第八章 假设检验

- § 8.1 假设检验
- § 8.2 正态总体均值的假设检验
- § 8.3 正态总体方差的假设检验
- § 8.6 分布拟合检验

第八章 假设检验

- §8.1 假设检验
- §8.2 正态总体均值的假设检验
- §8.3 正态总体方差的假设检验
- §8.6 分布拟合检验

§ 8.1 假设检验

- 参数估计：其目的对未知参量给出估计值及置信区间，一般情况下，参数估计是在总体形式已知的情况下，对未知参量的定量的估计问题
- 假设检验：其目的是对总体的某未知性质根据样本给出一个定性判断，这时总体的分布的函数形式未知，或只知其形式，但参数未知的情况
- 假设检验中，为推断总体的某些性质，首先提出某些关于总体的假设，然后根据样本对所提出的假设作出判断，是接受，还是拒绝

例如：提出总体期望服从泊松分布的假设，然后进行判断

提出正态总体期望为μ_0的假设，然后进行判断

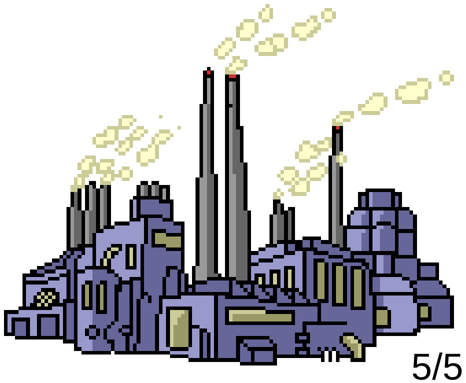
§ 8.1 假设检验

- 假设检验的基本思想和做法
  - 通常借助于直观分析和理论分析相结合的做法
  - 基本原理就是人们在实际问题中经常采用的所谓**小概率原理**:“**一个小概率事件在一次试验中几乎是不可能发生的**”
  - 假设检验的过程是要构造一个小概率事件，如果根据实际样本数据的计算，该小概率事件发生了，则拒绝原假设，否则接受原假设
- 下面结合实例来说明假设检验的基本思想.

§ 8.1 假设检验

- **实例**　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(公斤):
  - 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?

分析: 用 μ 和 σ 分别表示这一天袋装糖重总体 X 的均值和标准差



§ 8.1 假设检验

由长期实践可知, 标准差较稳定, 设 $\sigma = 0.015$,

则 $X \sim N(\mu,\ 0.015^2)$, 其中μ 未知.

问题: 根据样本值判断 $\mu = 0.5$ 还是 $\mu \neq 0.5$.

1° 提出两个对立假设 $H_0: \mu = \mu_0 = 0.5$ 和 $H_1: \mu \neq \mu_0$.

2° 结合合理法则，再利用已知样本作出判断是接受假设H_0(拒绝假设H_1), 还是拒绝假设H_0(接受假设H_1).

如果作出的判断是接受H_0, 则 $\mu = \mu_0$,

即认为机器工作是正常的, 否则, 认为是不正常的.

§ 8.1 假设检验

检验方法(即合理的法则)：对于未知参数，仍然从其点估计量开始讨论，将未知参数与其点估计量进行比较

由于要检验的假设涉及总体均值, 故可借助于样本均值来判断.

因为 $\overline{X}$ 是 μ 的无偏估计量

所以若 H_0 为真, 则 $|\overline{x}-\mu_0|$ 不应太大,

若过分大，则有理由怀疑H_0的正确性

当 H_0 为真时, $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$,

这里的检验统计量和分布均不含任何未知参数

衡量 $|\overline{x}-\mu_0|$ 的大小可归结为衡量 $\dfrac{|\overline{x}-\mu_0|}{\sigma/\sqrt{n}}$ 的大小,

于是可以选定一个适当的正数 k,

§ 8.1 假设检验

此即假定H_0正确时的小概率事件

当观察值$\bar{x}$满足$\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq k$时，拒绝假设$H_0$，

反之，当观察值$\bar{x}$满足$\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < k$时，接受假设$H_0$.

