

# 数理统计

## 第五章

### 参数假设检验

2026 年 5 月 11 日

## 1 5.1 假设检验的若干基本概念

- 5.1.1 原假设和备择假设
- 5.1.2 拒绝域、检验函数和检验统计量
- 5.1.3 两类错误与功效函数
- 5.1.4 检验水平和控制犯第一类错误概率的原则
- 5.1.5 求解假设检验问题的一般步骤
- 5.1.6 假设检验发展简史

## 5.1.1 原假设和备择假设

参数估计和假设检验是统计推断的两个主要形式。关于参数的点估计和区间估计的问题已在前两章中讨论，本章讨论假设检验问题。

## 5.1.1 原假设和备择假设

参数估计和假设检验是统计推断的两个主要形式。关于参数的点估计和区间估计的问题已在前两章中讨论，本章讨论假设检验问题。

对总体分布的某些特征所作的假设称为**统计假设**。

## 5.1.1 原假设和备择假设

参数估计和假设检验是统计推断的两个主要形式。关于参数的点估计和区间估计的问题已在前两章中讨论，本章讨论假设检验问题。

对总体分布的某些特征所作的假设称为**统计假设**。

使用样本对所作出的统计假设进行推断的方法和过程称为**假设检验**。

## 5.1.1 原假设和备择假设

根据总体分布形式是否已知，假设检验问题大致分为两大类。

## 5.1.1 原假设和备择假设

根据总体分布形式是否已知，假设检验问题大致分为两大类。

(1) **参数型假设检验**：即总体的分布形式已知（如正态、指数、二项分布等），总体分布依赖于未知参数（或参数向量） $\theta$ ，要检验的是有关未知参数的假设。

## 5.1.1 原假设和备择假设

根据总体分布形式是否已知，假设检验问题大致分为两大类。

(1) **参数型假设检验**：即总体的分布形式已知（如正态、指数、二项分布等），总体分布依赖于未知参数（或参数向量） $\theta$ ，要检验的是有关未知参数的假设。

例如，总体  $X \sim N(a, \sigma^2)$ ， $a$ 未知，检验

## 5.1.1 原假设和备择假设

根据总体分布形式是否已知，假设检验问题大致分为两大类。

(1) **参数型假设检验**：即总体的分布形式已知（如正态、指数、二项分布等），总体分布依赖于未知参数（或参数向量） $\theta$ ，要检验的是有关未知参数的假设。

例如，总体  $X \sim N(a, \sigma^2)$ ， $a$  未知，检验

$$H_0 : a = a_0 \text{ vs } H_1 : a \neq a_0 \text{ 或 } H_0 : a \leq a_0 \text{ vs } H_1 : a > a_0.$$

## 5.1.1 原假设和备择假设

(2) **非参数型假设检验**：如果总体分布形式未知，此时就需要有一种与总体分布族的具体数学形式无关的统计方法，称为非参数方法。

## 5.1.1 原假设和备择假设

(2) **非参数型假设检验**：如果总体分布形式未知，此时就需要有一种与总体分布族的具体数学形式无关的统计方法，称为非参数方法。

例如，检验一批数据是否来自某个已知的总体，就属于这类问题（第六章）。

## 5.1.1 原假设和备择假设

(2) **非参数型假设检验**：如果总体分布形式未知，此时就需要有一种与总体分布族的具体数学形式无关的统计方法，称为非参数方法。

例如，检验一批数据是否来自某个已知的总体，就属于这类问题（第六章）。

本章我们将主要讨论参数型假设检验问题。

## 5.1.1 原假设和备择假设

### 例 (5.1.1)

某工厂生产的一批产品要出厂上市销售。按规定次品率 $p$ 不得超过0.01, 今在其中随机不放回抽取100件, 经检验有3件次品, 问这批产品可否出厂?

关于这个问题, 在面前存在两种可能性:

## 5.1.1 原假设和备择假设

### 例 (5.1.1)

某工厂生产的一批产品要出厂上市销售。按规定次品率 $p$ 不得超过0.01, 今在其中随机不放回抽取100件, 经检验有3件次品, 问这批产品可否出厂?

关于这个问题, 在面前存在两种可能性:

$$\text{甲} : 0 < p \leq 0.01; \quad \text{乙} : 0.01 < p < 1.$$

要通过从这批产品中抽样来决定甲、乙两种可能性中哪个成立。

## 5.1.1 原假设和备择假设

这个问题常以下述方式提出，引进一个“假设”

$$H_0 : 0 < p \leq 0.01,$$

称为**原假设**或零假设。另一个可能是

## 5.1.1 原假设和备择假设

这个问题常以下述方式提出，引进一个“假设”

$$H_0 : 0 < p \leq 0.01,$$

称为**原假设**或零假设。另一个可能是

$$H_1 : 0.01 < p < 1,$$

称为**对立假设**或**备择假设**。

## 5.1.1 原假设和备择假设

这个问题常以下述方式提出，引进一个“假设”

$$H_0 : 0 < p \leq 0.01,$$

称为**原假设**或零假设。另一个可能是

$$H_1 : 0.01 < p < 1,$$

称为**对立假设**或**备择假设**。

目的是通过样本决定是否拒绝 $H_0$ 。

## 5.1.1 原假设和备择假设

可以形象地把问题写成

## 5.1.1 原假设和备择假设

可以形象地把问题写成

$$H_0 : 0 < p \leq 0.01 \text{ vs } H_1 : 0.01 < p < 1.$$

## 5.1.1 原假设和备择假设

可以形象地把问题写成

$$H_0 : 0 < p \leq 0.01 \text{ vs } H_1 : 0.01 < p < 1.$$

这里需要注意的是这个提法中将 $H_0$ 放在中心位置，它是检验的对象。

$H_0$ 和 $H_1$ 的位置不可颠倒。

## 5.1.1 原假设和备择假设

可以形象地把问题写成

$$H_0 : 0 < p \leq 0.01 \text{ vs } H_1 : 0.01 < p < 1.$$

这里需要注意的是这个提法中将 $H_0$ 放在中心位置，它是检验的对象。

$H_0$ 和 $H_1$ 的位置不可颠倒。

一般我们把有把握、不会轻易否定的命题放在“原假设”的位置。

## 5.1.1 原假设和备择假设

基于上述讨论，将总体参数的假设检验问题一般化：

## 5.1.1 原假设和备择假设

基于上述讨论，将总体参数的假设检验问题一般化：

设有参数分布族 $\{f(x, \theta) : \theta \in \Theta\}$ ，此处 $\Theta$ 为参数空间。 $X_1, \dots, X_n$ 是从上述分布族中抽取的简单样本。

