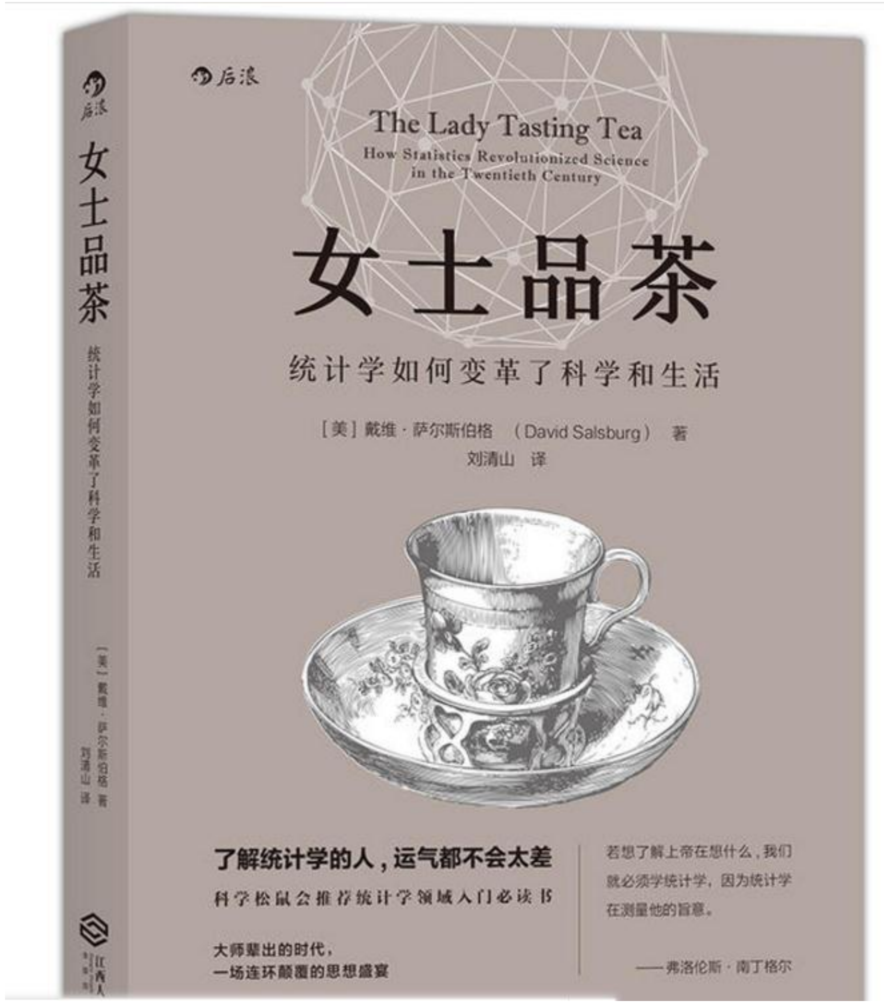
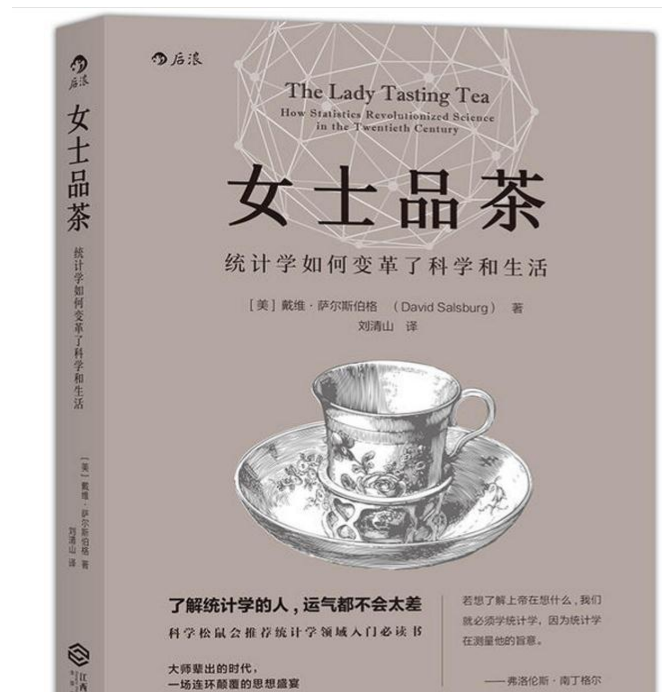
第八章　　假设检验

- 1 假设检验
- 2 正态总体均值的假设检验
- 3 正态总体方差的假设检验
- 4 置信区间与假设检验之间的关系
- 7 假设检验问题的*p*值法


女士品茶


后浪
The Lady Tasting Tea
How Statistics Revolutionized Science in the Twentieth Century
女士品茶
统计学如何变革了科学和生活
[美] 戴维·萨尔斯伯格 （David Salsburg） 著
刘清山 译
了解统计学的人，运气都不会太差
科学松鼠会推荐统计学领域入门必读书
大师辈出的时代，
一场连环颠覆的思想盛宴
若想了解上帝在想什么，我们就必须学统计学，因为统计学在测量他的旨意。
——弗洛伦斯·南丁格尔
后浪
女士品茶
统计学如何变革了科学和生活
[美] 戴维·萨尔斯伯格 著
刘清山 译

- 《女士品茶——20世纪统计学怎样变革了科学》，是美国统计学家萨尔斯伯格以“女士品茶问题”为切入点所著的一部关于统计学历史与变革的书，以一种全新的视角带领读者进入统计学的世界，体会统计学带给哲学观、宇宙观的变革.



女士品茶

- 20世纪20年代后期在英国剑桥一个夏日的下午，一群大学的绅士和他们的夫人以及来访者，正围坐在户外的桌旁享用下午的奶茶.

女士品茶

- 20世纪20年代后期在英国剑桥一个夏日的下午，一群大学的绅士和他们的夫人以及来访者，正围坐在户外的桌旁享用下午的奶茶.
- 奶茶调制时候可以先倒茶后倒牛奶（TM），也可以先倒牛奶后倒茶(MT).这时候，一名女士说她能区分这两种不同做法的调制出来的奶茶.
-

女士品茶

- 20世纪20年代后期在英国剑桥一个夏日的下午，一群大学的绅士和他们的夫人以及来访者，正围坐在户外的桌旁享用下午的奶茶.
- 奶茶调制时候可以先倒茶后倒牛奶（TM），也可以先倒牛奶后倒茶(MT).这时候，一名女士说她能区分这两种不同做法的调制出来的奶茶.
- “让我们检验这个命题吧.”为此Fisher开始规划实验.

女士品茶

- 他准备了10杯调好的奶茶，先倒茶与先倒奶都有.服务员一杯一杯地奉上，

女士品茶

- 他准备了10杯调好的奶茶，先倒茶与先倒奶都有.服务员一杯一杯地奉上，结果那位女士全部正确地分辨出来.

女士品茶

- 他准备了10杯调好的奶茶，先倒茶与先倒奶都有.服务员一杯一杯地奉上，结果那位女士全部正确地分辨出来.
- Fisher的想法是这样的，如果女士无此鉴别能力，只能猜，每次猜对的概率是0.5

女士品茶

- 他准备了10杯调好的奶茶，先倒茶与先倒奶都有.服务员一杯一杯地奉上，结果那位女士全部正确地分辨出来.
- Fisher的想法是这样的，如果女士无此鉴别能力，只能猜，每次猜对的概率是0.5

$$0.5^2 = 0.25$$

$$0.5^3 = 0.125$$

$$0.5^4 = 0.0625$$

$$0.5^5 = 0.03125$$

$$\cdots\cdots$$

$$0.5^{10} = 0.0009765625 \approx 0.001$$

女士品茶

- 他准备了10杯调好的奶茶，先倒茶与先倒奶都有.服务员一杯一杯地奉上，结果那位女士全部正确地分辨出来.
- Fisher的想法是这样的，如果女士无此鉴别能力，只能猜，每次猜对的概率是0.5

$0.5^2 = 0.25$

$0.5^3 = 0.125$

$0.5^4 = 0.0625$

$0.5^5 = 0.03125$

$\cdots\cdots$

$0.5^{10} = 0.0009765625 \approx 0.001$

小概率事件在一次试验中实际上几乎是不发生的。（称之为实际推断原理）.

女士品茶

- 他准备了10杯调好的奶茶，先倒茶与先倒奶都有.服务员一杯一杯地奉上，结果那位女士全部正确地分辨出来.
- Fisher的想法是这样的，如果女士无此鉴别能力，只能猜，每次猜对的概率是0.5

$$0.5^2 = 0.25$$

$$0.5^3 = 0.125$$

$$0.5^4 = 0.0625$$

$$0.5^5 = 0.03125$$

$$\cdots\cdots$$

$$0.5^{10} = 0.0009765625 \approx 0.001$$

Fisher的思维方式

如果试验结果与原有假设发生矛盾，拒绝原有假设。

“假设检验”思想：

为了检验一个假设是否成立，先假定它是成立的，然后看在这个假设成立的条件下，是否会导致不合理结果.