如何选取k呢，先看以下事实：

由于作出决策的依据是一个样本，当实际上H_0为真时，仍可能作出拒绝H_0的决策，这种可能性是无法消除的，这是一种错误。

§ 8.1 假设检验

- 因此自然希望将犯这类错误的概率控制在一定限度之内，即给出一个较小的数$\alpha(0<\alpha<1)$,使犯这类错误的概率不超过α，即使得：

$$P\{拒绝H_0|H_0为真\}\le\alpha$$

因为当H_0为真时前述的统计量$Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$,

及其分布，不含任何未知参数，

由最大允许错误概率α,令上式取等号得

$$P\{拒绝H_0\mid H_0为真\}=P\{\frac{\left|\overline{x}-\mu_0\right|}{\sigma/\sqrt{n}}\ge k\}=\alpha,$$

§ 8.1 假设检验

由标准正态分布分位点的定义得 $k = z_{\alpha/2}$，

当$\dfrac{|\overline{x}-\mu_0|}{\sigma/\sqrt{n}} \ge z_{\alpha/2}$时，拒绝$H_0$，$\dfrac{|\overline{x}-\mu_0|}{\sigma/\sqrt{n}} < z_{\alpha/2}$时，接受$H_0$.

前述的错误$P\{拒绝H_0 \mid 当H_0为真\}$，称为“弃真”类错误，也可记作：

$P_{\mu_0}\{拒绝H_0\}$：表示参数μ取μ_0时事件$\{拒绝H_0\}$的概率

或者

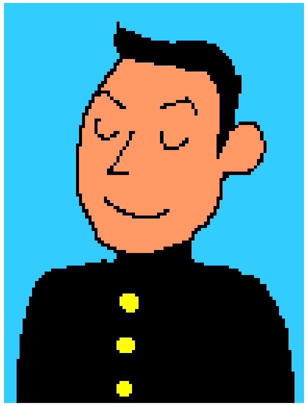
$P_{\mu\in H_0}\{拒绝H_0\}$：表示参数μ取H_0规定值时事件$\{拒绝H_0\}$的概率

§ 8.1 假设检验

在实例中若取定$\alpha = 0.05$,

则$k = z_{\alpha/2} = z_{0.025} = 1.96$,

又已知$n = 9$, $\sigma = 0.015$,



由样本算得 $\bar{x} = 0.511$，即有 $\dfrac{|\bar{x} - \mu_0|}{\sigma/\sqrt{n}} = 2.2 > 1.96$,

于是拒绝假设H_0，认为包装机工作不正常.

§ 8.1 假设检验

检验的合理性

以上所采取的检验法是符合实际推断原理的.

由于通常α总是取得很小,一般取$\alpha=0.01, \alpha=0.05,$

因而当H_0为真,即$\mu=\mu_0$时, $\left\{\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \geq z_{\alpha/2}\right\}$是一个

小概率事件,在一次实验当中几乎是不发生的,现竟然发生了,所以假设不正确,因而拒绝H_0

在假设检验中,数α称为显著性水平.

它与用户对H_0的正确的信心有关,信心越大,

α可以取得越大

§ 8.1 假设检验

- 总结以上实例：

- 在上例中，当样本容量n固定时，选定α后，数k可以确定，然后按照统计量Z= $\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ 的观察值的绝对值|z|是大于等于k，还是小于k来作出决策，

  k是检验上述假设的一个门槛值

  若$|z|=\left|\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geq k$，则称$\overline{x}$与$\mu_0$的差异是**显著的**，以至于小概率事件发生了，这时拒绝H_0，

  否则则称 $\overline{x}$与μ_0的差异是不显著的，这时接受H_0，

  选定的数α称为显著性水平，在α下对显著性判断统计量Z=$\left|\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|$ 称为检验统计量

§ 8.1 假设检验

假设检验的相关定义:

像上例中的假设检验问题可叙述成:

"在显著性水平α下，检验假设H_0：$\mu=\mu_0$，H_1：$\mu\neq\mu_0$"

或"在显著性水平α下，针对H_1检验H_0"

H_0称为原假设，或零假设，

H_1称为备择假设，(在原假设被拒绝后可供选择的假设)