## 5.1.1 原假设和备择假设

在参数假设检验问题中，感兴趣的是 $\theta$ 是否属于参数空间 $\Theta$ 的某个非空真子集 $\Theta_0$ 。则命题

## 5.1.1 原假设和备择假设

在参数假设检验问题中，感兴趣的是 $\theta$ 是否属于参数空间 $\Theta$ 的某个非空真子集 $\Theta_0$ 。则命题

$$H_0 : \theta \in \Theta_0$$

称为**零假设**或**原假设**。

## 5.1.1 原假设和备择假设

在参数假设检验问题中，感兴趣的是 $\theta$ 是否属于参数空间 $\Theta$ 的某个非空真子集 $\Theta_0$ 。则命题

$$H_0 : \theta \in \Theta_0$$

称为**零假设**或**原假设**。

其确切含义是：存在一个 $\theta_0 \in \Theta_0$ ，使得 $X$ 的分布为 $f(x, \theta_0)$ 。

## 5.1.1 原假设和备择假设

记  $\Theta_1 \subseteq \Theta - \Theta_0$ , 则命题

## 5.1.1 原假设和备择假设

记  $\Theta_1 \subseteq \Theta - \Theta_0$ ，则命题

$$H_1 : \theta \in \Theta_1$$

称为  $H_0$  的**对立假设**或**备择假设**。

## 5.1.1 原假设和备择假设

记  $\Theta_1 \subseteq \Theta - \Theta_0$ ，则命题

$$H_1 : \theta \in \Theta_1$$

称为  $H_0$  的**对立假设**或**备择假设**。

则假设检验问题表示为

## 5.1.1 原假设和备择假设

记  $\Theta_1 \subseteq \Theta - \Theta_0$ ，则命题

$$H_1 : \theta \in \Theta_1$$

称为  $H_0$  的**对立假设**或**备择假设**。

则假设检验问题表示为

$$H_0 : \theta \in \Theta_0 \text{ vs } H_1 : \theta \in \Theta_1.$$

## 5.1.1 原假设和备择假设

记  $\Theta_1 \subseteq \Theta - \Theta_0$ ，则命题

$$H_1 : \theta \in \Theta_1$$

称为  $H_0$  的**对立假设**或**备择假设**。

则假设检验问题表示为

$$H_0 : \theta \in \Theta_0 \text{ vs } H_1 : \theta \in \Theta_1.$$

但要注意一点，零假设总写在左边，作为中心位置；对立假设写在右边，作为陪衬地位。若没有对立假设，检验问题就无完整的意义。

## 5.1.1 原假设和备择假设

在上式中，若 $\Theta_0$ 和 $\Theta_1$ 只包含参数空间 $\Theta$ 中的一个点，则称为**简单假设**；  
否则，称为**复合假设**。

## 5.1.1 原假设和备择假设

在上式中，若 $\Theta_0$ 和 $\Theta_1$ 只包含参数空间 $\Theta$ 中的一个点，则称为**简单假设**；否则，称为**复合假设**。

例如，样本抽自 $N(a, \sigma^2)$ ， $\sigma^2$ 已知，则参数空间为

## 5.1.1 原假设和备择假设

在上式中，若 $\Theta_0$ 和 $\Theta_1$ 只包含参数空间 $\Theta$ 中的一个点，则称为**简单假设**；否则，称为**复合假设**。

例如，样本取自 $N(a, \sigma^2)$ ， $\sigma^2$ 已知，则参数空间为

$$\Theta = \{a : -\infty < a < \infty\}.$$

## 5.1.1 原假设和备择假设

令

$$H_0 : a = a_0 \text{ vs } H_1 : a \neq a_0,$$

## 5.1.1 原假设和备择假设

令

$$H_0 : a = a_0 \text{ vs } H_1 : a \neq a_0,$$

则 $H_0$ 为简单假设， $H_1$ 为复合假设。

## 5.1.1 原假设和备择假设

令

$$H_0 : a = a_0 \text{ vs } H_1 : a \neq a_0,$$

则 $H_0$ 为简单假设,  $H_1$ 为复合假设。

再如, 对假设

$$H'_0 : a \leq a_0 \text{ vs } H'_1 : a > a_0,$$

## 5.1.1 原假设和备择假设

令

$$H_0 : a = a_0 \text{ vs } H_1 : a \neq a_0,$$

则 $H_0$ 为简单假设,  $H_1$ 为复合假设。

再如, 对假设

$$H'_0 : a \leq a_0 \text{ vs } H'_1 : a > a_0,$$

则零假设 $H'_0$ 和对立假设 $H'_1$ 皆为复合假设。

## 5.1.2 拒绝域、检验函数和检验统计量

### 例 (5.1.2)

设  $\mathbf{X} = (X_1, \dots, X_n)$  为从正态总体  $N(a, \sigma^2)$  中抽取的随机样本，其中  $\sigma^2$  已知。考虑检验问题

$$H_0 : a = a_0 \text{ vs } H_1 : a \neq a_0,$$

此处  $a_0$  为给定的常数。

## 5.1.2 拒绝域、检验函数和检验统计量

这种检验问题的一种直观解决方法是：先求 $a$ 的一个估计量。

## 5.1.2 拒绝域、检验函数和检验统计量

这种检验问题的一种直观解决方法是：先求 $a$ 的一个估计量。

已知 $\bar{X}$ 是 $a$ 的一个优良估计（无偏，强相合）。

## 5.1.2 拒绝域、检验函数和检验统计量

这种检验问题的一种直观解决方法是：先求 $a$ 的一个估计量。

已知 $\bar{X}$ 是 $a$ 的一个优良估计（无偏，强相合）。

若 $|\bar{X} - a_0|$ 较大，就倾向于拒绝 $H_0$ ；

## 5.1.2 拒绝域、检验函数和检验统计量

这种检验问题的一种直观解决方法是：先求 $a$ 的一个估计量。

已知 $\bar{X}$ 是 $a$ 的一个优良估计（无偏，强相合）。

若 $|\bar{X} - a_0|$ 较大，就倾向于拒绝 $H_0$ ；

反之，如果 $|\bar{X} - a_0|$ 较小，就认为抽样结果与 $H_0$ 相接近，因而倾向于接受 $H_0$ 。

## 5.1.2 拒绝域、检验函数和检验统计量

再确定一个数 $A$ （临界值）。由 $X_1, \dots, X_n$ 算出样本均值 $\bar{X}$ 。

## 5.1.2 拒绝域、检验函数和检验统计量

再确定一个数 $A$ （临界值）。由 $X_1, \dots, X_n$ 算出样本均值 $\bar{X}$ 。

当 $|\bar{X} - a_0| > A$ 时就拒绝 $H_0$ ；当 $|\bar{X} - a_0| \leq A$ 时就接受 $H_0$ 。

## 5.1.2 拒绝域、检验函数和检验统计量

再确定一个数 $A$  (临界值)。由 $X_1, \dots, X_n$ 算出样本均值 $\bar{X}$ 。

当 $|\bar{X} - a_0| > A$ 时就拒绝 $H_0$ ；当 $|\bar{X} - a_0| \leq A$ 时就接受 $H_0$ 。

称

$$D = \{ \mathbf{X} = (X_1, \dots, X_n) : |\bar{X} - a_0| > A \}$$

为拒绝域，或否定域。

## 5.1.2 拒绝域、检验函数和检验统计量

再确定一个数 $A$ （临界值）。由 $X_1, \dots, X_n$ 算出样本均值 $\bar{X}$ 。

当 $|\bar{X} - a_0| > A$ 时就拒绝 $H_0$ ；当 $|\bar{X} - a_0| \leq A$ 时就接受 $H_0$ 。

称

$$D = \{ \mathbf{X} = (X_1, \dots, X_n) : |\bar{X} - a_0| > A \}$$

为拒绝域，或否定域。

即**拒绝域**是由样本空间 $\mathcal{X}$ 中一切使 $|\bar{X} - a_0| > A$ 的那些样本 $\mathbf{X} = (X_1, \dots, X_n)$ 构成。

