ch8：假设检验

一些定义

什么是假设

定义 8.1.1 假设（hypothesis）就是关于总体参数的一个陈述.

这个定义是颇为笼统的，但其重点在于假设作出的是关于总体的陈述．假设检验的目的就是依靠来自总体的样本去决定互补的两个假设哪个为真.

什么是原假设

定义 8.1.2 一个假设检验问题中两个互补的假设称为原假设（null hypothesis 译注：原假设也叫零假设）和备择假设（alternative hypothesis）．把它们分别记作 H_0 和 H_1.

举个例子

若 θ 表示一个总体参数，原假设和备择假设的一般格式是 $H_0: \theta \in \Theta_0$ 和 $H_1: \theta \in \Theta_0^C$，这里 Θ_0 是参数空间的某子集而 Θ_0^C 是它的补集．例如，如果 θ 表示一个病人服一种药以后血压的平均变化，而一项试验可能对于检验 $H_0: \theta=0$ 和 $H_1: \theta\neq 0$ 谁更合理感兴趣．这里原假设说的是在平均意义下这种药对于血压没有影响，而备择假设说的是有一些影响．H_0 所陈述的是治疗无效果，这是一种常见的情形，也正是这种情形，导致术语零假设．再看另一个例子，一位消费者可能对一个供应商的产品的次品比例感兴趣．如果 θ 表示次品的比例，此消费者可能想检验 $H_0: \theta \geqslant \theta_0$ 对 $H_1: \theta < \theta_0$．θ_0 值是最大可接受的次品比例，H_0 陈述的是次品的比例高得无法接受．这种问题叫做验收抽样（acceptance sampling）问题，它的假设关心的是一种产品的质量.

在一个假设检验问题中，试验者在观测到样本以后必须决定是接受 H_0 为真还是认为其为假而拒绝 H_0，即认为 H_1 为真.

注意

> ◁ Note: Θ_0 and Θ_1 are often called *Null* and *Alternative* space of parameter and the hypotheses are expressed as
>
> $$H_0 : \boldsymbol{\theta} \in \Theta_0 \quad vs \quad \boldsymbol{\theta} \in \Theta_1$$

什么是简单假设与复合假设

Definition

A hypothesis that completely specifies the distribution of $X_1, \cdots, X_n$ is called a *simple hypothesis* otherwise it is called *composite hypothesis*.

举个例子

▷ Example: $\theta_1 = \theta_2$, $\theta_1 = \theta_2 = 2$, $\theta_1 > \theta_2$.

简单原假设(simple null)：Θ_0只包含一个点，如$H_0 : \theta = \theta_0$

复杂原假设(composite null)：Θ_0只包含多个点，如$H_0 : \theta \le \theta_0$

注意

- After observing $X_1 = x_1, \cdots, X_n = x_n$, we need to decide which hypothesis, H_0 or H_1, we will accept. Let $\mathfrak{X}$ denote the set of all possible realization of $X_1, \cdots, X_n$. Testing function (rule) plays the same role as estimator in point estimation.

什么是测试函数与简单测试函数

1. A function $\phi : \mathfrak{X} \to [0, 1]$ is called a *testing function*.
2. If a testing function takes a values in $\{0, 1\}$, i.e. $\phi : \mathfrak{X} \to \{0, 1\}$, it is called a *simple testing function*.

◁ Note: The interpretation of definition 1 is that after observing $X_1 = x_1, \cdots, X_n = x_n$, reject H_0 with probability $\phi(x_1, \cdots, x_n)$ and accept H_0 with probability $1 - \phi(x_1, \cdots, x_n)$. This is called a randomized procedure.

什么是拒绝域与接受域

- $R_\phi = \{\mathbf{x} : \phi(\mathbf{x}) = 1\}$ is called the *rejection region* or *critical region*
- $A_\phi = \{\mathbf{x} : \phi(\mathbf{x}) = 0\}$ is called the *acceptance region*

什么是假设检验得两种错误

设 R 表示一个检验的拒绝区域．则当 $\theta\in\Theta_0$ 的时候，如果 $x\in R$，这个检验就会犯一个错误，所以犯第一类错误的概率是 $P_\theta(\mathbf{X}\in R)$．当 $\theta\in\Theta_0^C$ 的时候，犯第二类错误的概率是 $P_\theta(\mathbf{X}\in R^C)$．由 R 换到 R^C 有些容易混淆，但是如果我们认识到 $P_\theta(\mathbf{X}\in R^C)=1-P_\theta(\mathbf{X}\in R)$，就会看到 $P_\theta(\mathbf{X}\in R)$ 作为 θ 的函数，它包含着这个拒绝区域为 R 的检验的所有信息．我们有

表 8.3.1　假设检验中的两类错误

		判　决	
		接受 H_0	拒绝 H_0
真实情况	H_0	正确判决	第一类错误
	H_1	第二类错误	正确判决

$$P_\theta(\mathbf{X}\in R)=\begin{cases}\text{犯第一类错误的概率} & \text{如果 } \theta\in\Theta_0\\ 1-\text{犯第二类错误的概率} & \text{如果 } \theta\in\Theta_0^C\end{cases}$$

什么是功效函数

定义 8.3.1　一个拒绝区域为 R 的假设检验的功效函数（power function）是由 $\beta(\theta)=P_\theta(\mathbf{X}\in R)$ 所定义的函数．

理想的功效函数对于所有的 $\theta\in\Theta^0$ 函数值是 0 而对于所有的 $\theta\in\Theta_0^C$ 函数值是 1．除非在平凡情况，这种理想不可能达到．一个好的检验的功效函数在大多数的 $\theta\in\Theta_0^C$ 上接近于 1 而在大多数的 $\theta\in\Theta_0$ 上接近于 0．

The *power function* $\beta_\phi(\theta)$ of a test $\phi(\mathbf{x})$ is the function defined as

$$\beta_\phi(\theta)=P_\theta[\phi(\mathbf{X})=1]=E_\theta[\phi(\mathbf{X})]=P_\theta(\mathbf{X}\in R_\phi)$$

◁ Note:

- $\theta\in\Theta_0$, $\beta_\phi(\theta)=Pr[Type\ \ I\ \ error\ ]$.
  $\theta\in\Theta_1$, $\beta_\phi(\theta)=1-Pr[Type\ \ II\ \ error]$.
- $\sup_{\theta\in\Theta_0}\beta_\phi(\theta)$ is called the *size of the test* ϕ. Thus, any test such that $\sup_{\theta\in\Theta_0}\beta_\phi(\theta)=\alpha$ is called as a *size* α *test*.
- Test ϕ such that $\sup_{\theta\in\Theta_0}\beta_\phi(\theta)\le\alpha$ is called a *level* α *test*.

举个例子

▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim}$ exponential(θ).

$$H_0: \theta \geq 1 \quad vs \quad H_1: \theta < 1$$

Consider a test function

$$\phi(\mathbf{x}) = \begin{cases} 1 & \bar{x} < 1 \\ 0 & elsewhere. \end{cases}$$

$$n\bar{X} = \sum_{i=1}^{n} X_i \sim \Gamma(n, \theta)$$

$$\frac{2n\bar{X}}{\theta} \sim \Gamma(n, 2), \quad \text{i.e.} \quad \chi^2(2n)$$

最大功效检验

定义

> A test function $\phi[\mathbf{X} = (X_1, \cdots, X_n)]$ is said to be the *most powerful* test of size α for testing
>
> $$H_0: \theta = \theta_0 \quad vs \quad H_1: \theta = \theta_1$$
>
> if
>
> 1. $E_{\theta_0}[\phi(\mathbf{X})] = \alpha$, $[\beta_\phi(\theta_0) = \alpha.]$
> 2. for any other test function $\tilde{\phi}(\mathbf{X})$ with $E_{\theta_0}[\tilde{\phi}(\mathbf{X})] \leq \alpha$,
>
> $$E_{\theta_1}[\phi(\mathbf{X})] \geq E_{\theta_1}[\tilde{\phi}(\mathbf{X})], \quad [\beta_\phi(\theta_1) \geq \beta_{\tilde{\phi}}(\theta_1)]$$
>
> MP test has the smallest probability of type II error among all test rules with probability of type I error no bigger than α.

奈曼-皮尔逊引理

定义 8.3.11 的要求条件过强以至在很多实际问题中 UMP 检验不存在．但是在有 UMP 检验的问题中，一个 UMP 检验理应被考虑为该类中的最优检验．这样，我们希望如果 UMP 检验存在，就能够识别它们．下面的著名定理清楚地描述了在原假设和备择假设都只含有一个关于样本的概率分布［即 H_0 和 H_1 都是简单假设 (simple hypothesis)］的情况，哪些检验是 UMP 水平为 α 的检验．

定理 8.3.12［Neyman-Pearson（奈曼-皮尔逊）引理］ 考虑检验 $H_0: \theta=\theta_0$ 对 $H_1: \theta=\theta_1$，其中相应于 θ_i 的概率密度函数或概率质量函数是 $f(x|\theta_i)\,i=0, 1$，利用一个拒绝区域为 R 的检验，R 满足对某个 $k\geqslant 0$

$$\text{若 } f(\boldsymbol{x}|\theta_1)>kf(\boldsymbol{x}|\theta_0) \quad \text{则 } \boldsymbol{x}\in R$$

(8.3.1) 和

$$\text{若 } f(\boldsymbol{x}|\theta_1)<kf(\boldsymbol{x}|\theta_0) \quad \text{则 } \boldsymbol{x}\in R^{\mathrm{C}}$$

而且

(8.3.2) $$\alpha=P_{\theta_0}(\boldsymbol{X}\in R)$$

则有

a.（充分性）任意满足条件（8.3.1）和条件（8.3.2）的检验，是一个 UMP 水平为 α 的检验.

b.（必要性）如果存在一个满足条件（8.3.1）和条件（8.3.2）的检验，其中 $k>0$，则每一个 UMP 水平为 α 的检验是真实水平为 α 的检验［满足（8.3.2）］而且每一个 UMP 水平为 α 的检验必满足条件（8.3.1）除去在一个使 $P_{\theta_0}(\boldsymbol{X}\in A)=P_{\theta_1}(\boldsymbol{X}\in A)=0$ 的集合 A 上可能不满足.

