知常数。此时 σ^2 的无偏估计量为样本方差 S^2 ,

且有
$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

因此可取检验统计量为 $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$

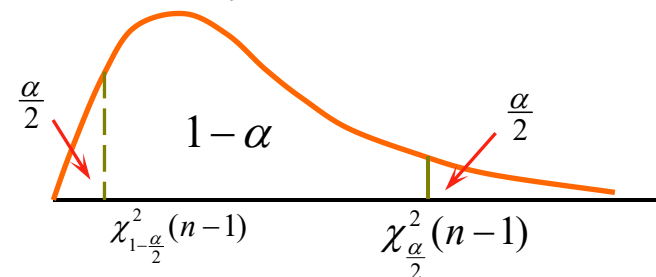
检验拒绝域形式为：

$$\frac{(n-1)S^2}{\sigma_0^2} \leq k_1, \text{ 或 } \frac{(n-1)S^2}{\sigma_0^2} \geq k_2.$$

在原假设成立时, $\frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$

$P\{\text{拒绝}H_0 \mid \text{当}H_0\text{为真}\}$

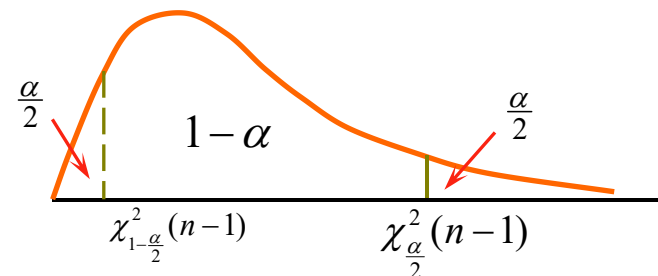
$$= P_{\sigma_0^2} \left\{ \frac{(n-1)S^2}{\sigma_0^2} \leq k_1, \text{ 或 } \frac{(n-1)S^2}{\sigma_0^2} \geq k_2 \right\} = \alpha$$



为计算方便，习惯上取

$$P \left\{ \frac{(n-1)S^2}{\sigma_0^2} \leq k_1 \right\} = \frac{\alpha}{2}, P \left\{ \frac{(n-1)S^2}{\sigma_0^2} \geq k_2 \right\} = \frac{\alpha}{2}$$

于是有 $k_1 = \chi^2_{1-\alpha/2}(n-1)$, $k_2 = \chi^2_{\alpha/2}(n-1)$ 。



拒绝域为：

$$\frac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{1-\alpha/2}(n-1), \text{ 或 } \frac{(n-1)S^2}{\sigma_0^2} \geq \chi^2_{\alpha/2}(n-1)$$

----- χ^2 检验法

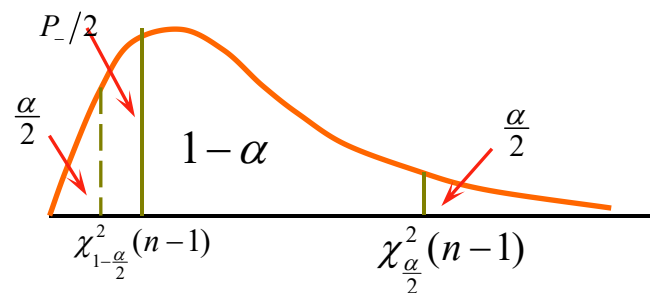
P_- 值计算:

$$\text{设 } p = P_{\sigma_0^2} \left\{ \frac{(n-1)S^2}{\sigma_0^2} \leq \frac{(n-1)s^2}{\sigma_0^2} \right\} = P \left\{ \chi^2(n-1) \leq \chi_0^2 \right\},$$

其中, 对样本观察值 x_1, \dots, x_n , $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

$$P_- = 2 \min(p, 1-p)$$



当 $P_- \leq \alpha$, 拒绝原假设,

当 $P_- > \alpha$, 接受原假设.

类似地，对于左边检验

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

拒绝域为：

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

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

当 $P_- \leq \alpha$ ，拒绝原假设，

当 $P_- > \alpha$ ，接受原假设.