## 5.1.2 拒绝域、检验函数和检验统计量

有了拒绝域，等价于将样本空间 $\mathcal{X}$ 分成不相交的两部分

$$D \text{ 和 } \bar{D} = \mathcal{X} - D.$$

## 5.1.2 拒绝域、检验函数和检验统计量

有了拒绝域，等价于将样本空间 $\mathcal{X}$ 分成不相交的两部分

$$D \text{ 和 } \bar{D} = \mathcal{X} - D.$$

一旦有了样本 $X$ ，当 $X \in D$ 时，就拒绝 $H_0$ ；

## 5.1.2 拒绝域、检验函数和检验统计量

有了拒绝域，等价于将样本空间 $\mathcal{X}$ 分成不相交的两部分

$$D \text{ 和 } \bar{D} = \mathcal{X} - D.$$

一旦有了样本 $X$ ，当 $X \in D$ 时，就拒绝 $H_0$ ；

当 $X \in \bar{D}$ 时，就接受 $H_0$ 。称 $\bar{D}$ 为接受域。

## 5.1.2 拒绝域、检验函数和检验统计量

有了拒绝域，等价于将样本空间 $\mathcal{X}$ 分成不相交的两部分

$$D \text{ 和 } \bar{D} = \mathcal{X} - D.$$

一旦有了样本 $X$ ，当 $X \in D$ 时，就拒绝 $H_0$ ；

当 $X \in \bar{D}$ 时，就接受 $H_0$ 。称 $\bar{D}$ 为接受域。

只要 $A$ 定下来了，则拒绝域（或接受域）也就确定了。

## 5.1.2 拒绝域、检验函数和检验统计量

因此，此问题中的检验可视为如下一种法则：

## 5.1.2 拒绝域、检验函数和检验统计量

因此，此问题中的检验可视为如下一种法则：

$$T : \begin{cases} \text{当} |\bar{X} - a_0| > A \text{时, 拒绝} H_0, \\ \text{当} |\bar{X} - a_0| \leq A \text{时, 接受} H_0. \end{cases}$$

## 5.1.2 拒绝域、检验函数和检验统计量

因此，此问题中的检验可视为如下一种法则：

$$T : \begin{cases} \text{当} |\bar{X} - a_0| > A \text{时, 拒绝} H_0, \\ \text{当} |\bar{X} - a_0| \leq A \text{时, 接受} H_0. \end{cases}$$

上式中的 $T$ 给出了一种法则：一旦有了样本，就可以在是否拒绝 $H_0$ 这两个结论中选一个。

## 5.1.2 拒绝域、检验函数和检验统计量

因此，此问题中的检验可视为如下一种法则：

$$T : \begin{cases} \text{当} |\bar{X} - a_0| > A \text{时, 拒绝} H_0, \\ \text{当} |\bar{X} - a_0| \leq A \text{时, 接受} H_0. \end{cases}$$

上式中的 $T$ 给出了一种法则：一旦有了样本，就可以在是否拒绝 $H_0$ 这两个结论中选一个。

称这样一种法则 $T$ 为检验问题 $H_0 : a = a_0$  vs  $H_1 : a \neq a_0$ 的一个检验。

## 5.1.2 拒绝域、检验函数和检验统计量

为便于数学上的处理，现引入检验函数 $\varphi(x)$ 的概念。 $\varphi(x)$ 与检验 $T$ 是一一对应的。

## 5.1.2 拒绝域、检验函数和检验统计量

为便于数学上的处理，现引入检验函数 $\varphi(x)$ 的概念。 $\varphi(x)$ 与检验 $T$ 是一一对应的。

在例5.1.2中，令

## 5.1.2 拒绝域、检验函数和检验统计量

为便于数学上的处理，现引入检验函数 $\varphi(\boldsymbol{x})$ 的概念。 $\varphi(\boldsymbol{x})$ 与检验 $T$ 是一一对应的。

在例5.1.2中，令

$$\varphi(\boldsymbol{x}) = \begin{cases} 1, & |\bar{x} - a_0| > A, \\ 0, & |\bar{x} - a_0| \leq A. \end{cases} \quad (1)$$

## 5.1.2 拒绝域、检验函数和检验统计量

为便于数学上的处理，现引入检验函数 $\varphi(\boldsymbol{x})$ 的概念。 $\varphi(\boldsymbol{x})$ 与检验 $T$ 是一一对应的。

在例5.1.2中，令

$$\varphi(\boldsymbol{x}) = \begin{cases} 1, & |\bar{x} - a_0| > A, \\ 0, & |\bar{x} - a_0| \leq A. \end{cases} \quad (1)$$

### 定义 (5.1.1 (检验函数))

由上式 (1) 给出的**检验函数** $\varphi(\boldsymbol{x})$ 是定义在样本空间 $\mathcal{X}$ 上，取值于 $[0, 1]$ 上的函数。它表示当有了样本 $\boldsymbol{X}$ 后，拒绝 $H_0$ 的概率。

## 5.1.2 拒绝域、检验函数和检验统计量

由上述定义可见，若  $\varphi(\mathbf{x}) = 1$ ，则以概率为 1 拒绝  $H_0$ ；

## 5.1.2 拒绝域、检验函数和检验统计量

由上述定义可见，若  $\varphi(\boldsymbol{x}) = 1$ ，则以概率为 1 拒绝  $H_0$ ；

若  $\varphi(\boldsymbol{x}) = 0$ ，则以概率为 0 拒绝  $H_0$ （即以概率为 1 接受  $H_0$ ）。

## 5.1.2 拒绝域、检验函数和检验统计量

由上述定义可见，若  $\varphi(\boldsymbol{x}) = 1$ ，则以概率为 1 拒绝  $H_0$ ；

若  $\varphi(\boldsymbol{x}) = 0$ ，则以概率为 0 拒绝  $H_0$ （即以概率为 1 接受  $H_0$ ）。

若  $\varphi(\boldsymbol{x})$  只取 0, 1 这两个值，则称这种检验为**非随机化检验**。

## 5.1.2 拒绝域、检验函数和检验统计量

由上述定义可见，若  $\varphi(\boldsymbol{x}) = 1$ ，则以概率为 1 拒绝  $H_0$ ；

若  $\varphi(\boldsymbol{x}) = 0$ ，则以概率为 0 拒绝  $H_0$ （即以概率为 1 接受  $H_0$ ）。

若  $\varphi(\boldsymbol{x})$  只取 0, 1 这两个值，则称这种检验为**非随机化检验**。

此时，拒绝域也可用检验函数表示如下

## 5.1.2 拒绝域、检验函数和检验统计量

由上述定义可见，若  $\varphi(\boldsymbol{x}) = 1$ ，则以概率为 1 拒绝  $H_0$ ；

若  $\varphi(\boldsymbol{x}) = 0$ ，则以概率为 0 拒绝  $H_0$ （即以概率为 1 接受  $H_0$ ）。

若  $\varphi(\boldsymbol{x})$  只取 0, 1 这两个值，则称这种检验为**非随机化检验**。

此时，拒绝域也可用检验函数表示如下

$$D = \{\boldsymbol{X} = (X_1, \dots, X_n) : \varphi(\boldsymbol{X}) = 1\}.$$

## 5.1.2 拒绝域、检验函数和检验统计量

若对某些样本  $\mathbf{X}$ ，有  $0 < \varphi(\mathbf{x}) < 1$ ，则称  $\varphi(\mathbf{x})$  为**随机化检验**。

