概率论与数理统计

第22讲

假设检验的基本概念和步骤

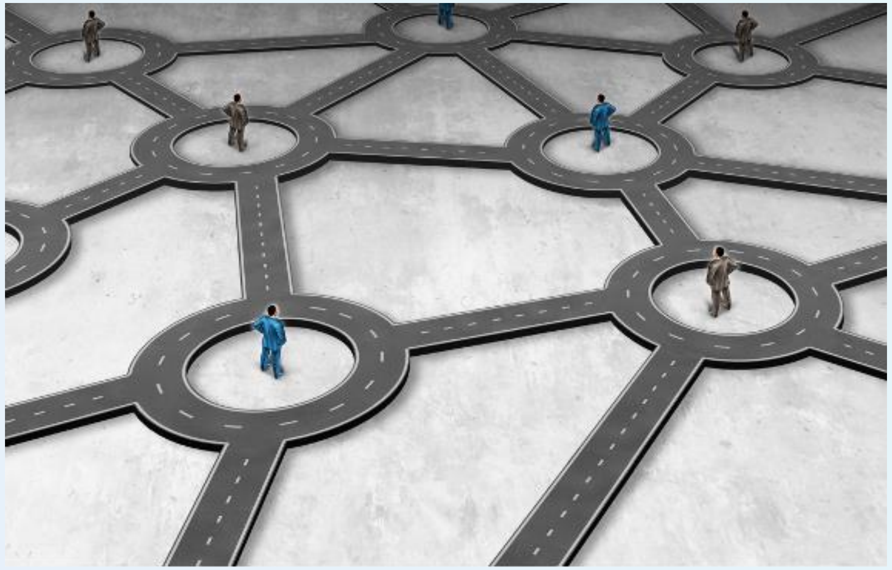
一. 理论依据

假设检验是统计推断的一个主要部分。在科学研究、日常工作甚至生活中经常对某一件事情提出疑问。

解决疑问的过程往往是先做一个和疑问相关的假设，然后在这个假设下去寻找有关的证据，如果得到的证据是和假设相矛盾的，就要否定这个假设。

假设检验是统计推断的另一个重要组成部分，分为参数假设检验和非参数假设检验。

一. 理论依据



参数假设检验所处理的问题是： 总体的分布类型已知，对总体分布中的未知参数提出某种假设，然后利用样本(即数据)对假设进行检验，最后根据检验的结果对所提出的假设作出成立与否的判断。

非参数假设检验是对总体分布函数的形式提出某种假设并利用样本进行检验。

一. 理论依据

假设检验的理论依据是实际推断原理(小概率原理)

——即认为小概率事件在一次试验中几乎(一般)是不会发生的。

这是人们在长期大量实践中总结出来的，并且广泛采用着的一个原理。

由大数定律 $\lim\limits_{n\to\infty} P\{|\frac{n_A}{n}-P(A)|<\varepsilon\}=1$ 可知n_A/n与$P(A)$相差不大。

如：当$P(A)$=0.01时，在100次试验中，A大概只发生1次，因而可以认为如果只做1次试验，A不会发生。

一. 理论依据

根据小概率原理： 如果实际观测到的数据在某假设下出现的可能性很小，则认为该假设是错误的。

统计假设检验是具有概率性质的反证法。

例1.设有一大批产品，从中随机不放回的取出5件产品检查，发现有4件次品，1件正品。问这批产品的次品率p是否是0.1?