> **Theorem (Neyman-Pearson Lemma)**
>
> *$X_1,\cdots,X_n$ has a joint pdf/pmf $f(\mathbf{x}|\theta)$, $\theta\in\Theta$. Consider the testing the hypotheses,*
>
> $$H_0:\theta=\theta_0 \quad vs \quad H_1:\theta=\theta_1$$
>
> *Then, for any $0\leq\alpha\leq 1$, there exist a MP test of size α given below;*
>
> $$\phi(\mathbf{x})=\begin{cases}1 & \text{if} \quad f(\mathbf{x}|\theta_1)>kf(\mathbf{x}|\theta_0),\\ \gamma & \text{if} \quad f(\mathbf{x}|\theta_1)=kf(\mathbf{x}|\theta_0),\\ 0 & \text{if} \quad f(\mathbf{x}|\theta_1)<kf(\mathbf{x}|\theta_0),\end{cases}$$
>
> *where the constants k and γ are chose to satisfy*
>
> $$E_{\theta_0}[\phi(\mathbf{X})]=\beta_\phi(\theta_0)=\alpha.$$

注意

> ◁ Note:
>
> 1. The MP test ϕ reject H_0 if the likelihood ratio
>
> $$L=\frac{f(\mathbf{x}|\theta_1)}{f(\mathbf{x}|\theta_0)}$$
>
>   is large.
> 2. In general, there may be more than one choice of k and γ that $\beta_\phi(\theta_0)=\alpha$. Then each is MP test of size α.
> 3. When $f(\mathbf{x}|\theta_1)/f(\mathbf{x}|\theta_0)$ has a continuous distribution under the null, H_0, $\gamma=0$ is usually taken and considered as the MP test of size α.

例

Example Let $X_1, \dots, X_n$ be i.i.d. $N(\theta, 1)$, and suppose $\theta_0 < \theta_1$. Find the most powerful size α test of

$$H_0 : \theta = \theta_0 \quad \text{vs.} \quad H_1 : \theta = \theta_1,$$

and the power of this test.

$$f(\boldsymbol{x}|\theta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i - \theta)^2\right).$$

$$f(\boldsymbol{x}|\theta_1) > k f(\boldsymbol{x}|\theta_0) \quad \text{iff}$$

$$\log f(\boldsymbol{x}|\theta_1) - \log f(\boldsymbol{x}|\theta_0) > \log k = k'.$$

$$\log f(\boldsymbol{x}|\theta_1) - \log f(\boldsymbol{x}|\theta_0) =$$
$$-\sum_{i=1}^{n}(x_i - \theta_1)^2/2 + \sum_{i=1}^{n}(x_i - \theta_0)^2/2$$

The last quantity exceeds k' iff

$$\bar{x} > \frac{(\theta_1^2 - \theta_0^2)/2 + k'/n}{(\theta_1 - \theta_0)}.$$

The test function of the most powerful test thus has the form

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \bar{x} > c \\ 0, & \bar{x} < c, \end{cases}$$

where c is determined from $E_{\theta_0}[\phi(\boldsymbol{X})] = \alpha$.

$$\begin{aligned} E_{\theta_0}[\phi(\boldsymbol{X})] &= P_{\theta_0}(\bar{X} > c) \\ &= P_{\theta_0}\left(\frac{\bar{X} - \theta_0}{1/\sqrt{n}} > \sqrt{n}(c - \theta_0)\right) = \alpha. \end{aligned}$$

This implies that $\sqrt{n}(c - \theta_0) = z$, or $c = \theta_0 + z_\alpha/\sqrt{n}$. So, we reject H_0 iff $\bar{x} > \theta_0 + z_\alpha/\sqrt{n}$.

Power of the test is

$$P_{\theta_1}\left(\bar{X} > \theta_0 + z_\alpha/\sqrt{n}\right) =$$

$$P_{\theta_1}\left(\frac{\bar{X} - \theta_1}{1/\sqrt{n}} > z_\alpha - \sqrt{n}(\theta_1 - \theta_0)\right) =$$

$$1 - \Phi\left(z_\alpha - \sqrt{n}(\theta_1 - \theta_0)\right).$$

例

▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$.

$$H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu = \mu_1 (< \mu_0)$$

Find the MP-test of size α.

$$L(\theta) = \left(\frac{1}{2\pi}\right)^{n/2} \exp\{-\sum_{i=1}^{n} \frac{(x_i - \theta)^2}{2\sigma^2}\}$$

$$\frac{L(\theta_1)}{L(\theta_0)} > k \Longrightarrow \quad \bar{x} < \frac{(2\sigma^2 \log k)/n - \theta_0^2 + \theta_1^2}{2(\theta_1 - \theta_0)}.$$

$$\alpha = P_{\theta_0}(\bar{X} < c). \quad \bar{X} < c = -\sigma z_\alpha/\sqrt{n} + \theta_0.$$

一致最大功效检验

定义

定义 8.3.11 设 $\mathcal{C}$ 是一个关于 H_0：$\theta\in\Theta_0$ 对 H_1：$\theta\in\Theta_0^C$ 的检验类. $\mathcal{C}$ 中一个功效函数为 $\beta(\theta)$ 的检验是一个一致最大功效 $\mathcal{C}$ 类检验［uniformly most powerful (UMP) class $\mathcal{C}$ test］，如果对每个 $\theta\in\Theta_0^C$ 与每个 $\mathcal{C}$ 中检验的功效函数 $\beta'(\theta)$，都有 $\beta(\theta)\geqslant\beta'(\theta)$.

Definition

Let $f(\mathbf{x}|\theta)$, $\theta\in\Theta$ be the joint pdf/pmf of $X_1,\cdots,X_n$. Let Θ_0 and Θ_1 be the nonempty disjoint subsets of Θ. A test rule $\phi(\mathbf{x})$ is said to be an *uniformly most powerful (UMP)* test of size α for testing

$$H_0:\theta\in\Theta_0 \quad vs \quad H_1:\theta\in\Theta_1$$

if

1. $\max_{\theta\in\Theta_0} E_\theta[\phi(\mathbf{X})]=\alpha$
2. for any other test $\tilde{\phi}(\mathbf{x})$ with $\max_{\theta\in\Theta_0} E_\theta[\tilde{\phi}(\mathbf{X})]\le\alpha$, we have

$$E_\theta[\phi(\mathbf{X})]\ge E_\theta[\tilde{\phi}(\mathbf{X})]$$

   for each $\theta\in\Theta_1$.

◁ Note:

1. A UMP test has the smallest probability of type II error for every $\theta\in\Theta_1$among all the test with size $\le\alpha$.
2. Condition 2 is a really strong requirement. Unlike the simple versus simple case, UMP test may not exist for composite H_0 and for composite H_1.
3. NP lemma can be used to show that UMP test does not exist or identify the UNP test if it exists. HOW ?

如何做

a. Fix $\theta_0\in\Theta_0$ appropriately (usually boundary of Θ_0).
b. Choose any $\theta_1\in\Theta_1$
c. Then find a MP test of size α, $\phi(\mathbf{x})$, for

$$H_0:\theta=\theta_0 \quad vs \quad H_1:\theta=\theta_1.$$

If

i $\phi(\mathbf{x})$ does not depend on θ_1
ii $\max_{\theta\in\Theta_0} E_\theta[\phi(\mathbf{X})]=\alpha$

then $\phi(\mathbf{x})$ is the UMP-test of size α.

MLR

> **Definition**
> Let $f(\mathbf{x}|\theta)$, $\theta \in \Theta$ be the joint pdf/pmf of $X_1, \cdots, X_n$. The family is said to have *Monotone Likelihood Ratio (MLR)* in a statistic $T(\mathbf{X})$ if, for all $\theta'' > \theta'$, $\theta'', \theta' \in \Theta$, there exist a nondecreasing function of T, g, such that
>
> $$L = \frac{f(\mathbf{x}|\theta'')}{f(\mathbf{x}|\theta')} = g_{\theta',\ \theta''}[T(\mathbf{x})]$$
>
> in a support of $\mathbf{x}$.
>
> $\triangleleft$ Note:
>
> - if $g_{\theta',\ \theta''}(x)$ is decreasing then $g_{\theta',\ \theta''}(-x)$ is increasing.
> - if $f(\mathbf{x}|\theta'') > 0$ and $f(\mathbf{x}|\theta') = 0$ then $L = \infty$.