类似地，对于右边检验

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

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

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

当 $P_- \leq \alpha$ ，拒绝原假设，

当 $P_- > \alpha$ ，接受原假设.

例6: 一个园艺科学家正在培养一个新品种的苹果,这种苹果除了口感好和颜色鲜艳以外,另一个重要特征是单个重量差异不大(对照品种的方差 $\sigma^2=7$)。为了评估新苹果,她随机挑选了25个测试重量(单位: 克),其样本方差为 $S^2 = 4.25$ 。
在 $\alpha=0.05$ 下检验新品种是否比对照品种方差小?

从资料来看想要支持的结论是: **新品种苹果的重量差异小**

解： $H_0 : \sigma^2 \geq 7, \quad H_1 : \sigma^2 < 7$

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

查表得： $\chi_{0.95}^2(24) = 13.848,$

计算得： $\frac{(25-1) \times 4.25}{7} = 14.57 > 13.848$

不拒绝原假设，即认为新品种的方差并不比对照组的小。

计算 $P_- = P\{\chi^2(24) \leq 14.57\} = 0.097032 > 0.05$

作出同样判断。

例2 一种摄影药品被其制造商声称其贮藏寿命是均值**180**天、标准差不多于**10**天的正态分布。某位使用者担心标准差可能超过**10**天。他随机选取**12**个样品并测试，得到样本标准差为**14**天。根据样本有充分证据证明标准差大于**10**天吗？
(取显著水平为**0.05**)

解： $H_0 : \sigma \leq 10, \quad H_1 : \sigma > 10$

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

查表得： $\chi_{0.05}^2(11) = 19.675,$

计算得： $\frac{(12-1) \times 14^2}{10^2} = 21.56 > 19.675.$

拒绝原假设，即有95%把握认为标准差大于10天。

$$P_- = P\{\chi^2(11) \leq 21.56\} = 0.028 < 0.05.$$

8.3 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

此时检验统计量 $\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$,

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

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

1. 当 σ_1^2 和 σ_2^2 已知时

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

则检验拒绝域为: $|Z| \geq z_{\alpha/2}$ --- z检验

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

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

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

首先利用合样本给出参数 σ^2 的无偏估计量

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

由情形1讨论知，可取检验统计量为：

$$T = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

根据抽样分布定理6.3.4知，在原假设成立时，

$$T \sim t(n_1 + n_2 - 2).$$

检验拒绝域为：

$$T = \frac{|\bar{X} - \bar{Y}|}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \geq t_{\alpha/2}(n_1 + n_2 - 2)$$

P_- 值为

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

——— 两样本精确 t 检验

其中

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

3. 当 $\sigma_1^2 \neq \sigma_2^2$ 且未知时

分别以两样本方差 S_1^2, S_2^2 作为 σ_1^2 和 σ_2^2 的无偏估计, 取检验统计量为

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

(i) 当两样本容量都充分大时，根据大数定律和中心极限定理，当原假设成立时，统计量 T 近似服从标准正态分布 $N(0, 1)$.

检验的拒绝域为 $|T| \geq z_{\alpha/2}$

P_- 值为 $P_- = P_{H_0} \{ |T| \geq |t_0| \} = 2P \{ Z \geq |t_0| \},$

其中 $Z \sim N(0, 1), t_0 = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$

(ii) 对于小样本情形，原假设成立时，统计量 T 近似服从 t 分布，自由度为

$$k = \min(n_1 - 1, n_2 - 1),$$

或更精确的近似自由度

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

则检验的拒绝域为 $|T| \geq t_{\alpha/2}(k)$

P_- 值为 $P_- = P_{H_0} \{|T| \geq |t_0|\} = 2P\{t(k) \geq |t_0|\}.$

-----两样本近似 t 检验

类似地， 我们可以给出左边检验

$$H_0 : \mu_1 \geq \mu_2, H_1 : \mu_1 < \mu_2$$

和右边检验

$$H_0 : \mu_1 \leq \mu_2, H_1 : \mu_1 > \mu_2$$

在上述三种情形下的检验规则。

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

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

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

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

右边检验 $H_0 : \mu_1 \leq \mu_2, H_1 : \mu_1 > \mu_2$

的拒绝域为：
$$T = \frac{(\bar{X} - \bar{Y})}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \geq t_\alpha(n_1 + n_2 - 2)$$

