第八章 假设检验

假设检验的基本思想

单个正态总体参数的假设检验

两个正态总体参数的假设检验

假设检验与区间估计

拟合优度检验

8.1 假设检验的基本思想

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1.1 体重指数BMI是目前国际上常用的衡量人体胖瘦程度以及是否健康的一个标准. 专家指出，健康成年人的BMI 取值应在 18.55- 24.99 之间.某种减肥药广告宣称，连续使用该种减肥药一个星期便可达到减肥的效果.为了检验其说法是否可靠,随机抽取9位试验者(要求BMI 指数超过25,年龄在20-25岁女生),

(一)问题的提出

先让每位女生记录没有服用减肥药前的体重，然后让每位女生服用该减肥药，服药期间，要求每位女生保持正常的饮食习惯，连续服用该减肥药1周后，再次记录各自的体重. 测得服减肥药前后的体重差值X(服药前体重−服药后体重)（单位：kg):

1.5，0.6，−0.3，1.1，−0.8，0，2.2，−1.0，1.4

设$X \sim N(\mu,0.36)$，μ未知，根据目前的样本资料能否认为该减肥药广告中的宣称是可靠的?

例1.2 一种饼干的包装盒上标注净重200g，假设包装盒的重量为定值，且设饼干净重服从$N(\mu,\sigma^2)$，μ, σ^2均未知. 现从货架上取来3盒，称得毛重（单位：g）为 233，215，221，根据这些数据是否可以认为这种包装饼干的标准差超过6g？

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，孟德尔分别得到频数315，108，101，32，这些数据提供充分证据证明该理论吗？

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

H_0：$\theta \leq \theta_0$(或$\theta=\theta_0$)，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

- 对例1.1的统计分析

设体重差值$X \sim N(\mu, 0.36)$，μ未知，考虑假设

$$H_0: \mu = 0, H_1: \mu > 0 \text{(右边检验)}$$

注意到 $\bar{X}$ 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，$\bar{X}$ 取值应偏小，即当 $\bar{X}$ 取值偏大时，应该拒绝原假设.

当$\bar{X} \geq c$时，拒绝原假设H_0，

当$\bar{X} < c$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$\bar{W}$称为**接受域**.

上述例子中，可取检验统计量为$\bar{X}$,拒绝域为

$$W=\left\{(X_1,\cdots,X_9):\bar{X}\geq c\right\}$$

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)

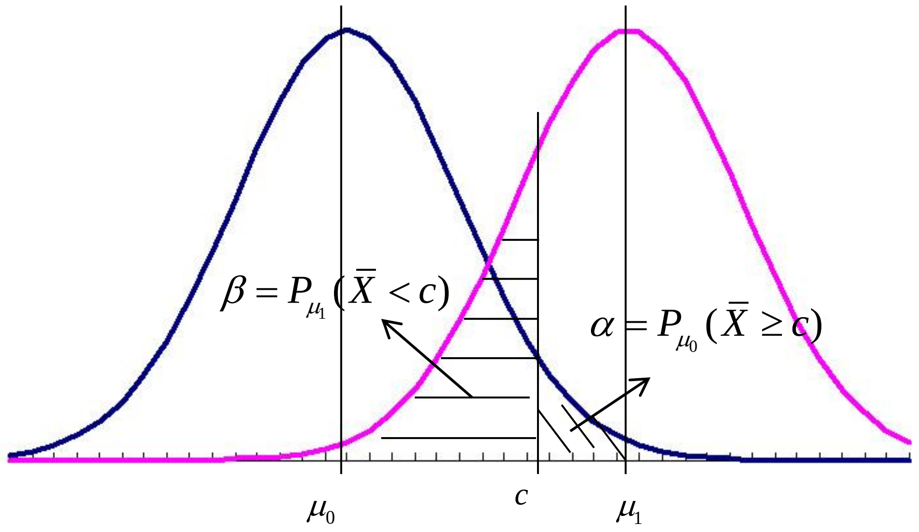
第II类错误：接受错误的原假设(存伪)

$\alpha = P\{$第I类错误$\}=P\{$拒绝$H_0 | H_0$是真实的$\}$,

$\beta = P\{$第II类错误$\}=P\{$接受$H_0 | H_0$是错误的$\}$.

例如：设总体$X \sim N(\mu,1)$,则$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu,\frac{1}{n})$,

$H_0:\mu=\mu_0, H_1:\mu=\mu_1(>\mu_0)$, 拒绝域：$\bar{X} \ge c$.