- **要进行的工作是根据样本，按上述检验方法做出决定在H_0和H_1之间接受其一**

§ 8.1 假设检验

- 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，则区域C称为拒绝域，拒绝域的边界点称为临界点

  上例中$|z|\geq z_{\alpha/2}$为拒绝域，$z_{\alpha/2}$为临界点

- 检验法则是根据样本作出的，总有可能作出错误的决策，
- 在H_0实际上为真时，可能犯拒绝H_0的错误，称这类“弃真”的错误为第I类错误
- 当H_0实际上不为真时，可能犯接受H_0的错误，称这类“取伪”的错误为第II类错误，犯第II类错误的概率记为

  $P\{$当H_0不真时接受$H_0\}$或 $P_{\mu\in H_1}\{$接受$H_0\}$

§ 8.1 假设检验

假设检验的两类错误

真实情况 (未知)	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不真	犯第II类错误	正确

§ 8.1 假设检验

- 在确定检测法则时应尽可能使犯两类错误的概率都小，但样本容量固定时减少犯弃真错误，往往会增加取伪，二者矛盾，除非增加样本容量。
- 一般情况下，我们总是控制犯第一类错误的概率，使他小于或等于α，α通常取0.1，0.05，0.01，0.005，等值，视情况而定
- 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，称为显著性检验
  - 即原假设为真却拒绝原假设，说明总体性质发生了显著的改变

§ 8.1 假设检验

- 在显著性水平α下，如果给出的两个对立的假设如下形式：

  H_0：$\mu=\mu_0$，H_1：$\mu\neq\mu_0$，

  则称为双边备择假设，称该类假设检验问题为双边假设检验

  其原因是：对于备择假设，μ可以小于μ_0，也可以大于μ_0，

- 有时我们只关心总体均值是否增大，例如试验新工艺以提高材料强度，总体均值越大越好。需要检验假设

  H_0：$\mu\leq\mu_0$，H_1：$\mu>\mu_0$，则称为右边检验问题

- 类似的有时需要检验假设，

  H_0：$\mu\geq\mu_0$，H_1：$\mu<\mu_0$，则称为左边检验问题

- 左边检验和右边检验统称为单边检验，检验的分类是依据备择假设的形式给出的

§ 8.1 假设检验

- **单边检验的拒绝域**
  - 这时原假设为真时被检参数是一个范围，而不是一个值
- 设总体X～$N(\mu, \sigma^2)$，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，来求如下检验问题的拒绝域

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

- 因H_0中的μ都比H_1中的要小，当H_1为真时观察值往往偏大，因此拒绝域的形式为

$\overline{x} \ge k$，k是某一正常数

§ 8.1 假设检验

确定k，与例1中的做法类似

$$P\{H_0\text{为真时拒绝}H_0\}=P_{\mu\in H_0}\{\text{拒绝}H_0\}$$

$$=P_{\mu\in H_0}\{\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\geq\frac{k-\mu_0}{\sigma/\sqrt{n}}\}$$

不等号成立是因为$\mu\leq\mu_0$

$$\leq P_{\mu\in H_0}\{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\geq\frac{k-\mu_0}{\sigma/\sqrt{n}}\}$$

注意：这里 $\overline{X}$的均值为μ而不是μ_0，所以放缩成μ后才能用正态分布。

要控制$P\{H_0\text{为真时拒绝}H_0\}\leq\alpha$

只需令$P_{\mu\in H_0}\{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\geq\frac{k-\mu_0}{\sigma/\sqrt{n}}\}=\alpha$

§ 8.1 假设检验

由于$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$,

得到$\dfrac{k-\mu_0}{\sigma/\sqrt{n}} = z_\alpha$, 即$k = \mu_0 + z_\alpha \dfrac{\sigma}{\sqrt{n}}$

即$\bar{x} \geq k = \mu_0 + z_\alpha \dfrac{\sigma}{\sqrt{n}}$, 即$z = \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \geq z_\alpha$

类似的有左边检验问题的拒绝域

$$z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha$$

§ 8.1 假设检验

➲ 处理参数的假设检验问题的步骤如下：

1. 根据实际问题的要求提出原假设H_0和备择假设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

- 其分布应与任何未知数无关，且统计量里不含其它未知参数
- 统计量的构造一般的从点估计量开始考虑

4. 按$P\{H_0为真时拒绝H_0\}\leq\alpha$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