女士品茶

- 如果该女士说对了9杯（或8杯），该如何进行判断呢？判断会发生错误吗？发生错误的概率是多少？
- Fisher 对这些细节做了周密的研究，提出一些新的概念，建立了一套可行的方法，形成了假设检验理论，为进一步发展假设检验理论与方法打下了牢固基础.

例8(P14) 某接待站在某一周曾接待过 12次来访,所有这 12 次接待都是在周二和周四进行的,问是否可以推断接待时间是有规定的.

例8(P14) 某接待站在某一周曾接待过 12次来访,所有这 12 次接待都是在周二和周四进行的,问是否可以推断接待时间是有规定的.

解 假设接待站的接待时间没有规定,且各来访者在一周的任一天中去接待站是等可能的.

故一周内接待 12 次来访共有7^{12}种.

12 次接待都是在周二和周四进行的共有2^{12}种.

故12 次接待都是在周二和周四进行的概率为

$$p = \frac{2^{12}}{7^{12}} = 0.0000003.$$

小概率事件在一次试验中实际上几乎是不发生的（称之为实际推断原理）， 从而可知接待时间是有规定的.

例1.　某车间用一台包装机包装葡萄糖. 包得的袋装糖重是一个随机变量, 它服从正态分布. 当机器正常时, 其均值为0.5公斤, 标准差为0.015公斤. 某日开工后为检验包装机是否正常, 随机地抽取它所包装的糖9袋, 称得净重为:

0.497	0.506	0.518	0.524	0.498
0.511	0.520	0.515	0.512	

问机器是否正常?

例1.　某车间用一台包装机包装葡萄糖.包得的袋装糖重是一个随机变量,它服从正态分布.当机器正常时,其均值为0.5公斤,标准差为0.015公斤.某日开工后为检验包装机是否正常,随机地抽取它所包装的糖9袋,称得净重为:

0.497	0.506	0.518	0.524	0.498
0.511	0.520	0.515	0.512	

问机器是否正常?

$$H_0: \mu = \mu_0 = 0.5 \qquad H_1: \mu \neq \mu_0$$

“假设检验”思想：

为了检验一个假设是否成立，先假定它是成立的，然后看在这个假设成立的条件下，是否会导致不合理结果.

如果假设H_0为真,

则观察值 $\bar{x}$与μ_0的偏差

$|\bar{x}-\mu_0|$一般不应太大.

检验法则:

$|\bar{x}-\mu_0| \geq c$ 拒绝H_0

如果假设H_0为真，

则观察值 $\bar{x}$与μ_0的偏差

$|\bar{x}-\mu_0|$一般不应太大.

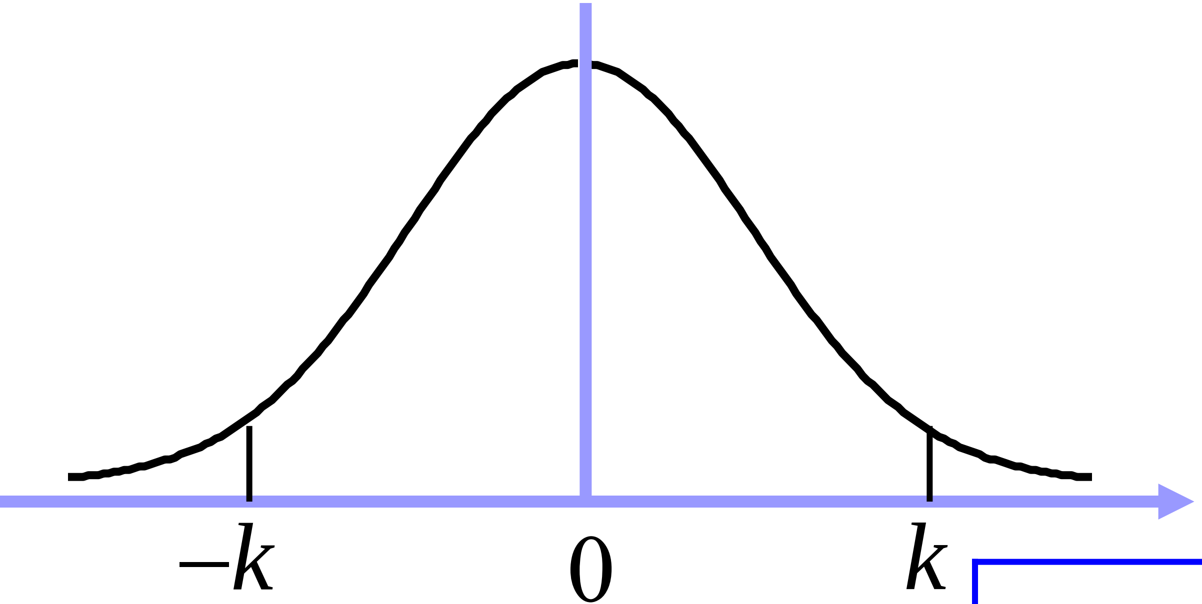
检验法则:

$$\left|\bar{x}-\mu_0\right| \geq c \quad \text{拒绝} H_0$$

$$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

检验法则:

$$\frac{\left|\bar{x}-\mu_0\right|}{\sigma/\sqrt{n}} \geq k \quad \text{拒绝} H_0$$



检验法则：

$$\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq k \quad 拒绝H_0$$

◆因为决策的依据是样本, 当实际上H_0为真时仍可能做出拒绝H_0的决策.

◆这是一种错误, 犯这种错误的概率记为

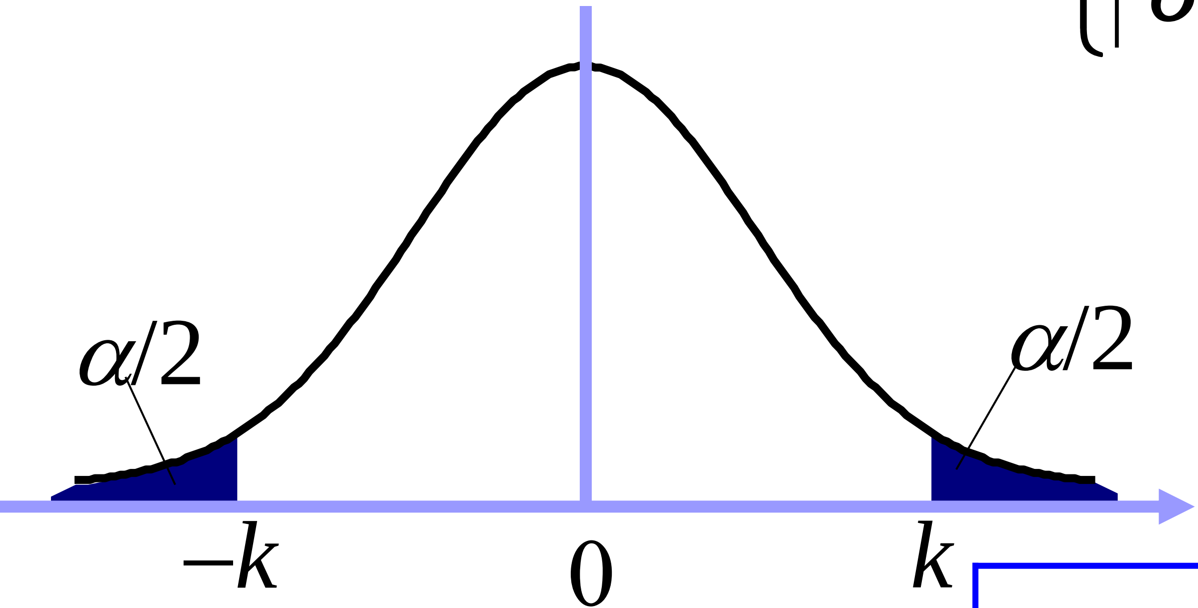
$$P\{当H_0为真拒绝H_0\}$$

$$P\{当H_0为真拒绝H_0\}\le\alpha. \qquad (1.1)$$

检验法则:

$$\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}}\ge k \quad 拒绝H_0$$

$$P\{\text{当}H_0\text{为真拒绝}H_0\} = P_{\mu_0}\left\{\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right| \geq k\right\} = \alpha.$$



检验法则:

$$\frac{\left|\overline{x}-\mu_0\right|}{\sigma/\sqrt{n}} \geq k \quad \text{拒绝}H_0$$

检验法则：

$$\frac{\left|\bar{x}-\mu_0\right|}{\sigma/\sqrt{n}} \geq z_{\frac{\alpha}{2}}$$

拒绝H_0

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

接受H_0

$\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$ 称为检验统计量

检验法则:

$$\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\frac{\alpha}{2}}$$ 拒绝 H_0

$$\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < z_{\frac{\alpha}{2}}$$ 接受 H_0

上例所采用的检验法则符合实际推断原理:

"**一个小概率事件在一次试验中几乎是不可能发生的**".