## 5.1.2 拒绝域、检验函数和检验统计量

若对某些样本 $\mathbf{X}$ ，有 $0 < \varphi(\mathbf{x}) < 1$ ，则称 $\varphi(\mathbf{x})$ 为**随机化检验**。

如在例5.1.1中，令 $X_i = 1$ ，若第 $i$ 个产品为次品；否则为0；

$i = 1, \dots, n$ 。

## 5.1.2 拒绝域、检验函数和检验统计量

若对某些样本 $\mathbf{X}$ ，有 $0 < \varphi(\mathbf{x}) < 1$ ，则称 $\varphi(\mathbf{x})$ 为**随机化检验**。

如在例5.1.1中，令 $X_i = 1$ ，若第 $i$ 个产品为次品；否则为0；

$i = 1, \dots, n$ 。

设 $\mathbf{X} = (X_1, \dots, X_n)$ 为样本，令 $T(\mathbf{X}) = \sum_{i=1}^{100} X_i$ 。

## 5.1.2 拒绝域、检验函数和检验统计量

当  $T(\mathbf{X}) < c$  时，认为这批产品合格，接受  $H_0$ ；

## 5.1.2 拒绝域、检验函数和检验统计量

当  $T(\mathbf{X}) < c$  时，认为这批产品合格，接受  $H_0$ ；

当  $T(\mathbf{X}) > c$  时，认为不合格，拒绝  $H_0$ 。

## 5.1.2 拒绝域、检验函数和检验统计量

当  $T(\mathbf{X}) < c$  时，认为这批产品合格，接受  $H_0$ ；

当  $T(\mathbf{X}) > c$  时，认为不合格，拒绝  $H_0$ 。

当  $T(\mathbf{X}) = c$  时，若规定拒绝  $H_0$ ，厂方觉得被拒绝的可能性大了，吃亏了。

## 5.1.2 拒绝域、检验函数和检验统计量

当  $T(\mathbf{X}) < c$  时，认为这批产品合格，接受  $H_0$ ；

当  $T(\mathbf{X}) > c$  时，认为不合格，拒绝  $H_0$ 。

当  $T(\mathbf{X}) = c$  时，若规定拒绝  $H_0$ ，厂方觉得被拒绝的可能性大了，吃亏了。

反之，若接受  $H_0$ ，买方（商店）接受不合格产品的可能性大了，也觉得吃亏。

## 5.1.2 拒绝域、检验函数和检验统计量

在双方僵持不下的情况下，下列折中方案是双方都可以接受的：

## 5.1.2 拒绝域、检验函数和检验统计量

在双方僵持不下的情况下，下列折中方案是双方都可以接受的：

定下一个数  $0 < r < 1$ ，规定当  $T(\mathbf{X}) = c$  时，作一次成功概率为  $r$  的随机试验，根据试验结果来决定拒绝还是接受这批产品。

## 5.1.2 拒绝域、检验函数和检验统计量

如取  $r = 1/2$ ，则可通过掷一枚均匀硬币来决定。

## 5.1.2 拒绝域、检验函数和检验统计量

如取  $r = 1/2$ ，则可通过掷一枚均匀硬币来决定。

规定若出现正面则拒绝  $H_0$ ，否则接受  $H_0$ 。

## 5.1.2 拒绝域、检验函数和检验统计量

如取  $r = 1/2$ ，则可通过掷一枚均匀硬币来决定。

规定若出现正面则拒绝  $H_0$ ，否则接受  $H_0$ 。

这样，当出现  $T(\mathbf{X}) = c$ ，双方都有1/2的可能，作出对自己不利的决定，双方都觉得合理，可以接受。

## 5.1.2 拒绝域、检验函数和检验统计量

如取  $r = 1/2$ ，则可通过掷一枚均匀硬币来决定。

规定若出现正面则拒绝  $H_0$ ，否则接受  $H_0$ 。

这样，当出现  $T(\mathbf{X}) = c$ ，双方都有1/2的可能，作出对自己不利的决定，双方都觉得合理，可以接受。

如果取  $r = 1/3$ ，试验可以通过在装有 2 个白球和 1 个黑球的盒子中随机摸球来决定。

## 5.1.2 拒绝域、检验函数和检验统计量

如取  $r = 1/2$ ，则可通过掷一枚均匀硬币来决定。

规定若出现正面则拒绝  $H_0$ ，否则接受  $H_0$ 。

这样，当出现  $T(\mathbf{X}) = c$ ，双方都有1/2的可能，作出对自己不利的决定，双方都觉得合理，可以接受。

如果取  $r = 1/3$ ，试验可以通过在装有 2 个白球和 1 个黑球的盒子中随机摸球来决定。

规定若摸到黑球（发生的概率为1/3）则拒绝  $H_0$ ，若摸到白球（发生的概率为2/3）则接受  $H_0$ 。

## 5.1.2 拒绝域、检验函数和检验统计量

这种随机化检验函数可表示为

## 5.1.2 拒绝域、检验函数和检验统计量

这种随机化检验函数可表示为

$$\varphi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) > c, \\ r, & T(\mathbf{x}) = c, \\ 0, & T(\mathbf{x}) < c. \end{cases} \quad (2)$$

## 5.1.2 拒绝域、检验函数和检验统计量

这种随机化检验函数可表示为

$$\varphi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) > c, \\ r, & T(\mathbf{x}) = c, \\ 0, & T(\mathbf{x}) < c. \end{cases} \quad (2)$$

在式 (1) 中的  $A$  或 (2) 中的  $c$  称为**临界值** (critical value), **检验统计量**是  $T(\mathbf{X}) = \sum_{i=1}^{100} X_i$ 。