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

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

思考题:

根据前面理论给出下列假设问题的检验.

$$H_0 : \mu_1 - \mu_2 = \delta, H_1 : \mu_1 - \mu_2 \neq \delta. \quad (\delta \text{为已知常数})$$

$$H_0 : \mu_1 - \mu_2 \geq \delta, H_1 : \mu_1 - \mu_2 < \delta. \quad (\delta \text{为已知常数})$$

$$H_0 : \mu_1 - \mu_2 \leq \delta, H_1 : \mu_1 - \mu_2 > \delta. \quad (\delta \text{为已知常数})$$

例7：某厂使用两种不同的原料A, B生产同一类型产品。各在一周的产品中取样分析。

取用原料A生产的样品220件，测得平均重量为2.46（公斤），样本标准差 $s=0.57$ （公斤）。

取用原料B生产的样品205件，测得平均重量为2.55（公斤），样本标准差为0.48（公斤）。

设两样本独立，来自两个方差相同的独立正态总体。问在水平0.05下能否认为用原料B的产品平均重量 μ_2 较用原料A的产品平均重量 μ_1 为大。

解：检验假设 $H_0: \mu_1 \geq \mu_2, H_1: \mu_1 < \mu_2$

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

$$n_1 = 220, \bar{x} = 2.46, s_1 = 0.57; n_2 = 205, \bar{y} = 2.55, s_2 = 0.48$$

$$t_{0.05}(423) \approx z_{0.05} = 1.645, s_w = 0.535, \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 0.097$$

$$\text{计算得: } \frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = -1.733 < -1.645, \text{ 从而拒绝原假设。}$$

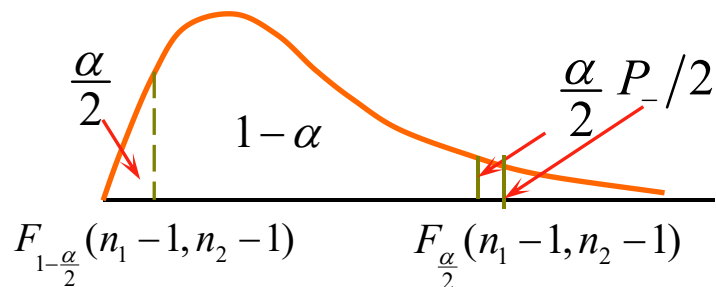
$$P_- = P\{t(423) \leq -1.733\} \approx \Phi(-1.733) = 0.042.$$

(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: $F \leq F_{1-\alpha}(n_1-1, n_2-1)$

$P_- = P(F \leq f_0)$, 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: $F \geq F_\alpha(n_1 - 1, n_2 - 1)$

$P_- = P(F \geq f_0)$, 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.

例7：两台机床生产同一个型号的滚珠，从甲机床生产的滚珠中抽取8个，从乙机床生产的滚珠中抽取9个，测得这些滚珠的直径(毫米)如下：

甲机床 15.0 14.8 15.2 15.4 14.9 15.1
15.2 14.8

乙机床 15.2 15.0 14.8 15.1 14.6 14.8
15.1 14.5 15.0

设两机床生产的滚珠直径分别为 X, Y ,
且 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$

(1) 检验假设 $H_0: \sigma_1^2 = \sigma_2^2, H_1: \sigma_1^2 \neq \sigma_2^2 (\alpha=0.1)$;

(2) 检验假设 $H_0: \mu_1 \leq \mu_2, H_1: \mu_1 > \mu_2 (\alpha=0.1)$;

(3) 检验假设 $H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2 (\alpha=0.1)$ 。

解：(1) 当 μ_1, μ_2 未知时，检验 $H_0: \sigma_1^2 = \sigma_2^2, H_1: \sigma_1^2 \neq \sigma_2^2$

的拒绝域为： $\frac{S_1^2}{S_2^2} \leq F_{1-\alpha/2}(n_1-1, n_2-1)$, 或 $\frac{S_1^2}{S_2^2} \geq F_{\alpha/2}(n_1-1, n_2-1)$

