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),

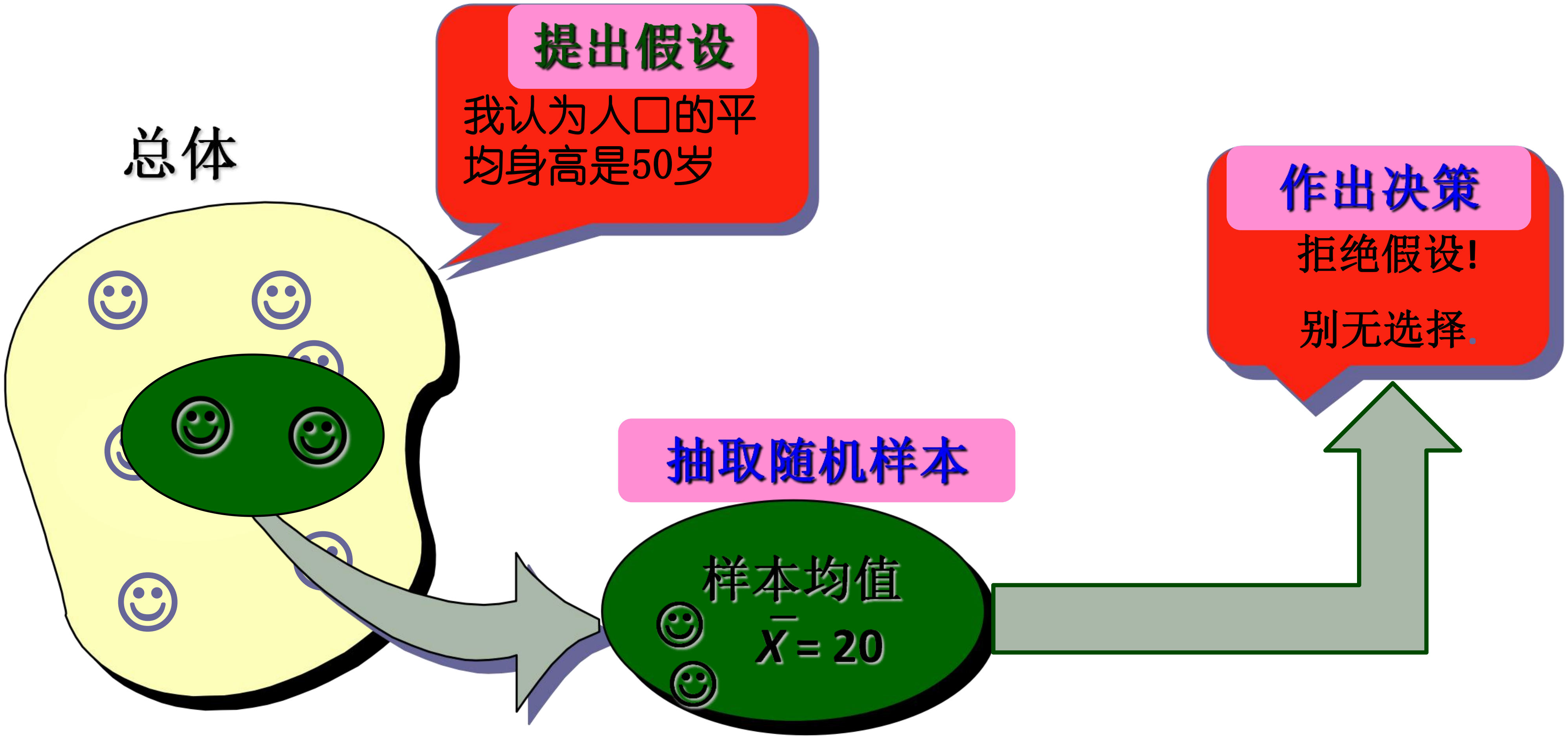
现根据样本检验均值是否与以往有显著差异？ (样本均值为6.4)

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?

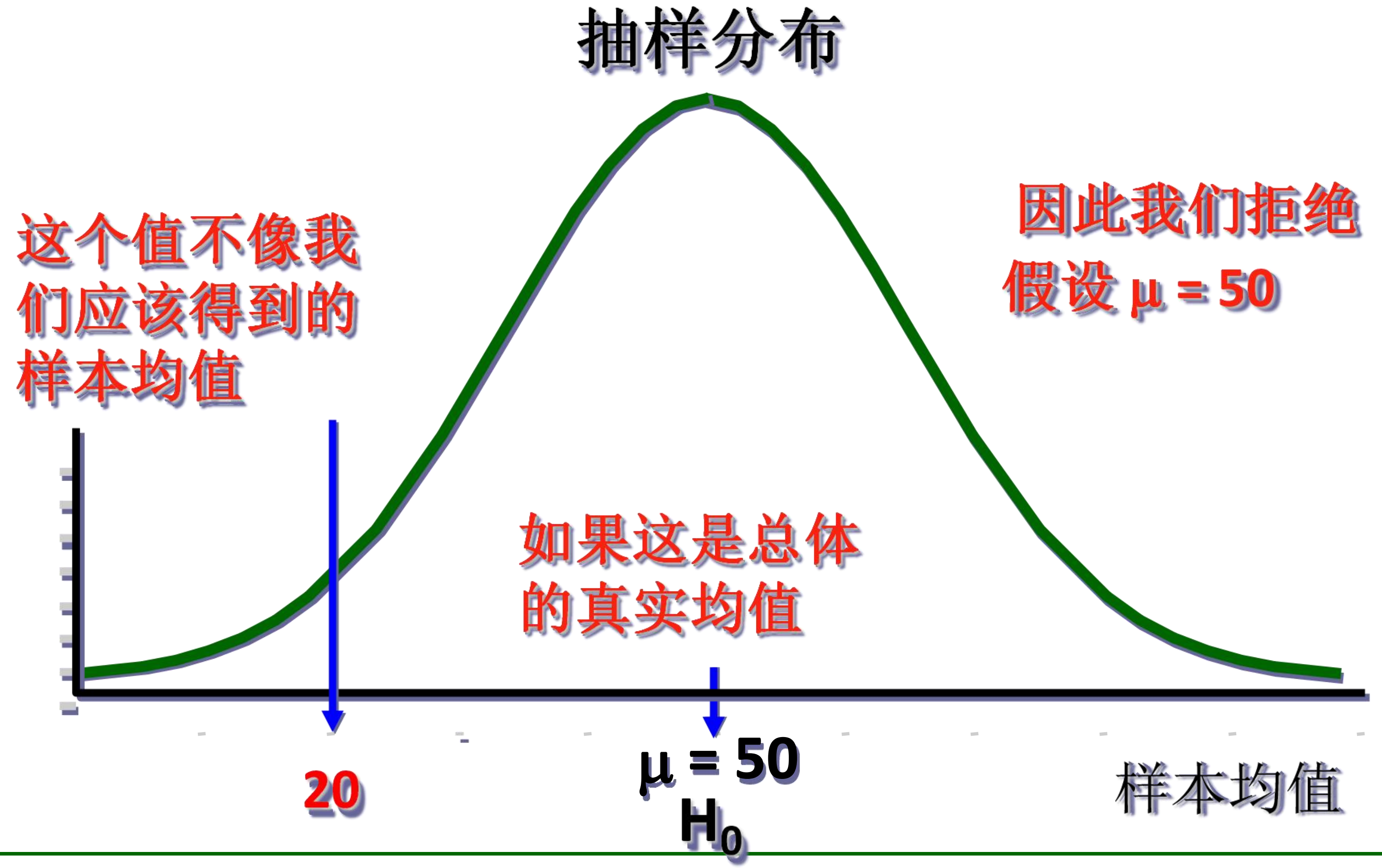
例3 孟德尔遗传理论断言，当两个品种的豆杂交时，圆的和黄的、起皱的和黄的、圆的和绿的、起皱的和绿的豆的频数将以比例9：3：3：1发生。在检验这个理论时，孟德尔分别得到频数315、101、108、32、这些数据提供充分证据拒绝该理论吗?