定理——如何找UMP

Theorem
Let $X_1, \cdots, X_n$ have joint pdf/pmf $f(\mathbf{x}|\theta)$, $\theta \in \Theta$. Assume the family has MLR in $T(\mathbf{X})$. Then

1. *A UMP test of size α for*

   $$H_0 : \theta \leq \theta_0 \quad vs \quad H_1 : \theta > \theta_0$$

   is given by

   $$\phi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) > k, \\ \gamma, & T(\mathbf{x}) = k, \\ 0, & T(\mathbf{x}) < k, \end{cases}$$

   where k and γ are determined by

   $$P_{\theta_0}[T(\mathbf{X}) > k] + \gamma P_{\theta_0}[T(\mathbf{X}) = k] = \alpha.$$

Theorem (-Continued)

2. *A UMP test of size α for*

$$H_0 : \theta \geq \theta_0 \quad vs \quad H_1 : \theta < \theta_0$$

is given by

$$\phi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) < k, \\ \gamma, & T(\mathbf{x}) = k, \\ 0, & T(\mathbf{x}) > k, \end{cases}$$

where k and γ are determined by

$$P_{\theta_0}[T(\mathbf{X}) < k] + \gamma P_{\theta_0}[T(\mathbf{X}) = k] = \alpha.$$

举个例子

▷ Example: $X_1, \cdots X_n \overset{iid}{\sim} f(x|\theta)$， $f(x|\theta) = c(\theta)h(x)\exp[w(\theta)t(x)]$

$$T^*(\mathbf{X}) = \sum_{i=1}^{n} t(X_i)$$

if $w(\theta)$ is an increasing function of θ $\quad let T(X)=T^*(X)$

if $w(\theta)$ is an increasing function of θ $\quad let T(X)=-T^*(X)$

$$H_0 : \theta \leq \theta_0 \quad vs \quad H_1 : \theta > \theta_0$$

$$\phi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) > k, \\ \gamma, & T(\mathbf{x}) = k, \\ 0, & T(\mathbf{x}) < k, \end{cases}$$

$$H_0 : \theta \geq \theta_0 \quad vs \quad H_1 : \theta < \theta_0$$

$$\phi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) < k, \\ \gamma, & T(\mathbf{x}) = k, \\ 0, & T(\mathbf{x}) > k, \end{cases}$$

例

Example Let $X_1, \ldots, X_n$ be i.i.d. $N(0, \theta)$. $\Theta = \{\theta : \theta > 0\}$. Find UMP test of

$$H_0 : \ \theta \geq \theta_0 \quad \text{vs.} \quad H_1 : \theta < \theta_0.$$

Check for the MLR property.

$$f(\boldsymbol{x}|\theta) = \exp\left(-\frac{1}{2\theta}\sum_{i=1}^{n} x_i^2 - \frac{n}{2}\log(2\pi\theta)\right)$$

$$\frac{f(\boldsymbol{x}|\theta'')}{f(\boldsymbol{x}|\theta')} \uparrow \quad \text{in} \quad \sum_{i=1}^{n} x_i^2 \quad \text{if} \quad \theta'' > \theta'$$

Since $1/\theta$ is a decreasing function of θ, Theorems 9 and 10 tell us that the UMP level α test of H_0 vs. H_1 has the form

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \text{if } \sum_{i=1}^{n} x_i^2 \leq c^* \\ 0, & \text{if } \sum_{i=1}^{n} x_i^2 > c^*. \end{cases}$$

The constant c^* is such that

$$P_{\eta_0}\left(\sum_{i=1}^{n} X_i^2 \leq c^*\right) = \alpha.$$

When $\theta = \theta_0$, we know that $X_i/\sqrt{\theta_0} \sim N(0,1)$, and so $X_i^2/\theta_0 \sim \chi_1^2$ and $\sum_{i=1}^{n} X_i^2/\theta_0 \sim \chi_n^2$.

It follows that $c^* = \chi_{n,\alpha}^2\,\theta_0$, where $\chi_{n,p}^2$ is the $100p$th percentile of the χ_n^2 distribution.

So, we have found the most powerful level α test of H_0' vs. H_1', and hence of H_0 vs. H_1.

$$\beta(\theta) = P_\theta\left[\chi_n^2 \leq \left(\frac{\theta_0}{\theta}\right)\chi_{n,\alpha}^2\right]$$

Limiting cases: $\theta = 0$ and $\theta = \infty$.

UMP检验不存在情况

Let the probability model be as in previous Example
Want to test

$$H_0 : \theta = \theta_0 \quad \text{vs.} \quad H_1 : \theta \neq \theta_0.$$

$$\Theta_0 = \{\theta_0\} \qquad \Theta_0^c = (0, \infty) \cap \{\theta_0\}^c$$

MP test of $H_0' : \theta = \theta_0$ vs. $H_1' : \theta = \theta_1$ for $\theta_1 > \theta_0$ is

$$\phi_1(\boldsymbol{x}) = \begin{cases} 1, & \text{if } \sum_{i=1}^n x_i^2 \geq c \\ 0, & \text{if } \sum_{i=1}^n x_i^2 < c. \end{cases}$$

MP test of $H_0'' : \theta = \theta_0$ vs. $H_1'' : \theta = \theta_2$ for $\theta_2 < \theta_0$ is

$$\phi_2(\boldsymbol{x}) = \begin{cases} 1, & \text{if } \sum_{i=1}^n x_i^2 \leq c_1 \\ 0, & \text{if } \sum_{i=1}^n x_i^2 > c_1. \end{cases}$$

Assume ϕ^* is UMP for testing H_0 vs. H_1. Then it is most powerful for $H_0' : \theta = \theta_0$ vs. $H_1' : \theta = \theta_1$ and hence agrees with ϕ_1 by N-P lemma.

Also, it must agree with ϕ_2 by the same logic. But $\phi_1 \neq \phi_2$, which yields a contradiction. Hence, there is no UMP test.

When a UMP test doesn't exist, one can look at a smaller class of tests and try to find the most powerful test within the smaller class.

Examples of such tests:

Class of unbiased tests A test is said to be unbiased if $\beta(\theta) \geq$ size of test for all $\theta \in \Theta_0^c$.

i.e. power function of a test satisfies

$$\beta(\theta) \leq \alpha \quad \text{if } \theta \in \Theta_0,$$
$$\beta(\theta) \geq \alpha \quad \text{if } \theta \in \Theta_1.$$

LRT

定义

假设检验的似然比方法与 7.2.2 节讨论的极大似然估计量有关，并且似然比检验就像极大似然估计那样应用广泛．设 X_1，…，X_n 是来自概率密度函数或概率质量函数为 $f(x|\theta)$（θ 可能是向量）的一组随机样本，回想似然函数的定义

$$L(\theta|x_1,\cdots,x_n)=L(\theta|\mathbf{x})=f(\mathbf{x}|\theta)=\prod_{i=1}^{n}f(x_i|\theta)$$

设 Θ 表示整个参数空间，似然比检验的定义如下：

定义 8.2.1 关于检验 $H_0:\theta\in\Theta_0$ 对 $H_1:\theta\in\Theta_0^c$ 的似然比检验统计量是

$$\lambda(\mathbf{x})=\frac{\sup\limits_{\Theta_0}L(\theta|\mathbf{x})}{\sup\limits_{\Theta}L(\theta|\mathbf{x})}$$

任何一个拒绝区域的形式为 $\{\mathbf{x}:\lambda(\mathbf{x})\leqslant c\}$ 的检验都叫做似然比检验（likelihood ratio test，简记为 LRT）．这里 c 是任意一个满足 $0\leqslant c\leqslant 1$ 的数．

似然比检验背后的原理在 $f(x|\theta)$ 是一个离散型随机变量的概率质量函数时可能最易于理解．在这种情况下，$\lambda(\mathbf{x})$ 的分子就是观测样本出现的最大概率，这里最大是指对参数取遍原假设范围的计算而言（见习题 8.4）．而 $\lambda(\mathbf{x})$ 的分母则是取遍所有可能的参数时观测样本出现的最大概率．如果存在备择假设中的参数值，使得样本出现的可能性比所有原假设下的参数对应的可能性更大的话，这两个最大值之比就小．在这种情况下 LRT 准则就说 H_0 应该被拒绝而接受 H_1 为真．选择数 c 的方法放在 8.3 节讨论．

Definition

Let $X_1,\cdots,X_n$ have joint pdf/pmf $f(\mathbf{x}|\theta)$, $\theta\in\Theta$. Let Θ_0 be a proper subset of Θ. Define the likelihood ratio

$$\lambda(\mathbf{x})=\frac{\sup_{\theta\in\Theta_0}f(\mathbf{x}|\theta)}{\sup_{\theta\in\Theta}f(\mathbf{x}|\theta)}.$$

Then the Likelihood Ratio Test (LRT) of size α for testing $H_0:\theta\in\Theta_0 \quad vs \quad H_1:\theta\in\Theta_0^c$ is

$$\phi(\mathbf{x})=\begin{cases}1, & \lambda(\mathbf{x})<k,\\ \gamma, & \lambda(\mathbf{x})=k,\\ 0, & \lambda(\mathbf{x})>k,\end{cases}$$

where k and γ satisfy $\sup_{\theta\in\Theta_0}E_\theta[\phi(\mathbf{x})]=\alpha$.

注意

◁ Note:

1. Let $\hat{\theta}_0$ be the MLE of θ under H_0 and $\hat{\theta}$ be the MLE of θ without any restriction. Then,

$$\lambda(\mathbf{x}) = \frac{f(\mathbf{x}|\hat{\theta}_0)}{f(\mathbf{x}|\hat{\theta})}.$$

2. $0 \leq \lambda(\mathbf{x}) \leq 1$.

▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$, σ^2 is known.