前面的检验问题常叙述成: 在**显著性水平**α下, 检验假设

$$H_0:\mu=\mu_0, \qquad H_1:\mu\neq\mu_0. \qquad (1.2)$$

H_0称为**原假设**或**零假设**, H_1称为**备择假设**.

当**检验统计量**取某个区域C中的值时, 我们拒绝原假设H_0, 则C称为**拒绝域**. 拒绝域的边界点称为**临界点**,

对 H_0 采取保护的态度

此药无效
此药有效

例 鉴定某药品的疗效，研发主管应该提出假设

H_0:　　　　　　　　　　，H_1:

对 H_0 采取保护的态度

此药无效
此药有效

例 鉴定某药品的疗效，研发主管应该提出假设

H_0：　　　　　　，H_1：

此人无罪
此人有罪

例 法庭审判时，对于被告人，法官应该提出假设

H_0：　　　　　　，H_1：

无罪推定

编辑

无罪推定（presumption of innocence），又可称为无罪类推（与有罪类推相对应），简单地说是指任何人在未经依法判决有罪之前，应视其无罪。除以上内容外，无罪推定还包括：被告人不负有证明自己无罪的义务，被告人提供证明有利于自己的证据的行为是行使辩护权的行为，不能因为被告人没有或不能证明自己无罪而认定被告人有罪。

无罪推定原则是现代法治国家刑事司法通行的一项重要原则，是国际公约确认和保护的基本人权，也是联合国在刑事司法领域制定和推行的最低限度标准之一。

1996年3月第一次修正后的《中华人民共和国刑事诉讼法》第12条明确规定："未经人民法院依法判决，对任何人都不得确定有罪"。虽然该规定中没有出现"推定"或"假定"无罪的规范性表述，但却含有无罪推定的精神。同时，在该法第162条第（3）项中还相应规定了罪疑从无原则，即："证据不足，不能认定被告人有罪的，应当作出证据不足、指控的犯罪不能成立的无罪判决。"具体体现在以下几个方面：

真实情况	所作决策	
	接受H_0	拒绝H_0
H_0为真	正确	犯第I类错误
H_0非真	犯第II类错误	正确

真实情况	所作决策	
	接受H_0	拒绝H_0
H_0为真	正确	犯第I类错误
H_0非真	犯第II类错误	正确

真实情况	所作决策	
	认为无罪	认为有罪
无罪	正确	犯第I类错误
有罪	犯第II类错误	正确

真实情况	所作决策	
	接受H_0	拒绝H_0
H_0为真	正确	犯第I类错误
H_0非真	犯第II类错误	正确

真实情况	所作决策	
	认为无罪	认为有罪
无罪	正确	犯第I类错误
有罪	犯第II类错误	正确

漏网之鱼

冤假错案

真实情况	所作决策	
	接受H_0	拒绝H_0
H_0为真	正确	犯第I类错误
H_0非真	犯第II类错误	正确

这种只对犯第I类错误的概率加以控制，
而不考虑犯第II类错误的概率的检验，
称为<u>显著性检验</u>.

在显著性水平α下, 检验假设

$$H_0:\mu=\mu_0, \qquad H_1:\mu\neq\mu_0. \qquad (1.2)$$

形如(1.2)式中的备择假设H_1, 表示μ可能大于也可能小于μ_0, 称为双边备择假设, 而称形如(1.2)式的假设检验为双边假设检验.

综上所述, 处理参数的假设检验问题步骤为:

1. 根据实际问题要求, 提出原假设H_0及备择假设 H_1;

2. 给定显著性水平α以及样本容量n;

3. 确定检验统计量以及拒绝域的形式;

4. 按$P\{$当H_0为真拒绝$H_0\}<=\alpha$求出拒绝域;

5. 取样, 根据样本观察值作出决策, 是接受H_0还是拒绝H_0.

练习1．已知某炼铁厂生产的铁水的含碳量在正常情况下服从正态分布$N(4.55, 0.12^2)$．现在测定了9炉铁水，测得其平均含碳量为4.49，若方差没有变化，可否认为现在生产的铁水的平均含碳量仍为4.55(取α=0.05)？

有时只关心总体均值是否增大. 例如试验新工艺以提高材料的强度。此时，我们需要检验假设

$$H_0: \mu \le \mu_0, \qquad H_1: \mu > \mu_0. \tag{1.3}$$

形如**(1.3)**的假设检验, 称为右边检验。

类似地, 有时需要检验假设

$$H_0: \mu \ge \mu_0, \qquad H_1: \mu < \mu_0. \tag{1.4}$$

形如**(1.4)**的假设检验, 称为左边检验.

设总体$X \sim N(\mu,\sigma^2)$, σ为已知, $X_1,X_2,\ldots,X_n$是来自X的样本. 给定显著性水平α. 来求检验问题

$$H_0:\mu \le \mu_0,\quad H_1:\mu > \mu_0 \qquad (1.3)$$

的拒绝域.

检验法则:

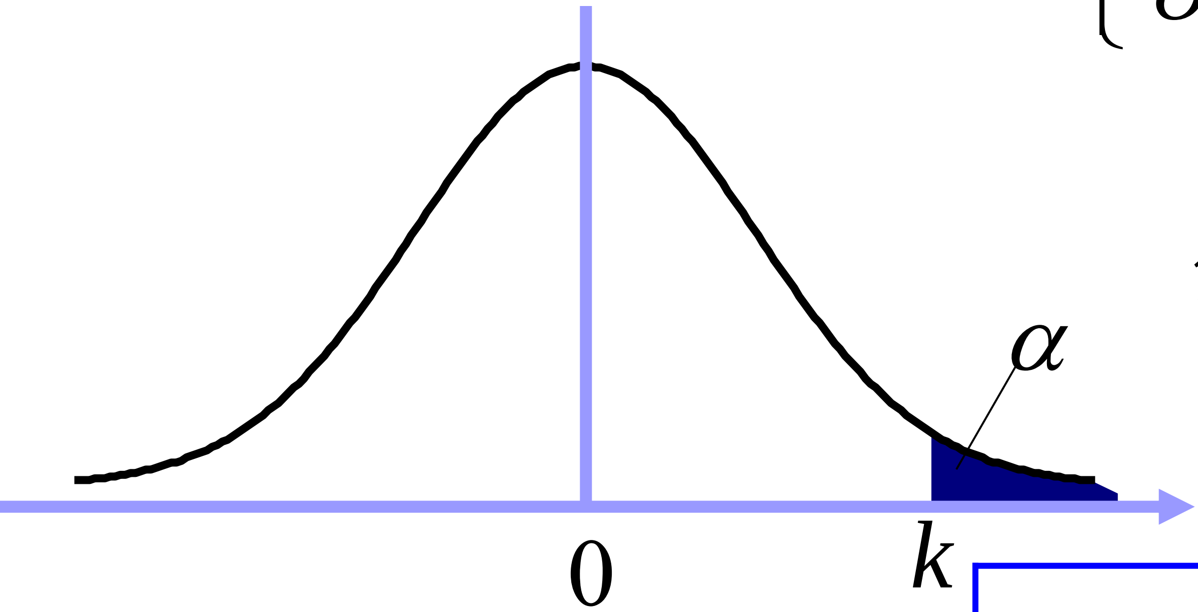
$\bar{x} \ge c$ 拒绝H_0

$$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

检验法则:

$$\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \ge k$$ 拒绝H_0

$$P\{\text{当}H_0\text{为真拒绝}H_0\} = P_{\mu_0}\left\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \geq k\right\} = \alpha.$$



由标准正态分布分位点的定义得:

$k=z_\alpha$.

检验法则:

$$\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \geq k \quad \text{拒绝} \boldsymbol{H_0}$$

例 公司从生产商购买牛奶。公司怀疑生产商在牛奶中掺水以谋利。通过测定牛奶的冰点，可以检测出牛奶是否掺水。天然牛奶的冰点温度近似服从正态分布，均值$\mu_0=-0.545$，标准差为$\sigma=0.008$.牛奶掺水可使冰点温度升高而接近于水的冰点温度，测得生产商提交的5批牛奶的冰点温度，其均值为-0.535.

问是否可以认为生产商在牛奶中掺水?

取显著性水平$\alpha=0.05$.

第二节 正态总体均值的假设检验

一、单个总体均值μ的检验

二、两个总体均值差的检验

三、基于成对数据的检验

一、单个总体$N(\mu,\sigma^2)$均值μ的检验

1. σ^2为已知，关于μ的检验

在上节中讨论过正态总体$N(\mu,\sigma^2)$

当σ^2为已知时，关于$\mu=\mu_0$的检验问题：

(1) 假设检验 $H_0:\mu=\mu_0, H_1:\mu\neq\mu_0$;

(2) 假设检验 $H_0:\mu\leq\mu_0, H_1:\mu>\mu_0$;

(3) 假设检验 $H_0:\mu\geq\mu_0, H_1:\mu<\mu_0$.