假设检验的过程

(提出假设→抽取样本→作出决策)



假设检验的基本思想


抽样分布
这个值不像我
们应该得到的
样本均值
因此我们拒绝
假设 μ = 50
如果这是总体
的真实均值
20
μ = 50
H0
样本均值

- **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：$\theta \geq \theta_0$，H_1：$\theta < \theta_0$(左边检验）

H_0：$\theta \leq \theta_0$，H_1：$\theta > \theta_0$（右边检验）

H_0：$\theta = \theta_0$，H_1：$\theta \neq \theta_0$（双边检验）

（二）检验统计量和拒绝域

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？(样本均值为6.4)

- 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σ^2=0. 36，考虑有关参数μ的假设：

H_0: $\mu = 6.0 \leftrightarrow H_1$: $\mu \neq 6.0$（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$T=T(X_1,\cdots,X_n)$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$\overline{W}$称为**接受域**.

上述例子中，可取检验统计量为$\bar{X}$, 拒绝域为

$$W=\left\{(x_1,\cdots,x_n):|\bar{x}-6.0|\geq C\right\}$$

(后面简单写为$\{|\bar{x}-6.0|\geq C\}$)

问题：(1）怎样确定C？ 基于什么准则?

(2）拒绝 H_0 是否就意味着H_0 是错的?

(3）接受 H_0 是否就意味着H_0 是对的?

分析:

$H_0: \mu = 6.0 \leftrightarrow H_1: \mu \neq 6.0$

拒绝条件： $|\bar{X} - 6.0| \geq C$

$\overline{X} - 6.0 = (\overline{X} - \mu) + (\mu - 6.0)$

$\bar{X} - \mu$: 取样（随机）误差

$\mu - 6.0$: 系统误差

由于取样（随机）误差，假设检验错误情况如下：

（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

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

犯第II类错误的概率

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

例1中，犯第I类错误的概率

$$\alpha(C)=P\{拒绝H_0 \mid H_0是真的\}$$

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

$$=P\{\frac{|\bar{X}-6.0|}{\sigma/\sqrt{n}}\geq\frac{C}{\sigma/\sqrt{n}} \mid \mu=6.0\}$$

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

$$\alpha(C)=2-2\Phi\left(\frac{C}{\sigma/\sqrt{n}}\right)$$ 关于C是单调减函数，

犯第II类错误的概率

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

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

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

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

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

关于C是单调增函数

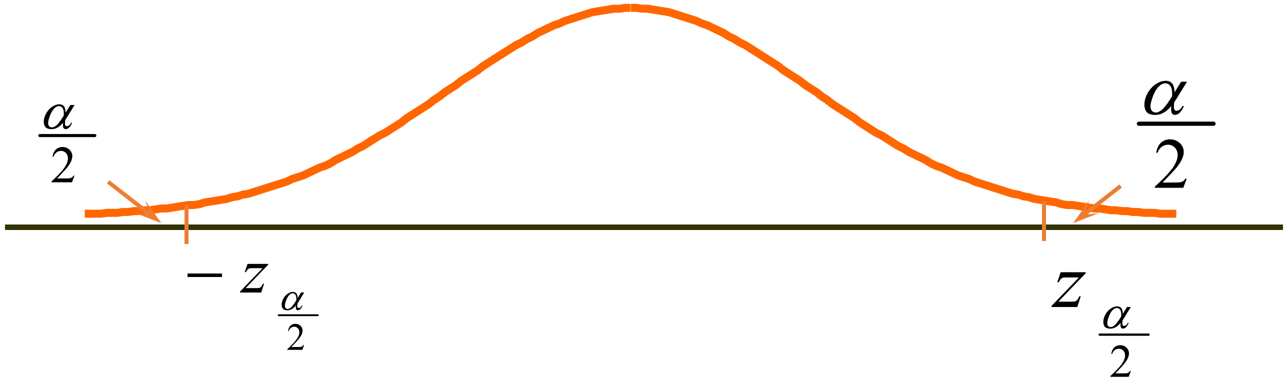
犯两类错误的概率相互制约！

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α=0.01, 0.05, 0.1等.



