Chpt.8 Statistical Inference: Hypothesis Testing

第八章 假设检验

上节回顾

假设检验：根据数据，对某一假设进行接受或拒绝的判断

（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；

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

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

（4）根据实际样本观测值作出判断.

8.2 正态总体均值的假设检验

（一）关于单个正态总体均值μ的假设检验

(1) σ^2已知，关于μ的检验

(2) σ^2未知，关于μ的检验

（二）关于两个正态总体均值差$\mu_1-\mu_2$的假设检验

(1) 已知σ_1，σ_2

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

(2) 未知σ_1，σ_2，但是知道$\sigma_1=\sigma_2$

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

（一）关于单个正态总体均值μ的假设检验

(1) σ^2已知，关于μ的检验

采用统计量 $Z=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$ 来确定拒绝域

（一）关于单个正态总体均值μ的假设检验

(1) σ^2已知，关于μ的检验

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

原假设H_0	备择假设H_1	在显著水平α下关于原假设H_0的拒绝域
$\mu=\mu_0$	$\mu\neq\mu_0$	$\lvert Z\rvert>z_{\alpha/2}$
$\mu\leq\mu_0$	$\mu>\mu_0$	$Z>z_\alpha$
$\mu\geq\mu_0$	$\mu<\mu_0$	$Z<-z_\alpha$

强调一下：没有H_0：$\mu\neq\mu_0$ 这个原假设。
换而言之，如果比较谁大谁小，则可将预期支持的观点作为备择假设，而原假设为“对抗观点”；如果比较是否相等，则只能使用$\mu=\mu_0$ 作为原假设，否则没办法构造合理的拒绝域。

练习题

例 2 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，可以检验出牛奶是否掺水. 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 $\sigma=0.008$ ℃. 牛奶掺水可使冰点温度升高而接近于水的冰点温度(0 ℃). 测得生产商提交的 5 批牛奶的冰点温度，其均值为 $\bar{x}=-0.535$ ℃，问是否可以认为生产商在牛奶中掺了水？取 $\alpha=0.05$.

解 按题意需检验假设

$$H_0: \mu \leqslant \mu_0 = -0.545\text{（即设牛奶未掺水），}$$
$$H_1: \mu > \mu_0\text{（即设牛奶已掺水）}$$

这是右边检验问题，其拒绝域如(1.6)式所示，即为

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \geqslant z_{0.05} = 1.645.$$

现在 $z = \dfrac{-0.535 - (-0.545)}{0.008/\sqrt{5}} = 2.7951 > 1.645$，$z$ 的值落在拒绝域中，所以我们在显著性水平 $\alpha = 0.05$ 下拒绝 H_0，即认为牛奶商在牛奶中掺了水. □

（一）关于单个正态总体均值的假设检验

(2) σ^2未知，关于μ的检验

检验统计量 $t=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$

（一）关于单个正态总体均值的假设检验

(2) σ^2未知，关于μ的检验

$$t=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}$$ ——t 检验

原假设H_0	备择假设H_1	在显著水平α下关于原假设H_0的拒绝域
$\mu=\mu_0$	$\mu\neq\mu_0$	$\|t\|>t_{\alpha/2}(n-1)$
$\mu\le\mu_0$	$\mu>\mu_0$	$t>t_{\alpha}(n-1)$
$\mu\ge\mu_0$	$\mu<\mu_0$	$t<-t_{\alpha}(n-1)$

例：设某次考试的考生成绩服从正态分布，从中随机地取36为考生的成绩，算得平均成绩为66.5分，标准差为15分。问在显著性水平为0.05下，是否可以认为这次考试全体考生的平均成绩为70分？

例2：设某次考试的考生成绩服从正态分布，从中随机地取36为考生的成绩，算得平均成绩为66.5分，标准差为15分。问在显著性水平为0.05下，是否可以认为这次考试全体考生的平均成绩为70分？

分析：题中没有假定σ已知，这是在σ未知下单正态总体均值μ的检验问题，应利用T检验法。

解：建立假设

$$H_0:\ \mu = 70 \leftrightarrow H_1:\ \mu \neq 70.$$

这里$n = 36, \mu_0 = 70, \bar{x} = 66.5,\ s = 15.$ 给定$\alpha = 0.05$，查附表$t_{\frac{\alpha}{2}}(n-1) = 2.030$，计算可得

$$t_0 = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} = \frac{66.5 - 70}{15/\sqrt{36}} =- 1.4$$