查表得： $F_{0.05}(7, 8) = 3.50, F_{0.95}(7, 8) = \frac{1}{F_{0.05}(8, 7)} = \frac{1}{3.73} = 0.268$

本题中 $n_1 = 8, \bar{x} = 15.05, S_1^2 = 0.0457; n_2 = 9, \bar{y} = 14.9, S_2^2 = 0.0575$

计算得： $0.268 < \frac{S_1^2}{S_2^2} = 0.795 < 3.50$

不拒绝原假设，故认为方差没有显著差异。

$P_- = 2P(F(7, 8) \leq 0.795) = 0.775 > 0.05$. 接受原假设.

$$n_1 = 8, \bar{x} = 15.05, S_1^2 = 0.0457; n_2 = 9, \bar{y} = 14.9, S_2^2 = 0.0575$$

$$(2) H_0: \mu_1 \leq \mu_2, H_1: \mu_1 > \mu_2$$

$$\text{的拒绝域为: } \frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \geq t_{\alpha}(n_1 + n_2 - 2)$$

$$t_{0.1}(15) = 1.3406, S_w = 0.228, \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 0.486$$

$$\text{计算得: } \frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = 1.354 > 1.3406, \text{ 从而拒绝原假设。}$$

$$P_- = P\{t(15) \geq 1.354\} = 0.098 < 0.1.$$

$$n_1 = 8, \bar{x} = 15.05, S_1^2 = 0.0457; n_2 = 9, \bar{y} = 14.9, S_2^2 = 0.0575$$

$$(3) H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$$

$$\text{的拒绝域为: } \frac{|\bar{X} - \bar{Y}|}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \geq t_{\frac{\alpha}{2}}(n_1 + n_2 - 2)$$

$$\text{计算得: } \frac{|\bar{X} - \bar{Y}|}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = 1.354 < t_{0.05}(15) = 1.7531, \text{ 从而接受原假设。}$$

$$P_- = 2P\{t(15) \geq 1.354\} = 0.196 > 0.1.$$

■ 在Excel中的实现----**FTSET**函数和**TTEST**函数

利用**FTSET**函数作方差齐性检验，再利用**TTEST**函数进行两样本的均值比较。

本例的分析步骤如下：

(1) 将两组数据输入**Excel** 表中，设数据区域分别为
A1:A8和**B1:B9**；

(2) 下拉菜单“插入”选项卡=>单击“函数”=>
在类别的下拉式菜单中选择“统计”=>选“**FTEST**”；

(3) 在“**Array1**”文本框中输入“**A1:A8**”，在
“**Array2**”文本框中输入“**B1:B9**”，
并点击**Enter**键，即显示**P_**值为“**0.7752**”，
因此认为两总体方差相同.

(4) 重新下拉菜单“插入”选项卡=>单击“函数”=>在类别的下拉式菜单中选择“统计”=>选“**TTEST**”;

(5) 在“**Array1**”文本框中输入“**A1:A8**”，在“**Array2**”文本框中输入“**B1:B9**”，“**Tails**”文本框中输入“1”（“1”代表单尾概率，“2”代表双尾概率），“**Type**”文本框中输入“2”（“1”代表成对数据的t检验，“2”代表方差齐性的两样本t检验，“3”代表异方差的两样本t检验）；

(6) 点击**Enter**键，即显示**P_**值为“**0.0979**”，因此在显著水平为**0.1**下，拒绝原假设 $H_0: \mu_1 \leq \mu_2$.

(7) 若在步骤 (5) 中的“**Tails**”文本框中输入“**2**”，并点击**Enter**键，即显示**P_**值为“**0.19587**”，因此在显著水平**0.1**下，接受原假设 $H_0: \mu_1 = \mu_2$.

8.4 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。

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

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

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

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

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.

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.

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 $\theta_0 \geq \hat{\theta}_U$.