§ 8.1 假设检验

例2：某工厂生产固体燃料推进器的燃烧率服从正态分布$N(\mu,\sigma^2)$，$\mu=40$cm/s，$\sigma=2$cm/s,现在用新方法生产了一批推进器，从中随机取n＝25只，测得燃烧率的样本均值为 $\bar{x}=$ 41.25cm/s，设新方法下总体均方差没变，问这批推进器的燃烧率较以往是否有显著的提高，取显著性水平$\alpha=0.05$

解：1° 提出原假设H_0和备择假设H_1；

H_0：$\mu\le\mu_0=40$，即假设新方法没有提高燃烧率

H_1：$\mu>\mu_0$，即假设新方法提高了燃烧率

2° 给定显著性水平$\alpha=0.05$

以及样本容量n＝25

§ 8.1 假设检验

3° 确定检验统计量以及拒绝域的形式

由例1，统计量为$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$～$N(0，1)$，拒绝域的形式为

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq k$$

4° 按$P\{H_0$为真时拒绝$H_0\}\leq\alpha$求出拒绝域

$$z=\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} \geq z_{0.05}=1.645$$

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

$$z=\frac{41.25-40}{2/\sqrt{25}}=3.125>1.645$$

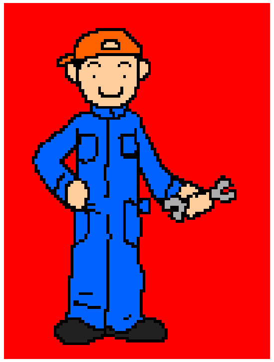
z落在拒绝域中，在显著性水平α下拒绝H_0，因此新方法有显著提高

第八章 假设检验

- §8.1 假设检验
- §8.2 正态总体均值的假设检验
- §8.3 正态总体方差的假设检验
- §8.6 分布拟合检验

§ 8.2 正态总体均值的假设检验

- 假设检验是针对弃真这一可能犯的错误人为设定一个界限，如果在这个界限内，认为原假设成立，否则的话，由于显著性水平取得很小，表明小概率事件发生，根据实际推断原理，原假设不成立。
- 尽管也可能犯第II类取伪的错误，这时尽管总体的性质发生了改变但没有发现，往往影响较小。
- 正态总体均值的检验分为三种情况
  - 单个正态总体
  - 两个正态总体
  - 成对数据



§ 8.2 正态总体均值的假设检验

(一)单个总体$N(\mu, \sigma^2)$均值μ 的检验

1° σ^2已知，关于μ 的检验(Z检验)

- 提出的假设，

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

单边：H_0：$\mu\leq\mu_0$，H_1：$\mu>\mu_0$，

H_0：$\mu\geq\mu_0$，H_1：$\mu<\mu_0$，

都是利用检验统计量

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

确定拒绝域的，当H_0为真时服从$N(0,1)$

这种检验法常称为**Z检验法**

§ 8.2 正态总体均值的假设检验

例1　某切割机在正常工作时, 切割每段金属棒的平均长度为10.5cm, 标准差是0.15cm, 今从一批产品中随机的抽取15段进行测量, 其结果如下:

10.4　10.6　10.1　10.4　10.5　10.3　10.3　10.2

10.9　10.6　10.8　10.5　10.7　10.2　10.7

假定切割的长度X服从正态分布, 且标准差没有变化, 试问该机工作是否正常? $(\alpha = 0.1)$

解　因为 $X \sim N(\mu, \sigma^2)$, $\sigma = 0.15$,

要检验假设

$$H_0 : \mu = 10.5, \quad H_1 : \mu \neq 10.5,$$

§ 8.2 正态总体均值的假设检验

$n = 15,\quad \overline{x} = 10.48,\quad \alpha = 0.1,$

则 $\dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} = \dfrac{10.48 - 10.5}{0.15 / \sqrt{15}} = -0.516,$

查表得 $z_{0.05} = 1.645,$

于是 $\left| \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} \right| = 0.516 < z_{0.05} = 1.645$

故接受H_0,认为该机工作正常

§ 8.2 正态总体均值的假设检验

2. σ^2未知, 关于μ的检验(t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$H_0:\mu=\mu_0, H_1:\mu\neq\mu_0$.