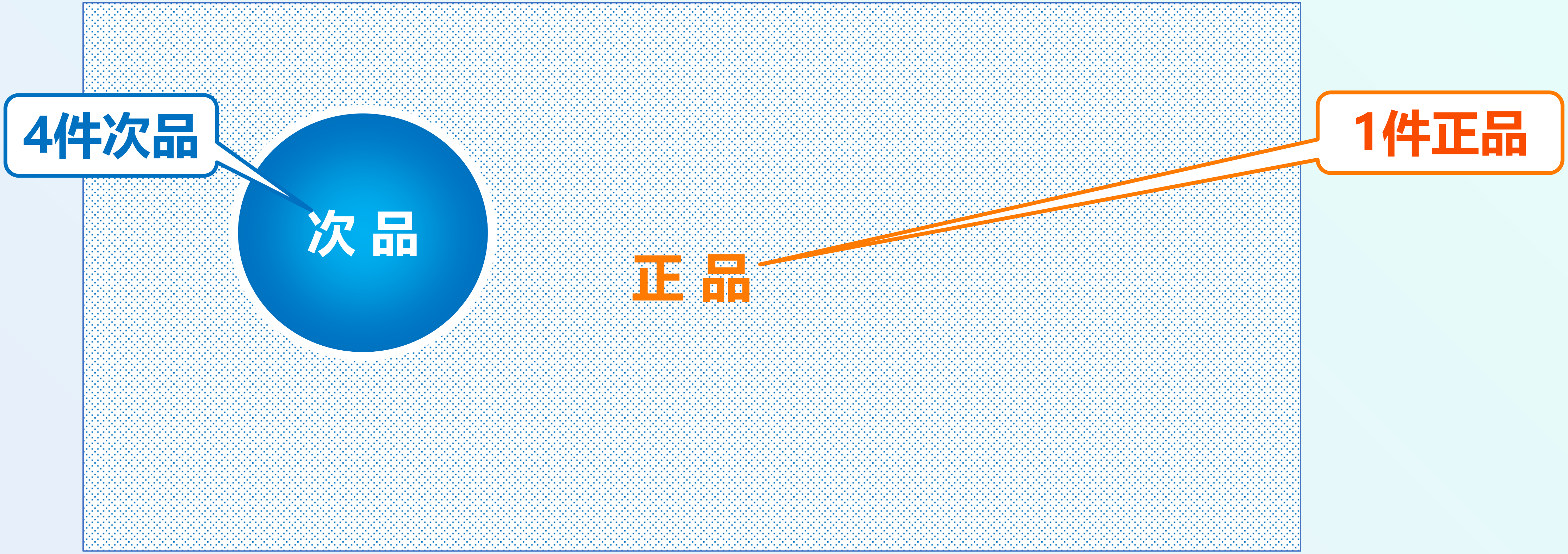
解：提出假设H_0：$p=0.1$。设A表示5件产品中有4件次品1件正品。

在H_0为真的条件下，计算可得

$$P(A)=C_5^4(0.1)^4 0.9=0.00045$$

即A为小概率事件，但在一次试验中A实际发生了，这与小概率原理矛盾。而这个矛盾是由H_0为真造成的。 故否定H_0，认为$p \neq 0.1$。

根据以上推导，否定H_0：$p=0.1$，但可能会犯错误。而且这种错误是不可避免的。



下面先从一个例子出发给出假设检验的基本概念

例2.某车间生产的钢管直径X服从正态分布$N(\mu,0.5^2)$，其中μ未知，生产的钢管平均直径达到100mm符合规定要求。现从一批钢管中随机抽取10根，测得其直径的平均值为 $\bar{x}$ =100.15mm，问该批生产的钢管是否符合规定要求?

已知：总体X服从正态分布$N(\mu,0.5^2)$，且μ 未知

符合规定要求的标准为：生产的钢管平均直径达到100mm。

$X_1, X_2, \ldots, X_{10}$为为抽取的10跟钢管的直径，即来自总体X的样本。

这组样本的一个观测值的平均值为100.15。 令μ_0=100，即μ_0为常数。

假设检验的做法分以下几步来叙述。

1. 建立假设 — 即提出一个关于总体分布的命题

钢管是否符合要求是指平均直径是$\mu=\mu_0$还是$\mu\neq\mu_0$，若相等就表示符合要求，否则就要进行停产检查。

根据以往经验，先假设符合要求，即$\mu=\mu_0$

称$\mu=\mu_0$为原假设或零假设，记为H_0: $\mu=\mu_0$

当原假设H_0不成立时，称μ的其它取值为备择假设或对立假设，记为H_1。

如：$\mu>\mu_0$，$\mu<\mu_0$，$\mu\neq\mu_0$ 本例中取$\mu\neq\mu_0$为备择假设，即H_1：$\mu\neq\mu_0$

