

FDU 数理统计 3. 假设检验

本文参考以下教材:

- 《数理统计讲义》(郑明, 陈子毅, 汪嘉冈) 第 3 章

3.1 基本概念

3.1.1 检验问题

设总体的可能分布族为 $\mathcal{F}_\xi = \{F_\xi(\theta) : \theta \in \Theta\}$ (或样本的可能分布族为 $\mathcal{F}_X = \{F_X(\theta) : \theta \in \Theta\}$)
设总体的真分布为 $F_\xi(\theta_{\text{true}})$ (或样本的真分布为 $F_X(\theta_{\text{true}})$)
设 Θ_0, Θ_1 是参数空间 Θ 的两个互不相交的非空子集.

- 我们称论断 " $\theta_{\text{true}} \in \Theta_0$ " 为**零假设** (null hypothesis), 记为 $H_0 : \theta_{\text{true}} \in \Theta_0$
- 我们称论断 " $\theta_{\text{true}} \in \Theta_1$ " 为**备择假设** (alternative hypothesis), 记为 $H_1 : \theta_{\text{true}} \in \Theta_1$

我们使用**简单假设**和**复合假设** (composite hypothesis) 的名称来区分 Θ_0, Θ_1 是否是单元素集.
例如 $\Theta_0 = \{\theta_0\}$ 的情况下, 零假设便可写为 $H_0 : \theta_{\text{true}} = \theta_0$
我们倾向于保护零假设, 因此有时会设置其为简单假设,
而备择假设通常不会设为简单假设, 否则本末倒置了.

- 对于一维实参数的情况 (即 $\Theta \subseteq \mathbb{R}$),
称 $\Theta_1 = \{\theta : \theta < \theta_0\}$ 和 $\Theta_1 = \{\theta : \theta > \theta_0\}$ 为**单侧 (one-side) 备择假设**
称 $\Theta_1 = \{\theta : \theta \neq \theta_0\}$ 或 $\Theta_1 = \{\theta : \theta < \theta_1 \text{ or } \theta > \theta_2\}$ 为**双侧 (two-side) 备择假设**
- 两种最简单的检验问题:
 - 都是简单假设: $H_0 : \theta = \theta_0 \leftrightarrow H_1 : \theta = \theta_1$
 - 都是单边假设: $H_0 : \theta \leq \theta_0 \leftrightarrow H_1 : \theta > \theta_0$

规定了零假设 $H_0 : \theta \in \Theta_0$ 和与之对立的备择假设 $H_1 : \theta \in \Theta_1$,
就确定一个假设检验问题, 记为 $H_0 : \theta \in \Theta_0 \leftrightarrow H_1 : \theta \in \Theta_1$
实施假设检验就是根据观测到的样本 X 对 H_0 和 H_1 成立的可能性做推断,
最终拒绝 H_0 或不拒绝 H_0 ("假设没有被推翻" 不代表可以 "接受假设")

- 若在零假设 H_0 成立的前提条件下, 样本 X 取到我们的观测值 x 的概率非常小,
则我们拒绝零假设 H_0
那么对于事件是否在一次试验中发生的问题, 多小的概率可以称为是 "小概率" 呢?
我们通常约定**显著水平** (significance level) α 为一个小的正数, 通常设为 0.05.
任何概率小于 α 的事件都将认为是在一次试验中不可能发生的事件.
注意与区间估计的**置信水平** $1 - \alpha$ 区分开.

假设检验一般分为以下步骤:

- 确定检验问题 $H_0 : \theta \in \Theta_0 \leftrightarrow H_1 : \theta \in \Theta_1$
- 选定一个对检验问题敏感的统计量 $T(X)$,
它在零假设 H_0 成立时的分布必须是已知的 (或者容易近似估计的),
因而可用于确定零假设 H_0 成立时统计量 $T(X)$ 的拒绝范围.
- 对于给定的显著水平 α , 确定统计量的拒绝范围 (违反零假设 H_0 的范围),
使得对于任意 $\theta \in \Theta_0$ 都有 $P_\theta\{T(X) \text{ takes a value in the rejection region}\} \leq \alpha$

- 然后我们根据样本 X 的观测值 x 作如下的判断:
若 $T(x)$ 落入拒绝范围内, 则**拒绝** (reject) 零假设 H_0 .