设$X_1,X_2,\cdots,X_n$为来自总体X的样本,

因为σ^2未知, 不能利用$\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$来确定拒绝域

因为S^2是σ^2的无偏估计, 故用S来取代σ,

即采用$t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$来作为检验统计量

§ 8.2 正态总体均值的假设检验

根据第六章§3定理三知，

当H_0为真时，$$\frac{\overline{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1),$$

由t分布上α分位点的定义知

$$P\left\{\left|\frac{\overline{X}-\mu_0}{S_n^*/\sqrt{n}}\right| \geq t_{\alpha/2}(n-1)\right\}=\alpha$$

拒绝域为$C=\{|t|=\left|\dfrac{\overline{x}-\mu_0}{s_n^*/\sqrt{n}}\right| \geq t_{\alpha/2}(n-1)\}$

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.

§ 8.2 正态总体均值的假设检验

例2 如果在例1中只假定切割的长度服从正态分布, 问该机切割的金属棒的平均长度有无显著变化? $(\alpha = 0.05)$

解 依题意 $X \sim N(\mu, \sigma^2)$, μ, σ^2均为未知,

要检验假设 $H_0: \mu = 10.5, \quad H_1: \mu \neq 10.5,$

$n = 15, \quad \bar{x} = 10.48, \quad \alpha = 0.05, \quad s = 0.237,$

$$|t| = \left|\frac{\bar{x} - \mu_0}{s / \sqrt{n}}\right| = \left|\frac{10.48 - 10.5}{0.237 / \sqrt{15}}\right| = 0.327,$$

查表得 $t_{\alpha/2}(n-1) = t_{0.025}(14) = 2.1448 > |t| = 0.327,$

故接受H_0, 认为金属棒的平均长度无显著变化

§ 8.2 正态总体均值的假设检验

(二) 两个正态总体均值差的检验(t检验)

1.已知方差时两正态总体均值的检验

利用z检验法检验.

设 $X_1, X_2, \cdots, X_{n_1}$ 为来自正态总体$N(\mu_1, \sigma_1^2)$的样本,

$Y_1, Y_2, \cdots, Y_{n_1}$ 为来自正态总体$N(\mu_2, \sigma_2^2)$的样本,

两样本独立又设 μ_1, μ_2均为未知 σ_1^2, σ_2^2已知,

需要检验假设: $H_0: \mu_1 = \mu_2,\quad H_1: \mu_1 \neq \mu_2,$

上述假设可等价的变为

$$H_0: \mu_1 - \mu_2 = 0,\quad H_1: \mu_1 - \mu_2 \neq 0,$$

§ 8.2 正态总体均值的假设检验

由于$\overline{X} \sim N(\mu_1, \dfrac{\sigma_1^2}{n_1})$,$\overline{Y} \sim N(\mu_2, \dfrac{\sigma_2^2}{n_2})$,且$\overline{X}$,$\overline{Y}$独立,

故 $\overline{X} - \overline{Y} \sim N(\mu_1 - \mu_2, \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2})$

取检验的统计量为

$$U = (\overline{X} - \overline{Y}) / \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

当H_0成立时,统计量$U \sim N(0,1)$

取显著性水平为α.

§ 8.2 正态总体均值的假设检验

由标准正态分布分位数的定义知

$$P\{|(\overline{X}-\overline{Y})/\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}|\geq z_{\alpha/2}\}=\alpha$$

故拒绝域为

$$\{|(\overline{x}-\overline{y})/\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}|\geq z_{\alpha/2}\}$$

§ 8.2 正态总体均值的假设检验

例1　卷烟厂向化验室送去A,B两种烟草,化验尼古丁的含量是否相同,从A,B中各随机抽取重量相同的5例进行化验,测得尼古丁的含量(单位:mg)分别为

A:　24　27　26　21　24

B:　27　28　23　31　26

据经验知,两种烟草的尼古丁含量均服从正态分布,且相互独立,A种的方差为5,B种的方差为8,取$\alpha=0.05$,问两种烟草的尼古丁含量是否有显著差异?

