3.3 假设检验

- 假设检验的基本思想及基本概念
- 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

前面一章我们介绍了统计推断的一个基本问题：参数估计问题。这一章介绍由样本推断总体的另一个问题：假设检验问题。

假设检验的基本思想及基本概念

一.问题的提出

实例：(1)设某厂生产的某型号的灯管,其寿命X $\sim N(\mu, 200^2)$

从过去较长一段时间来看：灯管平均寿命 $\mu = 1500h$

现采取新的工艺生产灯管，为了了解采用新工艺后，灯管的寿命有无显著增加，我们任意抽取25只进行测试，测得 $\bar{x}$=1675（h），问采取新工艺后，灯管的寿命有无显著提高？

这个问题是,在正态总体前提下,判断 “$\mu = 1500$ ” 是否成立？

(2)某种羊毛处理前后，各取得样本，测得含脂率(%)

处理前X:19,18,21,30,66,42,8,12,30,27

处理后Y:15,13,24, 7, 19,4, 8,20,

设羊毛含脂率服从正态分布，问处理前后平均含脂率有无显著变化？

这个问题是,在正态总体前提下,判断 “$EX = EY$ ” 是否成立

以上两个问题具有代表性,它们的共同点是对总体的分布或对总体的分布参数作出假设，然后通过样本观察值，去判断 “假设”是否成立？

把有关总体的分布或分布参数作出假设称为统计假设

把实例（1）用统计假设的形式提出来

第一个假设 $\mu = 1500$ 称为原假设，记为：$H_0: \mu = 1500$

相反的假设 $\mu \neq 1500$ 称为备择假设，记为：$H_1: \mu \neq 1500$

若由样本判别 H_0 成立，则认为采取新工艺后，灯管的寿命无显著提高;

若由样本判别 H_1 成立，则认为采取新工艺后，灯管的寿命有显著提高。

二.假设检验的基本原理

以实例（1）来说明假设检验的基本原理

具体要由样本判断是 H_0 成立，还是H_1 成立，因为$H_0(H_1)$ 是关于均值 μ 的假设，当然要计算 $\bar{X}=1675h$ 能否因为 $\bar{X}\neq 1500$,就断定 H_0 不成立呢？

不能轻率下结论，要知道引起 $\bar{x}$与 $\mu=1500$ 发生差异的可能是偶然因素（如工人技术水平差异等）引起的误差-随机误差，即使在同一工艺下，这种误差是不可避免的。亦可能是生产条件的改变（如采取新的工艺等）引起的误差-条件误差。

若差异是由随机误差引起的，则认为采取新工艺后，灯管的寿命无显著提高;

若差异是由**条件误差**引起的，则认为采取新工艺后，灯管的寿命有显著提高。

显然，如何区别是由**随机误差**与**条件误差**引起的差异，是解决假设检验问题的关键。

我们的具体做法是：假设差异是由一种误差引起，若导出矛盾，则可以肯定另一个误差存在（**罗辑思维方法-概率论中的“反证法”**）

(1)假设仅存在**随机误差**，即认为25只灯泡仍取自于原来的总体X~ $N(\mu, 200^2)$，即认为 H_0 成立

(2)在这个假设下，构造统计量

$$U=\frac{\overline{X}-1500}{\sqrt{\frac{200^2}{25}}} \overset{H_0}{\sim} N(0,1)$$

(3)利用统计量U,查N(0,1)分布表，使得

$$P\left(|U| > U_{\frac{\alpha}{2}}\right) = \alpha$$

具体 α 取多少，由实际问题来决定，一般取很小，保证 "$|U| > U_{\frac{\alpha}{2}}$" 为小概率事件。如 α =0.05 ,$U_{0.025} = 1.96$

(4)由样本观察值，算出U的值：$\overline{X} = 1675, U = 4.375$

因为"$|U| > 1.96$" 小概率事件，而 $|U| = 4.375 > 1.96$

即一次抽样后，小概率事件居然发生了。一般小概率事件，几乎是不可能发生的。这就使得我们不得不怀疑假设 H_0 成立的合理性，从而否定 H_0 成立，而认为 H_1 成立。

即认为差异是由条件误差引起的，则认为采取新工艺后，灯管的寿命有显著提高。

这里，我们最后作出统计推断的根据---**小概率原理**

称为**假设检验的基本原理，**即认为**小概率事件在一次试验中几乎不可能发生**。若一次试验中，小概率事件居然发生了，人们认为该事件的前提条件发生了变化，即认为前提与**小概率原理**矛盾。

在上面讨论中，牵涉到假设检验的**几个基本概念**

三.假设检验的基本概念

在上例中，我们取$\alpha = 0.05$，作出判断的根据是：

$$P\left(|U| > 1.96\right) = 0.05$$

对于给定的样本观察值，算出U的值：

$$|U| = 4.375$$

若 $|U|>1.96$ ，则拒绝 H_0，即否定 H_0。

若 $|U|\le 1.96$ ，则接受 H_0，即否定H_1 。

这里，把 $\alpha=0.05$ 称为检验水平（显著性水平）

而把 $|U|>1.96$ 所确定的区域，称为拒绝域（否定域），

$|U|\le 1.96$ 所确定的区域，称为接受域。

把-1.96,1.96称为临界值（点）（图示说明概念）

注：检验水平 α 是由实际问题而确定的，一般取得很小。

综合以上讨论，可以给出假设检验的一般步骤

四.假设检验的一般步骤:

1.由实际问题给出原假设 H_0 与备择假设H_1 ；

2.确定检验水平 α ；（检验水平α 据具体问题而定，一

对较重要的问题，希望作出否定原假设的结论要谨慎些，一般 α 该取得小一些）

3.选择一个合适的统计量 ξ，(如上例中的U),给出在 H_0 成立时的分布；

4. 对于给定的检验水平 α，取得临界值，从而定出拒绝域

确定拒绝域的根据是小概率原理，使 $P(\xi \in C) = \alpha$ 则

$"\xi \in C"$ 为拒绝域；

5. 由样本观察值，计算出统计量 ξ 的取值；

若 $"\xi \in C"$，则拒绝 H_0，即否定 H_0。

若 $"\xi \notin C"$，则接受 H_0，即否定 H_1。

五.两类不同的错误

由于样本的抽样随机性，每一个假设检验都会犯两种不同类型的错误。原假设H_0本来正确，而由于抽样的随机性，观察值落入了拒绝域，从而作出了拒绝H_0的决定，

我们称为**犯第一类错误**。类似 H_0 不正确，而由于抽样的随机性，观察值落入了接受域，从而H_0被接受了，

我们称为**犯第二类错误**。

一个检验的好坏，可以有犯两类错误的概率来度量，

显然： $P(H_0\text{拒绝}|H_0\text{为真})=\alpha$ **---犯第一类错误的概率**

$P(H_0\text{接受}|H_0\text{不真})=\beta$ **---犯第二类错误的概率**

一般在假设检验中，希望 α与β尽可能地小，但在

n固定时，我们无法使它们同时尽可能地小，它们是互相制约的。鉴于上述情况，统计学家奈曼与皮尔逊提出：在控制犯第一类错误的概率，即给定 $\alpha(0<\alpha<1)$，然后，寻找犯第二类错误的概率尽可能地小的检验，这就是假设检验的理论中著名的奈曼与皮尔逊原则。

下面就假设检验中所采用的**检验统计量**的形式的不同，介绍几种常见的参数的假设检验问题。

U-检验

在上一节例子中，我们对于正态总体下，总体均值 μ 的检验 $H_0: \mu = 1500$ 是借助于统计量

$$U = \frac{\overline{X} - \mu_0}{\sqrt{\frac{\sigma^2}{n}}} \overset{H_0}{\sim} N(0,1)$$

来进行检验的。今后，我们把借助于服从**标准正态分布** $N(0,1)$ 的统计量进行的检验，称为**U-检验。**

本节介绍**U-检验**的一般步骤及其它的应用

一.单个正态总体$(\sigma^2$已知$)$均值的检验

设 $X_1, X_2, \cdots, X_n$是取自于总体X$\sim N(\mu, \sigma^2)$ 的样本$(\sigma^2$已知$)$

检验： $H_0: \mu = \mu_0 \quad \Leftrightarrow H_1: \mu \neq \mu_0$

构造统计量： $$U = \frac{\overline{X} - \mu_0}{\sqrt{\dfrac{\sigma^2}{n}}} \overset{H_0}{\sim} N(0,1)$$

对于给定的检验水平α， 可以直接查标准正态分布表，

使得： $$P\left(|U| > U_{\frac{\alpha}{2}}\right) = \alpha$$

从而确定拒绝域： "$|U| > U_{\frac{\alpha}{2}}$"

从而对于给定的样本观察值 $x_1, x_2, \cdots, x_n$ ，计算U的值

若“$|U| > U_{\frac{\alpha}{2}}$”，则**拒绝**$H_0$ ，即**否定**H_0 。

若“$|U| \leq U_{\frac{\alpha}{2}}$”，则**接受** H_0 ，即**否定**H_1 。

二.两个正态总体 $\left(\sigma_1{}^2, \sigma_2{}^2\text{已知}\right)$ 均值的检验

设 $X_1, X_2, \cdots, X_n$ 与 $Y_1, Y_2, \cdots, Y_m$ 分别是来自正态总体 $X \sim N(\mu_1, \sigma_1^2)$ 与 $Y \sim N(\mu_2, \sigma_2^2)$ 的相互独立的简单随机样本. $\left(\sigma_1{}^2, \sigma_2{}^2\text{已知}\right)$

检验： $H_0: \mu_1 = \mu_2 \quad \Leftrightarrow H_1: \mu_1 \neq \mu_2$

构造统计量：

$$U = \frac{\overline{X} - \overline{Y}}{\sqrt{\dfrac{\sigma_1{}^2}{n} + \dfrac{\sigma_2{}^2}{m}}} \overset{H_0}{\sim} N(0,1)$$

对于给定的检验水平α， 可以直接查标准正态分布表，

使得： $P\left(|U| > U_{\frac{\alpha}{2}}\right) = \alpha$

从而确定拒绝域：“$|U| > U_{\frac{\alpha}{2}}$”

从而对于给定的样本观察值 $x_1, x_2, \cdots, x_n$ ，计算U的值

若“$|U| > U_{\frac{\alpha}{2}}$”，则**拒绝** H_0 ，即**否定** H_0。

若“$|U| \leq U_{\frac{\alpha}{2}}$”，则**接受** H_0，即**否定** H_1。