对于同一个假设检验问题, 不同的检验统计量和拒绝范围构成不同的检验法.

我们记统计量 $T(X)$ 的**拒绝范围**为 $\mathcal{T}_{\text{reject}}$

对应回**样本空间** \mathcal{X} 上, 我们记**拒绝域** (reject region) $\mathcal{X}_{\text{reject}} = \{x \in \mathcal{X} : T(x) \in \mathcal{T}_{\text{reject}}\}$

为了讨论方便, 我们引入**检验函数** (test function) 的概念:

$$\phi(X) = \begin{cases} 1, & X \in \mathcal{X}_{\text{reject}} \\ 0, & X \notin \mathcal{X}_{\text{reject}} \end{cases} = \begin{cases} 1, & T(X) \in \mathcal{T}_{\text{reject}} \\ 0, & T(X) \notin \mathcal{T}_{\text{reject}} \end{cases}$$

因此检验法显著水平就要求:

当对于任意 $\theta \in \Theta_0$ 都有 $\mathbb{E}_\theta[\phi(X)] \leq \alpha$ 成立时, 才不拒绝零假设 H_0

- 对于离散分布, 上述划分可能会失效,
此时需要引入一些随机化处理, 使 $\phi(X)$ 在 $[0, 1]$ 连续区间上取值.
这称为**随机的** (randomized) 检验法.
前文描述的是**非随机的** (non-randomized) 检验法.

3.1.2 两类错误

- 在零假设 H_0 成立时, 由于样本观测值 x 落入拒绝域 $\mathcal{X}_{\text{reject}}$ 而拒绝零假设 H_0 ,
这类错误称为**第一类错误** (error of type I, **拒真**, 合格产品被拒收了)
对于任意 $\theta \in \Theta_0$, $P_\theta\{\text{error of type 1}\} = P_\theta\{X \in \mathcal{X}_{\text{reject}}\} = \mathbb{E}_\theta[\phi(X)]$
- 在零假设 H_0 不成立时, 由于样本观测值 x 没有落入拒绝域 $\mathcal{X}_{\text{reject}}$ 而**没有拒绝**零假设 H_0 ,
这类错误称为**第二类错误** (error of type II, **受伪**, 不合格产品没有被拒收)
对于任意 $\theta \in \Theta_1$, $P_\theta\{\text{error of type 2}\} = P_\theta\{X \notin \mathcal{X}_{\text{reject}}\} = 1 - \mathbb{E}_\theta[\phi(X)]$
(数理统计讲义上将 "没有拒绝" 称为 "接受", 这点我不太认同)

表 3.1-1

客观情况 \ 采取决策	接受原假设 H_0	拒绝原假设 H_0
H_0 成立	正确决策	第一类错误
H_1 成立	第二类错误	正确决策

客观上不会两类错误不会同时发生 (即它们是互斥的),

理想的检验法要求发生两类错误的概率都要小.

但这个要求是自相矛盾的:

若将拒绝域 $\mathcal{X}_{\text{reject}}$ 缩小, 则可以减小第一类错误概率, 但必然导致第二类错误概率增加.

所以要同时限制两类错误的概率, 就可能需要增加样本量.

当样本量固定时, 我们首先要限制的是第一类错误概率.

我们对于 $\theta \in \Theta$ 定义检验法 (或检验函数 ϕ) 的**功效函数** (power function)

$$\gamma_\phi(\theta) = \mathbb{E}_\theta[\phi(X)] = P_\theta\{X \in \mathcal{X}_{\text{reject}}\}$$

它表征分布参数为 θ 时, 检验法拒绝零假设 H_0 的可能性.