## 5.1.2 拒绝域、检验函数和检验统计量

要定下  $A$  或  $c$  的值需要确定检验统计量的分布。

## 5.1.2 拒绝域、检验函数和检验统计量

要定下  $A$  或  $c$  的值需要确定检验统计量的分布。

此例中检验统计量是样本均值  $\bar{X}$ 。

## 5.1.2 拒绝域、检验函数和检验统计量

要定下  $A$  或  $c$  的值需要确定检验统计量的分布。

此例中检验统计量是样本均值  $\bar{X}$ 。

确定检验统计量的分布是解决假设检验问题的关键。

## 5.1.2 拒绝域、检验函数和检验统计量

要定下  $A$  或  $c$  的值需要确定检验统计量的分布。

此例中检验统计量是样本均值  $\bar{X}$ 。

确定检验统计量的分布是解决假设检验问题的关键。

当检验统计量的精确分布很难找到时，若其极限分布比较简单，可用极限分布代替精确分布，获得假设检验问题的近似解。

## 5.1.3 两类错误与功效函数

统计推断是以样本为依据的，由于样本的随机性，不能保证统计推断方法的绝对正确性，而只能以一定的概率去保证这种推断的可靠性。

## 5.1.3 两类错误与功效函数

统计推断是以样本为依据的，由于样本的随机性，不能保证统计推断方法的绝对正确性，而只能以一定的概率去保证这种推断的可靠性。

因此，在假设检验问题中可能出现下列**两类错误**：

## 5.1.3 两类错误与功效函数

统计推断是以样本为依据的，由于样本的随机性，不能保证统计推断方法的绝对正确性，而只能以一定的概率去保证这种推断的可靠性。

因此，在假设检验问题中可能出现下列**两类错误**：

(1) 零假设  $H_0$  本来是对的，由于样本的随机性，观察值落入拒绝域  $D$ ，错误地将  $H_0$  拒绝了，称为**弃真**。这时犯的**错误称为第一类错误 (type I error)**。

## 5.1.3 两类错误与功效函数

(2) 零假设  $H_0$  本来不对，由于样本的随机性，观察值落入接受域  $\bar{D}$ ，错误地将  $H_0$  接受了，称为**取伪**。这时犯的**错误称为第二类错误**（type II error）。

## 5.1.3 两类错误与功效函数

如在例5.1.1中，假设检验为  $H_0 : 0 < p \leq 0.01$  vs  $H_1 : 0.01 < p < 1$ ，确定了非随机化检验如下：

## 5.1.3 两类错误与功效函数

如在例5.1.1中，假设检验为  $H_0 : 0 < p \leq 0.01$  vs  $H_1 : 0.01 < p < 1$ ，确定了非随机化检验如下：

$$\varphi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) > 3, \\ 0, & T(\mathbf{x}) \leq 3. \end{cases}$$

### 5.1.3 两类错误与功效函数

如在例5.1.1中，假设检验为  $H_0 : 0 < p \leq 0.01$  vs  $H_1 : 0.01 < p < 1$ ，确定了非随机化检验如下：

$$\varphi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) > 3, \\ 0, & T(\mathbf{x}) \leq 3. \end{cases}$$

如果总体的真实次品率为  $p = 0.005 < 0.01$ （原假设为真），由于样本的随机性，抽样结果显示  $T(\mathbf{x}) = 5$ ，即样本落入拒绝域，这时犯第一类错误（弃真）。

## 5.1.3 两类错误与功效函数

但也有可能总体的真实次品率  $p = 0.03 > 0.01$ （备择假设为真），由于样本的随机性，抽样结果显示  $T(x) = 1$ ，即样本落入了接受域，这时犯第二类错误（取伪）。

## 5.1.3 两类错误与功效函数

应当注意，在每一具体场合，只会犯两类错误中的一个。

## 5.1.3 两类错误与功效函数

应当注意，在每一具体场合，只会犯两类错误中的一个。

当检验确定后，犯两类错误的概率也就确定了。我们希望犯两类错误的概率越小越好，但这一点很难做到。

## 5.1.3 两类错误与功效函数

应当注意，在每一具体场合，只会犯两类错误中的一个。

当检验确定后，犯两类错误的概率也就确定了。我们希望犯两类错误的概率越小越好，但这一点很难做到。

在样本大小  $n$  固定的前提下，二者不可兼得。

## 5.1.3 两类错误与功效函数

应当注意，在每一具体场合，只会犯两类错误中的一个。

当检验确定后，犯两类错误的概率也就确定了。我们希望犯两类错误的概率越小越好，但这一点很难做到。

在样本大小  $n$  固定的前提下，二者不可兼得。

那么，怎么去计算犯两类错误的概率呢？

## 5.1.3 两类错误与功效函数

### 定义 (5.1.2 功效函数)

设 $\varphi(\boldsymbol{x})$ 是 $H_0 : \theta \in \Theta_0$  vs  $H_1 : \theta \in \Theta_1$ 的一个检验函数, 则

## 5.1.3 两类错误与功效函数

### 定义 (5.1.2 功效函数)

设 $\varphi(\mathbf{x})$ 是 $H_0 : \theta \in \Theta_0$  vs  $H_1 : \theta \in \Theta_1$ 的一个检验函数, 则

$$\beta_\varphi(\theta) = P_\theta\{\text{用检验}\varphi\text{拒绝了}\mathcal{H}_0\} = E_\theta[\varphi(\mathbf{X})], \quad \theta \in \Theta$$

## 5.1.3 两类错误与功效函数

### 定义 (5.1.2 功效函数)

设 $\varphi(\mathbf{x})$ 是 $H_0 : \theta \in \Theta_0$  vs  $H_1 : \theta \in \Theta_1$ 的一个检验函数, 则

$$\beta_\varphi(\theta) = P_\theta\{\text{用检验}\varphi\text{拒绝了}\ H_0\} = E_\theta[\varphi(\mathbf{X})], \quad \theta \in \Theta$$

称为 $\varphi$ 的**功效函数** (power function), 也称为**效函数**或**势函数**。

## 5.1.3 两类错误与功效函数

若 $\varphi(x)$ 为非随机化检验，拒绝域为 $D$ ，则

## 5.1.3 两类错误与功效函数

若 $\varphi(x)$ 为非随机化检验，拒绝域为 $D$ ，则

$$\beta_{\varphi}(\theta) = P_{\theta}\{\mathbf{X} = (X_1, \dots, X_n) \in D\}, \theta \in \Theta.$$

## 5.1.3 两类错误与功效函数

若 $\varphi(x)$ 为非随机化检验，拒绝域为 $D$ ，则

$$\beta_{\varphi}(\theta) = P_{\theta}\{\mathbf{X} = (X_1, \dots, X_n) \in D\}, \theta \in \Theta.$$

因此功效函数表示当参数为 $\theta$ 时，拒绝 $H_0$ 的概率。

## 5.1.3 两类错误与功效函数

对例5.1.1, 当检验函数为

## 5.1.3 两类错误与功效函数

对例5.1.1, 当检验函数为

$$\varphi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) > c, \\ r, & T(\mathbf{x}) = c, \\ 0, & T(\mathbf{x}) < c. \end{cases}$$

时, 利用 $T(\mathbf{X}) = \sum_{i=1}^n X_i \sim B(n, \theta)$ ,  $0 < \theta < 1$ 可知检验的功效函数为

## 5.1.3 两类错误与功效函数

对例5.1.1, 当检验函数为

$$\varphi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) > c, \\ r, & T(\mathbf{x}) = c, \\ 0, & T(\mathbf{x}) < c. \end{cases}$$

时, 利用 $T(\mathbf{X}) = \sum_{i=1}^n X_i \sim B(n, \theta)$ ,  $0 < \theta < 1$ 可知检验的功效函数为

$$\beta_{\varphi}(\theta) = E_{\theta}[\varphi(\mathbf{X})] = P\left(\sum_{i=1}^{100} X_i > c\right) + rP\left(\sum_{i=1}^{100} X_i = c\right).$$