由于$|t_0| < t_{\frac{\alpha}{2}}(n-1) = 2.030$，故接受$H_0$，即认为平均成绩为70分。

（二）关于两个正态总体均值差的假设检验

有时，我们需要比较两个正态总体的参数（均值）是否存在显著差异。

- 两个农作物品种的产量
- 两种电子元件的使用寿命
- 两种加工工艺对产品质量的影响
- 两地区的气候差异等等

（二）关于两个正态总体均值差的假设检验

总体$X \sim N(\mu_1,\sigma_1^2)$，从中抽取容量为$n_1$的样本$X_1,X_2,\cdots,X_{n_1}$

总体 $Y \sim N(\mu_2,\sigma_2^2)$，从中抽取容量为$n_2$的样本 $Y_1,Y_2,\cdots,Y_{n_2}$

X, Y的样本均值及样本方差分别为 $\overline{X}$，S_1^2；$\overline{Y}$，S_2^2 。

考虑参数μ_1，μ_2的假设检验问题

（二）关于两个正态总体均值的假设检验

(1) 已知σ_1，σ_2，检验 H_0：$\mu_1-\mu_2=\delta$，H_1：$\mu_1-\mu_2\neq\delta$

为什么把假设做成如下形式：

H_0：$\mu_1—\mu_2=\delta$，H_1：$\mu_1—\mu_2\neq\delta$

这种形式几乎覆盖了多种关于μ_1、μ_2的关系，如：

[1] $\delta=0$，$\mu_1=\mu_2$

[1] $\delta>0$，$\mu_1>\mu_2$

[3] $\delta<0$，$\mu_1<\mu_2$

（二）关于两个正态总体均值的假设检验

(1) 已知σ_1，σ_2，检验 H_0：$\mu_1-\mu_2=\delta$，H_1：$\mu_1-\mu_2\neq\delta$

针对以上假设，选取一个统计量，它涉及假设，但不含有未知参数。

使用统计量 $$Z=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1)$$

拒绝域的形式为 $\left|\overline{X}-\overline{Y}-\delta\right|\geq C$

相对于显著水平α构造小概率事件

$$P\left(\,|Z|>z_{\alpha/2}\right)=\alpha$$

得到拒绝域为：$$|z|=\frac{\left|\overline{x}-\overline{y}-\delta\right|}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\geq z_{\alpha/2}$$ —Z检验

例3：据资料，已知某品种小麦每平方米产量(千克)的方差为 $\sigma^2=0.2$. 今在一块地上用A，B 两法，A法设12个样本点，得平均产量 $\overline{x}=1.5$ ；B 法设8个样本点，得平均产量 $\overline{y}=1.6$ ，试比较A、B两法的平均产量差异是否有统计意义 $(\alpha=0.05)$

例3：据资料，已知某品种小麦每平方米产量(千克)的方差为 $\sigma^2=0.2$. 今在一块地上用A，B 两法，A法设12个样本点，得平均产量 $\bar{x}=1.5$ ；B 法设8个样本点，得平均产量 $\bar{y}=1.6$ ，试比较A、B两法的平均产量差异是否有统计意义。$(\alpha=0.05)$

解 假设： $H_0:\mu_X=\mu_Y$， $H_1:\mu_X\neq\mu_Y$

因为：

$$\left|\frac{\bar{x}-\bar{y}}{\sqrt{\sigma^2/n_1+\sigma^2/n_2}}\right|=\left|\frac{1.5-1.6}{\sqrt{0.2/12+0.2/8}}\right|\approx 0.49<1.96=z_{0.025}$$

所以接受H_0假设，即认为 A、B两法的平均产量差异无统计意义

（二）关于两个正态总体均值的假设检验

(2) 未知σ_1，σ_2，但是知道$\sigma_1=\sigma_2$

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

使用统计量 $$T=\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$$

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

（二）关于两个正态总体均值的假设检验

(2) 未知σ_1，σ_2，但是知道$\sigma_1=\sigma_2$

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

提出关于μ_1，μ_2的不同的假设检验中的原假设H_0与备择假设H_1，在显著水平α下关于原假设H_0的拒绝域以及统计量的观测值落在拒绝域内的概率：

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

例4： 某灯泡厂在采用一种新工艺的前后，分别抽取10个灯泡进行寿命(单位：小时)检测，计算得到：

- 采用新工艺前，灯泡寿命的样本均值为2460，样本标准差为56；
- 采用新工艺后，灯泡寿命的样本均值为2550，样本标准差为48.

设灯泡的寿命服从正态分布

问题： 是否可以认为采用新工艺后灯泡的平均寿命有显著提高(取显著水平α=0.01)?