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta \neq \theta_0$$

Likelihood ratio tests are especially useful intwo situations:

(i) Two-sided tests

(ii) Tests in the presence of nuisance parameters

举个例子

Example (Likelihood ratio test for the mean in normal pdf)

$X_1,\cdots,X_n$ (iid) from a $N(\mu, \sigma^2)$ distribution, where $-\infty<\mu<\infty$, $\sigma>0$. Consider the hypotheses

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

where μ_0 is specified. The likelihood function

$$L = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left[-\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left[-\sum_{i=1}^{n}\frac{(x_i-\overline{x})^2}{2\sigma^2}\right]\exp\left[-\sum_{i=1}^{n}\frac{(\overline{x}-\mu)^2}{2\sigma^2}\right].$$

(1) σ (> 0) is known (p408, 8.37)

$$\Theta = \{\mu : -\infty < \mu < \infty\}; \quad \Theta_0 = \{\mu_0\}.$$

m.l.e in Θ : $\hat{\mu} = \overline{X}$

Restricted m.l.e in Θ_0 :$\hat{\mu} = \mu_0$

$$\lambda = \frac{L(\mu_0)}{L(\overline{X})} = \exp\{-(2\sigma^2)^{-1} n(\overline{X} - \mu_0)^2\}.$$

$$-2\ln\lambda = \frac{n(\overline{X} - \mu_0)^2}{\sigma^2} \triangleq Z^2.$$

$$\lambda \le \lambda_0 \Leftrightarrow |Z| \ge c = \sqrt{-2\ln\lambda_0}$$

$$Z = \frac{\overline{X} - \theta_0}{\sigma / \sqrt{n}} \overset{H_0}{\sim} N(0,1)$$

Therefore the reject region is

$$|Z| = \sqrt{-2\ln\lambda} \ge \Phi^{-1}(1 - \frac{\alpha}{2}).$$

(2) σ (> 0) is unknown (p408, 8.37,8.38)

$$\Theta = \{(\mu, \sigma^2): -\infty < \mu < \infty, \sigma^2 > 0\}$$

$$\Theta_0 = \{(\mu_0, \sigma^2): \sigma^2 > 0\}.$$

$$H_0 : \mu_1 = \mu_0, \sigma^2 > 0 \ \ vs \ \ H_1 : \mu \neq \mu_0, \sigma^2 > 0$$

m.l.e in Θ :

$$\hat{\mu} = \bar{X} \ and \ \hat{\sigma}^2 = (1/n)\sum_{i=1}^{n}(X_i - \bar{X})^2,$$

$$L(\hat{\mu}, \hat{\sigma}^2) = \frac{1}{(2\pi e \hat{\sigma}^2)^{n/2}}$$

m.l.e in Θ_0 :

$$\hat{\mu}^* = \mu_0 \ and \ \hat{\sigma}^{*2} = (1/n)\sum_{i=1}^{n}(X_i - \mu_0)^2,$$

$$L(\hat{\mu}^*, \hat{\sigma}^{*2}) = \frac{1}{(2\pi e \hat{\sigma}^{*2})^{n/2}}$$

Therefore, the likelihood ratio test statistic

$$\lambda = \left(\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{\sum_{i=1}^{n} (X_i - \mu_0)^2} \right)^{n/2} = \left(\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2 + n(\bar{X} - \mu_0)^2} \right)^{n/2}.$$

$$\lambda \le \lambda_0 \Leftrightarrow \lambda^{-2/n} = 1 + \frac{n(\bar{X} - \mu_0)^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2} \ge c,$$

$$\because T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \overset{H_0}{\sim} t(n-1)$$

$$\alpha = P\{|T| \ge t_{1-\alpha/2}(n-1)\} \Leftrightarrow \text{Reject } H_0 \text{ if } |T| \ge t_{1-\alpha/2}(n-1).$$

例

Example (Likelihood ratio test for the **variance** in normal pdf)

$X_1, \cdots, X_n$ (iid) from a $N(\mu, \sigma^2)$ distribution, where $-\infty < \mu < \infty$, $\sigma > 0$. Consider the hypotheses

$$H_0 : \sigma^2 = \sigma_0^2 \qquad H_1 : \sigma^2 \ne \sigma_0^2$$

where σ_0 is specified. The likelihood function

$$L = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp\left[-\sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2} \right]$$

1 μ is known

$$\Theta = \{\sigma^2 : \sigma^2 > 0\}, \quad \Theta_0 = \{\sigma_0^2\}$$

$$\text{m.l.e in } \Theta : \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

$$\text{m.l.e in } \Theta_0 : \ \hat{\sigma}_{(0)}^2 = \sigma_0^2$$

$$\lambda = \frac{L(\sigma_0^2)}{L(\hat{\sigma}^2)} = \left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{n/2} \exp\left[-\frac{n}{2}\left(\frac{\hat{\sigma}^2}{\sigma_0^2} - 1\right)\right] = \left(\frac{Q}{n}\right)^{n/2} \exp\left(-\frac{Q}{2} + \frac{n}{2}\right)$$

$$Q = \frac{n\hat{\sigma}^2}{\sigma_0^2} = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma_0}\right)^2 \overset{H_0}{\sim} \chi^2(n)$$

$$\frac{d(\ln \lambda(Q))}{dQ} = \frac{n}{2Q} - \frac{1}{2}$$

$$Q < 1/n, \ \frac{d(\ln \lambda)}{dQ} > 0, \ \lambda(Q)\uparrow; \quad Q > 1/n, \ \frac{d(\ln \lambda)}{dQ} < 0, \ \lambda(Q)\downarrow$$

Therefore the reject region

$$\lambda(Q) = \left(\frac{Q}{n}\right)^{n/2} \exp\left(-\frac{Q}{2} + \frac{n}{2}\right) \le k \Leftrightarrow Q \le c_1 \ or \ Q \ge c_2$$

Let $f(x)$ be pdf of $\chi^2(n)$. Then c_1, c_2 satisfy

$$\begin{cases} \int_{c_1}^{c_2} f(x)dx = 1 - \alpha \\ \lambda(c_1) = \lambda(c_2) \end{cases} \Rightarrow \begin{cases} \int_{c_1}^{c_2} f(x)dx = 1 - \alpha \\ c_1^{\ n/2} e^{-c_1/2} = c_2^{\ n/2} e^{-c_2/2} \end{cases}$$

For convenience, we take c_1, c_2 as $\chi^2_{\alpha/2}(n)$, $\chi^2_{1-\alpha/2}(n)$

2 μ is unknown

$$\Theta = \{(\mu, \sigma^2): -\infty < \mu < \infty, \sigma^2 > 0\}$$

$$\Theta_0 = \{(\mu, \sigma_0^2): -\infty < \mu < \infty\}.$$

m.l.e in Θ : $\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$,

m.l.e in Θ_0 : $\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2_{(0)} = \sigma_0^2$,

$$\lambda = \frac{L(\bar{X}, \sigma_0^2)}{L(\bar{X}, \hat{\sigma}^2)} = \left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{n/2} \exp\left[-\frac{n}{2}\left(\frac{\hat{\sigma}^2}{\sigma_0^2} - 1\right)\right] = \left(\frac{Q}{n}\right)^{n/2} \exp\left(-\frac{Q}{2} + \frac{n}{2}\right)$$

$$Q = \frac{n\hat{\sigma}^2}{\sigma_0^2} = \frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i - \bar{X})^2 \overset{H_0}{\sim} \chi^2(n-1)$$

Similarly, the reject region

$$\lambda(Q) = \left(\frac{Q}{n}\right)^{n/2} \exp\left(-\frac{Q}{2} + \frac{n}{2}\right) \le k \Leftrightarrow Q \le c_1 \text{ or } Q \ge c_2$$

Let $f(x)$ be pdf of $\chi^2(n-1)$. Then c_1, c_2 satisfy

$$\begin{cases} \int_{c_1}^{c_2} f(x)dx = 1-\alpha \\ \lambda(c_1) = \lambda(c_2) \end{cases} \Rightarrow \begin{cases} \int_{c_1}^{c_2} f(x)dx = 1-\alpha \\ c_1^{n/2} e^{-c_1/2} = c_2^{n/2} e^{-c_2/2} \end{cases}$$

For convenience, we take c_1, c_2 as $\chi^2_{\alpha/2}(n-1)$, $\chi^2_{1-\alpha/2}(n-1)$

p-values

定义

做完假设检验之后，必须用具有统计意义的方式报告出结论. 一种报告假设检验结果的方法是报告检验所用的真实水平 α，以及拒绝或者接受 H_0 的判决. 检验的真实水平携带着重要的信息. 如果 α 小，判拒绝 H_0 是相当令人信服的，但是如果 α 大，判拒绝 H_0 就不是很令人信服了，这是因为检验作出的这个判决不正确的概率也大. 另一种报告假设检验结果的方法是报告一种叫做 p-值的统计量的值.

定义 8.3.26 p-值（p-value）$p(\boldsymbol{X})$是一个满足对每一个样本点 $\boldsymbol{x}$，都有 $0\leqslant p(\boldsymbol{x})\leqslant 1$ 的检验统计量，如果 $p(\boldsymbol{X})$的值小则可作为 H_1 为真的证据. 一个 p-值称为是有效的，如果对于每一个 $\theta\in\Theta_0$ 和每一个 $0\leqslant\alpha\leqslant 1$，都有

$$P_\theta(p(\boldsymbol{X})\leqslant\alpha)\leqslant\alpha. \tag{8.3.8}$$

如果 $p(\boldsymbol{X})$是一个有效的 p-值，基于 $p(\boldsymbol{X})$易构建出一个水平为 α 的检验. 根据式（8.3.8），当且仅当 $p(\boldsymbol{X})\leqslant\alpha$ 时拒绝 H_0 的检验就是一个水平为 α 的检验. 通过 p-值报告检验结果的一个优点是每位读者能够选择他或她认为适当的 α，然后拿报告的 $p(\boldsymbol{x})$去和 α 比较，并且知道这些数据导致接受还是拒绝 H_0. 此外，p-值越小，就越强烈地拒绝 H_0. 因此，p-值以一个更连续的尺度报告出一个检验的结论，它胜于仅分成两种决策结果的接受 H_0 或拒绝 H_0.

定理

定理 8.3.27 设 $W(\boldsymbol{X})$是这样一个检验统计量，如 W 的值大则可作为 H_1 为真的依据. 对于每个样本点 $\boldsymbol{x}$，定义

$$p(\boldsymbol{x})=\sup_{\theta\in\Theta_0} P_\theta(W(\boldsymbol{X})\geqslant W(\boldsymbol{x})) \tag{8.3.9}$$

则 $p(\boldsymbol{X})$是一个有效的 p-值.