## 5.1.3 两类错误与功效函数

知道了检验 $\varphi(\boldsymbol{x})$ 的功效函数 $\beta_{\varphi}(\theta)$ 后，就可以计算犯两类错误的概率。

## 5.1.3 两类错误与功效函数

知道了检验 $\varphi(x)$ 的功效函数 $\beta_\varphi(\theta)$ 后，就可以计算犯两类错误的概率。

若以  $\alpha_\varphi^*(\theta)$  和  $\gamma_\varphi^*(\theta)$  分别记犯第一、第二类错误的概率，则犯第一类错误的概率可以表示为

### 5.1.3 两类错误与功效函数

知道了检验 $\varphi(x)$ 的功效函数 $\beta_\varphi(\theta)$ 后，就可以计算犯两类错误的概率。

若以 $\alpha_\varphi^*(\theta)$ 和 $\gamma_\varphi^*(\theta)$ 分别记犯第一、第二类错误的概率，则犯第一类错误的概率可以表示为

$$\alpha_\varphi^*(\theta) = \begin{cases} \beta_\varphi(\theta), & \theta \in \Theta_0, \\ 0, & \theta \in \Theta_1. \end{cases}$$

## 5.1.3 两类错误与功效函数

犯第二类错误的概率可以表示为

## 5.1.3 两类错误与功效函数

犯第二类错误的概率可以表示为

$$\gamma_{\varphi}^*(\theta) = \begin{cases} 0, & \theta \in \Theta_0, \\ 1 - \beta_{\varphi}(\theta), & \theta \in \Theta_1. \end{cases}$$

## 5.1.3 两类错误与功效函数

如在例5.1.1中,  $\Theta = (0, 1)$ , 则检验问题

## 5.1.3 两类错误与功效函数

如在例5.1.1中,  $\Theta = (0, 1)$ , 则检验问题

$$H_0 : 0 < \theta \leq 0.01 \text{ vs } H_1 : 0.01 < \theta < 1$$

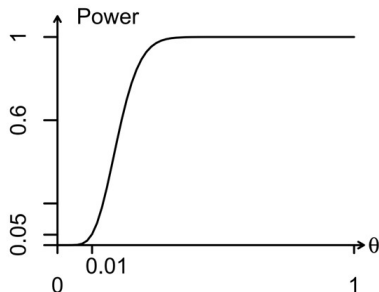
的一个较好的功效函数的形状如下图所示

## 5.1.3 两类错误与功效函数

如在例5.1.1中,  $\Theta = (0, 1)$ , 则检验问题

$$H_0 : 0 < \theta \leq 0.01 \quad vs \quad H_1 : 0.01 < \theta < 1$$

的一个较好的功效函数的形状如下图所示



## 5.1.3 两类错误与功效函数

一个好的检验  $\varphi(x)$ ，犯两类错误的概率都应较小。

## 5.1.3 两类错误与功效函数

一个好的检验  $\varphi(x)$ ，犯两类错误的概率都应较小。

也就是功效函数  $\beta_\varphi(\theta)$  在  $\Theta_0$  中应尽可能地小，在  $\Theta_1$  中应尽可能地大。

## 5.1.3 两类错误与功效函数

一个好的检验  $\varphi(x)$ ，犯两类错误的概率都应较小。

也就是功效函数  $\beta_\varphi(\theta)$  在  $\Theta_0$  中应尽可能地小，在  $\Theta_1$  中应尽可能地大。

犯两类错误的概率完全由功效函数决定。如果两个检验有同一功效函数，则此两个检验在性质上也完全相同。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

要使犯第一类错误的概率减小，就要缩小拒绝域，使接受域增大，这必然导致犯第二类错误概率增大，反之亦然。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

要使犯第一类错误的概率减小，就要缩小拒绝域，使接受域增大，这必然导致犯第二类错误概率增大，反之亦然。

因此，Neyman 和 Pearson 提出了一条**原则**，就是限制犯第一类错误概率的原则。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

要使犯第一类错误的概率减小，就要缩小拒绝域，使接受域增大，这必然导致犯第二类错误概率增大，反之亦然。

因此，Neyman 和 Pearson 提出了一条**原则**，就是限制犯第一类错误概率的原则。

即在保证犯第一类错误的概率不超过指定数值  $\alpha$  ( $0 < \alpha < 1$ ，通常取较小的数) 的检验中，寻找犯第二类错误概率尽可能小的检验。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

若记

$$S_\alpha = \{\varphi : \beta_\varphi(\theta) \leq \alpha, \theta \in \Theta_0\},$$

## 5.1.4 检验水平和控制犯第一类错误概率的原则

若记

$$S_\alpha = \{\varphi : \beta_\varphi(\theta) \leq \alpha, \theta \in \Theta_0\},$$

$S_\alpha$ 表示由所有犯第一类错误的概率都不超过 $\alpha$ 的检验函数构成的类。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

若记

$$S_\alpha = \{\varphi : \beta_\varphi(\theta) \leq \alpha, \theta \in \Theta_0\},$$

$S_\alpha$ 表示由所有犯第一类错误的概率都不超过 $\alpha$ 的检验函数构成的类。

只考虑 $S_\alpha$ 中的检验，在 $S_\alpha$ 中挑选“犯第二类错误的概率尽可能小的检验”，这种法则称为**控制犯第一类错误概率的法则**。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

根据 Neyman-Pearson 原则，在原假设 $H_0$ 为真时，作出错误决定（即拒绝 $H_0$ ）的概率受到了控制。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

根据 Neyman-Pearson 原则，在原假设 $H_0$ 为真时，作出错误决定（即拒绝 $H_0$ ）的概率受到了控制。

这表明，原假设 $H_0$ 受到保护，不至于轻易被拒绝。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

根据 Neyman-Pearson 原则，在原假设 $H_0$ 为真时，作出错误决定（即拒绝 $H_0$ ）的概率受到了控制。

这表明，原假设 $H_0$ 受到保护，不至于轻易被拒绝。

所以在具体问题中，往往将有把握、不能轻易拒绝的命题作为原假设 $H_0$ ，而把没有把握的、不能轻易肯定的命题作为对立假设。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

根据 Neyman-Pearson 原则，在原假设 $H_0$ 为真时，作出错误决定（即拒绝 $H_0$ ）的概率受到了控制。

这表明，原假设 $H_0$ 受到保护，不至于轻易被拒绝。

所以在具体问题中，往往将有把握、不能轻易拒绝的命题作为原假设 $H_0$ ，而把没有把握的、不能轻易肯定的命题作为对立假设。

因此原假设 $H_0$ 和对立假设 $H_1$ 的地位是不平等的，不能相互调换。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

### 定义 (5.1.3 (检验水平))

设  $\varphi$  是假设检验问题  $H_0: \theta \in \Theta_0$  vs  $H_1: \theta \in \Theta_1$  的一个检验, 而  $0 \leq \alpha \leq 1$ 。

如果  $\varphi$  犯第一类错误的概率总不超过  $\alpha$  (或等价地说,  $\varphi$  满足  $\beta_\varphi(\theta) \leq \alpha$ , 一切  $\theta \in \Theta_0$ ), 则称  $\alpha$  是检验  $\varphi$  的一个水平, 而  $\varphi$  称为**显著性水平为  $\alpha$  的检验**, 简称水平为  $\alpha$  的检验。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