讨论中都是利用 H_0 为真时服从 $N(0,1)$ 分布的统计量 $Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ 来确定拒绝域的,这种检验法称为 $\boxed{Z\text{ 检验法.}}$

2. σ^2为未知, 关于μ的检验

设总体$X \sim N(\mu,\sigma^2)$,其中μ,σ^2未知,显著性水平为α

检验假设$H_0:\mu=\mu_0, H_1:\mu\neq\mu_0$。

原检验法则：

$$\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq k \quad \text{拒绝} H_0$$

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

检验法则：

$$\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq k \quad \text{拒绝}\boldsymbol{H_0}$$

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

不是统计量

检验法则：

$$\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq k \quad \text{拒绝}\boldsymbol{H_0}$$

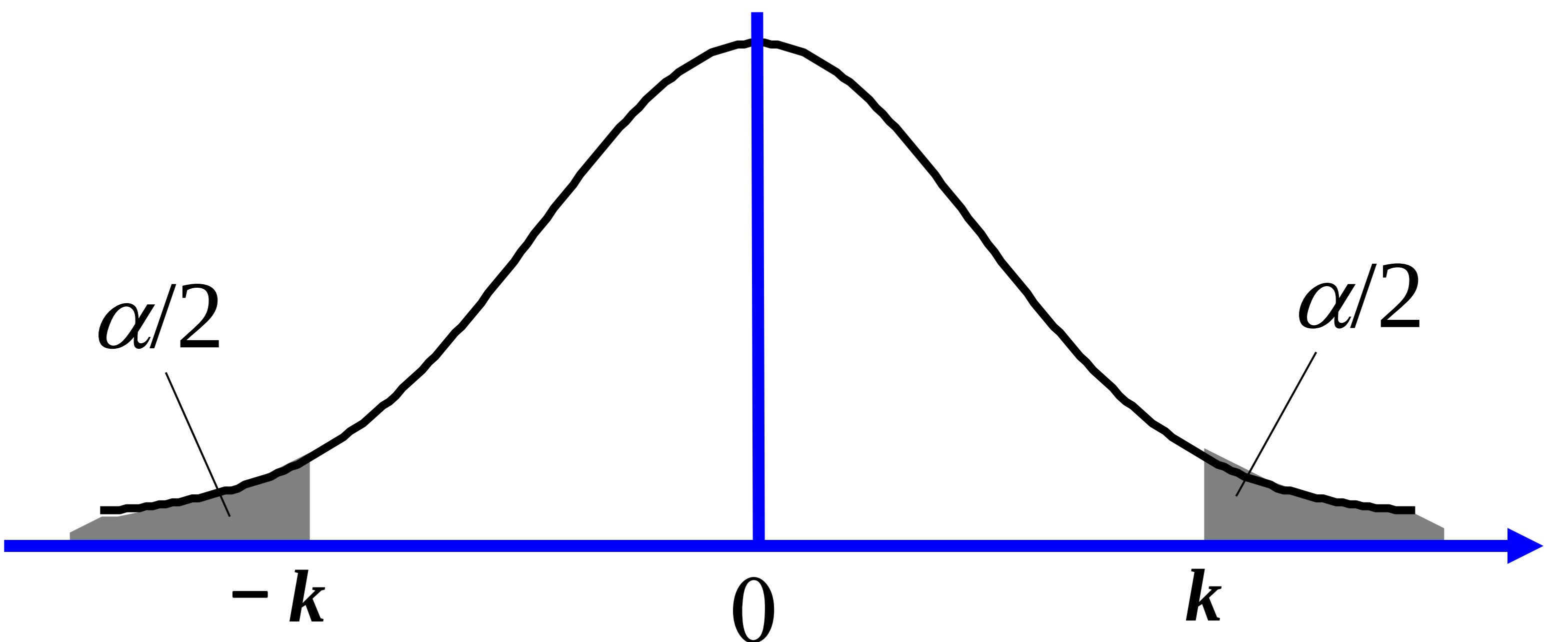
$$\frac{\bar{\boldsymbol{X}}-\boldsymbol{\mu_0}}{\boldsymbol{\sigma}/\sqrt{\boldsymbol{n}}} \sim N(\mathbf{0,1}) \qquad \frac{\bar{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1)$$

不是统计量

检验法则：

$$\frac{\left|\bar{x}-\mu_0\right|}{s/\sqrt{n}} \geq k \quad \text{拒绝} H_0$$

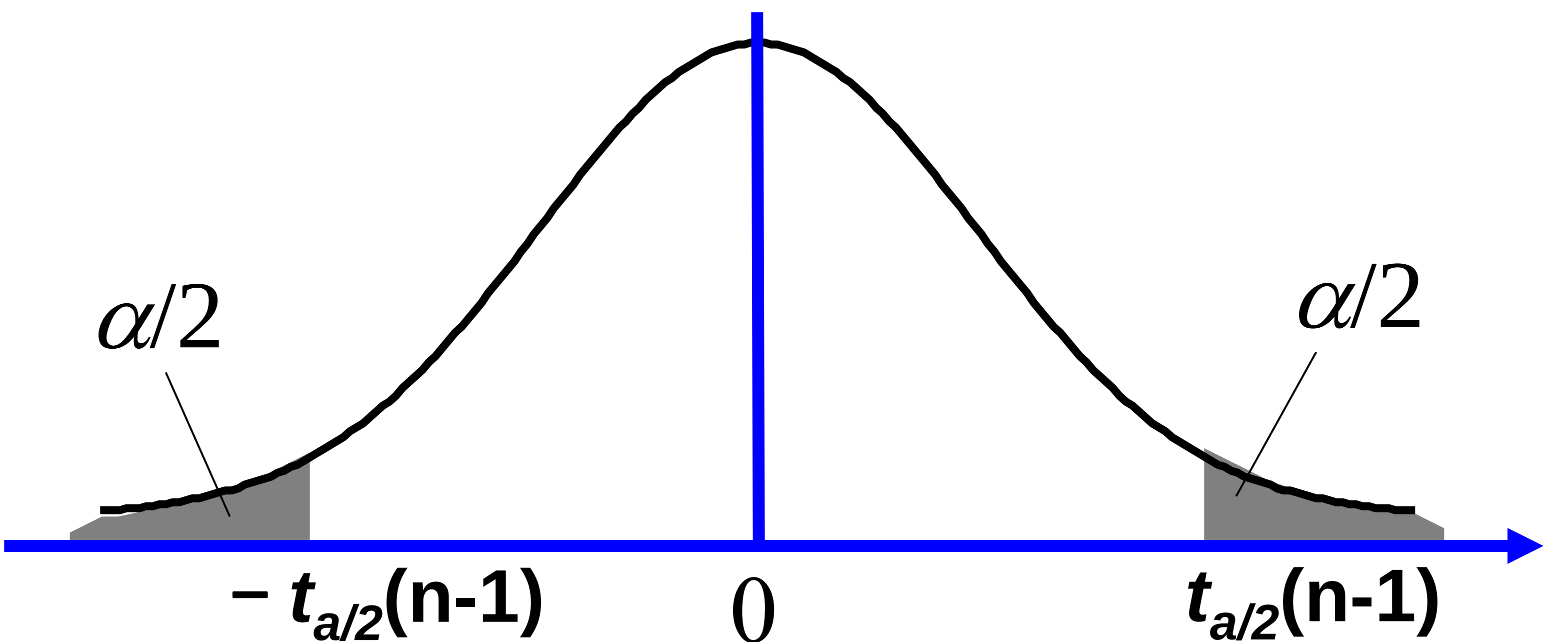
$$P\{\text{当}H_0\text{为真拒绝}H_0\} = P_{\mu_0}\left\{\left|\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\right| \geq k\right\} = \alpha$$



检验法则：

$$\frac{|\bar{x}-\mu_0|}{s/\sqrt{n}} \geq k \quad \text{拒绝} H_0$$

$$P\{\text{当}H_0\text{为真拒绝}H_0\} = P_{\mu_0}\left\{\left|\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\right| \geq k\right\} = \alpha$$



检验法则：

$$|t| = \frac{|\bar{x} - \mu_0|}{s/\sqrt{n}} \geq t_{\alpha/2}(n-1) \qquad \text{拒绝} H_0$$

拒绝域为$|t|=\left|\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}\right|\geq t_{\alpha/2}(n-1)$.

上述利用 t 统计量得出的检验法称为 t 检验法.

在实际中，正态总体的方差常为未知，所以我们常用 t 检验法来检验关于正态总体均值的检验问题.

例 某种电子元件的寿命X(以小时计)服从正态分布, μ,σ^2 均为未知. 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等于225(小时)?

解 依题意需检验假设

$$H_0:\mu=\mu_0=225,\ \ H_1:\mu\neq 225,$$

取 $\alpha=0.05$, $n=16$, $\bar{x}=241.5$, $s=98.7259$,

二、两个正态总体均值差的检验（方差已知）

给定显著性水平为 α, 并设 $X_1, X_2, \cdots, X_{n_1}$ 为第一个总体 $N(\mu_1, \sigma_1^2)$ 的样本, $Y_1, Y_2, \cdots, Y_{n_2}$ 为第二个总体 $N(\mu_2, \sigma_2^2)$ 的样本, $\bar{X}, \bar{Y}$ 分别是第一、二个总体的样本均值, S_1^2, S_2^2 分别是第一、二个总体的样本方差.