单侧置信限与单边假设检验的关系：

(1) 若 $\hat{\theta}_L$ 是 θ 的置信水平为 $1 - \alpha$ 的单侧置信下限，
则当 $\theta_0 \geq \hat{\theta}_L$ 时，接受右边检验 $H_0 : \theta \leq \theta_0, H_1 : \theta > \theta_0$
中的原假设 H_0 ，反之，拒绝原假设。

(2) 若 $\hat{\theta}_U$ 是 θ 的置信水平为 $1 - \alpha$ 的单侧置信上限，
则当 $\theta_0 \leq \hat{\theta}_U$ 时，接受左边检验 $H_0 : \theta \geq \theta_0, H_1 : \theta < \theta_0$
中的原假设 H_0 ，反之，拒绝原假设。

例：为了研究某种止痛药的副作用，调查了服用该种止痛药的440名患者，发现有23名出现了“反症状”，那么是否有足够的理由说明在服用该种止痛药的病人中，出现“反症状”的比例低于10%？

解 这是有关二点分布 $B(1; p)$ 中参数 p 的检验问题.

由题意知, 考虑左边检验

$$H_0 : p \geq 0.1, \quad H_1 : p < 0.1$$

根据7.5节有关二点分布 $B(1; p)$ 中参数 p 的区间估计的理论, 得 p 的置信水平为 $1 - \alpha$ 的单侧置信上限为

$$\hat{p}_U = \frac{1}{2a} \left(-b + \sqrt{b^2 - 4ac} \right)$$

其中 $a = n + z_{\alpha}^2$, $b = -(2n\bar{X} + z_{\alpha}^2)$, $c = n\bar{X}^2$.

将已知资料 $n = 440$, $\bar{x} = 23 / 440 = 0.0523$,
并取 $\alpha = 0.05$, 代入计算得 $\hat{p}_U \approx 0.0726$

由于 $p_0 = 0.1 > \hat{p}_U = 0.0726$,因此, 作出拒绝原假设的判断, 即认为服用该种止痛药的病人中, 出现“反症状”的比例低于10%.

正态总体均值、方差的置信区间与假设检验

	待估参数	原假设	枢轴量	检验统计量	分布	置信区间	拒绝域
一个正态总体	μ (σ^2 已知)	$\mu = \mu_0$ (σ^2 已知)	$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$	$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$N(0, 1)$	$\frac{ \bar{X} - \mu }{\sigma/\sqrt{n}} < z_{\alpha/2}$	$\frac{ \bar{X} - \mu_0 }{\sigma/\sqrt{n}} \geq z_{\alpha/2}$
	μ (σ^2 未知)	$\mu = \mu_0$ (σ^2 未知)	$\frac{\bar{X} - \mu}{S/\sqrt{n}}$	$\frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$t(n-1)$	$\frac{ \bar{X} - \mu }{S/\sqrt{n}} < t_{\alpha/2}(n-1)$	$\frac{ \bar{X} - \mu_0 }{S/\sqrt{n}} \geq t_{\alpha/2}(n-1)$
	σ^2 (μ 未知)	$\sigma^2 = \sigma_0^2$ (μ 未知)	$\frac{(n-1)S^2}{\sigma^2}$	$\frac{(n-1)S^2}{\sigma_0^2}$	$\chi^2(n-1)$	$\frac{\chi^2_{1-\alpha/2}(n-1)}{\sigma^2} < \frac{(n-1)S^2}{\sigma^2}$	$\frac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{1-\alpha/2}(n-1)$ 或 $\frac{(n-1)S^2}{\sigma_0^2} \geq \chi^2_{\alpha/2}(n-1)$
两个正态总体	$\mu_1 - \mu_2$ ($\sigma_1^2 = \sigma_2^2 = \sigma^2$)	$\mu_1 = \mu_2$ ($\sigma_1^2 = \sigma_2^2 = \sigma^2$)	$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$\frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$t(n_1 + n_2 - 2)$	$\frac{ (\bar{X} - \bar{Y}) - (\mu_1 - \mu_2) }{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < t_{\alpha/2}(n_1 + n_2 - 2)$	$\frac{ \bar{X} - \bar{Y} }{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \geq t_{\alpha/2}(n_1 + n_2 - 2)$
	$\frac{\sigma_1^2}{\sigma_2^2}$	$\sigma_1^2 = \sigma_2^2$	$\frac{S_1^2}{S_2^2} / \frac{\sigma_1^2}{\sigma_2^2}$	$\frac{S_1^2}{S_2^2}$	$F(n_1 - 1, n_2 - 1)$	$\frac{F_{1-\alpha/2}(n_1 - 1, n_2 - 1)}{S_2^2 / \sigma_2^2} < \frac{S_1^2}{S_2^2} / \frac{\sigma_1^2}{\sigma_2^2} < \frac{F_{\alpha/2}(n_1 - 1, n_2 - 1)}{S_2^2 / \sigma_2^2}$	$\frac{S_1^2}{S_2^2} \leq F_{1-\alpha/2}(n_1 - 1, n_2 - 1)$ 或 $\frac{S_1^2}{S_2^2} \geq F_{\alpha/2}(n_1 - 1, n_2 - 1)$