证明 固定 $\theta\in\Theta_0$. 设 $F_\theta(w)$表示 $-W(\boldsymbol{X})$的经验分布函数. 定义

$$p_\theta(\boldsymbol{x})=P_\theta(W(\boldsymbol{X})\geqslant W(\boldsymbol{x}))=P_\theta(-W(\boldsymbol{X})\leqslant -W(\boldsymbol{x}))=F_\theta(-W(\boldsymbol{x}))$$

因而随机变量 $p_\theta(\boldsymbol{X})$等于 $F_\theta(-W(\boldsymbol{X}))$. 因此，经过概率积分变换或者习题 2.10，$p_\theta(\boldsymbol{X})$的分布随机地大于或等于一个 $U(0，1)$ 分布. 就是说，对于每个 $0\leqslant\alpha\leqslant 1$，$P_\theta(p_\theta(\boldsymbol{X})\leqslant\alpha)\leqslant\alpha$. 因为对于每个 $\boldsymbol{x}$，$p(\boldsymbol{x})=\sup_{\theta'\in\Theta_0} p_{\theta'}(\boldsymbol{x})\geqslant p_\theta(\boldsymbol{x})$，

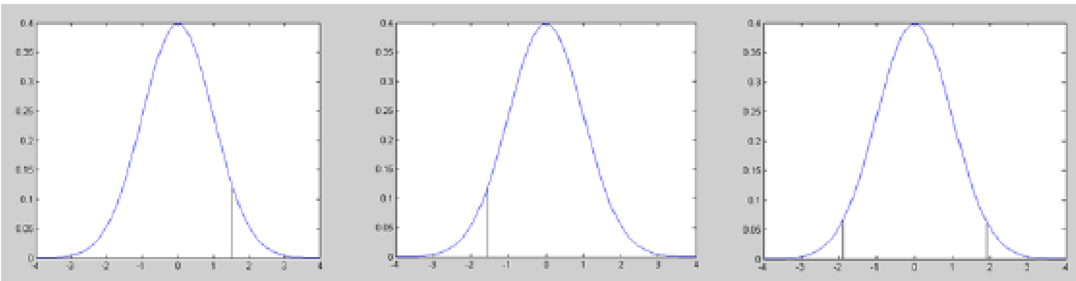
$$P_\theta(p(\boldsymbol{X})\leqslant\alpha)\leqslant P_\theta(p_\theta(\boldsymbol{X})\leqslant\alpha)\leqslant\alpha$$

对于每一个 $\theta\in\Theta_0$ 和每一个 $0\leqslant\alpha\leqslant 1$ 都成立；$p(\boldsymbol{x})$是一个有效的 p-值. ■

举个例子

Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$, where σ^2 is known. Suppose that we are testing $H_0 : \mu \le \mu_0$ against $H_1 : \mu > \mu_0$ and we observe $\frac{(\bar{x} - \mu_0)}{\sigma / \sqrt{n}} = 3$, then the p-value is $P(Z > 3) = .0013$.

If $\frac{(\bar{x} - \mu_0)}{\sigma / \sqrt{n}} = 2$, then the p-value is $P(Z > 2) = .0228$. If $\frac{(\bar{x} - \mu_0)}{\sigma / \sqrt{n}} = 1.5$, then the p-value is $P(Z > 1.5) = 0.0668$



$H_1 : \mu > \mu_0$ $H_1 : \mu < \mu_0$ $H_1 : \mu \ne \mu_0$

$P_{\mu_0}\{Z > z\}$ $P_{\mu_0}\{Z < z\}$ $P_{\mu_0}\{|Z| < |z|\}$

贝叶斯检验

介绍

假设检验也可以在 Bayes 模型里面被系统地描述. 回顾 7.2.3 节，一个 Bayes 模型不仅包括抽样分布 $f(\boldsymbol{x}|\theta)$而且还包括先验分布 $\pi(\theta)$，先验分布反映了在抽样前试验者关于参数的看法.

Bayes 范式规定利用 Bayes 定理把样本信息与先验信息结合以得到后验分布 $\pi(\theta|\boldsymbol{x})$. 所有关于 θ 的推断都基于后验分布进行.

在一个假设检验问题中，后验分布可以被用来计算 H_0 和 H_1 为真的概率. 记住，$\pi(\theta|\boldsymbol{x})$ 是一个随机变量的概率分布. 因此，后验概率 $P(\theta \in \Theta_0 | \boldsymbol{x}) = P(H_0$ 为真$|\boldsymbol{x})$与 $P(\theta \in \Theta_0^C | \boldsymbol{x}) = P(H_1$ 为真 $| \boldsymbol{x})$ 都可以计算出来.

概率 $P(H_0$ 为真 $| \boldsymbol{x})$ 与 $P(H_1$ 为真$|\boldsymbol{x})$ 对于经典统计学家是没有意义的. 经典统计学家把 θ 考虑为一个固定的数. 因而，一个假设或是真或是假. 如果 $\theta \in \Theta_0$，

那么对于所有的 $\boldsymbol{x}$ 值都有 $P(H_0$ 为真 $|\boldsymbol{x})=1$ 与 $P(H_1$ 为真 $|\boldsymbol{x})=0$. 如果 $\theta\in\Theta_0^C$，那么这些值则反过来. 由于这些概率是未知的（因为 θ 未知）并且不依赖于样本 $\boldsymbol{x}$，所以它们不被经典统计学家所用. 在一个假设检验问题的 Bayes 表述中，这些概率是依赖于样本 $\boldsymbol{x}$ 的，并且能给出关于 H_0 和 H_1 的真实性的有用信息.

Bayes 假设检验者利用后验分布进行假设检验，一种可能的方法是：如果 $P(\theta\in\Theta_0|\boldsymbol{X})\geqslant P(\theta\in\Theta_0^C|\boldsymbol{X})$ 就接受 H_0 为真否则就拒绝 H_0. 用以前各节的术语，检验统计量即样本的一个函数，在这里就是 $P(\theta\in\Theta_0^C|\boldsymbol{X})$，而拒绝区域就是 $\left\{x: P(\theta\in\Theta_0^C|\boldsymbol{x})>\frac{1}{2}\right\}$. 还有另外一种利用后验分布的方法，就是如果 Bayes 假设检验者希望防止错误地拒绝 H_0，那么他只有在 $P(\theta\in\Theta_0^C|\boldsymbol{X})$ 超过某个大的数，譬如 0.99 的时候才可能拒绝 H_0.

举个例子

Let $X_1,\ldots,X_n$ be i.i.d. $N(\mu,\sigma^2)$ (σ^2 known) and $\pi(\mu)\sim N\left(\theta,\tau^2\right)$.

We have shown that $\pi\left(\mu|\mathbf{x}\right)\sim N\left(\dfrac{n\tau^2\overline{x}+\sigma^2\theta}{n\tau^2+\sigma^2},\dfrac{\sigma^2\tau^2}{n\tau^2+\sigma^2}\right)$.

We test $H_0:\mu\le\mu_0$ vs. $H_1:\mu>\mu_0$.

We reject if $P\left(\mu>\mu_0|\mathbf{x}\right)>1/2$.

If $\mu_0 > \dfrac{n\tau^2\bar{x}+\sigma^2\theta}{n\tau^2+\sigma^2}$ then the area under the curve to the right of μ_0 is less than 1/2.

If $\mu_0 < \dfrac{n\tau^2\bar{x}+\sigma^2\theta}{n\tau^2+\sigma^2}$ then the area under the curve to the right of μ_0 is greater than 1/2.

So we reject H_0 if

$$\mu_0 < \frac{n\tau^2\bar{x}+\sigma^2\theta}{n\tau^2+\sigma^2} \Leftrightarrow n\tau^2\mu_0+\sigma^2\mu_0 < n\tau^2\bar{x}+\sigma^2\theta$$

$$\Leftrightarrow \mu_0+\frac{\sigma^2(\mu_0-\theta)}{n\tau^2} < \bar{x}.$$

(Note, if our prior mean was $\theta=\mu_0$, then a priori we would be putting equal weight on H_0 and H_1. We would reject H_0 if $\bar{x}>\mu_0$.)

习题

8.1-8.2

8.1 在 1000 次抛硬币中，出现 560 次正面和 440 次反面．假设这枚硬币是均匀的合理吗？并证明你的结论.

8.2 假定在某城市里某年中发生车祸的次数遵从一个 Poisson 分布．以往每年的平均事故次数是 15，并且今年是 10 次．声称事故率下降理由充足吗?

8.1 Let $X = \#$ of heads out of 1000. If the coin is fair, then $X \sim \text{binomial}(1000, 1/2)$. So

$$P(X \geq 560) = \sum_{x=560}^{1000} \binom{1000}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{n-x} \approx .0000825,$$

where a computer was used to do the calculation. For this binomial, $\mathrm{E}\,X = 1000p = 500$ and $\mathrm{Var}\,X = 1000p(1-p) = 250$. A normal approximation is also very good for this calculation.

$$P\{X \geq 560\} = P\left\{\frac{X - 500}{\sqrt{250}} \geq \frac{559.5 - 500}{\sqrt{250}}\right\} \approx P\{Z \geq 3.763\} \approx .0000839.$$

Thus, if the coin is fair, the probability of observing 560 or more heads out of 1000 is very small. We might tend to believe that the coin is not fair, and $p > 1/2$.

8.2 Let $X \sim \text{Poisson}(\lambda)$, and we observed $X = 10$. To assess if the accident rate has dropped, we could calculate

$$P(X \leq 10 | \lambda = 15) = \sum_{i=0}^{10} \frac{e^{-15} 15^i}{i!} = e^{-15}\left[1+15+\frac{15^2}{2!}+\cdots+\frac{15^{10}}{10!}\right] \approx .11846.$$

This is a fairly large value, not overwhelming evidence that the accident rate has dropped. (A normal approximation with continuity correction gives a value of .12264.)