检验假设 $\boldsymbol{H_0: \mu_1 - \mu_2 = \delta, H_1: \mu_1 - \mu_2 \neq \delta}$。

二、两个正态总体均值差 的检验（方差未知但相等）

设 $X_1, X_2, \cdots, X_n$ 为来自正态总体 $N(\mu_1, \sigma^2)$ 的样本，$Y_1, Y_2, \cdots, Y_n$ 为来自正态总体 $N(\mu_2, \sigma^2)$ 的样本，且设两样本独立. 注意两总体的方差相等.

又设 $\overline{X}, \overline{Y}$ 分别是总体的样本均值，S_1^2, S_2^2 是样本方差，μ_1, μ_2, σ^2 均为未知，

检验假设 $H_0: \mu_1 - \mu_2 = \delta, H_1: \mu_1 - \mu_2 \neq \delta$。

确定拒绝域

检验法则：

$$\left|\frac{(\bar{x}-\bar{y})-\delta}{s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right| \geq k \qquad \text{拒绝} H_0$$

$$P\{H_0\text{为真,拒绝}H_0\}$$

$$= P_{\mu_1-\mu_2=\delta}\left\{\left|\frac{(\bar{X}-\bar{Y})-\delta}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right| \geq k\right\} = \alpha$$

检验法则：

$$\left|\frac{(\bar{x}-\bar{y})-\delta}{s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right| \geq t_{\alpha/2}(n_1+n_2-2) \quad \text{拒绝}H_0$$

常用$\delta=0$的情况.

例 3 有两台光谱仪 I_x, I_y，用来测量材料中某种金属的含量，为鉴定它们的测量结果有无显著的差异，制备了 9 件试块（它们的成分、金属含量、均匀性等各不相同），现在分别用这两台仪器对每一试块测量一次，得到 9 对观察值如下.

$x(\%)$	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
$y(\%)$	0.10	0.21	0.52	0.32	0.78	0.59	0.68	0.77	0.89

问能否认为这两台仪器的测量结果有显著的差异（取 $\alpha=0.01$）？

（三）基于成对数据的检验（t 检验）

有时为了比较两种产品、两种仪器、两种方法等的差异，我们常在相同的条件下做对比试验，得到一批成对的观察值. 然后分析观察数据作出推断. 这种方法常称为**逐对比较法**.

例 3 有两台光谱仪 I_x，I_y，用来测量材料中某种金属的含量，为鉴定它们的测量结果有无显著的差异，制备了 9 件试块（它们的成分、金属含量、均匀性等各不相同），现在分别用这两台仪器对每一试块测量一次，得到 9 对观察值如下.

$x(\%)$	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
$y(\%)$	0.10	0.21	0.52	0.32	0.78	0.59	0.68	0.77	0.89

问能否认为这两台仪器的测量结果有显著的差异（取 $\alpha=0.01$）？

（三）基于成对数据的检验（t 检验）

有时为了比较两种产品、两种仪器、两种方法等的差异，我们常在相同的条件下做对比试验，得到一批成对的观察值. 然后分析观察数据作出推断. 这种方法常称为**逐对比较法**.

例 3 有两台光谱仪 I_x, I_y，用来测量材料中某种金属的含量，为鉴定它们的测量结果有无显著的差异，制备了 9 件试块（它们的成分、金属含量、均匀性等各不相同），现在分别用这两台仪器对每一试块测量一次，得到 9 对观察值如下.

$x(\%)$	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
$y(\%)$	0.10	0.21	0.52	0.32	0.78	0.59	0.68	0.77	0.89
$d=x-y(\%)$	0.10	0.09	−0.12	0.18	−0.18	0.11	0.12	0.13	0.11

问能否认为这两台仪器的测量结果有显著的差异（取 $\alpha=0.01$）？

第三节 正态总体方差的假设检验

一、单个总体的情况

二、两个总体的情况

一、单个总体 $N(\mu,\sigma^2)$ 的情况

设总体 $X \sim N(\mu,\sigma^2)$, μ,σ^2均为未知,

$X_1, X_2,\cdots, X_n$ 为来自总体 X 的样本,

(1) 要求检验假设: $H_0:\sigma^2 = \sigma_0{}^2$, $H_1:\sigma^2 \neq \sigma_0{}^2$,

一、单个总体 $N(\mu,\sigma^2)$ 的情况

设总体 $X \sim N(\mu,\sigma^2)$, μ,σ^2 均为未知,

$X_1, X_2, \cdots, X_n$ 为来自总体 X 的样本,

(1) 要求检验假设: $H_0:\sigma^2=\sigma_0^2$, $H_1:\sigma^2\neq\sigma_0^2$,

其中 σ_0 为已知常数. 设显著性水平为 α,

当 H_0 为真时,

比值 $\dfrac{s^2}{\sigma_0^2}$ 在1附近摆动, 不应过分大于1或过分小于1,

检验法则：

$$\frac{(n-1)s^2}{\sigma_0^2} \le k_1 \text{或} \frac{(n-1)s^2}{\sigma_0^2} \ge k_2 \qquad \text{拒绝} H_0$$

$P\{H_0\text{为真},\text{拒绝}H_0\}$

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

为了计算方便, 习惯上取

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

故得 $k_1 = \chi^2_{1-\alpha/2}(n-1), \quad k_2 = \chi^2_{\alpha/2}(n-1).$

检验法则:

$$\frac{(n-1)s^2}{\sigma_0^2} \le \chi^2_{1-\alpha/2}(n-1) \text{或} \frac{(n-1)s^2}{\sigma_0^2} \ge \chi^2_{\alpha/2}(n-1) \text{拒绝} H_0$$

上述检验法称为χ^2检验法.

例1 某厂生产的某种型号的电池，其寿命长期以来服从方差σ^2=5000 (小时^2) 的正态分布，现有一批这种电池，从它生产情况来看，寿命的波动性有所变化. 现随机的取26只电池，测出其寿命的样本方差 s^2=9200(小时^2). 问根据这一数据能否推断这批电池的寿命的波动性较以往的有显著的变化? $(\alpha = 0.02)$

解 要检验假设 $H_0 : \sigma^2 = 5000, \quad H_1 : \sigma^2 \neq 5000,$

$$n = 26, \quad \alpha = 0.02, \quad \sigma_0{}^2 = 5000,$$

$$\chi^2_{\alpha/2}(n-1) = \chi^2_{0.01}(25) = 44.314,$$

$$\chi^2_{1-\alpha/2}(n-1) = \chi^2_{0.99}(25) = 11.524,$$

拒绝域为: $\frac{(n-1)s^2}{\sigma_0^{\ 2}} \le 11.524$, 或 $\frac{(n-1)s^2}{\sigma_0^{\ 2}} \ge 44.314$.

因为 $\frac{(n-1)s^2}{\sigma_0^{\ 2}} = \frac{25 \times 9200}{5000} = 46 > 44.314$,

所以拒绝 H_0，认为这批电池的寿命波动性较以往有显著的变化.

二、两个总体$N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$的情况

设 $X_1,X_2,\cdots,X_n$ 为来自正态总体$N(\mu_1,\sigma_1{}^2)$的样本,

$Y_1,Y_2,\cdots,Y_n$为来自正态总体$N(\mu_2,\sigma_2{}^2)$的样本,

且设两样本独立, 其样本方差为 $S_1{}^2, S_2{}^2$.

又设 $\mu_1, \mu_2, \sigma_1{}^2, \sigma_2{}^2$均为未知,

需要检验假设: $H_0: \sigma_1{}^2 = \sigma_2{}^2$, $H_1: \sigma_1{}^2 \neq \sigma_2{}^2$,

当 H_0 为真时，$E(S_1^2)=\sigma_1^2=\sigma_2^2=E(S_2^2)$，

观察值 $\dfrac{s_1^2}{s_2^2}$ 不应该过分大于1或过分小于1

检验法则：

$$\frac{s_1^2/s_2^2}{\sigma_1^2/\sigma_2^2}\le k_1 \text{或} \frac{s_1^2/s_2^2}{\sigma_1^2/\sigma_2^2}\ge k_2 \text{拒绝} H_0$$

$P\{H_0$ 为真, 拒绝 $H_0\}$

$$= P_{\sigma_1^2=\sigma_2^2}\left\{\left(\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \le k_1\right) \cup \left(\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \ge k_2\right)\right\} = \alpha.$$

为了计算方便, 习惯上取

$$P_{\sigma_1^2=\sigma_2^2}\left\{\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \le k_1\right\} = \frac{\alpha}{2} \quad P_{\sigma_1^2=\sigma_2^2}\left\{\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \ge k_2\right\} = \frac{\alpha}{2}$$