§ 8.2 正态总体均值的假设检验

解 以X和Y分别表示A,B两种烟草的尼古丁含量

则$X \sim N(\mu_1,\sigma_1^2),Y \sim N(\mu_2,\sigma_2^2)$,且$X,Y$独立.

欲检验假设: $H_0:\mu_1=\mu_2,\qquad H_1:\mu_1\neq\mu_2$

现已知$\sigma_1^2=5,\sigma_2^2=8,n_1=n_2=5$.由所给数据求得

$$\bar{x}=24.4,\qquad \bar{y}=27$$

$$u=(\bar{x}-\bar{y})/\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}=\frac{24.4-27}{\sqrt{\frac{5}{5}+\frac{8}{5}}}=-1.612$$

对$\alpha=0.05$,查正态分布表得$u_{\alpha/2}=1.96$,由于

$|z|=1.612<1.96$,故接受原假设H_0.

§ 8.2 正态总体均值的假设检验

2.未知方差时两正态总体均值的检验

利用*t*检验法检验具有相同方差的两正态总体均值差的假设.

设 $X_1, X_2, \cdots, X_{n_1}$ 为来自正态总体$N(\mu_1, \sigma^2)$的样本, $Y_1, Y_2, \cdots, Y_{n_2}$ 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

§8.2 正态总体均值的假设检验

取显著性水平为α.

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

引入统计量

$$T = \frac{(\overline{X} - \overline{Y})}{S_w \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}},\ \text{其中}\ S_w^2 = \frac{(n_1 - 1)S_1^{*2} + (n_2 - 1)S_2^{*2}}{n_1 + n_2 - 2}.$$

当H_0为真时，根据第六章§3定理四的推论2知,

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

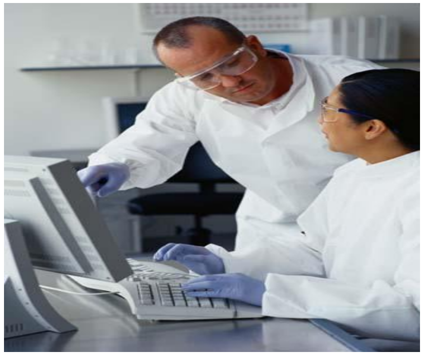
§ 8.2 正态总体均值的假设检验

对给定的α由t分布的分位表可查得$t_{\alpha/2}(n_1+n_2-2)$.

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

故拒绝域为

$$W_1=\{\frac{|(\bar{x}-\bar{y})|}{s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\geq t_{\alpha/2}(n_1+n_2-2)\}$$



§ 8.2 正态总体均值的假设检验

例2 有甲、乙两台机床加工相同的产品, 从这两台机床加工的产品中随机地抽取若干件, 测得产品直径(单位:*mm*)为

机床甲: 20.5, 19.8, 19.7, 20.4, 20.1, 20.0, 19.0, 19.9

机床乙: 19.7, 20.8, 20.5, 19.8, 19.4, 20.6, 19.2,

试比较甲、乙两台机床加工的产品直径有无显著差异? 假定两台机床加工的产品直径都服从正态分布, 且总体方差相等. $(\alpha = 0.05)$

解 依题意, 两总体 X 和 Y 分别服从正态分布

$N(\mu_1, \sigma^2)$和$N(\mu_2, \sigma^2)$, μ_1, μ_2, σ^2均为未知,

§ 8.2 正态总体均值的假设检验

需要检验假设 $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$.

$$n_1 = 8, \quad \bar{x} = 19.925, \quad s_1^{*2} = 0.216,$$

$$n_2 = 7, \quad \bar{y} = 20.000, \quad s_2^{*2} = 0.397,$$

且 $s_w^2 = \dfrac{(8-1)s_1^{*2} + (7-1)s_2^{*2}}{8+7-2} = 0.547$,

查表可知 $t_{0.05}(13) = 2.160$,

$|t| = \left| \dfrac{\bar{x} - \bar{y}}{s_w \sqrt{\dfrac{1}{8} + \dfrac{1}{7}}} \right| = 0.265 < 2.160$, 所以接受 H_0,

即甲、乙两台机床加工的产品直径无显著差异.