- 当 $\theta \in \Theta_0$ 时 $\gamma_\phi(\theta)$ 就是 θ 对应的第一类错误概率.
- 当 $\theta \in \Theta_1$ 时 $1 - \gamma_\phi(\theta)$ 就是 θ 对应的第二类错误概率.

例如正态分布 (方差已知) 的两点检验 $H_0 : \mu = \mu_0 \leftrightarrow H_1 : \mu = \mu_1$ (其中 $\mu_0 < \mu_1$)

- 一旦拒绝域缩小 (即第一个子图中两条决策边界远离对称中心 $\mu = \mu_0$), 则第一类错误概率 (图中记为 α) 会减小, 但第二类错误概率 (图中记为 β) 会增大.
- 其实我觉得这张图 H_1 应当扩充为 $H_1: \mu \in \{\mu_1, \mu_2\}$, 其中 $\mu_2 < \mu_0 < \mu_1$ 这样第一张子图左侧的拒绝域部分代表着 H_1 中 $\mu = \mu_2$ 的情况比 $H_0: \mu = \mu_0$ 更可能成立.

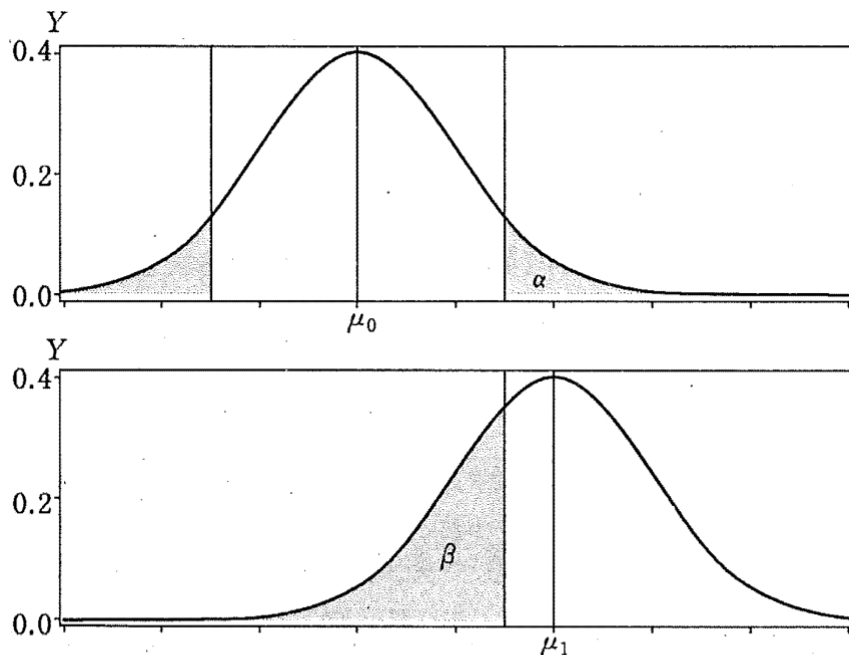


图 3.1-1 正态均值检验的两类错误

因此检验法显著水平 α 就要求:

当对于任意 $\theta \in \Theta_0$ 都有**第一类错误概率** $\gamma_\phi(\theta) = \mathbb{E}_\theta[\phi(X)] \leq \alpha$ 成立时, 才不拒绝零假设 H_0 也就是说, 检验法的**显著水平** α 是其第一类错误概率的一个**上界**.

上式也可写为:

当 $\sup_{\theta \in \Theta_0} \gamma_\phi(\theta) = \sup_{\theta \in \Theta_0} \mathbb{E}_\theta[\phi(X)] \leq \alpha$ 时, 才不拒绝零假设 H_0

我们称 $\sup_{\theta \in \Theta_0} \gamma_\phi(\theta) = \sup_{\theta \in \Theta_0} \mathbb{E}_\theta[\phi(X)]$ 为检验法的**实际水平** (size of test)

(数理统计讲义 例 3.1.14)

规定工厂产品次品率不得超过 6%.