犯两类错误的概率相互制约

例1.1中，犯第I类错误的概率

$$\alpha(c) = P\{拒绝H_0 \mid H_0是真的\}$$

$$= P\{\bar{X} \geq c \mid \mu = 0\}$$

$$= P\{\frac{\bar{X}}{\sigma/\sqrt{n}} \geq \frac{c}{\sigma/\sqrt{n}} \mid \mu = 0\}$$

当H_0成立（即$\mu = 0$)时，$\frac{\bar{X}}{\sigma/\sqrt{n}} \sim N(0,1)$

$\alpha(c) = 1 - \Phi\left(\frac{c}{\sigma/\sqrt{n}}\right)$ **是c的单调减函数,**

犯第II类错误的概率

$$\beta(c) = P\{接受H_0 | H_0是假的\}$$

$$= P\{\bar{X} < c | \mu > 0\}$$

$$= P\left\{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{c-\mu}{\sigma/\sqrt{n}} \middle| \mu > 0\right\}$$

$$= \Phi\left\{\frac{c-\mu}{\sigma/\sqrt{n}}\right\}, \quad (\mu > 0)$$ **是*c*的单调增函数,**

在给定样本容量*n*时，不可能找到*c*，使$\alpha(c)$和$\beta(c)$都尽可能小.

Neyman-Pearson原则：

首先控制犯第I类错误的概率不超过常数α(称为**显著水平**)，$0<\alpha<1$，再寻找检验，使得犯第II类错误的概率尽可能小.

常取$\alpha=0.01$，0.05，0.1等.

在例**1.1**中，若取显著水平***α*=0.05**，则有

$$\alpha(c) = 1 - \Phi\left(\frac{c}{\sigma/\sqrt{n}}\right) \le 0.05$$

计算得

$$c \ge z_{0.05}\sigma/\sqrt{n} = 1.645 \times 0.6/3 = 0.329.$$

由于犯第**II**类错误的概率***β*(*c*)** 关于***c***是单调增函数，根据**Neyman-Pearson**原则，应取

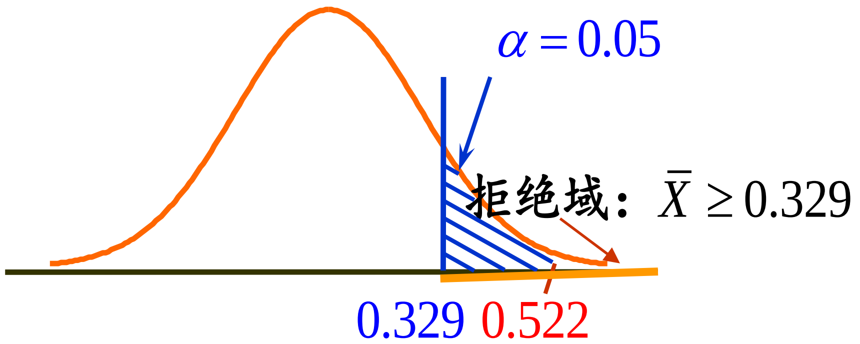
$$c = 0.329$$

因此拒绝域为

$$W = \left\{(X_1, \cdots, X_9) : \bar{X} \ge 0.329\right\}$$

根据实际样本资料，得 $\bar{x}=0.522>0.329,$

即，样本落在拒绝域中.



当原假设H_0成立时，样本落在拒绝域的概率不超过0.05，是小概率事件。

根据实际推断原理，有充分的理由拒绝原假设，认为厂家的宣传是可靠的.

根据上述检验规则，犯第I类错误的概率

$$\alpha(0.329)=0.05=\alpha$$

犯第II类错误的概率

$$\beta(0.329)=\Phi\left\{\frac{0.329-\mu}{0.6/\sqrt{9}}\right\}=\Phi\left\{\frac{0.329-\mu}{0.2}\right\}, \qquad (\mu>0)$$

例如，当$\mu=0.529$时，犯第II类错误的概率

$$\beta=\Phi\left\{\frac{0.329-0.529}{0.2}\right\}=\Phi(-1)=0.1587.$$

（四）*P*_值与统计显著性

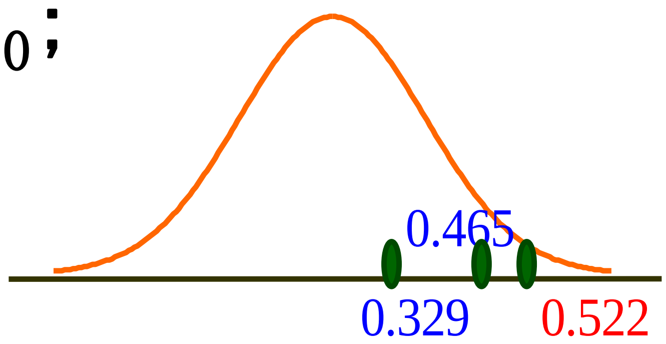
在例1.1中看到，取显著水平**α=0.05**，拒绝域

$W=\left\{\bar{X}\geq 0.329\right\}$，$\bar{x}=0.522>0.329$，拒绝$H_0$;

若取显著水平**α=0.01**，拒绝域 $W=\{\bar{X}\geq 0.465\}$,

$\bar{x}=0.522>0.465$，依然拒绝H_0;

那么，拒绝H_0的最小的值是多少？最小的显著水平又是多少？



$P_$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.