§ 8.2 正态总体均值的假设检验

(三)基于配对数据的检验（t检验）

有时为了比较两种产品，两种仪器，或两种试验方法等的差异，我们常常在相同的条件下做对比试验，得到一批成对（配对）的观测值，然后对观测数据进行分析。作出推断，这种方法常称为配对分析法。

例3 比较甲，乙两种橡胶轮胎的耐磨性，今从甲，乙两种轮胎中各随机地抽取8个，其中各取一个组成一对。再随机选择8架飞机，将8对轮胎随机地搭配给8架飞机，做耐磨性实验

飞行一段时间的起落后，测得轮胎磨损量（单位：mg）数据如下：

§ 8.2 正态总体均值的假设检验

轮胎甲：4900，5220，5500，6020，6340，7660，8650，4870

轮胎乙；4930，4900，5140，5700，6110，6880，7930，5010

解：用X及Y分别表示甲，乙两种轮胎的磨损量

假定 $X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$，其中

$\sigma_1^2 = \sigma_2^2$，欲检验假设

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

下面分两种情况讨论：

§ 8.2 正态总体均值的假设检验

（1）实验数据不配对分析：

将两种轮胎的数据看作来自两个总体的样本观测值，这种方法称为不配对分析法。欲检验假设

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

我们选择统计量

$$t=\frac{\overline{X}-\overline{Y}}{\sqrt{(n_1-1)S^2+(n_2-1)S^2}}\sqrt{\frac{n_1n_2(n_1+n_2-2)}{n_1+n_2}}$$

§ 8.2 正态总体均值的假设检验

由样本数据及$n_1=n_2=8$可得

$$\bar{x}=6145,\quad \bar{y}=5825$$

$$S_1^2=1633900\times 8/7\qquad S_2^2=1053875\times 8/7$$

$$t=320/619.7\approx 0.516$$

对给定的 $\alpha=0.05$查自由度为16-2=14的t分布表

得临界值$t_{\alpha/2}(16-2)=t_{0.025}(14)=2.145$

由于$|t|=0.516<2.145=t_{0.025}(14)$，因而接受$H_0$，即认为这两种轮胎的耐磨性无显著差异

§ 8.2 正态总体均值的假设检验

（2）实验数据配对分析：记 $Z=X-Y$，则

$E(Z)=\mu_1-\mu_2 \overset{def}{=\!=\!=} d$，$D(Z)=\sigma_Z^2$，由正态分布的可加性知，Z服从正态分布 $N(d,2\sigma^2)$ 。

于是，对 μ_1 与 μ_2 是否相等的检验

就变对 $d=0$ 的检验，这时我们可采用关于一个正态总体均值的 t 检验法。将甲，乙两种轮胎的数据对应相减得Z的样本值为：

-30，320，360，320，230，780，720，-140

§ 8.2 正态总体均值的假设检验

计算得样本均值 $\overline{Z}=\frac{1}{8}\sum_{i=1}^{8}Z_i=320$

$$S^2=\sum_{i=1}^{8}(Z_i-\overline{Z})^2/7=102200$$

$$t=(\overline{Z}-0)/\sqrt{S^2/8}=320\times\sqrt{8}/\sqrt{102200}\approx 2.83$$

对给定 α=0.05，查自由度为8-1=7的t分布表

得临界值 $t_{0.025}(7)=2.365$

由于t=2.83>2.365，因而否定H_0，即认为这种轮胎的耐磨性有显著差异。

§ 8.2 正态总体均值的假设检验

- 以上是在同一检验水平α=0.05下采用不同方法的分析结果
- 方法不同所得结果也不一致，到底哪个结果正确呢？下面作一简要分析。
- 因为我们将8对轮胎随机地搭配给8架飞机作轮胎耐磨性试验，两种轮胎不仅对试验数据产生影响，而且不同的飞机也对试验数据产生干扰，因此试验数据配对分析，消除了飞机本身对数据的干扰，突出了比较两种轮胎之间耐磨性的差异。对试验数据不做配对分析，轮胎之间和飞机之间对数据的影响交织在一起，两组样本不独立。用两个独立正态总体的t检验法是不合适的

本章作业

- P_{218}：2，3，5