今在一批产品中任取 50 件发现有 8 件次品, 请问这批产品是否能够出厂?

- 我们不能根据 $\frac{8}{50} = 0.16 > 0.06$ 直接判断次品率 > 0.06
考虑两种**极端情况**: (总体分布和抽样分布天差地别)
 - 这批产品 (假设其数量成百上千) 只有 8 个次品, 但恰好抽样时全被抽取;
 - 这批产品 (假设其数量成百上千) 只有 42 个合格品, 但恰好抽样时全被抽取;

因此我们无法直接根据抽样的**点估计**说明次品率 > 0.06 还是 < 0.06

记这批产品的次品率为 π

- 零假设 $H_0: \pi \leq 0.06$
- 备择假设 $H_1: \pi > 0.06$

记 $X_i = \begin{cases} 1 & \text{if the } i\text{-th product is defective} \\ 0 & \text{otherwise} \end{cases}$

我们知道对于 Bernoulli 分布族 \bar{X} 或 $\sum_{i=1}^n X_i$ 是参数 π 的充分统计量, 因此可以基于 $\sum_{i=1}^n X_i$ 进行推断.

平均意义上, $\sum_{i=1}^n X_i$ 的均值是 50π

- 当 π 比较小时, $\sum_{i=1}^n X_i$ 取值倾向于比较小;
- 当 π 比较大时, $\sum_{i=1}^n X_i$ 取值倾向于比较大;

在零假设 $H_0: \pi \leq 0.06$ 成立时, 我们有:

$$P_{\pi} \left\{ \sum_{i=1}^{50} X_i \geq 8 \right\} \leq P_{0.06} \left\{ \sum_{i=1}^{50} X_i \geq 8 \right\} = 0.00938$$

现有样本观测值 $\sum_{i=1}^{50} x_i = 8$

如果零假设 $H_0: \pi \leq 0.06$ 成立的话, 上述样本观测出现的概率是非常小的, 几乎不可能在一次试验中出现, 因此我们拒绝零假设 $H_0: \pi \leq 0.06$

下面我们考虑显著水平 $\alpha = 0.05$ 时 $T(X) = \sum_{i=1}^{50} X_i$ 的拒绝范围:

(计算概率 $P_{H_0}\{T(X) \text{ is more extreme than the observed } T(x)\}$, 作为 $P_{H_0}\{\text{observing } T(x)\}$ 的上界)

- $P_{0.06}\{\sum_{i=1}^{50} X_i \geq 6\} = 0.07764 > 0.05 = \alpha$
- $P_{0.06}\{\sum_{i=1}^{50} X_i \geq 7\} = 0.02892 < 0.05 = \alpha$

因此对于任意 $\pi \leq 0.06$ 我们都有 $P_{\pi}\{\sum_{i=1}^{50} X_i \geq 7\} \leq P_{0.06}\{\sum_{i=1}^{50} X_i \geq 7\} < 0.05$

我们可以设定 $T(X) = \sum_{i=1}^{50} X_i$ 的拒绝范围 $\mathcal{R}_{\text{reject}} = \{t \in \mathbb{N} : t \geq 7\}$

因此样本 X 的拒绝域为 $\mathcal{X}_{\text{reject}} = \{x : T(x) = \sum_{i=1}^{50} x_i \geq 7\}$

检验函数 $\phi(X) = \begin{cases} 1 & X \in \mathcal{X}_{\text{reject}} \\ 0 & X \notin \mathcal{X}_{\text{reject}} \end{cases} = \begin{cases} 1 & T(X) \in \mathcal{R}_{\text{reject}} \\ 0 & T(X) \notin \mathcal{R}_{\text{reject}} \end{cases}$

其功效函数为:

$$\begin{aligned} \gamma_{\phi}(\pi) &= P_{\pi} \left\{ \sum_{i=1}^{50} X_i \geq 7 \right\} \\ &= P_{\pi}\{B(50, \pi) \geq 7\} \\ &= \sum_{i=7}^{50} \binom{50}{i} \pi^i (1-\pi)^{50-i} \\ &= \frac{1}{\beta(7, 44)} \int_0^{\pi} t^6 (1-t)^{43} dt \end{aligned}$$

其中 Beta 函数 $\beta(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, 最后一步实际上变换为 Beta(7, 44) 分布 (方便计算).