备择假设就是H_0不对的时候可供选择的命题。

在该例中，两个命题为：H_0: $\mu=\mu_0$　　H_1： $\mu\neq\mu_0$

2. 寻找检验统计量

假设检验问题就是利用样本提供的信息(即数据)来检验原假设H_0成立还是备择假设H_1成立。

做法是：先假定原假设H_0成立，根据样本构造一个合适的小概率事件。根据实际推断原理，如果在一次具体抽样中这个小概率事件发生了，此时与小概率原理相矛盾，说明原假设H_0成立的假定是错误的；若小概率事件没有发生，则无法拒绝原假设H_0。此时，需要根据实际问题做进一步的研究。

二. 基本概念

由于样本所含信息较分散，因此需要构造一个统计量$T(X_1,\dots,X_n)$，然后利用该统计量进行判断，称该统计量为**检验统计量**。

在该例中，由于要检验的假设是关于总体均值μ的，选择样本均值为检验统计量,即 $T(X_1,X_2,\cdots,X_{10})=\bar{X}$。

假定H_0成立，即： $\mu=\mu_0=100$

此时$\bar{X}$的观测值$\bar{x}$应接近100，如果相差较大，就有理由怀疑H_0不真。

即：当H_0： $\mu=\mu_0=100$成立时，$\bar{x}$与μ_0相差较大，是小概率事件

如今观测值100.15与100相差是否较大？或者讲 $\bar{x}$ 与100相差多大才算相差较大，从而拒绝H_0呢？

因此这里需要一个界限，记为c，当$\bar{x}$与100相差大于c就拒绝H_0，否则就接受H_0。即：

当 $|\bar{x}-100|\geq c$ 就拒绝H_0； 当 $|\bar{x}-100|< c$ 就接受H_0。

即当H_0成立时，$\{|\bar{X}-100|\geq c\}$是小概率事件，如果在一次试验中发生了，就有理由怀疑H_0的正确性而拒绝H_0。

3. 检验法则

当$T(x_1,\ldots, x_{10})\in C$ 时，拒绝H_0，否则接受H_0。

令$W=\{(x_1,\ldots, x_{10}):\ T(x_1,\ldots, x_{10})\in C\}$

称其为检验的拒绝域，它的边界点称为检验的临界点

令$A=\{(x_1,\ldots, x_{10}):\ T(x_1,\ldots, x_{10})\notin C\}$

称其为检验的接受域。

4. 显著水平与临界值

由于作出决策的依据是一个样本，当实际H_0为真时仍有可能作出拒绝H_0的判断，这是一种错误。我们无法排除这类错误，因此自然希望将犯这类错误的概率控制在一定的限度内，即给出一个较小的数α $(0<\alpha<1)$，使

$$P\{拒绝H_0 | H_0为真\} \le \alpha，即P\{W| H_0\} \le \alpha$$

确定检验的临界点

其中α称为检验的显著性水平。

由上式可知，当H_0为真时，W为小概率事件。

如果样本观测值落入了W中，说明小概率事件在一次试验中发生了，根据小概率原理，应该拒绝H_0。

例2. 当 $|\bar{x}-100|\geq c$ 就拒绝 H_0； 当 $|\bar{x}-100|< c$ 就接受 H_0

假定 H_0 成立时 $\bar{X}\sim N(100, 0.5^2/10)$

对给定的显著性水平 α=0.05，可由下式确定 c 的值。

$$P\{|\bar{X}-100|\geq c \mid H_0\}=0.05$$

$$P\{|\bar{X}-100|\geq c \mid H_0\}=P\{|\bar{X}-100|\geq c \mid \mu=100\}$$

$$=P\{\frac{|\bar{X}-100|}{\sqrt{0.5^2/10}}\geq \frac{c}{\sqrt{0.5^2/10}} \mid \mu=100\}=0.05$$