故得 $k_1 = F_{1-2/\alpha}(n_1 - 1, n_2 - 1)$

$$k_2 = F_{2/\alpha}(n_1 - 1, n_2 - 1)$$

第八章　　假设检验

- 1 假设检验
- 2 正态总体均值的假设检验
- 3 正态总体方差的假设检验
- 4 置信区间与假设检验之间的关系

作业

2题　10题　13题　17题

P187 某种电子元件的寿命X(以小时计)服从正态分布, μ,σ^2 均为未知. 现测得16只元件的寿命如下:

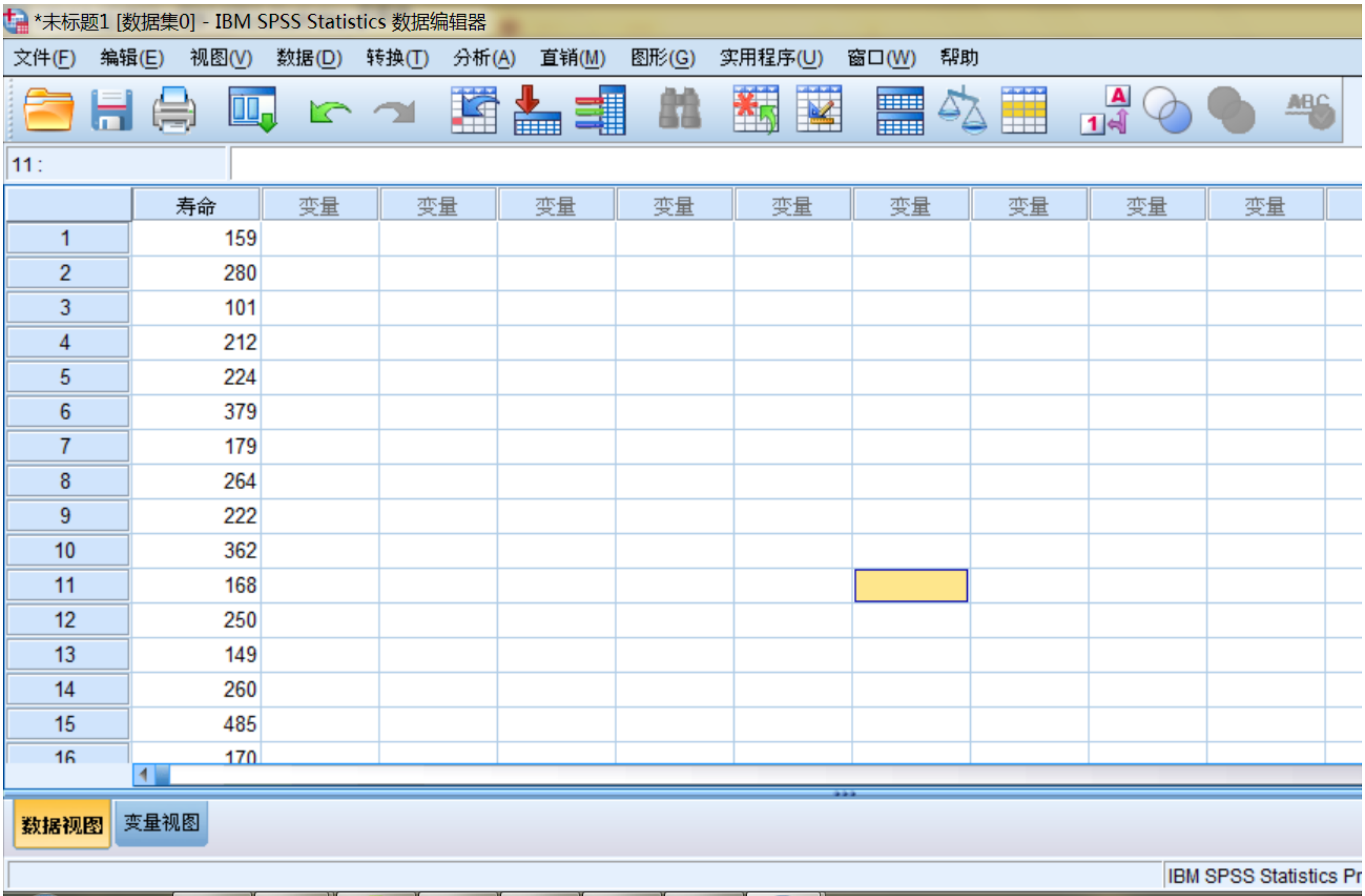
159	280	101	212	224	379	179	264
222	362	168	250	149	260	485	170

问是否有理由认为元件的平均寿命等于225(小时)?

解 依题意需检验假设

$$H_0:\mu=\mu_0=225,\ \ H_1:\mu\neq 225,$$

取 $\alpha=0.05,\ \ n=16,\ \ \bar{x}=241.5,\ \ s=98.7259,$


*未标题1 [数据集0] - IBM SPSS Statistics 数据编辑器
文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M) 图形(G) 实用程序(U) 窗口(W) 帮助
11 :
寿命 变量 变量 变量 变量 变量 变量 变量 变量 变量
1 159
2 280
3 101
4 212
5 224
6 379
7 179
8 264
9 222
10 362
11 168
12 250
13 149
14 260
15 485
16 170
数据视图
变量视图
IBM SPSS Statistics Pr


单样本 T 检验
检验变量(T):
寿命
选项(O)...
Bootstrap(B)...
检验值(V):
225
确定
粘贴(P)
重置(R)
取消
帮助

单个样本统计量

	N	均值	标准差	均值的标准误
寿命	16	241.50	98.726	24.681

单个样本检验

	检验值 = 225					
					差分的 95% 置信区间	
	t	df	Sig.(双侧)	均值差值	下限	上限
寿命	.669	15	.514	16.500	-36.11	69.11

第八章　　假设检验

- 1 假设检验
- 2 正态总体均值的假设检验
- 3 正态总体方差的假设检验
- 4 置信区间与假设检验之间的关系
- 7 假设检验问题的*p*值法

第7节 假设检验问题的p值法

◆ 以上讨论的假设检验方法称为临界值法。

例1　设总体X服从$N(\mu,\sigma^2)$,μ未知，$\sigma^2=100$，现有样本$x_1,x_2,\ldots,x_{52}$,算得样本均值为62. 75. 现在来检验假设

$$H_0:\mu\le\mu_0=60,H_1:\mu>60$$

采用Z检验法，检验统计量为 $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

检验统计量的观察值为

$$z_0 = \frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{62.75-60}{10/\sqrt{52}} = 1.983$$

看z_0是否落入拒绝域 $\{z \geq z_\alpha\}$

其实就是要比较z_0和z_α

采用Z检验法，检验统计量为 $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

检验统计量的观察值为

$$z_0=\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}=\frac{62.75-60}{10/\sqrt{52}}=1.983$$

看z_0是否落入拒绝域 $\{z \geq z_\alpha\}$

其实就是要比较z_0和z_α

$$P\{Z \geq z_0\}=P\{Z \geq 1.983\}=1-\Phi(1.983)=0.0238$$

这个概率称为Z检验法的右边检验的p值。

$$p\text{值} = P\left\{Z \geq z_0\right\} = \mathbf{0.0238}$$

是原假设H_0可被拒绝的最小显著性水平。

$$p\text{值} = P\{Z \geq z_0\} = 0.0238$$

是原假设H_0可被拒绝的最小显著性水平。

对于任意给定的显著性水平α,

(1)若 p值 $\leq \alpha$,则在显著性水平α下拒绝H_0

(2)若 p值 $> \alpha$,则在显著性水平α下接受H_0

正态总体$N(\mu,\sigma^2)$均值的检验中

假设检验$H_0:\mu\le\mu_0, H_1:\mu>\mu_0$;

假设检验$H_0:\mu\ge\mu_0, H_1:\mu<\mu_0$.