### 定义 (5.1.3 (检验水平))

设  $\varphi$  是假设检验问题  $H_0: \theta \in \Theta_0$  vs  $H_1: \theta \in \Theta_1$  的一个检验, 而  $0 \leq \alpha \leq 1$ 。

如果  $\varphi$  犯第一类错误的概率总不超过  $\alpha$  (或等价地说,  $\varphi$  满足  $\beta_\varphi(\theta) \leq \alpha$ , 一切  $\theta \in \Theta_0$ ), 则称  $\alpha$  是检验  $\varphi$  的一个水平, 而  $\varphi$  称为**显著性水平为  $\alpha$  的检验**, 简称水平为  $\alpha$  的检验。

按照这一定义, 检验的水平不唯一。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

若  $\alpha$  为检验  $\varphi$  的水平，而  $\alpha < \alpha' < 1$ ，则  $\alpha'$  也是检验  $\varphi$  的水平。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

若  $\alpha$  为检验  $\varphi$  的水平，而  $\alpha < \alpha' < 1$ ，则  $\alpha'$  也是检验  $\varphi$  的水平。

为避免这一问题，有时称一个检验的最大水平为其真实水平，也就是

## 5.1.4 检验水平和控制犯第一类错误概率的原则

若  $\alpha$  为检验  $\varphi$  的水平，而  $\alpha < \alpha' < 1$ ，则  $\alpha'$  也是检验  $\varphi$  的水平。

为避免这一问题，有时称一个检验的最大水平为其真实水平，也就是

$$\text{检验}\varphi\text{的真实水平} = \sup_{\theta \in \Theta_0} \{\beta_{\varphi}(\theta)\}.$$

## 5.1.4 检验水平和控制犯第一类错误概率的原则

若  $\alpha$  为检验  $\varphi$  的水平，而  $\alpha < \alpha' < 1$ ，则  $\alpha'$  也是检验  $\varphi$  的水平。

为避免这一问题，有时称一个检验的最大水平为其真实水平，也就是

$$\text{检验}\varphi\text{的真实水平} = \sup_{\theta \in \Theta_0} \{\beta_{\varphi}(\theta)\}.$$

习惯上，把  $\alpha$  取得比较小且标准化，如  $\alpha = 0.01, 0.05$  或  $0.10$  等。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

水平 $\alpha$ 的选取，对检验的性质有很大影响。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

水平 $\alpha$ 的选取，对检验的性质有很大影响。

如果水平选得很低，那么容许犯第一类错误的概率很小，而为了达到这一点势必大大缩小拒绝域，使接受域扩大，从而增加了犯第二类错误的可能性。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

水平 $\alpha$ 的选取，对检验的性质有很大影响。

如果水平选得很低，那么容许犯第一类错误的概率很小，而为了达到这一点势必大大缩小拒绝域，使接受域扩大，从而增加了犯第二类错误的可能性。

反之，若水平选得高，则拒绝域扩大，接受域缩小，从而犯第二类错误的概率将相应地降低。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

这样看来，水平的选择不是一个数学问题，而是一个必须从实际角度来考虑的问题。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

这样看来，水平的选择不是一个数学问题，而是一个必须从实际角度来考虑的问题。

一般地，有以下几个因素影响水平的选择。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

(1) 当一个检验涉及双方利益时，水平的选定常是双方协议的结果。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

(1) 当一个检验涉及双方利益时，水平的选定常是双方协议的结果。

例如，商店向工厂进货，检验其次品率是否超过0.01。假设原假设和备择假设为

## 5.1.4 检验水平和控制犯第一类错误概率的原则

(1) 当一个检验涉及双方利益时，水平的选定常是双方协议的结果。

例如，商店向工厂进货，检验其次品率是否超过0.01。假设原假设和备择假设为

$$H_0 : p \leq 0.01 \text{ vs } H_1 : p > 0.01.$$

## 5.1.4 检验水平和控制犯第一类错误概率的原则

(1) 当一个检验涉及双方利益时，水平的选定常是双方协议的结果。

例如，商店向工厂进货，检验其次品率是否超过0.01。假设原假设和备择假设为

$$H_0 : p \leq 0.01 \text{ vs } H_1 : p > 0.01.$$

拒绝域形式为

## 5.1.4 检验水平和控制犯第一类错误概率的原则

(1) 当一个检验涉及双方利益时，水平的选定常是双方协议的结果。

例如，商店向工厂进货，检验其次品率是否超过0.01。假设原假设和备择假设为

$$H_0 : p \leq 0.01 \text{ vs } H_1 : p > 0.01.$$

拒绝域形式为

$$D = \{X = (X_1, \dots, X_n) : |T(X) - a_0| > A\}.$$

## 5.1.4 检验水平和控制犯第一类错误概率的原则

因为临界值 $A$ 与水平 $\alpha$ 有关，若水平选得低，临界值较大，拒绝域变小，接受域变大，则可能有较多的次品被商店接受（犯第二类错误概率变大）；

## 5.1.4 检验水平和控制犯第一类错误概率的原则

因为临界值 $A$ 与水平 $\alpha$ 有关，若水平选得低，临界值较大，拒绝域变小，接受域变大，则可能有较多的次品被商店接受（犯第二类错误概率变大）；

反之，若水平定得高，拒绝域变大，临界值较小，则将有较多的合格品被商店拒收（犯第一类错误概率变大）。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

因为临界值 $A$ 与水平 $\alpha$ 有关，若水平选得低，临界值较大，拒绝域变小，接受域变大，则可能有较多的次品被商店接受（犯第二类错误概率变大）；

反之，若水平定得高，拒绝域变大，临界值较小，则将有较多的合格品被商店拒收（犯第一类错误概率变大）。

因此，水平定的大小涉及商店和工厂双方利益，应由双方商定。有时还要采取随机化的方法，使双方利益达到平衡。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

(2) 犯两种错误的后果一般在性质上有很大的不同。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

(2) 犯两种错误的后果一般在性质上有很大的不同。

如果犯第一类错误的后果在性质上很严重，就力求在合理的范围内尽量减少犯这种错误的可能性。这时，相应的水平就取得更低一些。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

(2) 犯两种错误的后果一般在性质上有很大的不同。

如果犯第一类错误的后果在性质上很严重，就力求在合理的范围内尽量减少犯这种错误的可能性。这时，相应的水平就取得更低一些。

例如，制药厂要生产一种新药代替旧药治疗某种疾病，安排了一些试验，要对新旧药物疗效作出检验。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

由于旧药已经长期临床使用，有一定的疗效。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

由于旧药已经长期临床使用，有一定的疗效。

而新药尚未经长期临床使用，一旦效果不好，将危及患者的生命安全，造成严重的后果。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

由于旧药已经长期临床使用，有一定的疗效。

而新药尚未经长期临床使用，一旦效果不好，将危及患者的生命安全，造成严重的后果。

所以在进行检验时，将原假设 $H_0$ 设为“旧药不比新药差”。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