$$\frac{c}{\sqrt{0.5^2/10}}=z_{0.025}=1.96 \quad \text{即} \quad c=z_{0.025}\sqrt{0.5^2/10}=0.3099$$

例2. 对给定的显著性水平α=0.05，拒绝域为$|\bar{x}-100|\geq 0.3099$

根据现有样本观测值 $\bar{x}=100.15$，而 $|100.15-100|=0.15<0.3099$

即，样本没有落入拒绝域，认为该批生产的钢管符合规定要求。

假设检验中的基本概念

(1) 假设：关于总体分布的某个命题

原假设：指需要检验的假设，记为H_0。

备择假设：在拒绝原假设后，可供选择的一个命题。

它可以是原假设对立面的全体，也可以是其中的一部分，记为H_1。

给定H_0和H_1就等于给定一个检验问题，记为检验问题（H_0, H_1）。

(2) 检验

根据从总体中抽取的样本，并集中样本中的有关信息，对假设H_0的真伪进行判断，称为检验。需要构造检验统计量。

(3) 检验法则：就是构建拒绝域和接受域

拒绝域： 使原假设H_0被拒绝的样本观测值所组成的区域，记为W。

接受域： 保留原假设H_0的样本观测值所组成的区域，记为$A=S-W$。

(4) 显著性水平

$$P\{拒绝H_0 \mid H_0为真\}=P\{W \mid H_0\}\le \alpha$$

α称为检验的显著性水平，为给定的常数值。利用此式可以确定拒绝域中的临界值。

假设检验的基本步骤

1 第一步，根据实际问题的要求和已知信息提出原假设H_0和备择假设H_1。

2 第二步，构造检验统计量

设概率密度$f(x,\theta)$的表达式已知，通常是基于参数θ的最大似然估计构造一个检验统计量$T=T(X_1, X_2, \ldots, X_n)$，并在原假设$H_0$成立下，确定$T$的精确分布或渐近分布。

3 第三步，确定拒绝域的形式

根据原假设H_0和备择假设H_1的形式，分析并提出拒绝域的形式，以及待定的临界点。

4 第四步，给定显著性水平α的值，确定临界点c。

由$P\{$拒绝$H_0 | H_0$为真$\} = P\{W | H_0\} \le \alpha$出发，以此确定临界点$c$，从而也就确定了拒绝域。

5 第五步，根据样本观察值作出是否拒绝H_0的判断。

若样本观察值$(x_1, x_2, \ldots, x_n)$落入拒绝域，即$(x_1, x_2, \ldots, x_n) \in W$，则拒绝$H_0$；否则无法拒绝$H_0$。判断的基本原理是实际推断原理即小概率原理。

三. 假设检验的两类错误

第一类错误：

原假设H_0为真，但由于样本的随机性，使样本观测值落入拒绝域，从而作出拒绝H_0的结论，这类错误称为第一类错误，它发生的概率称为犯第一类错误的概率，也称为 **“拒真概率”** 。

第二类错误：

原假设H_0为假，但由于样本的随机性，使样本观测值落入接受域，从而作出保留H_0的结论，这类错误称为第二类错误，它发生的概率称为犯第二类错误的概率，也称为 **“取伪概率”** 。