当H_0成立时：$Z=\dfrac{\bar{X}-6.0}{\sqrt{\frac{0.6^2}{9}}}\sim N(0,1)$

$$\Rightarrow P\left(\frac{\left|\bar{X}-6.0\right|}{0.6/3}\geq z_{\alpha/2}\right)=\alpha$$

$\Rightarrow$ 拒 绝 条 件 为： $\dfrac{\left|\bar{X}-6.0\right|}{0.2}\geq z_{\alpha/2}$

若取α=0.05，则拒绝域为$\{|\bar{x}-6|\geq 0.392|\}$

$\because \bar{x}=6.4, |\bar{x}-6|=0.4>0.392.$

即$\bar{x}$落在拒绝域内，$\bar{x}$ 与$\mu=6$的差异显著，因此

拒绝原假设，认为干燥时间的均值与以往有显著差异。

犯第I类错误的概率$\alpha(0.392)=0.05$

犯第II类错误的概率$\beta(0.392) = P\{$接受$H_0 | H_0$是假的$\}$

$$= P\{|\bar{X} - 6.0| < 0.392 | \mu \neq 6.0\}$$

$$= P\{5.608 < \bar{X} < 6.392 | \mu \neq 6.0\}$$

$$= \Phi\left\{\frac{6.392 - \mu}{0.2}\right\} - \Phi\left\{\frac{5.608 - \mu}{0.2}\right\}, \quad \mu \neq 6.0$$

例当 μ=5.4时，

$$\beta=\Phi\left\{\frac{6.392-5.4}{0.2}\right\}-\Phi\left\{\frac{5.608-5.4}{0.2}\right\}$$

$$=\Phi\ \ (4.96)-\Phi\ \ (1.04)\approx 1.00-0.85=0.15$$

（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：

$$P\_ = P_{H_0}\left(|\bar{X}-6.0| \geq 0.4\right) = P_{H_0}\left(\left|\frac{\bar{X}-6.0}{0.2}\right| \geq 2\right)$$

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

$|\bar{X}-6.0| \geq 0.4$是小概率事件.

作出拒绝原假设的判断.

$P_$值与显著水平α的关系：

若$P_ \le \alpha$，则拒绝原假设，此时称检验结果在水平α下是统计显著的.

若$P_ > \alpha$，则接受原假设，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.

注:H_0与H_1地位不等

由于控制犯第1类错误，因此错误拒绝H_0是小概率事件，也就是说H_0不会轻易被拒绝掉。因此如果落在拒绝域，则说明明已有了显著的差异，从而拒绝H_0.

H_0选取：不能轻易拒绝的，后果严重的，或维持现状的

如：H_0：新技术未提高效益 $\leftrightarrow$ H_1：新技术提高效益

假设检验中的两类错误
（决策结果）

H_0: 无罪

假设检验就好像一场审判过程

统计检验过程

陪审团审判		
裁决	实际情况	
	无罪	有罪
无罪	正确	错误
有罪	错误	正确

H_0 检验		
决策	实际情况	
	H_0为真	H_0为假
接受H_0	$1 - \alpha$	第二类错误(β)
拒绝H_0	第一类错误(α)	功效($1-\beta$)

8.2 单个正态总体参数的假设检验

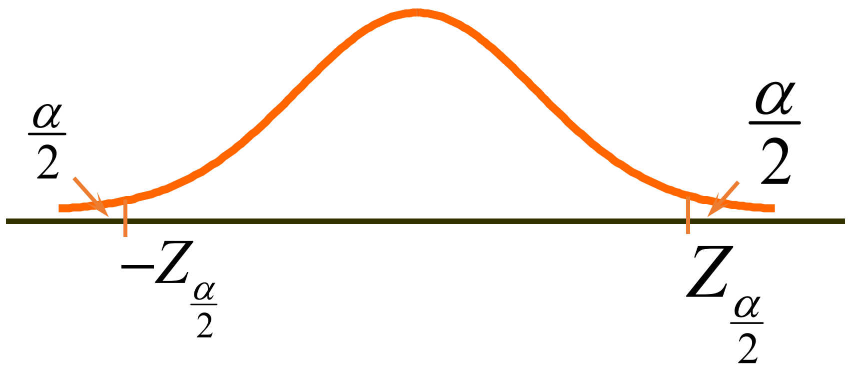
设样本$X_1, X_2, \cdots, X_n$来自正态总体$N\left(\mu, \sigma^2\right)$,

$\overline{X}$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$H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$



取检验统计量为 $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$

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$x_1,\cdots,x_n$,记检验统计量Z的取值为

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

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

当P_{-}小于显著水平α时，拒绝原假设，

否则，接受原假设.

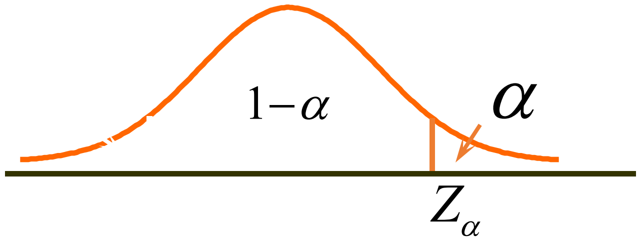
右边检验 $H_0: \mu \le \mu_0, H_1: \mu > \mu_0$

检验统计量为 $\bar{X}$, 检验拒绝域的形式为 $\bar{X} - \mu_0 \ge k$.

当 $\mu \in \mathrm{H}_0$, 即 $\mu \le \mu_0$ 时， $\because \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$,

$\therefore$ 犯第一类错误概率为 $P_\mu\left\{\bar{X} - \mu_0 \ge k\right\}$

$= 1 - \Phi(\frac{k + \mu_0 - \mu}{\sigma/\sqrt{n}}) = \Phi(\frac{\mu - \mu_0 - k}{\sigma/\sqrt{n}})$ 是 μ 的增函数.



$\therefore$ 当 $\mu = \mu_0$ 时，犯第一类错误概率最大。

故只要 $\Phi(\frac{-k}{\sigma/\sqrt{n}}) = \alpha$, 即 $\frac{-k}{\sigma/\sqrt{n}} = -z_\alpha$ 便可，

因此，拒绝域为： $\boxed{\{z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \ge z_\alpha\}}$.