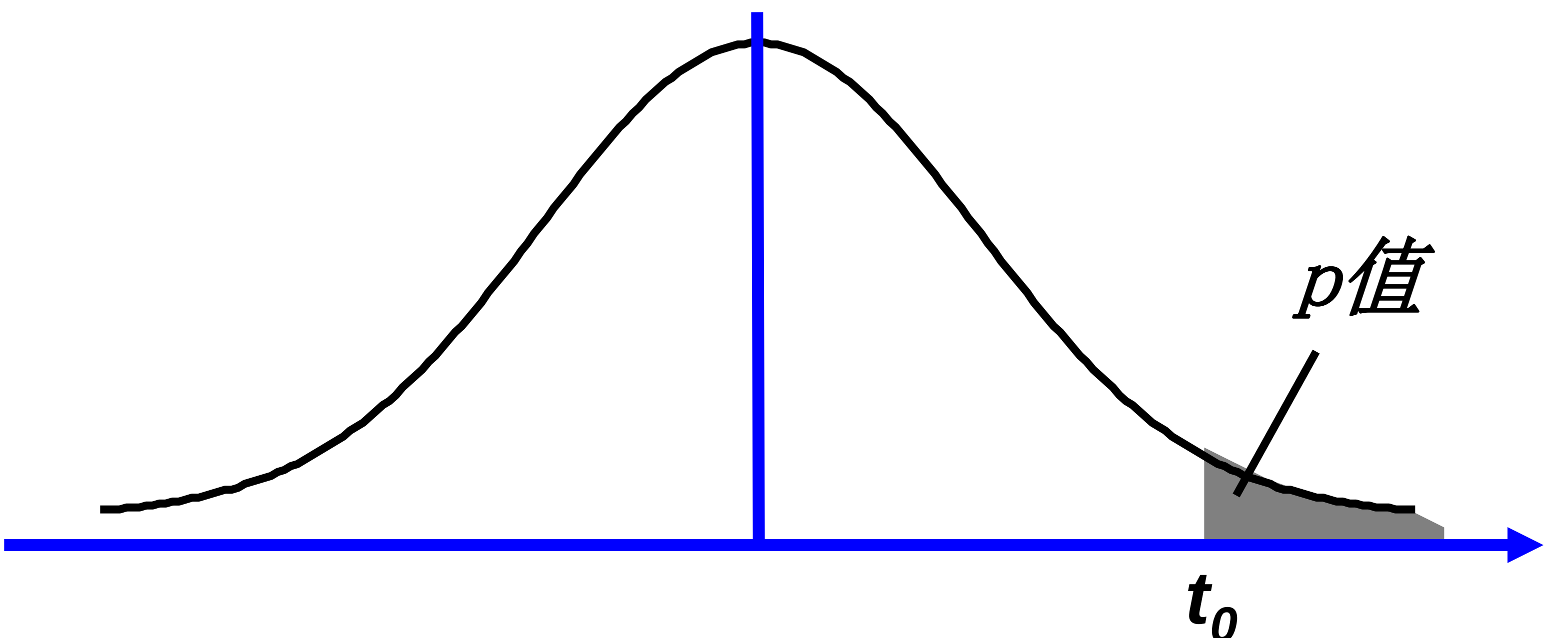
假设检验$H_0:\mu=\mu_0, H_1:\mu\ne\mu_0$;

σ^2未知时，可采用检验统计量

$$t=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$$

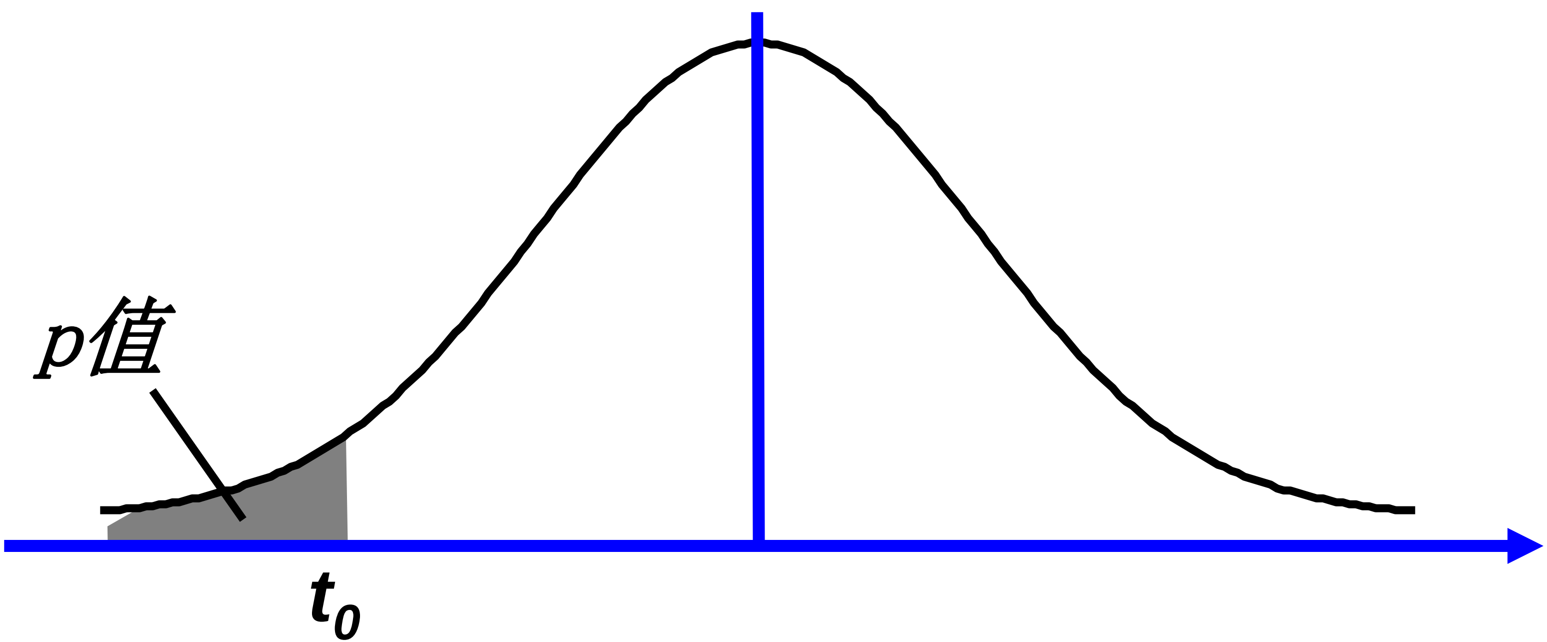
假设检验 $H_0: \mu \le \mu_0,\ H_1: \mu > \mu_0$;

$$p\text{值} = P_{\mu_0}\{t \ge t_0\} = t_0\text{右侧尾部面积}$$



假设检验 $H_0: \mu \geq \mu_0, H_1: \mu < \mu_0$;

p值$= P_{\mu_0}\{t \leq t_0\} = t_0$左侧尾部面积

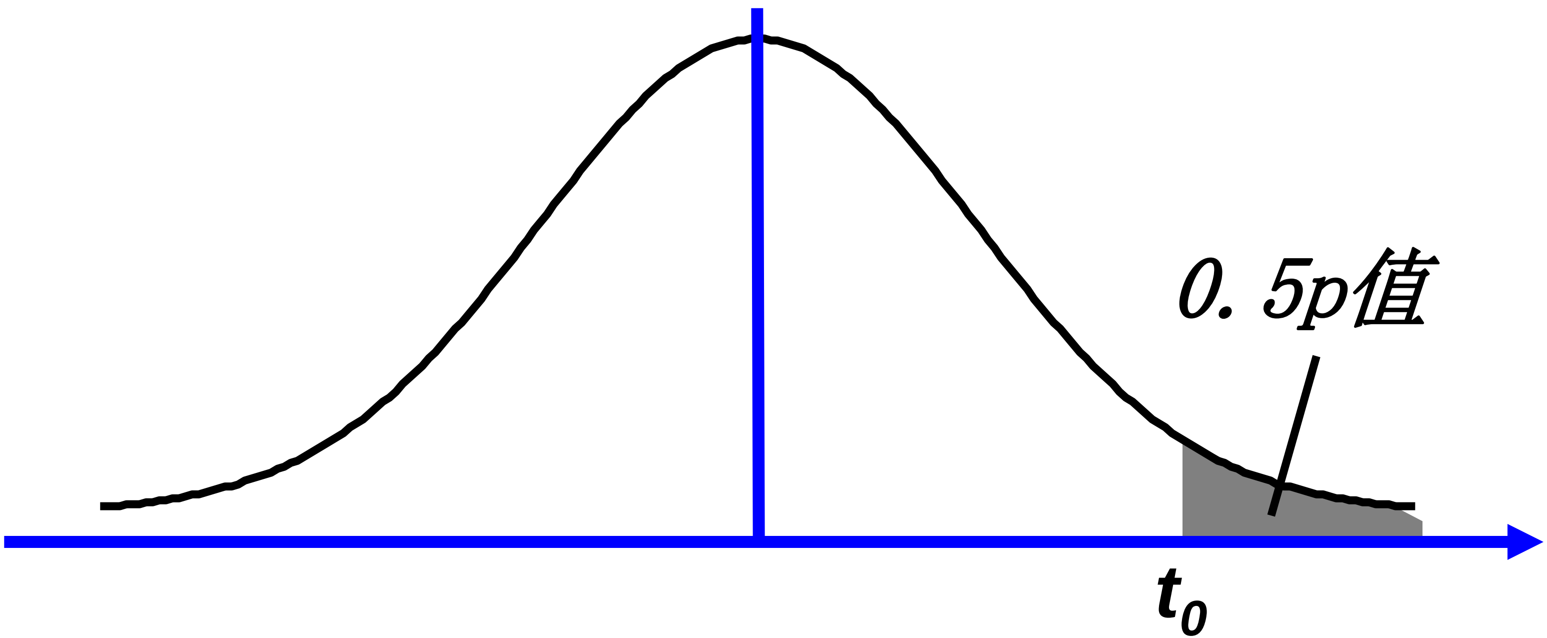


假设检验 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$;