三. 假设检验的两类错误

所作判断 / 真实情况	接受 H_0	拒绝 H_0
H_0 为真	正 确	第一类错误（弃真）
H_0 为假	第二类错误（取伪）	正 确

易知： 第一类错误的概率=“拒真概率” $= P(W \mid H_0)$

第二类错误的概率=“取伪概率”$=P(A \mid H_1) = P(\overline{W} \mid H_1)$

例3. 某车间生产的钢管直径X服从正态分布$N(\mu,0.5^2)$，其中μ未知，生产的钢管平均直径达到100mm符合规定要求。现从一批钢管中随机抽取10根，测得其直径的平均值为$\bar{x}$=100.15mm。已知假设检验问题为：

$$H_0: \mu=\mu_0=100 \qquad H_1: \mu\neq\mu_0$$

且已知该检验的拒绝域为 $|\bar{x}-100|\geq 0.3099$

求该检验的犯两类错误的概率。

解：先求犯第一类错误的概率

易知当H_0成立 $\bar{X}\sim N(100, 0.5^2/10)$

三. 假设检验的两类错误

$H_0: \mu = \mu_0 = 100 \quad H_1: \mu \neq \mu_0$ **拒绝域为:** $|\bar{x} - 100| \geq 0.3099 \quad \bar{X} \sim N(100, 0.5^2/10)$

$$P\{W \mid H_0\} = P\{|\bar{X} - 100| \geq 0.3099 \mid \mu = \mu_0 = 100\}$$

$$= P\{\frac{|\bar{X} - 8000|}{\sqrt{0.5^2/10}} \geq \frac{0.3099}{\sqrt{0.5^2/10}} \mid \mu = 100\}$$

$$= P\{\frac{|\bar{X} - 8000|}{\sqrt{0.5^2/10}} \geq 1.96 \mid \mu = 100\}$$

$$= P\{\frac{|\bar{X} - 8000|}{\sqrt{0.5^2/10}} \geq z_{0.025} \mid \mu = 100\}$$

$N(0,1)$

$$= 0.05 = \alpha$$

三. 假设检验的两类错误

H_0: $\mu = \mu_0 = 100$ H_1: $\mu \neq \mu_0$ **拒绝域为:** $|\bar{x} - 100| \geq 0.3099$

第二类错误的概率: H_1 **成立** $\bar{X} \sim N(\mu, 0.5^2/10)$

$$P\{\bar{W} \mid H_1\} = P\{|\bar{X} - 100| < 0.3099 \mid \mu \neq 100\}$$

$$= P\{100 - 0.3099 < \bar{X} < 100 + 0.3099 \mid \mu \neq 100\}$$

$$= \Phi\left(\frac{100 + 0.3099 - \mu}{\sqrt{0.5^2/10}}\right) - \Phi\left(\frac{100 - 0.3099 - \mu}{\sqrt{0.5^2/10}}\right)$$

$$= \Phi\left(\frac{100 - \mu}{\sqrt{0.5^2/10}} + \frac{0.3099}{\sqrt{0.5^2/10}}\right) - \Phi\left(\frac{100 - \mu}{\sqrt{0.5^2/10}} - \frac{0.3099}{\sqrt{0.5^2/10}}\right)$$

$$= \Phi\left(\frac{100-\mu}{\sqrt{0.5^2/10}} + \frac{0.3099}{\sqrt{0.5^2/10}}\right) - \Phi\left(\frac{100-\mu}{\sqrt{0.5^2/10}} - \frac{0.3099}{\sqrt{0.5^2/10}}\right)$$

$$= \Phi\left(\frac{100-\mu}{\sqrt{0.5^2/10}} + 1.96\right) - \Phi\left(\frac{100-\mu}{\sqrt{0.5^2/10}} - 1.96\right)$$

$$= \Phi\left(\frac{100-\mu}{\sqrt{0.5^2/10}} + z_{0.025}\right) - \Phi\left(\frac{100-\mu}{\sqrt{0.5^2/10}} - z_{0.025}\right)$$

$$= \Phi\left(\frac{100-\mu}{\sqrt{0.5^2/10}} + z_{\alpha/2}\right) - \Phi\left(\frac{100-\mu}{\sqrt{0.5^2/10}} - z_{\alpha/2}\right)$$