8.6

8. 6　设我们有两组独立样本：X_1，…，X_n 来自指数分布 EXPO（θ），Y_1，…，Y_m 来自指数分布 EXPO（μ）.

（a）求 $H_0: \theta = \mu$ 对 $H_1: \theta \neq \mu$ 的 LRT.

（b）证明：（a）中的检验能够基于统计量

$$T = \frac{\sum X_i}{\sum X_i + \sum Y_i}$$

（c）求 H_0 为真时 T 的分布.

8.6 a.

$$\begin{aligned}
\lambda(\mathbf{x}, \mathbf{y}) &= \frac{\sup_{\Theta_0} L(\theta|\mathbf{x}, \mathbf{y})}{\sup_{\Theta} L(\theta|\mathbf{x}, \mathbf{y})} = \frac{\sup_\theta \prod_{i=1}^n \frac{1}{\theta} e^{-x_i/\theta} \prod_{j=1}^m \frac{1}{\theta} e^{-y_j/\theta}}{\sup_{\theta,\mu} \prod_{i=1}^n \frac{1}{\theta} e^{-x_i/\theta} \prod_{j=1}^m \frac{1}{\mu} e^{-y_j/\mu}} \\
&= \frac{\sup_\theta \frac{1}{\theta^{m+n}} \exp\left\{-\left(\sum_{i=1}^n x_i + \sum_{j=1}^m y_j\right)\Big/\theta\right\}}{\sup_{\theta,\mu} \frac{1}{\theta^n} \exp\left\{-\sum_{i=1}^n x_i/\theta\right\} \frac{1}{\mu^m} \exp\left\{-\sum_{j=1}^m y_j/\mu\right\}}.
\end{aligned}$$

Differentiation will show that in the numerator $\hat{\theta}_0 = (\sum_i x_i + \sum_j y_j)/(n+m)$, while in the denominator $\hat{\theta} = \bar{x}$ and $\hat{\mu} = \bar{y}$. Therefore,

$$\begin{aligned}\lambda(\mathbf{x},\mathbf{y}) &= \frac{\left(\frac{n+m}{\sum_i x_i+\sum_j y_j}\right)^{n+m} \exp\left\{-\left(\frac{n+m}{\sum_i x_i+\sum_j y_j}\right)\left(\sum_i x_i+\sum_j y_j\right)\right\}}{\left(\frac{n}{\sum_i x_i}\right)^n \exp\left\{-\left(\frac{n}{\sum_i x_i}\right)\sum_i x_i\right\}\left(\frac{m}{\sum_j y_j}\right)^m \exp\left\{-\left(\frac{m}{\sum_j y_j}\right)\sum_j y_j\right\}} \\ &= \frac{(n+m)^{n+m}}{n^n m^m}\frac{\left(\sum_i x_i\right)^n\left(\sum_j y_j\right)^m}{\left(\sum_i x_i+\sum_j y_j\right)^{n+m}}.\end{aligned}$$

And the LRT is to reject H_0 if $\lambda(\mathbf{x},\mathbf{y}) \le c$.

b.

$$\lambda = \frac{(n+m)^{n+m}}{n^n m^m}\left(\frac{\sum_i x_i}{\sum_i x_i+\sum_j y_j}\right)^n\left(\frac{\sum_j y_j}{\sum_i x_i+\sum_j y_j}\right)^m = \frac{(n+m)^{n+m}}{n^n m^m}T^n(1-T)^m.$$

Therefore λ is a function of T. λ is a unimodal function of T which is maximized when $T = \frac{n}{m+n}$. Rejection for $\lambda \le c$ is equivalent to rejection for $T \le a$ or $T \ge b$, where a and b are constants that satisfy $a^n(1-a)^m = b^n(1-b)^m$.

c. When H_0 is true, $\sum_i X_i \sim \text{gamma}(n,\theta)$ and $\sum_j Y_j \sim \text{gamma}(m,\theta)$ and they are independent. So by an extension of Exercise 4.19b, $T \sim \text{beta}(n,m)$.

8.10

8.10 设 X_1，…，X_n 是 iid Poisson 分布 Poisson(λ) 的，其中 $\lambda\sim$gamma(α，β)，这个先验分布是 Poisson 分布的共轭族. 在习题 7.24 中求过 λ 的后验分布，包括后验均值与后验方差. 现在考虑 H_0：$\lambda\leq\lambda_0$ 对 H_1：$\lambda>\lambda_0$ 的 Bayes 检验.

(a) 计算 H_0 和 H_1 的后验概率表达式.

(b) 如果 $\alpha=\frac{5}{2}$ 和 $\beta=2$，则先验分布是自由度是 5 的 χ^2 分布. 解释 χ^2 表如何能够用来进行一个 Bayes 检验.

8.10 Let $Y = \sum_i X_i$. The posterior distribution of $\lambda|y$ is gamma $(y+\alpha, \beta/(\beta+1))$.

a.

$$P(\lambda \le \lambda_0|y) = \frac{(\beta+1)^{y+\alpha}}{\Gamma(y+\alpha)\beta^{y+\alpha}}\int_0^{\lambda_0} t^{y+\alpha-1}e^{-t(\beta+1)/\beta}\,dt.$$

$P(\lambda > \lambda_0|y) = 1 - P(\lambda \le \lambda_0|y)$.

b. Because $\beta/(\beta+1)$ is a scale parameter in the posterior distribution, $(2(\beta+1)\lambda/\beta)|y$ has a gamma$(y+\alpha, 2)$ distribution. If 2α is an integer, this is a $\chi^2_{2y+2\alpha}$ distribution. So, for $\alpha = 5/2$ and $\beta = 2$,

$$P(\lambda \le \lambda_0|y) = P\left(\frac{2(\beta+1)\lambda}{\beta} \le \frac{2(\beta+1)\lambda_0}{\beta}\middle| y\right) = P(\chi^2_{2y+5} \le 3\lambda_0).$$

8.12

8.12 对于从一个均值为 μ 和已知的方差为 σ^2 的正态总体中抽取的样本量为 1，4，16，64，100 的样本，画出以下假设的 LRT 的功效函数图形，取 α=0.05.

(a) H_0：$\mu\leq0$ 对 H_1：$\mu>0$

(b) H_0：$\mu=0$ 对 H_1：$\mu\neq0$

hint：求功效函数就好

8.12 a. For $H_0: \mu \le 0$ vs. $H_1: \mu > 0$ the LRT is to reject H_0 if $\bar{x} > c\sigma/\sqrt{n}$ (Example 8.3.3). For $\alpha = .05$ take $c = 1.645$. The power function is

$$\beta(\mu) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > 1.645 - \frac{\mu}{\sigma/\sqrt{n}}\right) = P\left(Z > 1.645 - \frac{\sqrt{n}\mu}{\sigma}\right).$$

Note that the power will equal .5 when $\mu = 1.645\sigma/\sqrt{n}$.

b. For $H_0: \mu = 0$ vs. $H_A: \mu \ne 0$ the LRT is to reject H_0 if $|\bar{x}| > c\sigma/\sqrt{n}$ (Example 8.2.2). For $\alpha = .05$ take $c = 1.96$. The power function is

$$\beta(\mu) = P\left(-1.96 - \sqrt{n}\mu/\sigma \le Z \le 1.96 + \sqrt{n}\mu/\sigma\right).$$

In this case, $\mu = \pm 1.96\sigma/\sqrt{n}$ gives power of approximately .5.

8.14

8.14 对于 Bernoulli（p）变量的随机样本 X_1，…，X_n，欲检验

$$H_0: p=0.49 \text{ 对 } H_1: p=0.51$$

利用中心极限定理近似地决定让两类错误的概率大约都是 0.01 所需要的样本量.使用的检验是如果 $\sum_{i=1}^{n} X_i$ 过大就拒绝 H_0.

8.14 The CLT tells us that $Z = (\sum_i X_i - np)/\sqrt{np(1-p)}$ is approximately $\mathrm{n}(0,1)$. For a test that rejects H_0 when $\sum_i X_i > c$, we need to find c and n to satisfy

$$P\left(Z > \frac{c-n(.49)}{\sqrt{n(.49)(.51)}}\right) = .01 \quad \text{and} \quad P\left(Z > \frac{c-n(.51)}{\sqrt{n(.51)(.49)}}\right) = .99.$$

We thus want

$$\frac{c-n(.49)}{\sqrt{n(.49)(.51)}} = 2.33 \quad \text{and} \quad \frac{c-n(.51)}{\sqrt{n(.51)(.49)}} = -2.33.$$

Solving these equations gives $n = 13{,}567$ and $c = 6{,}783.5$.

8.15

8.15 证明：基于来自正态总体 $\mathrm{n}(0, \sigma^2)$ 的一组随机样本 X_1，…，X_n，关于 $H_0: \sigma=\sigma_0$ 对 $H_1: \sigma=\sigma_1$（其中 $\sigma_0<\sigma_1$）的最大功效检验是

$$\phi\left(\sum X_i^2\right)=\begin{cases}1 & \text{若} \sum X_i^2 > c \\ 0 & \text{若} \sum X_i^2 \leqslant c\end{cases}$$

对于一个给定的犯第一类错误概率值 α，给出决定 c 值的表达式.