取$\bar{x}=0.522, W=\left\{\bar{X}\geq 0.522\right\}$，则

$$P\_=P\{\bar{X}\geq\bar{x}=0.522\mid\mu=0\}=1-\Phi(\frac{0.522}{0.6/\sqrt{9}})=0.0045$$

$$P\_=0.0045<0.05=\alpha$$

概率这么小的事件！竟然发生了！拒绝原假设！

用***P*_**值计算的优点是知道拒绝的概率大小.

$P_$值与显著水平α的关系：

若$P_ \le \alpha$，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

临界值法：

（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。

处理假设检验问题的基本步骤

***P*_值法：**

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

$\bar{X}$和S^2分别为样本均值和方差,显著水平为 α,

考虑关于μ和σ^2的检验问题.

(一)均值μ的检验

(1)σ^2已知——Z检验

双边假设检验问题

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

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

拒绝域形式 $|Z|\geq c.$ 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 $z_0=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$，则有

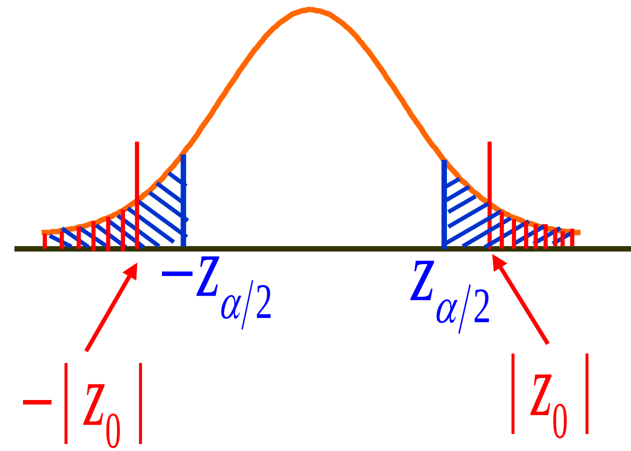
$$P\_=P_{H_0}\{|Z|\ge|z_0|\}$$

$$=2P_{H_0}\{Z\ge|z_0|\}$$

$$=2(1-\Phi(|z_0|)).$$

当$P_\le\alpha$时，拒绝原假设，

当$P_>\alpha$时，接受原假设.



红色区域概率值：$P_$值

蓝色区域概率值：α

$P_$值$\le\alpha$，拒绝H_0.

左边假设检验问题

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

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

拒绝域形式为 $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq c. \Rightarrow c = -z_\alpha$

$$P\{\text{第I类错误}\} = P\left\{\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq c \,\middle|\, \mu \geq \mu_0\right\}$$

$$= P\left\{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq c + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}} \,\middle|\, \mu \geq \mu_0\right\} = \Phi\left(c + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right) \leq \alpha$$

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

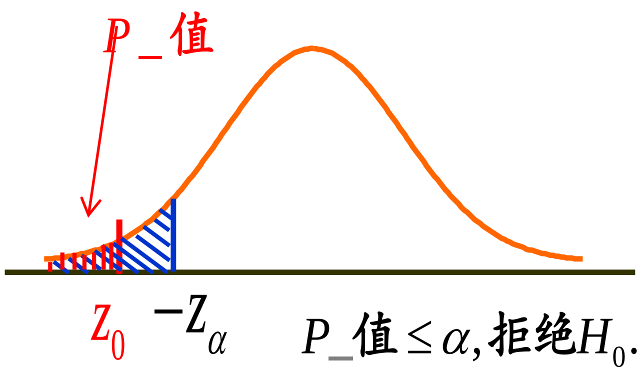
$P_$值的计算 对给定的样本观察值 $x_1,\cdots,x_n$, $z_0=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$.

$$P\_=\sup_{\mu\ge\mu_0}P_{H_0}\{Z\le z_0\}=\sup_{\mu\ge\mu_0}P\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\le z_0\mid \mu\ge\mu_0\}$$

$$=\sup_{\mu\ge\mu_0}P\{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\le z_0+\frac{\mu_0-\mu}{\sigma/\sqrt{n}}\mid \mu\ge\mu_0\}$$

$$=\sup_{\mu\ge\mu_0}\Phi(z_0+\frac{\mu_0-\mu}{\sigma/\sqrt{n}})$$

$$=\Phi(z_0).$$



$P_$值 $\le\alpha$, 拒绝 H_0.