三. 假设检验的两类错误

我们总希望所用的检验方法尽量少犯错误，但不能完全排除犯错误的可能性。理想的检验方法应使犯两类错误的概率都很小。

第一类错误的概率 $P\{W \mid H_0\} = \alpha$

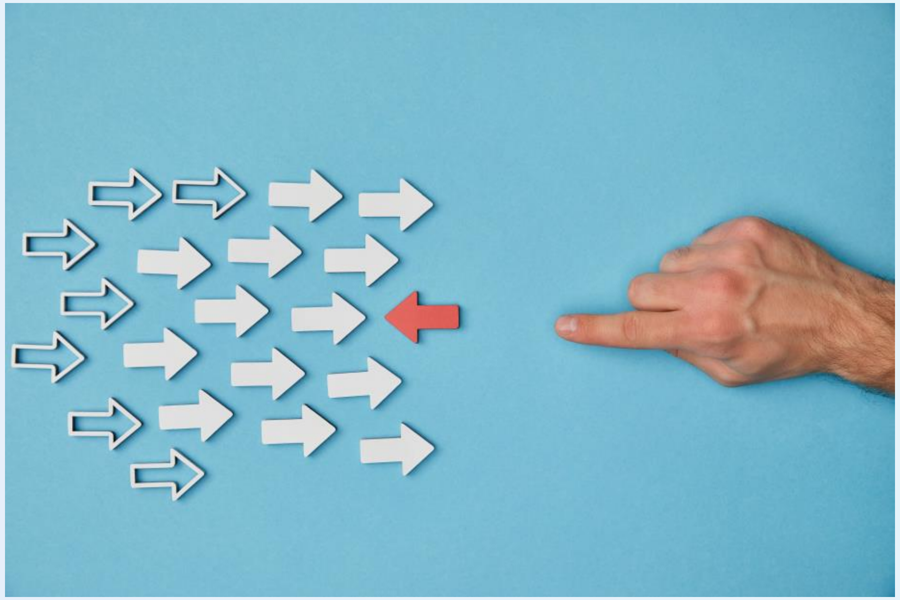
第二类错误的概率 $P\{\bar{W} \mid H_1\} = \Phi\left(\dfrac{100-\mu}{\sqrt{0.5^2/10}} + z_{\alpha/2}\right) - \Phi\left(\dfrac{100-\mu}{\sqrt{0.5^2/10}} - z_{\alpha/2}\right)$