P_-值的计算

$$P_- = \sup_{\mu \le \mu_0} P_{H_0}\left\{Z \ge z_0\right\}$$

$$= P\left\{Z \ge z_0 \mid \mu = \mu_0\right\}$$

$$= 1 - \Phi(z_0).$$

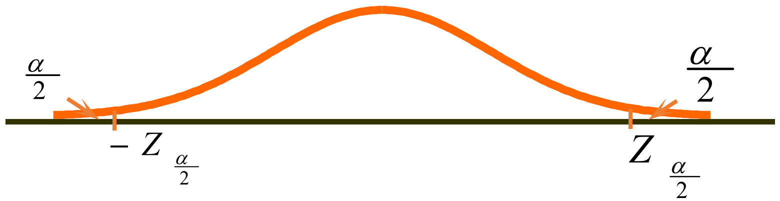
总结：σ^2已知检验μ （Z检验法）

$$z_0=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}},\quad \text{当}\mu=\mu_0\text{时，}\ Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$$

（1）$H_0:\mu=\mu_0 \leftrightarrow H_1:\mu\neq\mu_0$

拒绝域： $|Z|\geq z_{\alpha/2}$

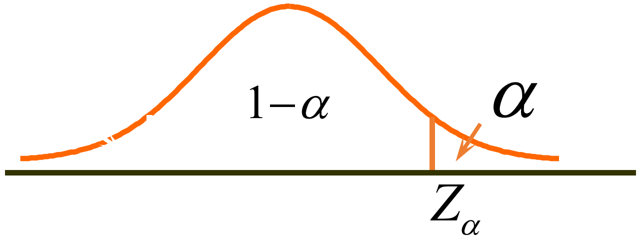
$$P\_=P_{\mu_0}\{|Z|\geq|z_0|\}=2(1-\Phi(|z_0|)).$$



(2)$H_0:\mu\leq\mu_0 \leftrightarrow H_1:\mu>\mu_0$

拒绝域： $Z\geq z_\alpha$

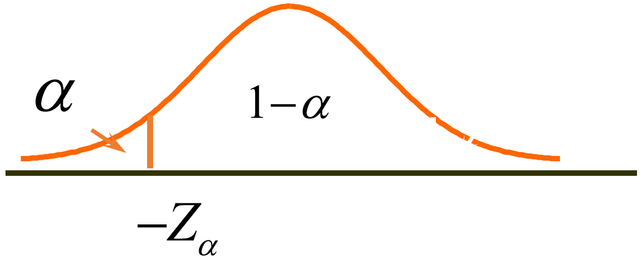
$$P\_=P_{\mu_0}\{Z\geq z_0\}=1-\Phi(z_0)$$



(3)$H_0:\mu\geq\mu_0,\ H_1:\mu<\mu_0$

拒绝域： $Z\leq -z_\alpha$

$$P\_=P_{\mu_0}\{Z\leq z_0\}=\Phi(z_0)$$

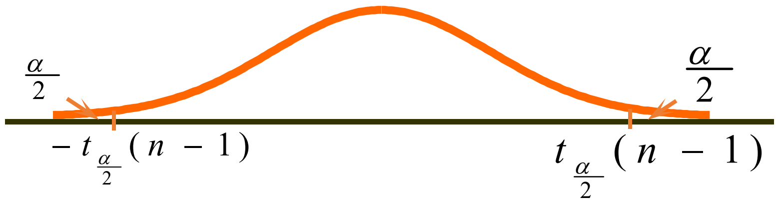


σ^2未知检验μ（t检验法）

$$t_0=\frac{\bar{x}-\mu_0}{s/\sqrt{n}},\quad \text{当}\mu=\mu_0\text{时，}\quad t=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$$

（1） $H_0:\mu=\mu_0 \leftrightarrow H_1:\mu\neq\mu_0$

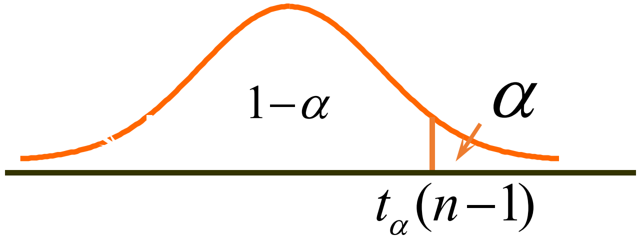
拒绝域： $|t|\geq t_{\alpha/2}(n-1)$



$$P\_=P_{\mu_0}\{|t|\geq|t_0|\}=2P(t(n-1)\geq|t_0|).$$

(2)$H_0:\mu\leq\mu_0 \leftrightarrow H_1:\mu>\mu_0$

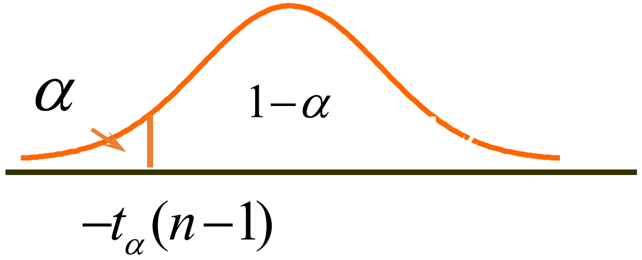
拒绝域： $t\geq t_\alpha(n-1)$



$$P\_=P_{\mu_0}\{t\geq t_0\}=P(t(n-1)\geq t_0)$$

(3)$H_0:\mu\geq\mu_0,H_1:\mu<\mu_0$

拒绝域： $t\leq -t_\alpha(n-1)$



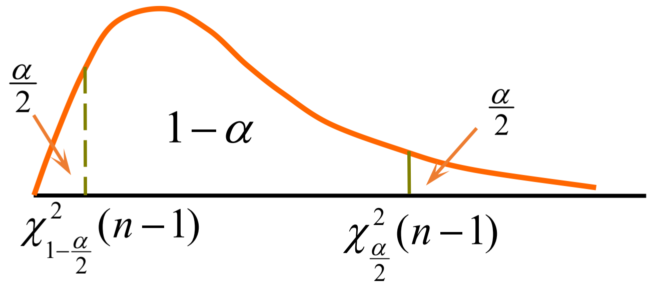
$$P\_=P_{\mu_0}\{t\leq t_0\}=P(t(n-1)\leq t_0)$$

μ未知检验σ^2（χ^2检验法）

当$\sigma^2=\sigma_0^2$时，$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$

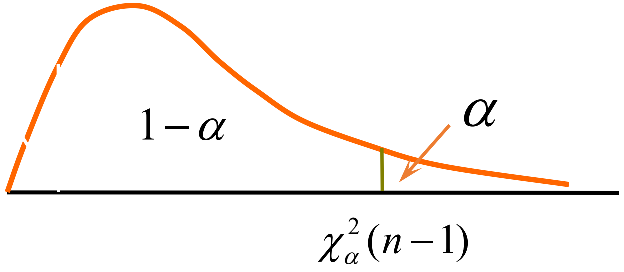
（1）$H_0:\sigma^2=\sigma_0^2 \leftrightarrow H_1:\sigma^2\neq\sigma_0^2$