思考题：比较

双边假设问题

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

右边假设检验问题

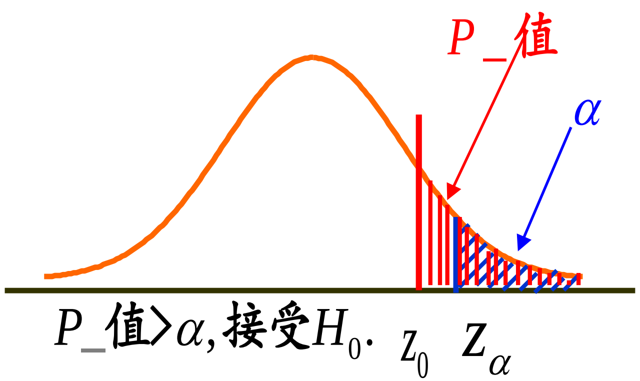
$H_0:\mu\le\mu_0, H_1:\mu>\mu_0$， 其中$\mu_0$是已知的常数

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

拒绝域为 $W=\left\{Z=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\ge z_\alpha\right\}.$

P_值的计算 对给定的样本观察值$x_1,\cdots,x_n$， $z_0=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}.$

$$
\begin{aligned}
P_- &= \sup_{\mu\le\mu_0} P_{H_0}\{Z\ge z_0\}\\
&= \sup_{\mu\le\mu_0}[1-\Phi(z_0+\frac{\mu_0-\mu}{\sigma/\sqrt{n}})]\\
&= 1-\Phi(z_0).
\end{aligned}
$$



(2)σ^2未知——t检验

双边假设检验问题

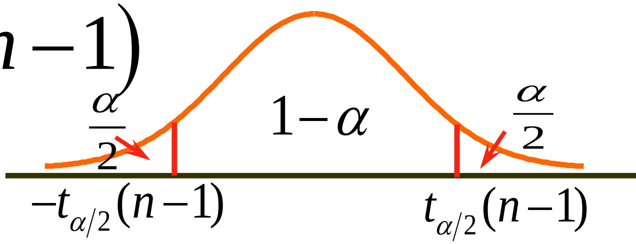
$$H_0:\mu=\mu_0, H_1:\mu\neq\mu_0$$

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

采用 $T=\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}$ 作检验统计量.

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

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



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　对给定样本值$x_1,\cdots,x_n$，记$t_0=\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}$，

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

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

左边假设检验问题

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

拒绝域为

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

对给定样本值$x_1, \cdots, x_n$，记$t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$， $P_$值为

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

右边假设检验问题

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

拒绝域为

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

对给定样本值$x_1, \cdots, x_n$, 记$t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$， $P_$值为

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

例2.1　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：

659	780	601	712	724	879	679	764
722	862	668	750	649	760	985	670

问是否有理由认为元件的平均寿命大于725（小时）？（取显著水平为0.05）

解：按题意需检验

$$H_0: \mu \le \mu_0 = 725, \qquad H_1: \mu > 725.$$

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

$n = 16, t_{0.05}(15) = 1.7531. \quad \bar{x} = 741.5, s = 98.7259$

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725小时。

由*Excel*可计算$P_$值为

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

接受原假设，即认为元件的平均寿命不大于725小时.

判断结果与前面一致！

例2.2　要求某种产品的平均长度不得低于10米，生产者从一批这种产品中随机抽取25件，测得其平均长度为9.5米，标准差为1米. 已知这批产品的长度服从正态分布。试在显著性水平0.05下确定这批产品是否合格？

解：按题意需检验

$$H_0: \mu \geq \mu_0 = 10, \qquad H_1: \mu < 10.$$

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

$n = 25, t_{0.05}(24) = 1.7109. \quad \bar{x} = 9.5, s = 1,$

计算得：$t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = -2.5 < -1.7109 = -t_{0.05}(24).$

t_0落在拒绝域内，故拒绝原假设，

认为这批产品的平均长度小于10米，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！

(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

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 $W=\left\{|T|\geq \mathrm{t}_{\alpha/2}(n-1)\right\}$,

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

$P_$值为

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

例2.3 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$\bar{D}, S_D^2$,

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

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

计算得：$t_0 = \dfrac{\bar{d}}{s_d \big/ \sqrt{n}} = -0.142, \left| t_0 \right| < t_{0.025}\left(9\right)$

样本没有落在拒绝域内，不拒绝原假设H_0.

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

- 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)参数σ^2的假设检验(设μ未知)

双边假设检验问题

$H_0:\sigma^2=\sigma_0^2, H_1:\sigma^2\neq\sigma_0^2,$ 其中σ_0^2是已知常数。

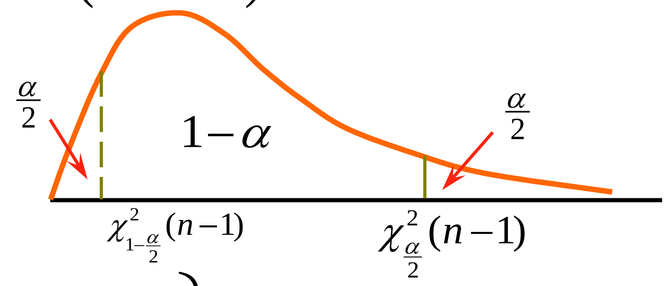
注意到S^2是σ^2的无偏估计量，取检验统计量为

$$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}$$

检验拒绝域形式为:

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

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



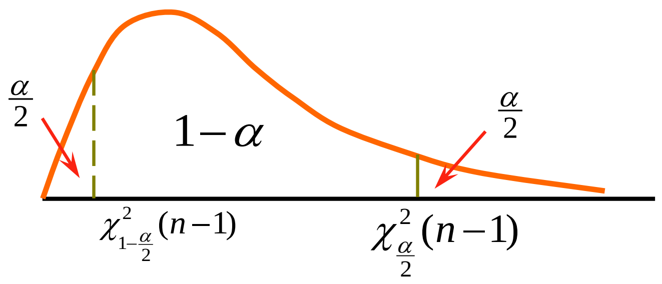
$$P\{\text{拒绝}H_0 \mid \text{当}H_0\text{为真}\}$$

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

为计算方便，习惯上取

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

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



拒绝域为：

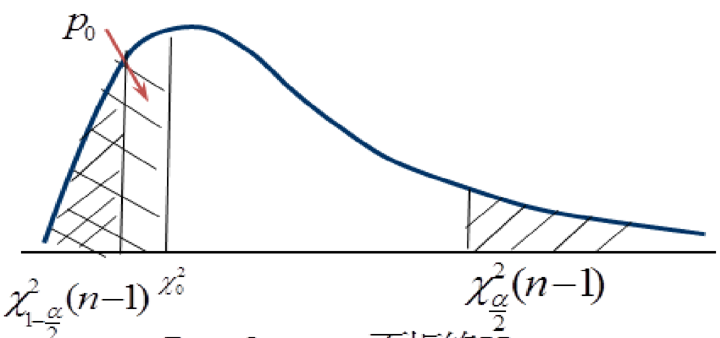
$$\boxed{\frac{(n-1)S^2}{\sigma_0^2} \le \chi^2_{1-\alpha/2}(n-1), \text{或} \frac{(n-1)S^2}{\sigma_0^2} \ge \chi^2_{\alpha/2}(n-1)}$$

----χ^2检验法

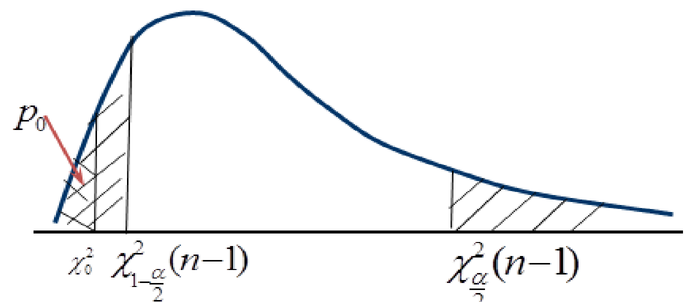
$P_$值计算：对样本观察值$x_1,\cdots,x_n$，记$\chi_0^2=\dfrac{(n-1)s^2}{\sigma_0^2}$，

令$p_0=P_{\sigma_0^2}\left\{\dfrac{(n-1)S^2}{\sigma_0^2}\le\chi_0^2\right\}=P\left\{\chi^2(n-1)\le\chi_0^2\right\}$，

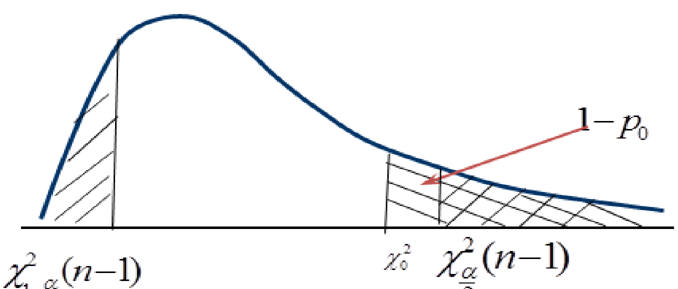
$P\_=2\min(p_0,1-p_0)$.