$P\{W \mid H_0\}=\alpha$ 减小，导致 $z_{\alpha/2}$ 增大，$P\{\bar{W} \mid H_1\}$ 增大。

在一般情形，当样本容量固定时，减小一类错误概率会导致另一类错误概率的增加。

三. 假设检验的两类错误

要同时降低两类错误的概率，或者要在第一类错误的概率不变的条件下降低第二类的错误概率，需要增加样本容量。



一般来说，总是控制犯第一类错误的概率，使它不大于α，α的大小视具体情况而定，然后在此约束下根据对犯第二类错误概率的实际要求选取适当的样本容量。

三. 假设检验的两类错误

但具体实行这一原则还会有许多理论和实际上的困难，因而有时把这一原则简化成只对犯第一类错误的概率加以控制，而不考虑犯第二类错误的概率，称这种检验为显著性检验。

——控制第一类错误的原则

假设检验的显著性水平α实际上就是犯第一类错误的最大概率。

关于零假设与备择假设的选取

H_0与H_1地位应平等，但在控制犯第一类错误的概率的原则下，使得采取拒绝H_0的决策变得较慎重，即H_0得到特别的保护。

因而，通常把有把握的、不能轻易改变的或存在已久的状态作为原假设。

另外，由于在做假设检验时，控制的是犯第一类错误的概率，因此选取后果严重的错误作为第一类错误，这也是选取原假设和备择假设的一个原则。

三. 假设检验的两类错误

在实际问题中，情况比较复杂，如何选取原假设和备择假设没有一个绝对的标准，只能在实践工作中积累经验，根据实际情况去判断如何选取。

例4.设总体$X\sim N(\mu,1)$，其中μ为未知参数，$X_1, X_2, \ldots, X_9$为来自总体X的一个样本，考虑假设检验问题H_0：$\mu=2, H_1:\mu=3$，若检验的拒绝域由$W=\{(X_1,\cdots,X_9):\bar{X}\geq 2.6\}$所确定。求该检验犯第一类错误的概率$\alpha$和第二类错误的概率$\beta$。

解：当原假设H_0: $\mu=2$成立时，$\bar{X}\sim N(2,\frac{1}{9})$

$$\alpha=P\{W\mid H_0\}=P\{\bar{X}\geq 2.6\mid \mu=2\}=P\{\frac{\bar{X}-2}{1/\sqrt{9}}\geq\frac{2.6-2}{1/\sqrt{9}}\}=1-P\{\frac{\bar{X}-2}{1/\sqrt{9}}<0.6\sqrt{9}\}$$

$$=1-\Phi(0.6\sqrt{9})=1-\Phi(1.8)=1-0.9641=0.0459$$

假设检验问题： $H_0:\mu=2, H_1:\mu=3$，**拒绝域** $W=\{(X_1,\cdots,X_9):\bar{X}\geq 2.6\}$。

求第二类错误的概率β。

解： 当原假设H_1: μ =3成立时，$\bar{X}\sim N(3,\frac{1}{9})$

$$\beta=P\{\bar{W}\mid H_1\}=P\{\bar{X}<2.6\mid\mu=3\}=P\{\frac{\bar{X}-3}{1/\sqrt{9}}<\frac{2.6-3}{1/\sqrt{9}}\}=P\{\frac{\bar{X}-3}{1/\sqrt{9}}<-0.4\sqrt{9}\}$$

$$=\Phi(-1.2)=1-\Phi(1.2)=1-0.8849=0.1151$$

例5.设总体$X\sim\pi(\lambda)$，其中$\lambda>0$为未知参数，X_1, X_2, X_3为来自总体X的一个样本，考虑假设检验问题H_0: $\lambda=3$, H_1: $\lambda=1/3$，若检验的拒绝域由$W=\{X_1+X_2+X_3\le 1.5\}$所确定.求该检验犯第一类错误的概率α和第二类错误的概率β。

解：当H_0:λ=3成立时，$X_1+X_2+X_3\sim\pi(9)$

犯第一类错误的概率为：

$$\alpha = P\{W \mid \lambda=3\}=P\{X_1+X_2+X_3\le 1.5 \mid \lambda=3\}$$

$$= P\{X_1+X_2+X_3=0\}+P\{X_1+X_2+X_3=1\} = \frac{9^0 e^{-9}}{0!}+\frac{9^1 e^{-9}}{1!} = 10e^{-9}$$

例6.假设检验问题H_0:λ=3，H_1:λ=1/3，拒绝域为$W=\{X_1+X_2+X_3\le 1.5\}$。求第二类错误的概率β。

解：当H_1:λ =1/3成立时，$X_1+X_2+X_3 \sim \pi(1)$

$$\beta = P\{\overline{W} \mid \lambda = 1/3\} = P\{X_1 + X_2 + X_3 > 1.5 \mid \lambda = 1/3\}$$

$$= 1 - P\{X_1 + X_2 + X_3 = 0\} - P\{X_1 + X_2 + X_3 = 1\}$$

$$= 1 - \frac{1^0 e^{-1}}{0!} - \frac{1^1 e^{-1}}{1!} \quad = 1 - 2e^{-1}$$

作业：6，7

第22讲

谢谢观看