拒绝域：$\chi^2\le\chi^2_{1-\alpha/2}(n-1)$或 $\chi^2\ge\chi^2_{\alpha/2}(n-1)$



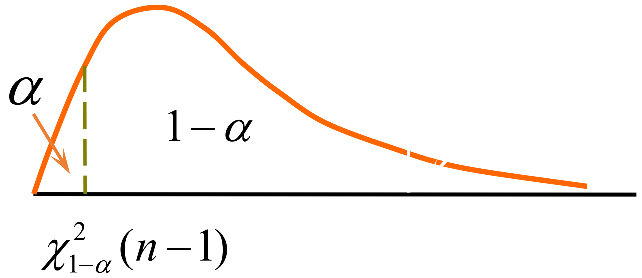
(2)$H_0:\sigma^2\le\sigma_0^2 \leftrightarrow H_1:\sigma^2>\sigma_0^2$

拒绝域： $\chi^2\ge\chi^2_\alpha(n-1)$



(3)$H_0:\sigma^2\ge\sigma_0^2 \leftrightarrow H_1:\sigma^2<\sigma_0^2$

拒绝域：$\chi^2\le\chi^2_{1-\alpha}(n-1)$



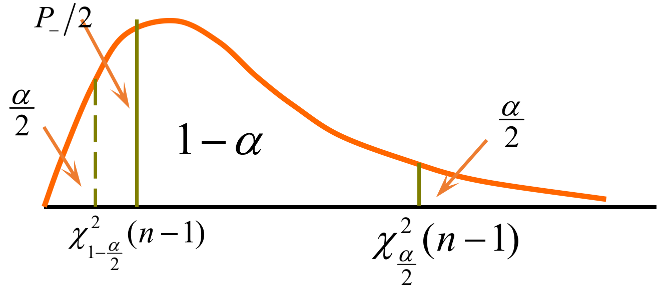
$P_$值计算：

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

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

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

$$P\_ = 2\min(p, 1-p)$$



当$P_ \le \alpha$，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：$$\frac{(n-1)S^2}{\sigma_0^2} \leq \chi_{1-\alpha}^2(n-1);$$

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

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

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

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

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

当$P_{_}\le\alpha$，拒绝原假设，

当$P_{_}>\alpha$，接受原假设.

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^2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

$$H_0: \mu \le \mu_0 = 225, \qquad H_1: \mu > 225.$$

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

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

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

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

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

- 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

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

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

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

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

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

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

因此拒绝原假设，判断结果与前面一致！

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：$\dfrac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{1-\alpha}(n-1)$

查表得：$\chi^2_{0.95}\left(24\right) = 13.848,$

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

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$P_ = P\{\chi^2(24) \leq 14.57\} = 0.06729 > 0.05$

作出同样判断。

（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 $(X_1, Y_1), \cdots, (X_n, Y_n)$,

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

可以看成来自正态总体$N(\mu_d, \sigma_d^2)$的样本

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(xi)	23	35	29	42	39	29	37	34	35	28
种子B(yi)	26	39	35	40	38	24	36	27	41	27
di=xi-yi	-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$D_1, D_2, ..., D_n$的样本均值和样本方差记为$\overline{D}, S_D^2,$

拒绝域为：$\dfrac{\left|\overline{D}\right|}{S_D \big/ \sqrt{n}} \geq t_{\alpha/2}\left(n-1\right),$

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

计算得：$\dfrac{\left|\overline{d}\right|}{s_d \big/ \sqrt{n}} = 0.142 < 2.2622$

接受原假设H_0，认为两种种子的产量没有显著差异。

- 在Execl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Execl 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。

8.3 两个正态总体$N\left(\mu_1,\sigma_1^2\right),N\left(\mu_2,\sigma_2^2\right)$的检验

$X_1,X_2,\cdots,X_{n_1}$来自$N\left(\mu_1,\sigma_1^2\right)$,

$Y_1,Y_2,\cdots,Y_{n_2}$来自$N\left(\mu_2,\sigma_2^2\right)$,

$\overline{X},\overline{Y},S_1^2,S_2^2$分别为第一,二个

总体的样本均值和方差,

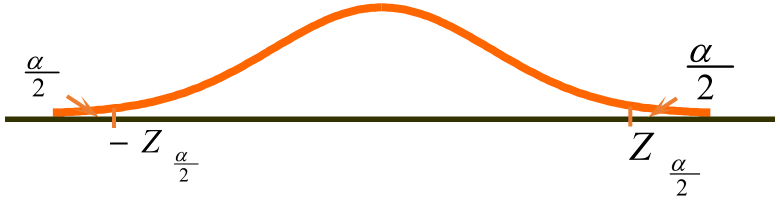
显著性水平为α.

σ_1^2,σ_2^2已知检验$\mu_1-\mu_2$ （Z检验法）

当$\mu_1-\mu_2=\delta$时，$Z=\dfrac{\bar{X}-\bar{Y}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\sim N(0,1)$

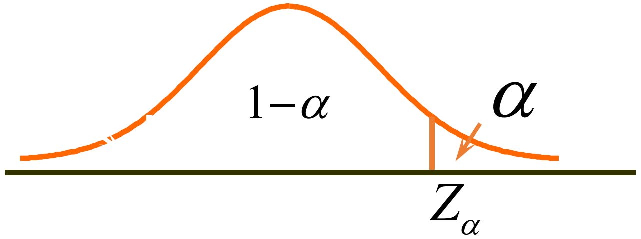
(1) $H_0:\mu_1-\mu_2=\delta$

拒绝域： $|Z|\geq z_{\alpha/2}$



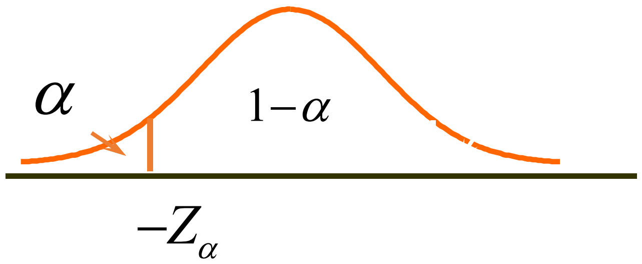
(2) $H_0:\mu_1-\mu_2\leq\delta$

拒绝域： $Z\geq z_\alpha$



(3) $H_0:\mu_1-\mu_2\geq\delta$

拒绝域： $Z\leq -z_\alpha$

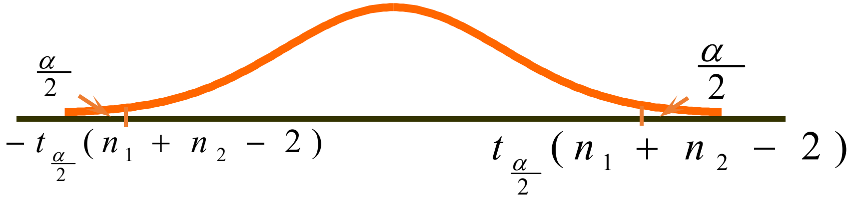


$\sigma_1^2 = \sigma_2^2$未知检验$\mu_1 - \mu_2$ （t检验法）

当$\mu_1 - \mu_2 = \delta$时，$t = \dfrac{\bar{X} - \bar{Y} - \delta}{S_w\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$