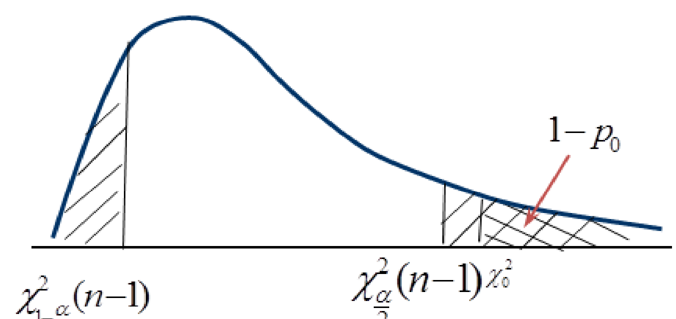
$P\_=2p_0>\alpha$,不拒绝H_0



$P\_=2p_0\le\alpha$,拒绝H_0



$P\_=2(1-p_0)>\alpha$,不拒绝H_0



$P\_=2(1-p_0)\le\alpha$,拒绝H_0

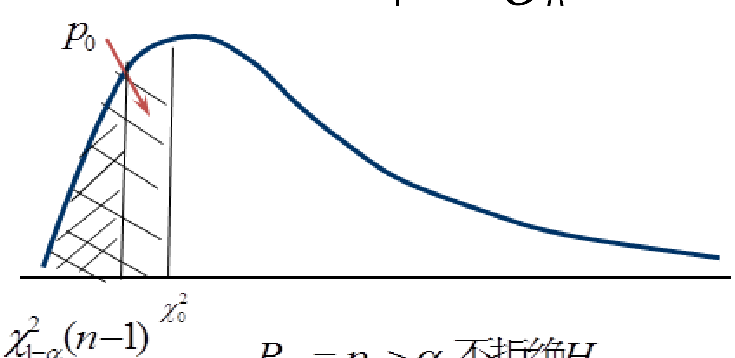
左边假设检验问题

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

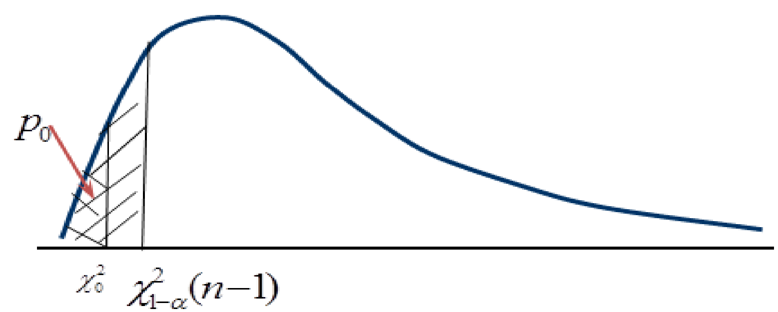
拒绝域为： $$\frac{(n-1)S^2}{\sigma_0^2} \le \chi^2_{1-\alpha}(n-1);$$

$P_$值计算：对样本观察值$x_1,\cdots,x_n$，记$\chi_0^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$，

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



$P\_=p_0>\alpha$,不拒绝H_0



$P\_=p_0\le\alpha$,拒绝H_0

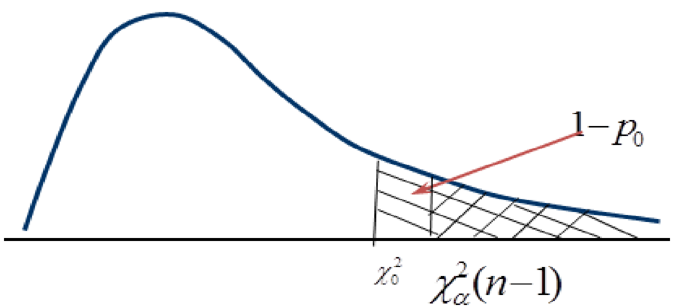
右边假设检验问题

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

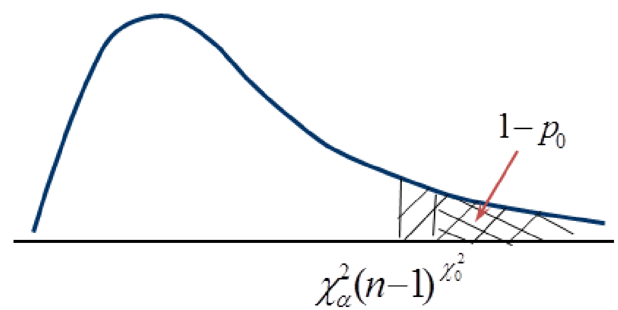
拒绝域为： $$\frac{(n-1)S^2}{\sigma_0^2} \ge \chi_\alpha^2(n-1);$$

$P_$值计算：对样本观察值$x_1,\cdots,x_n$，记$\chi_0^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$，

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



$P\_=1-p_0>\alpha$,不拒绝H_0



$P\_=1-p_0\le\alpha$,拒绝H_0

例1.2 一种饼干的包装盒上标注净重200g，假设包装盒的重量为定值，且设饼干净重服从$N(\mu,\sigma^2)$，μ, σ^2均未知. 现从货架上取来3盒，称得毛重(单位：g)为 233，215，221，根据这些数据是否可以认为这种包装饼干的标准差超过6g？

解: $H_0: \sigma^2 \le 36, \quad H_1: \sigma^2 > 36,$

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

查表得: $\chi_{0.05}^2\left(2\right) = 5.991,$

计算得: s^2=84, $\chi_0^2 = \dfrac{\left(3-1\right)\times 84}{36} = 4.667 < 5.991$

不能拒绝原假设，即认为标准差不超过6g.

计算$P_ = P\{\chi^2(2) > 4.667\} = 0.097 > 0.05$

作出同样判断.

8.3两个正态总体参数的假设检验

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

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

取显著水平为α，考虑两个均值比较和

两个方差比较的检验问题.