8.15 From the Neyman-Pearson lemma the UMP test rejects H_0 if

$$\frac{f(x \mid \sigma_1)}{f(x \mid \sigma_0)} = \frac{(2\pi\sigma_1^2)^{-n/2} e^{-\Sigma_i x_i^2/(2\sigma_1^2)}}{(2\pi\sigma_0^2)^{-n/2} e^{-\Sigma_i x_i^2/(2\sigma_0^2)}} = \left(\frac{\sigma_0}{\sigma_1}\right)^n \exp\left\{\frac{1}{2}\sum_i x_i^2 \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)\right\} > k$$

for some $k \ge 0$. After some algebra, this is equivalent to rejecting if

$$\sum_i x_i^2 > \frac{2\log\left(k\left(\sigma_1/\sigma_0\right)^n\right)}{\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)} = c \qquad \left(\text{because } \frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} > 0\right).$$

This is the UMP test of size α, where $\alpha = P_{\sigma_0}(\sum_i X_i^2 > c)$. To determine c to obtain a specified α, use the fact that $\sum_i X_i^2/\sigma_0^2 \sim \chi_n^2$. Thus

$$\alpha = P_{\sigma_0}\left(\sum_i X_i^2/\sigma_0^2 > c/\sigma_0^2\right) = P\left(\chi_n^2 > c/\sigma_0^2\right),$$

so we must have $c/\sigma_0^2 = \chi^2_{n,\alpha}$, which means $c = \sigma_0^2 \chi^2_{n,\alpha}$.

8.17

8.17 设 X_1，…，X_n 是 iid 的，具有贝塔分布概率密度函数 beta（μ，1）；Y_1，…，Y_m 是 iid 的，具有概率密度函数 beta（θ，1）．还假定这两组样本是相互独立的．

（a）求：关于 H_0：$\mu=\theta$ 对 H_1：$\mu\neq\theta$ 的 LRT.

（b）证明（a）中的检验可以基于统计量

$$T=\frac{\sum\log X_i}{\sum\log X_i+\sum\log Y_i}$$

（c）求：当 H_0 真时 T 的分布，然后说明怎样得到一个真实水平为 $\alpha=0.10$ 的检验.

8.17 a. The likelihood function is

$$L(\mu,\theta|\mathbf{x},\mathbf{y})=\mu^n\left(\prod_i x_i\right)^{\mu-1}\theta^n\left(\prod_j y_j\right)^{\theta-1}.$$

Maximizing, by differentiating the log-likelihood, yields the MLEs

$$\hat{\mu}=-\frac{n}{\sum_i\log x_i}\quad\text{and}\quad\hat{\theta}=-\frac{m}{\sum_j\log y_j}.$$

Under H_0, the likelihood is

$$L(\theta|\mathbf{x},\mathbf{y})=\theta^{n+m}\left(\prod_i x_i\prod_j y_j\right)^{\theta-1},$$

and maximizing as above yields the restricted MLE,

$$\hat{\theta}_0=-\frac{n+m}{\sum_i\log x_i+\sum_j\log y_j}.$$

The LRT statistic is

$$\lambda(\mathbf{x},\mathbf{y})=\frac{\hat{\theta}_0^{m+n}}{\hat{\mu}^n\hat{\theta}^m}\left(\prod_i x_i\right)^{\hat{\theta}_0-\hat{\mu}}\left(\prod_j y_j\right)^{\hat{\theta}_0-\hat{\theta}}.$$

b. Substituting in the formulas for $\hat{\theta}$, $\hat{\mu}$ and $\hat{\theta}_0$ yields $\left(\prod_i x_i\right)^{\hat{\theta}_0-\hat{\mu}}\left(\prod_j y_j\right)^{\hat{\theta}_0-\hat{\theta}}=1$ and

$$\lambda(\mathbf{x},\mathbf{y})=\frac{\hat{\theta}_0^{m+n}}{\hat{\mu}^n\hat{\theta}^m}=\frac{\hat{\theta}_0^n}{\hat{\mu}^n}\frac{\hat{\theta}_0^m}{\hat{\theta}^m}=\left(\frac{m+n}{m}\right)^m\left(\frac{m+n}{n}\right)^n(1-T)^mT^n.$$

This is a unimodal function of T. So rejecting if $\lambda(\mathbf{x},\mathbf{y})\leq c$ is equivalent to rejecting if $T\leq c_1$ or $T\geq c_2$, where c_1 and c_2 are appropriately chosen constants.

c. Simple transformations yield $-\log X_i\sim\text{exponential}(1/\mu)$ and $-\log Y_i\sim\text{exponential}(1/\theta)$. Therefore, $T=W/(W+V)$ where W and V are independent, $W\sim\text{gamma}(n,1/\mu)$ and $V\sim\text{gamma}(m,1/\theta)$. Under H_0, the scale parameters of W and V are equal. Then, a simple generalization of Exercise 4.19b yields $T\sim\text{beta}(n,m)$. The constants c_1 and c_2 are determined by the two equations

$$P(T\leq c_1)+P(T\geq c_2)=\alpha\quad\text{and}\quad(1-c_1)^mc_1^n=(1-c_2)^mc_2^n.$$

8.19

8.19 随机变量 X 具有概率密度函数 $f(x)=e^{-x}$，$x>0$. 对随机变量 $Y=X^{\theta}$ 取得一次观测，而需要构建关于 H_0：$\theta=1$ 对 H_1：$\theta=2$ 的一个检验. 求：UMP 水平为$\alpha=0.10$ 的检验并计算犯第二类错误的概率.

8.19 The pdf of Y is

$$f(y|\theta)=\frac{1}{\theta}y^{(1/\theta)-1}e^{-y^{1/\theta}},\quad y>0.$$

By the Neyman-Pearson Lemma, the UMP test will reject if

$$\frac{1}{2}y^{-1/2}e^{y-y^{1/2}}=\frac{f(y|2)}{f(y|1)}>k.$$

To see the form of this rejection region, we compute

$$\frac{d}{dy}\left(\frac{1}{2}y^{-1/2}e^{y-y^{1/2}}\right)=\frac{1}{2}y^{-3/2}e^{y-y^{1/2}}\left(y-\frac{y^{1/2}}{2}-\frac{1}{2}\right)$$

which is negative for $y<1$ and positive for $y>1$. Thus $f(y|2)/f(y|1)$ is decreasing for $y\le 1$ and increasing for $y\ge 1$. Hence, rejecting for $f(y|2)/f(y|1)>k$ is equivalent to rejecting for $y\le c_0$ or $y\ge c_1$. To obtain a size α test, the constants c_0 and c_1 must satisfy

$$\alpha=P(Y\le c_0|\theta=1)+P(Y\ge c_1|\theta=1)=1-e^{-c_0}+e^{-c_1}\quad\text{and}\quad\frac{f(c_0|2)}{f(c_0|1)}=\frac{f(c_1|2)}{f(c_1|1)}.$$

Solving these two equations numerically, for $\alpha=.10$, yields $c_0=.076546$ and $c_1=3.637798$. The Type II error probability is

$$P(c_0<Y<c_1|\theta=2)=\int_{c_0}^{c_1}\frac{1}{2}y^{-1/2}e^{-y^{1/2}}\,dy=-e^{-y^{1/2}}\Big|_{c_0}^{c_1}=.609824.$$

8.20

8.20 设一个随机变量 X 在 H_0 和 H_1 之下的概率质量函数由下表给出

x	1	2	3	4	5	6	7
$f(x\|H_0)$	0.01	0.01	0.01	0.01	0.01	0.01	0.94
$f(x\|H_1)$	0.06	0.05	0.04	0.03	0.02	0.01	0.79

利用 Neyman-Pearson 引理求 H_0 对 H_1 的真实水平为 $\alpha=0.04$ 的最大功效检验. 计算这个检验犯第二类错误的概率.

8.20 By the Neyman-Pearson Lemma, the UMP test rejects for large values of $f(x|H_1)/f(x|H_0)$. Computing this ratio we obtain

x	1	2	3	4	5	6	7
$\frac{f(x\|H_1)}{f(x\|H_0)}$	6	5	4	3	2	1	.84

The ratio is decreasing in x. So rejecting for large values of $f(x|H_1)/f(x|H_0)$ corresponds to rejecting for small values of x. To get a size α test, we need to choose c so that $P(X\le c|H_0)=\alpha$. The value $c=4$ gives the UMP size $\alpha=.04$ test. The Type II error probability is $P(X=5,6,7|H_1)=.82$.

8.22

8.22 设 X_1，…，X_n 是 iid Bernoulli（p）的.

（a）求：$H_0: p=\frac{1}{2}$ 对 $H_1: p=\frac{1}{4}$ 的真实水平为 $\alpha=0.0547$ 的最大功效检验. 求这个检验的功效.

（b）关于检验 $H_0: p\leqslant\frac{1}{2}$ 对 $H_1: p>\frac{1}{2}$，求："若 $\sum_{i=1}^{10} X_i>6$，就拒绝 H_0" 的检验的真实水平并且勾画这个检验的功效函数略图.

（c）对于什么样的水平 α，一定存在一个（a）中假设的 UMP 检验?