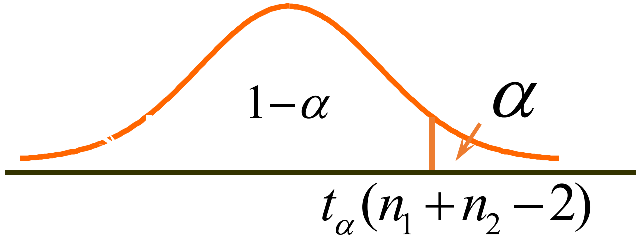
(1) $H_0: \mu_1 - \mu_2 = \delta$

拒绝域：$|t| \ge t_{\alpha/2}(n_1 + n_2 - 2)$



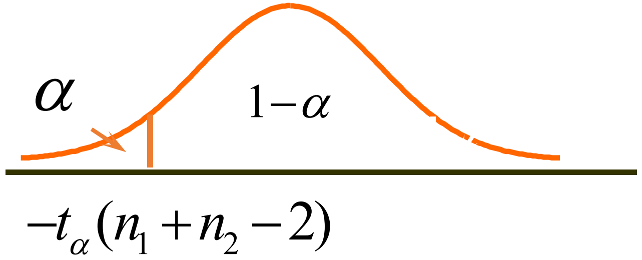
(2) $H_0: \mu_1 - \mu_2 \le \delta$

拒绝域：$t \ge t_\alpha(n_1 + n_2 - 2)$



(3) $H_0: \mu_1 - \mu_2 \ge \delta$

拒绝域：$t \le -t_\alpha(n_1 + n_2 - 2)$



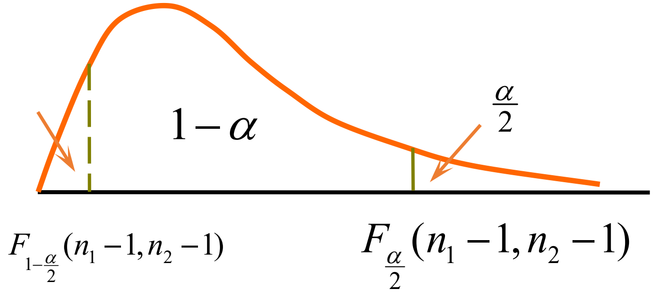
μ_1,μ_2未知检验σ_1^2/σ_2^2

(F检验法)

当$\sigma_1^2=\sigma_2^2$时，$F=\dfrac{S_1^2}{S_2^2}\sim F(n_1-1,n_2-1)$

(1) $H_0:\sigma_1^2=\sigma_2^2 \leftrightarrow H_1:\sigma_1^2\neq\sigma_2^2$

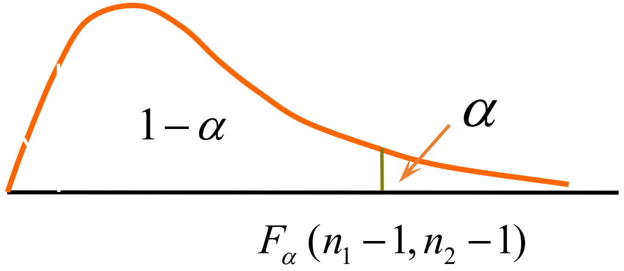
拒绝域：$F\le F_{1-\alpha/2}(n_1-1,n_2-1)$或$F\ge F_{\alpha/2}(n_1-1,n_2-1)$



$\frac{\alpha}{2}$

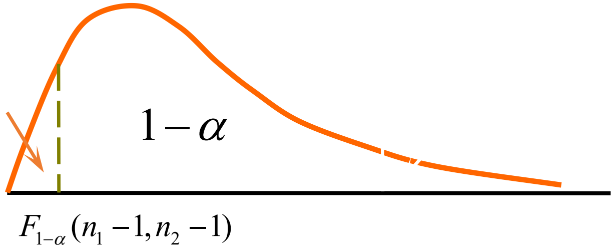
(2) $H_0:\sigma_1^2\le\sigma_2^2 \leftrightarrow H_1:\sigma_1^2>\sigma_2^2$

拒绝域：$F\ge F_\alpha(n_1-1,n_2-1)$



(3) $H_0:\sigma_1^2\ge\sigma_2^2 \leftrightarrow H_1:\sigma_1^2<\sigma_2^2$

拒绝域：$F\le F_{1-\alpha}(n_1-1,n_2-1)$



α

例：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的产品平均重量较用原料A的为大。

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

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

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

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

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

例： 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$X \sim N\left(\mu_1,\sigma_1^2\right),Y \sim N\left(\mu_2,\sigma_2^2\right)$

(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：$\dfrac{S_1^2}{S_2^2}\leq F_{1-\alpha/2}(n_1-1,n_2-1)$, 或 $\dfrac{S_1^2}{S_2^2}\geq F_{\alpha/2}(n_1-1,n_2-1)$

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

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：$0.268<\dfrac{S_1^2}{S_2^2}=0.795<3.50$

不拒绝原假设，故认为方差没有显著差异。

令 $p=P(F(7,8)\leq 0.795)=0.3875$

$P_{\_}=2\min(p,1-p)=0.775>0.05.$ 接受原假设.

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

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

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

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

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

$P\_=P\{t(15)\ge 1.354\}=0.098<0.1.$

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

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

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

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

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

- 在Execl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Execl 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 $H_0: \mu_1 \le \mu_2$.

（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$H_0 : \mu_1 = \mu_2$.

（3） 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”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“0.482994”.

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$X_1, X_2, \ldots, X_n$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

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

接受域为

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

$$= \left\{ \bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} < \mu_0 < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right\}$$

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 $\theta_0 \le \hat{\theta}_L$或$\theta_0 \ge \hat{\theta}_U$.

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.