8.5 拟合优度检验

前面介绍的各种检验都是在总体服从正态分布前提下，对参数进行假设检验的。

实际中可能遇到这样的情形，总体服从何种理论分布并不知道，要求我们直接对总体分布提出一个假设。

例如，要检验在计算机上产生随机数的一个程序。指令该程序产生0到9之间的100个单个数字。观察整数的频数如下表。那么以0.05的显著性水平，有充分的理由相信该批整数不是均匀产生的吗？

整数	0	1	2	3	4	5	6	7	8	9
频数	11	8	7	7	10	10	8	11	14	14

例如，从1500到1931年的432年间，每年爆发战争的次数可以看作一个随机变量，据统计，这432年间共爆发了299次战争，具体数据如下：

战争次数 X	0	1	2	3	4
发生 X 次战争的年数	223	142	48	15	4

通常假设每年爆发战争的次数服从泊松分布。

那么上面的数据是否有充分的理由推翻每年爆发战争的次数服从泊松分布假设？

记 $F(x)$ 为总体 X 的未知的分布函数，设 $F_0(x)$ 是形式已知但可能含有若干个未知参数的分布函数，需检验假设

$$H_0 : F(x) = F_0(x) \quad \forall x \in R$$

注：若总体 X 为离散型，则 H_0 相当于

$$H_0 : \text{总体} X \text{的分布律为 } P\{X = t_i\} = p_i, i = 1, 2, \dots$$

若总体 X 为连续型，则 H_0 相当于

$$H_0 : \text{总体} X \text{的概率密度为 } f(x)。$$

———拟合优度检验问题

注意：在拟合优度检验中，一般地，把想要支持结论放在原假设。

(一) Pearson χ^2 检验

基本原理和步骤:

1. 在 H_0 下, 总体 X 取值的全体分成 k 个两两不相交的子集 A_1, \dots, A_k .
2. 以 $n_i (i = 1, \dots, k)$ 记样本观察值 x_1, \dots, x_n 中落在 A_i 的个数 (实际频数).

3. 当 H_0 为真且 $F_0(x)$ 完全已知时，计算事件 A_i 发生的概率 $p_i = P_{F_0}(A_i), i = 1, \dots, k$;

当 $F_0(x)$ 含有 r 个未知参数时，先利用极大似然法估计 r 个未知参数，然后求得 p_i 的估计 \hat{p}_i .

此时称 np_i (或 $n\hat{p}_i$)为理论频数.

4. 统计量

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^k \frac{n_i^2}{np_i} - n$$

$$(\text{或} \chi^2 = \sum_{i=1}^k \frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i} = \sum_{i=1}^k \frac{n_i^2}{n\hat{p}_i} - n)$$

反映了实际频数与理论频数的综合偏差，当 H_0 成立时， χ^2 的取值偏小，因此检验的拒绝域形式为： $\chi^2 \geq c$.

定理： 若 n 充分大，则当 H_0 为真时，统计量 χ^2 近似服从 $\chi^2(k-r-1)$ 分布，其中 k 为分类数， r 为 $F_0(x)$ 中含有的未知参数个数.