当 $t_0 > 0$ 时

$$p\text{值} = 2\times(t_0\text{右侧尾部面积})$$

$$= P_{\mu_0}\{|t| \geq t_0\} = P_{\mu_0}\{(t \leq -t_0) \cup (t \geq t_0)\}$$

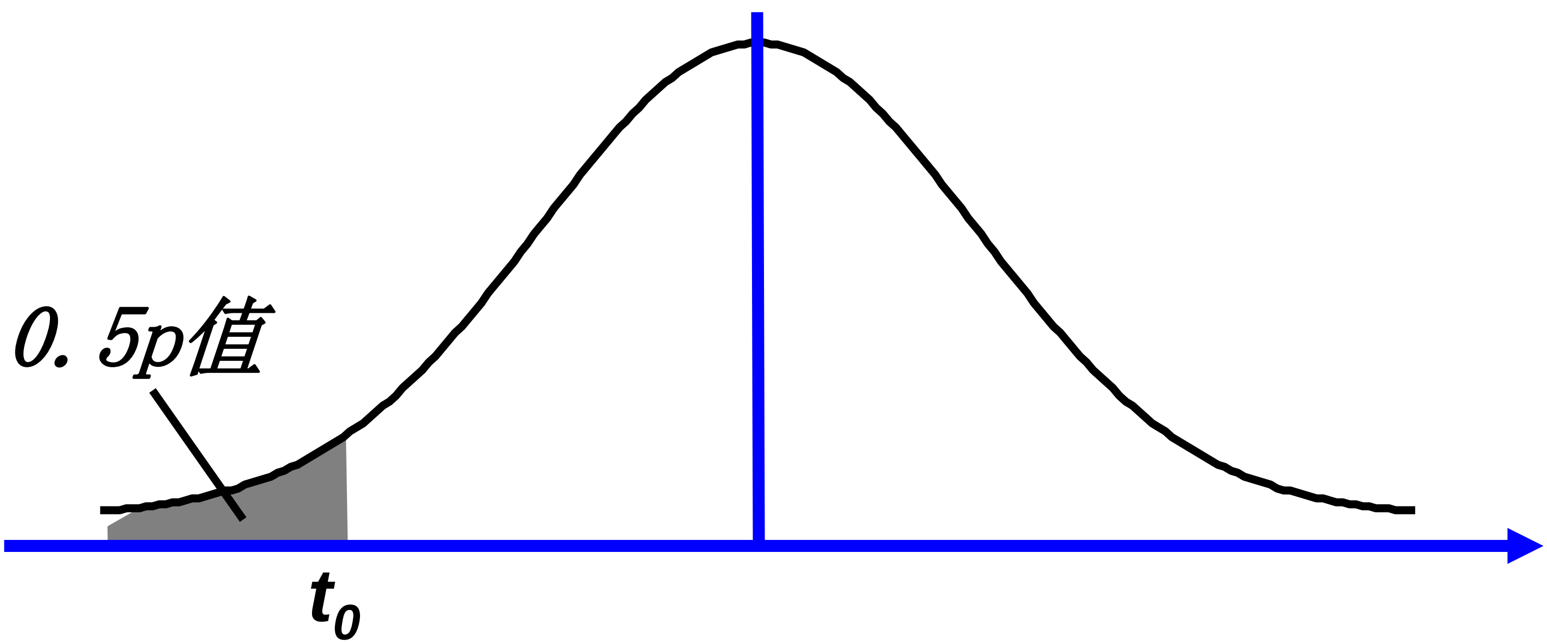


假设检验 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$;

当 $t_0 < 0$ 时

$$p\text{值} = 2 \times (t_0\text{左侧尾部面积})$$

$$= P_{\mu_0}\{|t| \geq -t_0\} = P_{\mu_0}\{(t \leq t_0) \cup (t \geq -t_0)\}$$



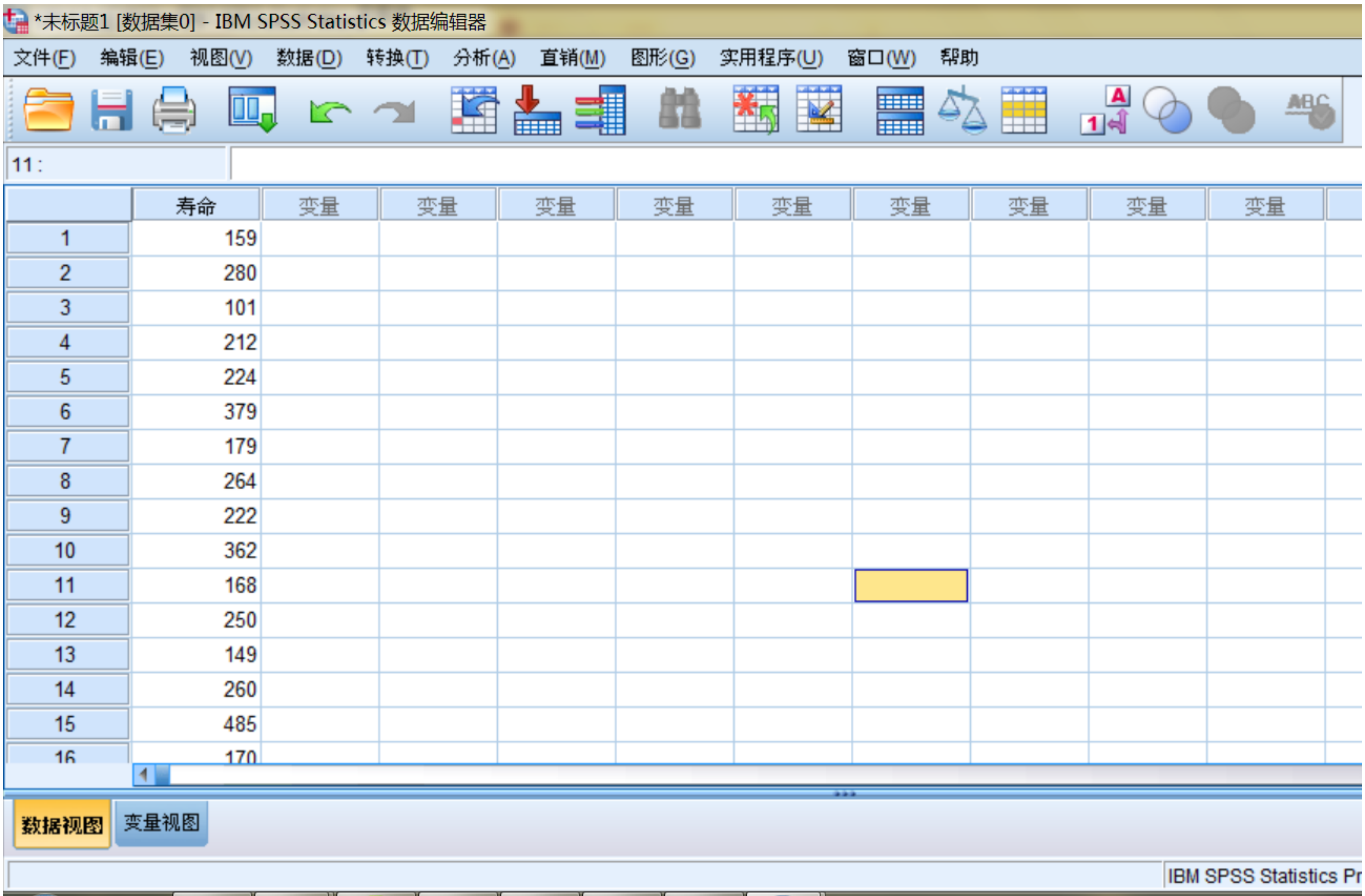
若p值≤ 0.01　称拒绝H_0的依据很强，或称检验是高度显著的。

若$0.01 < p$值≤ 0.05　称拒绝H_0的依据是强的，或称检验是显著的。

若$0.05 < p$值≤ 0.1　称拒绝H_0的依据是弱的，或称检验是不显著的。

若p值> 0.1　一般没有理由拒绝H_0

P187 例1



*未标题1 [数据集0] - IBM SPSS Statistics 数据编辑器
文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M) 图形(G) 实用程序(U) 窗口(W) 帮助
11 :
寿命
变量
1 159
2 280
3 101
4 212
5 224
6 379
7 179
8 264
9 222
10 362
11 168
12 250
13 149
14 260
15 485
16 170
数据视图
变量视图
IBM SPSS Statistics Pr


单样本 T 检验
检验变量(T):
寿命
选项(O)...
Bootstrap(B)...
检验值(V):
225
确定
粘贴(P)
重置(R)
取消
帮助

单个样本统计量

	N	均值	标准差	均值的标准误
寿命	16	241.50	98.726	24.681

单个样本检验

	检验值 = 225					
					差分的 95% 置信区间	
	t	df	Sig.(双侧)	均值差值	下限	上限
寿命	.669	15	.514	16.500	-36.11	69.11

第八章　　假设检验

- 1 假设检验
- 2 正态总体均值的假设检验
- 3 正态总体方差的假设检验
- 4 置信区间与假设检验之间的关系
- 7 假设检验问题的*p*值法