解：设采用新工艺前后的灯泡寿命分别为

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

因为未知σ_1，σ_2，暂时先假定$\sigma_1 = \sigma_2$ (将在以后检验)，

检验假设

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

取统计量

$$T = \overline{X} - \overline{Y} \Big/ S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \sim t(n_1 + n_2 - 2)$$

其中$n_1 = n_2 = 10$， $\bar{x} = 2460$， $\bar{y} = 2550$

$$s_1^2 = 56^2 \qquad s_2^2 = 48^2$$

$$s_w=\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}=\sqrt{\frac{9\times 56^2+9\times 48^2}{20-2}}\approx 52.15$$

$$T\approx\frac{2460-2550}{52.15\sqrt{\frac{1}{10}+\frac{1}{10}}}\approx -3.86$$

查表得

$$t_\alpha(n_1+n_2-2)=t_{0.01}(18)=2.55$$

因为$T=-3.86<-t_{0.01}(18)=-2.55$，所以拒绝原假设$H_0$，

接受备择假设H_1

即在显著水平α=0.01下，认为采用新工艺后灯泡的平均寿命有显著提高.

（三）成对数据的检验

对于某些情况下，我们提出一种方法，需要知道这个方法是否有效。统计中每个样本具有各自的随机性，需要研究A-B对应状态下非独立样本的假设检验

书上P187. 例4：我们希望知道人对于红光和绿光的反应时间是否有差异。实验在点亮红光或绿光的同时，启动计时器，要求受试者见到红光或绿光点亮时，就按下按钮，切断计时器，这就能测得反应时间。结果如下：

红光(x)	0.3	0.23	0.41	0.53	0.24	0.36	0.38	0.51
绿光(y)	0.43	0.32	0.58	0.46	0.27	0.41	0.38	0.61
d=x-y	-0.13	-0.09	-0.17	0.07	-0.03	-0.05	0	-0.1

不同于之前的两个正态分布的比对：第一行的8个数据不能看成某个总体的样本，因为不同的实验条件，如人的反应速度，发出红光的机器等随机因素，导致不满足同分布的条件。

同样，第二行的8个数据也不能看成某个总体的样本。

每列都是一个受检样本的个例，具有各列的独立性和上下相关性。

（三）成对数据的检验

对于n对相互独立的观测结果 $(X_1, Y_1), \dots, (X_n, Y_n)$.

$X_1, \dots, X_n$不是某个总体的样本，$Y_1, \dots, Y_n$也不是某个总体的样本。

令$D_i = X_i - Y_i$。则可以得到以下结论：

- $D_1, \dots, D_n$之间相互独立。
- D_i的变化来自于同一个因素，则D_i服从同一分布。

假设$D_i \sim N(\mu, \sigma^2)$，可以看成是某一正态总体的样本。

则根据以前的知识点，可以做以下几种假设检验：

1） $H_0: \mu = 0, H_1: \mu \neq 0$

2） $H_0: \mu \leq 0, H_1: \mu > 0$

3） $H_0: \mu \geq 0, H_1: \mu < 0$

（三）成对数据的检验

红光(x)	0.3	0.23	0.41	0.53	0.24	0.36	0.38	0.51
绿光(y)	0.43	0.32	0.58	0.46	0.27	0.41	0.38	0.61
d=x-y	-0.13	-0.09	-0.17	0.07	-0.03	-0.05	0	-0.1

这里的原假设是什么呢？

我们发现d 负数更多，所以猜想 D<0(X普遍小于Y)，所以做假设检验

$$H_0: \mu \geq 0, \qquad H_1: \mu < 0$$

$n = 8, \bar{x}_d =- 0.0625, s_d = 0.0765$,根据此数据做$t$检验

$$\frac{\bar{x}_d - 0}{s_d/\sqrt{8}} =- 2.311 <- t_{0.05}(7) =- 1.8946$$

落入拒绝域$(-\infty, -1.8946)$，拒绝H_0，接受信备择假设$H_1: \mu_0 < 0$，即X<Y，**人眼对红光反应时间更短**

正态分布下对于μ检验的总结

	σ情况	H_0	检验方法	统计量	拒绝域	选择该统计量、分布及拒绝域的原因
单样本	已知	$\mu = \mu_0$	z检验			
		$\mu \geq \mu_0$				
		$\mu \leq \mu_0$				
	未知	$\mu = \mu_0$	t检验			
		$\mu \geq \mu_0$				
		$\mu \leq \mu_0$				
两个样本	已知	$\mu_1 - \mu_2 = \delta$	z检验			
		$\mu_1 - \mu_2 \geq \delta$				
		$\mu_1 - \mu_2 \leq \delta$				
	未知，但相等	$\mu_1 - \mu_2 = \delta$	t检验			
		$\mu_1 - \mu_2 \geq \delta$				
		$\mu_1 - \mu_2 \leq \delta$				
对比样本	不讨论	$\mu = \mu_0$	t检验			
		$\mu \geq \mu_0$				
		$\mu \leq \mu_0$				