(一)比较两个总体均值的检验

双边假设检验问题

$$H_0:\mu_1=\mu_2,\ H_1:\mu_1\neq\mu_2.$$

1. 当σ_1^2和σ_2^2已知时

取检验统计量为 $Z=\dfrac{\bar{X}-\bar{Y}}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$

拒绝域的形式为 $|Z|\geq c.$

当H_0成立时，$Z\sim N(0,1).$

检验拒绝域为：$|Z|\geq z_{\alpha/2}$ $--$z检验

$P_$值计算：对样本观察值$x_1,\cdots,x_{n_1},y_1,\cdots,y_{n_2}$,

记$z_0=\dfrac{\bar{x}-\bar{y}}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$.

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

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，

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

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

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

---两样本精确t检验

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

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

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

（i）当两样本容量都充分大时，若原假设成立，则统计量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 $P\_ = P_{H_0}\{|T| \geq |t_0|\} \approx 2P\{Z \geq |t_0|\}$,

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

（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

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

则检验的拒绝域为 $|T|\geq t_{\alpha/2}(k)$

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

——两样本近似t检验

类似地，　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 $H_0:\mu_1\ge\mu_2, H_1:\mu_1<\mu_2$

检验拒绝域为：$T=\dfrac{\left(\bar{X}-\bar{Y}\right)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\le -t_\alpha(n_1+n_2-2).$

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

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

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

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

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

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

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

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

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

$n_1 = 220, \quad \bar{x} = 2.46, \quad s_1 = 0.57;$

$n_2 = 205, \quad \bar{y} = 2.55, \quad s_2 = 0.48;$

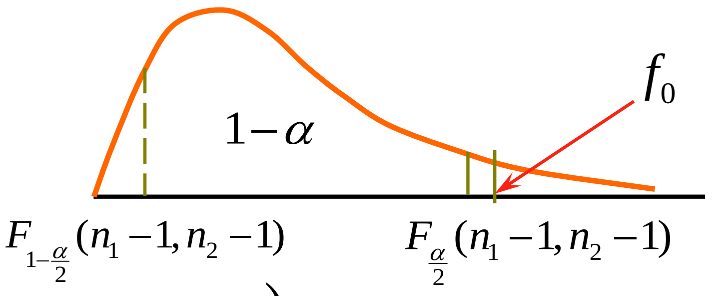
$t_{0.05}(423) \approx z_{0.05} = 1.645, s_w = 0.535, \sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}} = 0.097$

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

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

(二)比较两个总体方差的检验(设μ_1,μ_2未知)

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



取检验统计量为 $F=\dfrac{S_1^2}{S_2^2}.$

在原假设成立时， $F\sim F\left(n_1-1,n_2-1\right)$

检验拒绝域为：

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

$$P\_=2\min\{P(F\ge f_0),P(F\le f_0)\}$$

$P\_\le\alpha$,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 $H_0:\sigma_1^2 \geq \sigma_2^2, H_1:\sigma_1^2 < \sigma_2^2$

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

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3.2 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$H_0:\mu_1\le\mu_2, H_1:\mu_1>\mu_2(\alpha=0.1)$;

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

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

$n_1=8,\ \bar{x}=15.05,\ s_1^2=0.0457$；$n_2=9,\ \bar{y}=14.9,\ s_2^2=0.0575;$

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。

$$P\_=2P(F(7,8)\le 0.795)=0.775>0.1.$$

$n_1 = 8,\ \bar{x} = 15.05,\ s_1^2 = 0.0457;\ n_2 = 9,\ \bar{y} = 14.9,\ s_2^2 = 0.0575;$

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

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

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

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

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

$\dfrac{\left|\bar{X}-\bar{Y}\right|}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}=1.354<t_{0.05}(15)=1.7531$, 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 $H_0\colon \mu_1 \le \mu_2$.

（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$H_0\colon \mu_1 = \mu_2$.

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$X_1, X_2, ..., X_n$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

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

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

显著水平为α的检验拒绝域$\frac{|\bar{X} - \mu_0|}{\sigma / \sqrt{n}} \geq z_{\alpha/2}$

接受域$\frac{|\bar{X} - \mu_0|}{\sigma / \sqrt{n}} < z_{\alpha/2} \Leftrightarrow \bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} < \mu_0 < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$

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

思考题：

$N\left(\mu_1,\sigma_1^2\right)$的样本为$X_1,X_2,\cdots,X_{n_1}$和$N\left(\mu_2,\sigma_2^2\right)$的样本为$Y_1,Y_2,\cdots,Y_{n_2}$,

两个样本独立，在置信水平为1-α下有下列置信区间：

(1)$\sigma_1^2=\sigma_2^2$时，$\mu_1-\mu_2$的置信区间$\left(\left(\bar{X}-\bar{Y}\right)\pm t_{\frac{\alpha}{2}}\left(n_1+n_2-2\right)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$

(2)$\dfrac{\sigma_1^2}{\sigma_2^2}$的置信区间$\left(\frac{S_1^2}{S_2^2}\frac{1}{F_{\frac{\alpha}{2}}\left(n_1-1,n_2-1\right)},\frac{S_1^2}{S_2^2}\frac{1}{F_{1-\frac{\alpha}{2}}\left(n_1-1,n_2-1\right)}\right)$.