如假设旧药的治疗效果为 $\theta_0$ ，假设原假设和备择假设为

## 5.1.4 检验水平和控制犯第一类错误概率的原则

如假设旧药的治疗效果为 $\theta_0$ ，假设原假设和备择假设为

$$H_0 : \theta \leq \theta_0 \text{ vs } H_1 : \theta > \theta_0.$$

## 5.1.4 检验水平和控制犯第一类错误概率的原则

如假设旧药的治疗效果为 $\theta_0$ ，假设原假设和备择假设为

$$H_0 : \theta \leq \theta_0 \text{ vs } H_1 : \theta > \theta_0.$$

拒绝域的形式为  $D = \{X : |T(X) - a_0| > A\}$ ，如下图阴影部分。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

如假设旧药的治疗效果为 $\theta_0$ ，假设原假设和备择假设为

$$H_0 : \theta \leq \theta_0 \text{ vs } H_1 : \theta > \theta_0.$$

拒绝域的形式为  $D = \{X : |T(X) - a_0| > A\}$ ，如下图阴影部分。



## 5.1.4 检验水平和控制犯第一类错误概率的原则

且使检验水平 $\alpha$ 定得更小一些，临界值 $A$ 变大，拒绝域变小，接受域变大，这样使 $H_0$ 被拒绝的可能性大大减小了（犯第一类错误的概率变小）。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

且使检验水平 $\alpha$ 定得更小一些，临界值 $A$ 变大，拒绝域变小，接受域变大，这样使 $H_0$ 被拒绝的可能性大大减小了（犯第一类错误的概率变小）。

这样就保证了“原假设被拒绝、新药被接受的检验”是十分严格的。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

(3) 一般来说, 试验者在试验前对问题的情况总不是一无所知的。他对问题的了解使他对零假设是否能成立就有了一定的看法, 这种看法可能影响到他对水平的选择。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

(3) 一般来说, 试验者在试验前对问题的情况总不是一无所知的。他对问题的了解使他对零假设是否能成立就有了一定的看法, 这种看法可能影响到他对水平的选择。

比方说, 一个物理学家根据某种理论推定随机变量 $X$ 应有分布 $F$ , 而他打算将这一理论付诸检验。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

(3) 一般来说, 试验者在试验前对问题的情况总不是一无所知的。他对问题的了解使他对零假设是否能成立就有了一定的看法, 这种看法可能影响到他对水平的选择。

比方说, 一个物理学家根据某种理论推定随机变量  $X$  应有分布  $F$ , 而他打算将这一理论付诸检验。

即假设  $H_0$ : 新分布 *vs*  $H_1$ : 旧分布。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

很明显，如果他对这一理论很有信心，他将非常倾向于认为假设能成立（将有把握、不能轻易拒绝的命题放在原假设），这时只有很有力的证据才可能使他认为这一假设不对。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

很明显，如果他对这一理论很有信心，他将非常倾向于认为假设能成立（将有把握、不能轻易拒绝的命题放在原假设），这时只有很有力的证据才可能使他认为这一假设不对。

相应地，他将把检验水平取得低一些（临界值变大，拒绝域变小，接受域变大，犯第一类错误概率小）。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

在实际问题中，零假设被拒绝，常常意味着推翻一种理论或用新的方法来代替一直使用的标准方法。在大多数情况下，人们希望这样做时有相当大的根据。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

在实际问题中，零假设被拒绝，常常意味着推翻一种理论或用新的方法来代替一直使用的标准方法。在大多数情况下，人们希望这样做时有相当大的根据。

可见，Neyman-Pearson 控制犯第一类错误的原则，在零假设的选择中有很大的实际意义，而决不单纯是一个数学问题。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

若水平  $\alpha$  很小，原假设  $H_0$  不会轻易被拒绝（临界值大，接受域大）。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

若水平  $\alpha$  很小，原假设  $H_0$  不会轻易被拒绝（临界值大，接受域大）。

(a) 如果样本落入了拒绝域，作出“拒绝原假设  $H_0$ （即接受  $H_1$ ）”的结论就比较可靠。因为，此时只会犯第一类错误，且其概率很小。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

若水平  $\alpha$  很小，原假设  $H_0$  不会轻易被拒绝（临界值大，接受域大）。

(a) 如果样本落入了拒绝域，作出“拒绝原假设  $H_0$ （即接受  $H_1$ ）”的结论就比较可靠。因为，此时只会犯第一类错误，且其概率很小。

(b) 反之，如果样本落入接受域，作出“接受原假设  $H_0$ ”的结论未必可靠（因为选取的  $\alpha$  水平小，影响了接受域的大小）。

## 5.1.4 检验水平和控制犯第一类错误概率的原则

若水平  $\alpha$  很小，原假设  $H_0$  不会轻易被拒绝（临界值大，接受域大）。

(a) 如果样本落入了拒绝域，作出“拒绝原假设  $H_0$ （即接受  $H_1$ ）”的结论就比较可靠。因为，此时只会犯第一类错误，且其概率很小。

(b) 反之，如果样本落入接受域，作出“接受原假设  $H_0$ ”的结论未必可靠（因为选取的  $\alpha$  水平小，影响了接受域的大小）。

这只能表明，在所选定的水平下没有充分根据拒绝  $H_0$ ，但绝不意味着有充分根据说明它正确。因为此时会犯第二类错误，其概率可能很大。所以更恰当的结论应该是“不足以拒绝原假设  $H_0$ ”。

## 5.1.5 求解假设检验问题的一般步骤

求解假设检验问题的一般步骤：

## 5.1.5 求解假设检验问题的一般步骤

求解假设检验问题的一般步骤:

- (1) 根据问题的要求提出原假设  $H_0$  和备择假设  $H_1$ 。

## 5.1.5 求解假设检验问题的一般步骤

求解假设检验问题的一般步骤:

- (1) 根据问题的要求提出原假设  $H_0$  和备择假设  $H_1$ 。
- (2) 导出拒绝域的形式, 确定检验统计  $T(\mathbf{X})$ , 其中临界值  $A$  待定。

## 5.1.5 求解假设检验问题的一般步骤

求解假设检验问题的一般步骤：

- (1) 根据问题的要求提出原假设  $H_0$  和备择假设  $H_1$ 。
- (2) 导出拒绝域的形式，确定检验统计  $T(\mathbf{X})$ ，其中临界值  $A$  待定。
- (3) 选取适当水平，利用检验统计量的分布求出临界值  $A$ 。

## 5.1.5 求解假设检验问题的一般步骤

求解假设检验问题的一般步骤:

- (1) 根据问题的要求提出原假设  $H_0$  和备择假设  $H_1$ 。
- (2) 导出拒绝域的形式, 确定检验统计  $T(\mathbf{X})$ , 其中临界值  $A$  待定。
- (3) 选取适当水平, 利用检验统计量的分布求出临界值  $A$ 。
- (4) 由样本  $\mathbf{X}$  算出检验统计量  $T(\mathbf{X})$  的具体值, 代入拒绝域中, 与临界值相比较, 作出接受 (不拒绝) 或者拒绝原假设的结论。

## 5.1.6 假设检验发展简史

详见教材。