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）	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	$\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$N(0,1)$	$\dfrac{\lvert\bar{X}-\mu\rvert}{\sigma/\sqrt{n}}<z_{\alpha/2}$	$\dfrac{\lvert\bar{X}-\mu_0\rvert}{\sigma/\sqrt{n}}\geq z_{\alpha/2}$
	μ （σ^2未知）	$\mu=\mu_0$ （σ^2未知）	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	$\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$t(n-1)$	$\dfrac{\lvert\bar{X}-\mu\rvert}{S/\sqrt{n}}<t_{\alpha/2}(n-1)$	$\dfrac{\lvert\bar{X}-\mu_0\rvert}{S/\sqrt{n}}\geq t_{\alpha/2}(n-1)$
	σ^2 (μ未知)	$\sigma^2=\sigma_0^2$ (μ未知)	$\dfrac{(n-1)S^2}{\sigma^2}$	$\dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi^2(n-1)$	$\chi^2_{1-\alpha/2}(n-1)<\dfrac{(n-1)S^2}{\sigma^2}<\chi^2_{\alpha/2}(n-1)$	$\dfrac{(n-1)S^2}{\sigma_0^2}\leq\chi^2_{1-\alpha/2}(n-1)$ 或$\dfrac{(n-1)S^2}{\sigma_0^2}\geq\chi^2_{\alpha/2}(n-1)$
两个正态总体	$\mu_1-\mu_2$ $(\sigma_1^2=\sigma_2^2=\sigma^2)$	$\mu_1=\mu_2$ $(\sigma_1^2=\sigma_2^2=\sigma^2)$	$\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$\dfrac{\bar{X}-\bar{Y}}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$t(n_1+n_2-2)$	$\dfrac{\lvert(\bar{X}-\bar{Y})-(\mu_1-\mu_2)\rvert}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}<t_{\alpha/2}(n_1+n_2-2)$	$\dfrac{\lvert\bar{X}-\bar{Y}\rvert}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\geq t_{\alpha/2}(n_1+n_2-2)$
	$\dfrac{\sigma_1^2}{\sigma_2^2}$	$\sigma_1^2=\sigma_2^2$	$\dfrac{S_1^2}{S_2^2}\Big/\dfrac{\sigma_1^2}{\sigma_2^2}$	$\dfrac{S_1^2}{S_2^2}$	$F(n_1-1,n_2-1)$	$F_{1-\alpha/2}(n_1-1,n_2-1)<\dfrac{S_1^2}{S_2^2}\Big/\dfrac{\sigma_1^2}{\sigma_2^2}<F_{\alpha/2}(n_1-1,n_2-1)$	$\dfrac{S_1^2}{S_2^2}\leq F_{1-\alpha/2}(n_1-1,n_2-1)$ 或$\dfrac{S_1^2}{S_2^2}\geq F_{\alpha/2}(n_1-1,n_2-1)$

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5.1 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5.2 一淘宝店主搜集了一年中每天的订单数X，除去春节期间及双十一前后外，按330天计，具体数据如下:

订单数X	0	1	2	3	4	5	6	7
天数	3	6	21	46	48	61	52	42

订单数X	8	9	10	11	12	13	16
天数	27	11	6	4	1	1	1

通常认为每天的订单数服从泊松分布，以上的数据是否支持这个结论?

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

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

注：若总体X为离散型随机变量，则H_0相当于

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

若总体X为连续型随机变量，则H_0相当于

H_0:总体X的概率密度函数为$f(x)$.

—————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。

拟合优度检验的基本原理和步骤：

1. 在H_0下，将总体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$n_i (i=1,...,k)$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$p_i = P_{F_0}(A_i), i = 1, \ldots, k$;

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

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

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

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5.2 一淘宝店主搜集了一年中每天的订单数 X，除去春节期间及双十一前后外，按330天计，具体数据如下:

订单数X	0	1	2	3	4	5	6	7
天数	3	6	21	46	48	61	52	42

订单数X	8	9	10	11	12	13	16
天数	27	11	6	4	1	1	1

通常认为每天的订单数服从泊松分布，以上的数据是否支持这个结论?

解：$H_0: X \sim P(\lambda)$，λ未知，总订单数为1749，

所以，平均每天订单数$\hat{\lambda} = \bar{X} = 1749/330 = 5.3.$

订单数大于**10**的进行合并，对订单数为***i***

(*i*=0,1,…,10,11+)的概率值进行估计：

需注意

!

$$\hat{p}_i = \frac{\hat{\lambda}^i e^{-\hat{\lambda}}}{i!}, i = 0,1,\ldots,10, \quad \hat{p}_{11} = \sum_{j=11}^{\infty} \frac{\hat{\lambda}^j e^{-\hat{\lambda}}}{j!} = 1 - \sum_{i=0}^{10} \hat{p}_i.$$

理论频数：$n\hat{p}_i, i = 0,1,\ldots,10,11, n\hat{p}_0 = 1.65 < 5,$

将$x = 0$与$x = 1$合并. 最后共有**11**类，具体结果为

订单数X	0	1	2	3	4	5
天数	3	6	21	46	48	61
	9					
概率估计	0.005	0.026	0.070	0.124	0.164	0.174
理论频数	1.65	8.73	23.13	40.87	54.16	57.41
	10.23					

订单数X	6	7	8	9	10	≥11
天数	52	42	27	11	6	7
概率估计	0.154	0.116	0.077	0.045	0.024	0.021
理论频数	50.71	38.39	25.44	14.98	7.94	6.60

检验统计量的值为

$$\chi^2 = \sum_{i=1}^{k} \frac{{n_i}^2}{n\hat{p}_i} - n = \sum_{i=1}^{11} \frac{{n_i}^2}{n\hat{p}_i} - 330 = 3.97$$

即在显著性水平 $\alpha = 0.05$ 下临界值

$${\chi_\alpha}^2(k-r-1) = \chi^2_{0.05}(11\text{-}1\text{-}1) = 16.92$$

于是， $3.97 < 16.92$, 不拒绝原假设。

$$P\_ = P(\chi^2(9) \geq 3.97) = 0.913.$$

例1.3 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，孟德尔收集了556个观察数据，分别得到频数为315，108，101，32，这些数据提供充分证据证明该理论吗？

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

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

$$p_3 = P(X=3) = \frac{3}{16}, p_4 = P(X=4) = \frac{1}{16}.$$

豆子状态x	1	2	3	4
实测频数 n_i	315	108	101	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