为什么说，均值差的置信区间包含0，说明均值之间没有显著差异？

方差比的置信区间包含1，说明方差之间没有显著差异？

这是因为(1)

$$0 \in \left(\left(\bar{X} - \bar{Y}\right) \pm t_{\frac{\alpha}{2}}\left(n_1 + n_2 - 2\right) S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

$$\Leftrightarrow \left(\bar{X} - \bar{Y}\right) + t_{\frac{\alpha}{2}}\left(n_1 + n_2 - 2\right) S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} > 0,$$

$$\left(\bar{X} - \bar{Y}\right) - t_{\frac{\alpha}{2}}\left(n_1 + n_2 - 2\right) S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < 0,$$

$$\Leftrightarrow \frac{\left|\bar{X} - \bar{Y}\right|}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < t_{\frac{\alpha}{2}}\left(n_1 + n_2 - 2\right),$$

恰好是$H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$样本落在接受域的情形.

$$(2)\ \frac{S_1^2}{S_2^2}\frac{1}{F_{\frac{\alpha}{2}}\left(n_1-1,n_2-1\right)}<1<\frac{S_1^2}{S_2^2}\frac{1}{F_{1-\frac{\alpha}{2}}\left(n_1-1,n_2-1\right)}$$

$$\Leftrightarrow F_{1-\frac{\alpha}{2}}\left(n_1-1,n_2-1\right)<\frac{S_1^2}{S_2^2}<F_{\frac{\alpha}{2}}\left(n_1-1,n_2-1\right)$$

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

样本落在接受域的情形.

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$P\{X = t_i\} = p_i, i = 1, 2, \ldots$

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$A_1,\ldots,A_k$.

2. 以$n_i\,(i=1,\ldots,k)$记样本观察值$x_1,\ldots,x_n$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

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

$$\chi^2=\sum_{i=1}^{k}\frac{n_i^2}{n\hat{p}_i}-n\geq\chi_\alpha^2(k-r-1),$$ （有r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$\boldsymbol{n\geq 50}$，$\boldsymbol{np_i}$ **(或** $\boldsymbol{n\hat{p}_i)\geq 5}$，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

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

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？

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

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

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

豆子状态x	1	2	3	4
实测频数n_i	315	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$\Delta=(4100.5-2100.5)/10=200$,

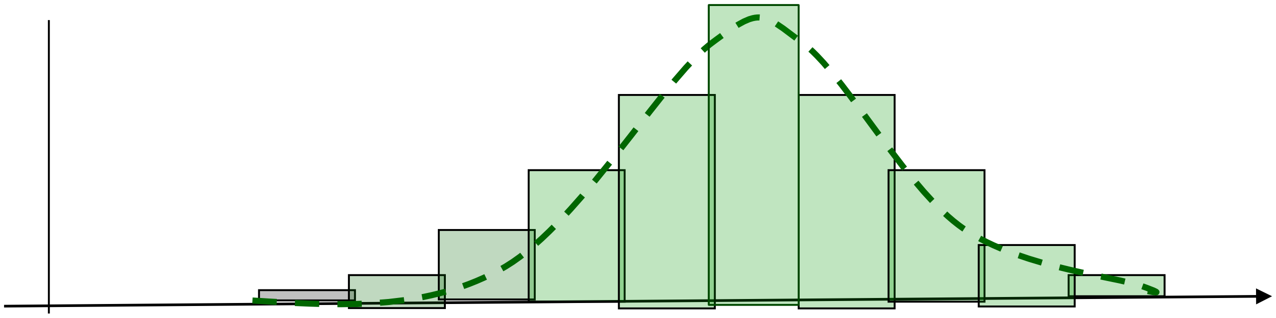
Δ称为组距，小区间的端点称为组限，建立

下表：

组 限	频数	频率	累计频率
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

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

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

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

A_i		n_i	$\hat{p}_i$	$n\hat{p}_i$	$n_i^2/n\hat{p}_i$
A_1	$x \le 2300.5$	3	0.0144	2.419 } 8.181	
A_2	$2300.5 < x \le 2500.5$	5	0.0343	5.762	7.822
A_3	$2500.5 < x \le 2700.5$	13	0.0809	13.591	12.435
A_4	$2700.5 < x \le 2900.5$	22	0.1447	24.310	19.910
A_5	$2900.5 < x \le 3100.5$	28	0.1972	33.130	23.665
A_6	$3100.5 < x \le 3300.5$	39	0.2047	34.390	44.228
A_7	$3300.5 < x \le 3500.5$	28	0.1616	27.149	28.878
A_8	$3500.5 < x \le 3700.5$	21	0.0972	16.330	27.006
A_9	$3700.5 < x \le 3900.5$	7	0.0445	7.470 } 10. 923	10.922
A_{10}	$3900.5 < x < \infty$	2	0.0205	3.453	$\Sigma=174.866$

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

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

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！