8.22 a. From Corollary 8.3.13 we can base the test on $\sum_i X_i$, the sufficient statistic. Let $Y = \sum_i X_i \sim \text{binomial}(10, p)$ and let $f(y|p)$ denote the pmf of Y. By Corollary 8.3.13, a test that rejects if $f(y|1/4)/f(y|1/2) > k$ is UMP of its size. By Exercise 8.25c, the ratio $f(y|1/2)/f(y|1/4)$ is increasing in y. So the ratio $f(y|1/4)/f(y|1/2)$ is decreasing in y, and rejecting for large value of the ratio is equivalent to rejecting for small values of y. To get $\alpha = .0547$, we must find c such that $P(Y \leq c|p = 1/2) = .0547$. Trying values $c = 0, 1, \ldots$, we find that for $c = 2$, $P(Y \leq 2|p = 1/2) = .0547$. So the test that rejects if $Y \leq 2$ is the UMP size $\alpha = .0547$ test. The power of the test is $P(Y \leq 2|p = 1/4) \approx .526$.

b. The size of the test is $P(Y \geq 6|p = 1/2) = \sum_{k=6}^{10} \binom{10}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{10-k} \approx .377$. The power function is $\beta(\theta) = \sum_{k=6}^{10} \binom{10}{k} \theta^k (1-\theta)^{10-k}$

c. There is a nonrandomized UMP test for all α levels corresponding to the probabilities $P(Y \leq i|p = 1/2)$, where i is an integer. For $n = 10$, α can have any of the values 0, $\frac{1}{1024}$, $\frac{11}{1024}$, $\frac{56}{1024}$, $\frac{176}{1024}$, $\frac{386}{1024}$, $\frac{638}{1024}$, $\frac{848}{1024}$, $\frac{968}{1024}$, $\frac{1013}{1024}$, $\frac{1023}{1024}$, and 1.

8.29

8.29 设 X 是来自一个 Cauchy 分布 Cauchy（θ）的一次观测.

（a）证明：这个族没有 MLR.

（b）证明：检验

$$\phi(x)=\begin{cases}1 & \text{若 } 1<x<3\\ 0 & \text{其他}\end{cases}$$

在它的真实水平上是 $H_0: \theta=0$ 对 $H_1: \theta=1$ 的最大功效检验. 计算犯第一类和第二类错误的概率.

（c）证明或驳斥：（b）中的检验是关于 $H_0: \theta\leqslant 0$ 对 $H_1: \theta>0$ 的 UMP 检验. 就一般 Cauchy 位置参数族而言，关于它的 UMP 检验能够讲什么?

8.29 a. Let $\theta_2 > \theta_1$. Then

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = \frac{1+(x-\theta_1)^2}{1+(x-\theta_2)^2} = \frac{1+(1+\theta_1)^2/x^2 - 2\theta_1/x}{1+(1+\theta_2)^2/x^2 - 2\theta_2/x}.$$

The limit of this ratio as $x \to \infty$ or as $x \to -\infty$ is 1. So the ratio cannot be monotone increasing (or decreasing) between $-\infty$ and ∞. Thus, the family does not have MLR.

b. By the Neyman-Pearson Lemma, a test will be UMP if it rejects when $f(x|1)/f(x|0) > k$, for some constant k. Examination of the derivative shows that $f(x|1)/f(x|0)$ is decreasing for $x \leq (1-\sqrt{5})/2 = -.618$, is increasing for $(1-\sqrt{5})/2 \leq x \leq (1+\sqrt{5})/2 = 1.618$, and is decreasing for $(1+\sqrt{5})/2 \leq x$. Furthermore, $f(1|1)/f(1|0) = f(3|1)/f(3|0) = 2$. So rejecting if $f(x|1)/f(x|0) > 2$ is equivalent to rejecting if $1 < x < 3$. Thus, the given test is UMP of its size. The size of the test is

$$P(1 < X < 3|\theta = 0) = \int_1^3 \frac{1}{\pi}\frac{1}{1+x^2}\,dx = \frac{1}{\pi}\arctan x\Big|_1^3 \approx .1476.$$

The Type II error probability is

$$1 - P(1 < X < 3|\theta = 1) = 1 - \int_1^3 \frac{1}{\pi} \frac{1}{1+(x-1)^2}\, dx = 1 - \frac{1}{\pi}\arctan(x-1)\Big|_1^3 \approx .6476.$$

c. We will not have $f(1|\theta)/f(1|0) = f(3|\theta)/f(3|0)$ for any other value of $\theta \neq 1$. Try $\theta = 2$, for example. So the rejection region $1 < x < 3$ will not be most powerful at any other value of θ. The test is not UMP for testing $H_0: \theta \leq 0$ versus $H_1: \theta > 0$.

8.33

8.33 设 X_1，…，X_n 是来自 $U(\theta, \theta+1)$ 分布的一组随机样本．欲检验 H_0：$\theta=0$ 对 H_1：$\theta>0$，使用检验

如果 $Y_n \geqslant 1$ 或 $Y_1 \geqslant k$ 就拒绝 H_0

其中 k 是一个常数，$Y_1=\min\{X_1, \cdots, X_n\}$，$Y_n=\max\{X_1, \cdots, X_n\}$．

(a) 确定 k 以使检验具有真实水平 α.

(b) 求 (a) 中检验的功效函数的表达式.

(c) 证明：这个检验是一个 UMP 真实水平为 α 的检验.

(d) 求 n 和 k 的值使得 UMP 水平为 0.10 的检验的功效在 $\theta>1$ 至少为 0.8.

8.33 a. From Theorems 5.4.4 and 5.4.6, the marginal pdf of Y_1 and the joint pdf of (Y_1, Y_n) are

$$\begin{aligned} f(y_1|\theta) &= n(1-(y_1-\theta))^{n-1}, \quad \theta < y_1 < \theta+1, \\ f(y_1, y_n|\theta) &= n(n-1)(y_n - y_1)^{n-2}, \quad \theta < y_1 < y_n < \theta + 1. \end{aligned}$$

Under H_0, $P(Y_n \geq 1) = 0$. So

$$\alpha = P(Y_1 \geq k|0) = \int_k^1 n(1-y_1)^{n-1}\, dy_1 = (1-k)^n.$$

Thus, use $k = 1 - \alpha^{1/n}$ to have a size α test.

b. For $\theta \le k-1$, $\beta(\theta) = 0$. For $k-1 < \theta \le 0$,

$$\beta(\theta) = \int_k^{\theta+1} n(1-(y_1-\theta))^{n-1}\,dy_1 = (1-k+\theta)^n.$$

For $0 < \theta \le k$,

$$\begin{aligned}\beta(\theta) &= \int_k^{\theta+1} n(1-(y_1-\theta))^{n-1}\,dy_1 + \int_\theta^k \int_1^{\theta+1} n(n-1)(y_n-y_1)^{n-2}\,dy_n\,dy_1 \\ &= \alpha + 1 - (1-\theta)^n.\end{aligned}$$

And for $k < \theta$, $\beta(\theta) = 1$.

c. (Y_1, Y_n) are sufficient statistics. So we can attempt to find a UMP test using Corollary 8.3.13 and the joint pdf $f(y_1, y_n|\theta)$ in part (a). For $0 < \theta < 1$, the ratio of pdfs is

$$\frac{f(y_1,y_n|\theta)}{f(y_1,y_n|0)} = \begin{cases} 0 & \text{if } 0 < y_1 \le \theta,\ y_1 < y_n < 1 \\ 1 & \text{if } \theta < y_1 < y_n < 1 \\ \infty & \text{if } 1 \le y_n < \theta+1, \theta < y_1 < y_n. \end{cases}$$

For $1 \le \theta$, the ratio of pdfs is

$$\frac{f(y_1,y_n|\theta)}{f(y_1,y_n|0)} = \begin{cases} 0 & \text{if } y_1 < y_n < 1 \\ \infty & \text{if } \theta < y_1 < y_n < \theta+1. \end{cases}$$

For $0 < \theta < k$, use $k' = 1$. The given test always rejects if $f(y_1,y_n|\theta)/f(y_1,y_n|0) > 1$ and always accepts if $f(y_1,y_n|\theta)/f(y_1,y_n|0) < 1$. For $\theta \ge k$, use $k' = 0$. The given test always rejects if $f(y_1,y_n|\theta)/f(y_1,y_n|0) > 0$ and always accepts if $f(y_1,y_n|\theta)/f(y_1,y_n|0) < 0$. Thus the given test is UMP by Corollary 8.3.13.

d. According to the power function in part (b), $\beta(\theta) = 1$ for all $\theta \ge k = 1-\alpha^{1/n}$. So these conditions are satisfied for any n.

8.39

8.39 设 (X_1, Y_1), …, (X_n, Y_n) 是一组来自参数为 μ_X, μ_Y, σ_X^2, σ_Y^2, ρ 的二元正态分布的随机样本. 我们的兴趣在于检验

$$H_0: \mu_X = \mu_Y \text{ 对 } H_1: \mu_X \ne \mu_Y$$

(a) 证明：随机变量 $W_i = X_i - Y_i$ 是 iid $\mathrm{n}(\mu_W, \sigma_W^2)$ 的，其中 $\mu_W = \mu_X - \mu_Y$，$\sigma_W^2 = \sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_X\sigma_Y$.

(b) 证明上面的假设能够用以下统计量检验

$$T_W = \frac{\overline{W}}{\sqrt{\frac{1}{n}S_W^2}}$$

其中 $\overline{W} = \frac{1}{n}\sum_{i=1}^n W_i$ 而 $S_W^2 = \frac{1}{(n-1)}\sum_{i=1}^n (W_i - \overline{W})^2$. 进一步证明，在 H_0 下，$T_W \sim$ 自由度为 $n-1$ 的 t 分布.（这个检验被称为配对的 t 检验.）

8.39 a. From Exercise 4.45c, $W_i = X_i - Y_i \sim \mathrm{n}(\mu_W, \sigma_W^2)$, where $\mu_X - \mu_Y = \mu_W$ and $\sigma_X^2 + \sigma_Y^2 - \rho\sigma_X\sigma_Y = \sigma_W^2$. The W_is are independent because the pairs (X_i, Y_i) are.

b. The hypotheses are equivalent to $H_0: \mu_W = 0$ vs $H_1: \mu_W \ne 0$, and, from Exercise 8.38, if we reject H_0 when $|\,\overline{W}\,| > t_{n-1,\alpha/2}\sqrt{S_W^2/n}$, this is the LRT (based on $W_1, \ldots, W_n$) of size α. (Note that if $\rho > 0$, $\operatorname{Var} W_i$ can be small and the test will have good power.)