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

因此没有充分的理由否定该理论.

$$P\_ = P(\chi^2(3) \geq 0.47) = 0.925.$$

例5.3 从某医院收集到168名新生女婴的体重数据 (单位:g)，试检验这些数据是否来自正态总体(取α=0. 1)

2 880	2 440	2 700	3 500	3 500	3 600	3 080	3 860	3 200	3 100	3 180	3 200
3 300	3 020	3 040	3 420	2 900	3 440	3 000	2 620	2 720	3 480	3 320	3 000
3 120	3 180	3 220	3 160	3 940	2 620	3 120	2 520	3 060	2 620	3 400	2 160
2 960	2 980	3 000	3 020	3 760	3 500	3 060	3 160	2 700	3 500	3 080	3 100
2 860	3 500	3 000	2 520	3 660	3 200	3 140	3 100	3 520	3 640	3 500	2 940
3 620	2 860	3 300	3 800	2 140	3 080	3 420	2 900	3 650	3 400	2 900	2 980
3 000	2 880	3 400	3 400	3 380	3 820	3 240	2 640	3 020	2 520	2 400	3 420
3 640	2 700	2 700	3 500	3 440	3 240	3 120	2 800	3 300	2 920	2 900	3 400
3 300	3 260	2 540	3 200	3 200	3 300	4 000	3 400	3 400	2 700	2 700	2 920
3 300	3 140	2 300	2 200	3 160	2 700	2 900	3 180	3 400	3 160	2 440	3 640
2 620	3 100	2 980	3 200	3 100	3 260	3 100	3 160	3 540	3 100	2 840	3 660
2 820	3 140	3 800	3 000	2 800	2 660	3 600	3 760	2 540	2 780	2 760	2 380
3 500	3 300	3 200	3 400	3 460	3 220	3 100	3 120	3 280	2 560	2 940	2 840
3 400	3 420	3 400	3 500	3 740	2 820	3 100	2 820	3 880	2 500	3 400	3 540

解 为粗略了解数据的分布情况，先画出直方图.

步骤如下：

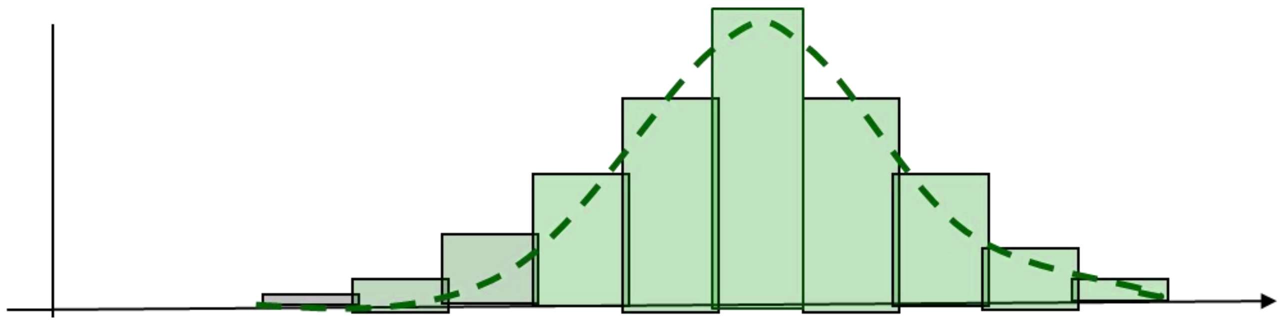
1.找出数据的最小值、最大值为2150, 4058，取区间[2100.5, 4100.5],它能覆盖[2150, 4058]；

2.将区间[2100.5 , 4100.5]等分为10个小区间，小区间的长度$\Delta=(4100.5-2100.5)/10=200$，$\Delta$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
2100.5-2300.5	3	0.0179	0.0179
2300.5-2500.5	5	0.0298	0.0476
2500.5-2700.5	13	0.0774	0.1250
2700.5-2900.5	22	0.1310	0.2560
2900.5-3100.5	28	0.1667	0.4226
3100.5-3300.5	39	0.2321	0.6548
3300.5-3500.5	28	0.1667	0.8214
3500.5-3700.5	21	0.1250	0.9464
3700.5-3900.5	7	0.0417	0.9881
3900.5-4100.5	2	0.0119	1.0000

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

$\hat{\mu} = 3127.56,\ \hat{\sigma}^2 = 378.428^2.$

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

例如

$$\hat{p}_1 = \hat{P}(A_1) = \hat{P}\{X \le 2300.5\}$$

$$= \Phi\left(\frac{2300.5 - 3127.56}{378.428}\right)$$

$$= \Phi(-2.1855) = 0.0144,$$

A_i	n_i	$\hat{p}_i$	$n\hat{p}_i$	$n_i^2/n\hat{p}_i$
$A_1 \quad x\le 2300.5$ $A_2 \quad 2300.5<x\le 2500.5$	$\left.\begin{array}{c}3\\5\end{array}\right\}$	$\left.\begin{array}{c}0.0144\\0.0343\end{array}\right\}$	$\left.\begin{array}{l}2.419\\5.762\end{array}\right\}8.181$	7.822
$A_3 \quad 2500.5<x\le 2700.5$	13	0.0809	13.591	12.435
$A_4 \quad 2700.5<x\le 2900.5$	22	0.1447	24.310	19.910
$A_5 \quad 2900.5<x\le 3100.5$	28	0.1972	33.130	23.665
$A_6 \quad 3100.5<x\le 3300.5$	39	0.2047	34.390	44.228
$A_7 \quad 3300.5<x\le 3500.5$	28	0.1616	27.149	28.878
$A_8 \quad 3500.5<x\le 3700.5$	21	0.0972	16.330	27.006
$A_9 \quad 3700.5<x\le 3900.5$ $A_{10} \quad 3900.5<x<\infty$	$\left.\begin{array}{c}7\\2\end{array}\right\}$	$\left.\begin{array}{c}0.0445\\0.0205\end{array}\right\}$	$\left.\begin{array}{l}7.470\\3.453\end{array}\right\}10.923$	10.922 Σ=174.866

$$\chi^2 = 174.866 - 168 = 6.866$$

$$\chi^2_{0.1}(k-r-1) = \chi^2_{0.1}(8-2-1) = 9.236 > 6.866$$

故在水平**0.1**下接受H_0, 认为数据来自正态总体。

Pearsonχ^2拟合优度检验的缺点：

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！