即在显著性水平 α 下拒绝域为

$$\chi^2 = \sum_{i=1}^k \frac{n_i^2}{np_i} - n \geq \chi_{\alpha}^2(k-1), \quad (\text{没有参数需要估计})$$

$$\chi^2 = \sum_{i=1}^k \frac{n_i^2}{n\hat{p}_i} - n \geq \chi_{\alpha}^2(k-r-1), \quad (\text{有} r \text{个参数需要估计})$$

注： χ^2 拟合检验使用时必须注意 n 要足够大， np_i (或 $n\hat{p}_i$) 不能太小。根据实践，要求 $n \geq 50$ ， np_i (或 $n\hat{p}_i$) ≥ 5 ，否则应适当合并相邻的类，以满足要求。

例: 从1500到1931年的432年间，每年爆发战争的次数可以看作一个随机变量，据统计，这432年间共爆发了299次战争，具体数据如下：

战争次数 X	0	1	2	3	4
发生 X 次战争的 年数	223	142	48	15	4

通常假设每年爆发战争的次数服从泊松分布。

那么上面的数据是否有充分的理由推翻每年爆发战争的次数服从泊松分布假设？

解： $H_0: X \sim \pi(\lambda)$, λ 未知, $\hat{\lambda} = \bar{X} = 299/432 = 0.69$.

$$\hat{p}_i = \frac{\hat{\lambda}^i e^{-\hat{\lambda}}}{i!}, i = 0, 1, 2, 3, \quad \hat{p}_4 = \sum_{j=4}^{\infty} \frac{\hat{\lambda}^j e^{-\hat{\lambda}}}{j!} = 1 - \sum_{i=0}^3 \hat{p}_i.$$

战争次数x	0	1	2	3	≥ 4
实测频数 n_i	223	142	48	15	4
概率估计 \hat{p}_i	0.502	0.346	0.119	0.027	0.006
理论频数 $n\hat{p}_i$	217	149	51	12	3

$\underbrace{\hspace{10em}}_{15}$

检验统计量的观察值为

$$\chi^2 = \sum_{i=1}^k \frac{n_i^2}{n\hat{p}_i} - n = \frac{223^2}{217} + \frac{142^2}{149} + \frac{48^2}{51} + \frac{19^2}{15} - 432 = 1.74$$

即在显著性水平 $\alpha = 0.05$ 下临界值

$$\chi_{\alpha}^2(k-r-1) = \chi_{0.05}^2(4-1-1) = 5.991$$

于是， $1.74 < 5.991$, 不能拒绝原假设。

例2 孟德尔遗传理论断言，当两个品种的豆杂交时，圆的和黄的、起皱的和黄的、圆的和绿的、起皱的和绿的豆的频数将以比例 $9:3:3:1$ 发生。在检验这个理论时，孟德尔分别得到频数 315 、 101 、 108 、 32 、这些数据提供充分证据拒绝该理论吗？

解：定义 $X = \begin{cases} 1, & \text{若豆子是圆的和黄的} \\ 2, & \text{若豆子是起皱的和黄的} \\ 3, & \text{若豆子是圆的和绿的} \\ 4, & \text{若豆子是起皱的和绿的} \end{cases}$

$$H_0 : p_1 = P(X = 1) = \frac{9}{16}, p_2 = P(X = 2) = \frac{3}{16},$$
$$p_3 = P(X = 3) = \frac{3}{16}, p_4 = P(X = 4) = \frac{1}{16}.$$

豆子状态x	1	2	3	4
实测频数 n_i	315	101	108	32
概率 p_i	9/16	3/16	3/16	1/16
理论频数 np_i	312.75	104.25	104.25	34.75

$$\chi^2 = \sum_{i=1}^4 \frac{n_i^2}{np_i} - n = 0.47 < \chi_{0.05}^2(3) = 7.815,$$

因此没有充分的理由否定该理论.

例3 下面列出了84个伊特拉斯坎(Etruscan)人男子的头颅的最大宽度(mm)，试检验这些数据是否来自正态总体（取 $\alpha=0.1$ ）

141	148	132	138	154	142	150	146	155	158	150	140
147	148	144	150	149	145	149	158	143	141	144	144
126	140	144	142	141	140	145	135	147	146	141	136
140	146	142	137	148	154	137	139	143	140	131	143
141	149	148	135	148	152	143	144	141	143	147	146
150	132	142	142	143	153	149	146	149	138	142	149
142	137	134	144	146	147	140	142	140	137	152	145

解 为粗略了解数据的分布情况，先画出直方图。

步骤如下：

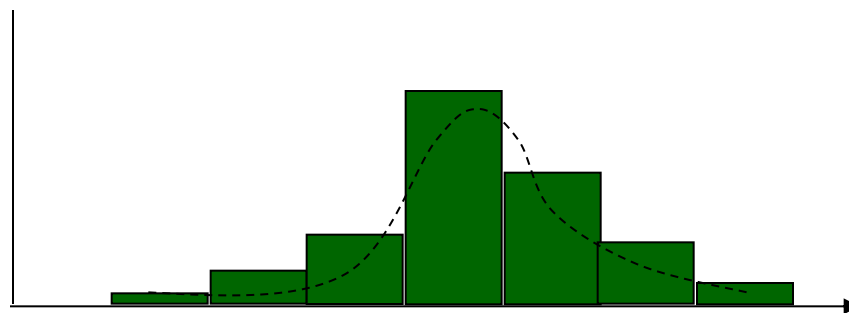
1.找出数据的最小值、最大值为126、158，取区间 $[124.5, 159.5]$ ，它能覆盖 $[126, 158]$ ；

2.将区间 $[124.5, 159.5]$ 等分为7个小区间，小区间的长度 $\Delta = (159.5 - 124.5) / 7 = 5$ ， Δ 称为组距，小区间的端点称为组限，建立下表：

组 限	频数	频率	累计频率
124.5-129.5	1	0.0119	0.0119
129.5-134.5	4	0.0476	0.0595
134.5-139.5	10	0.1191	0.1786
139.5-144.5	33	0.3929	0.5715
144.5-149.5	24	0.2857	0.8572
149.5-154.5	9	0.1071	0.9524
154.5-159.5	3	0.0357	1

3.自左向右在各小区间上作以 $n_i/n \Delta$ 为高的小矩形
如下图，即为直方图。

注：直方图的小区间可以不等长，但小区间的长度不能太大，否则平均化作用突出，淹没了密度的细节部分；也不能太小，否则受随机化影响太大，产生极不规则的形状。



从本例的直方图看，有一个峰，中间高，两头低，较对称，样本象来自正态总体。于是检验

$$H_0 : X \sim N(\mu, \sigma^2)$$

其中 μ, σ^2 未知，其最大似然估计分别为 $\hat{\mu} = 143.8, \hat{\sigma}^2 = 6.0^2$.

计算每一事件 A_i 的概率估计值 $\hat{p}_i = \hat{P}(A_i)$.

例如

$$\begin{aligned}\hat{p}_1 &= \hat{P}(A_1) = \hat{P}\{X \leq 129.5\} \\ &= \Phi\left(\frac{129.5 - 143.8}{6.0}\right) \\ &= \Phi(-2.383) = 0.0087,\end{aligned}$$

A_i	n_i	\hat{p}_i	$n\hat{p}_i$	$n_i^2 / n\hat{p}_i$
A_1 $x \leq 129.5$	1	0.0087	0.73	5.09
A_2 $129.5 < x \leq 134.5$	4	0.0519	4.36	
A_3 $134.5 < x \leq 139.5$	10	0.1752	14.72	6.79
A_4 $139.5 < x \leq 144.5$	33	0.3120	26.21	41.55
A_5 $144.5 < x \leq 149.5$	24	0.2811	23.61	24.40
A_6 $149.5 < x \leq 154.5$	9	0.1336	11.22	14.37
A_7 $154.5 < x < \infty$	3	0.0375	3.15	
				$\Sigma = 87.67$

$$\chi^2 = 87.67 - 84 = 3.67$$

$$\chi_{0.1}^2(k - r - 1) = \chi_{0.1}^2(2) = 4.605 > 3.67$$

故在水平0.1下接受 H_0 , 认为数据来自正态总体。

Pearson χ^2 检验的缺点:

对于连续性随机变量，检验统计量的取值依赖于区间的划分，影响检验的功效。

适用于离散型随机变量的分布检验！