$\xi \sim \text{Beta}(7, 44)$ 的概率密度函数便是 $f_{\xi}(t) = \frac{1}{\beta(a, b)} t^{a-1} (1-t)^{b-1} I_{(0,1)}(t)$

注意到功效函数 $\gamma_{\phi}(\pi)$ 关于 π 是递增函数.

(这在 3.2.2 节我们会更深入地讨论)

- 当 $\pi \leq 0.06$ ($H_0: \pi \leq 0.06$) 时,
第一类错误概率 $P_{\pi}\{\text{Type-1 Error}\} = \gamma_{\phi}(\pi) \leq \gamma_{\phi}(0.06) \approx 0.02892 < 0.05 = \alpha$
- 当 $\pi > 0.06$ ($H_1: \pi > 0.06$) 时,
第二类错误概率 $P_{\pi}\{\text{Type-2 Error}\} = 1 - \gamma_{\phi}(\pi)$ 关于 π 递减.

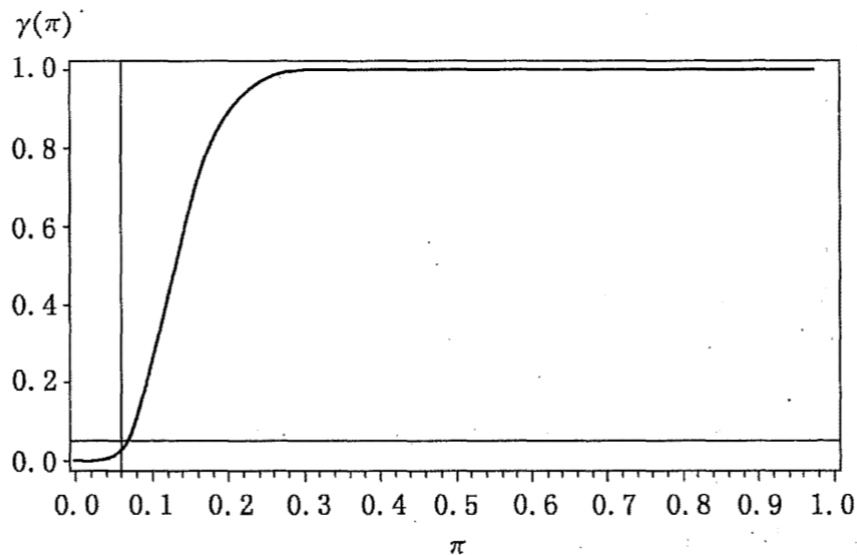


图 3.1-2 概率检验的功效函数图

(数理统计讲义 例 3.1.15)

糖厂使用包装机将糖以 100kg 的标准重量打包.

根据以往的经验, 其称重的标准差 $\sigma = 1.15\text{kg}$.

某日抽检 9 包, 其重量为 99.3, 98.7, 100.5, 101.2, 98.3, 99.7, 99.5, 102.1, 99.8.

请问包装机工作是否正常?

我们有以下疑问:

- 糖包重量是否服从正态分布? (这个涉及总体分布类型的问题我们暂不考虑)
- 在假设其总体分布服从方差为 $(1.15)^2$ 的正态分布的前提条件下, 其均值 μ 是否为 100?

设糖包重量的可能分布族为 $\{N(\mu, 1.15^2) : \mu > 0\}$

- 零假设 $H_0 : \mu = 100$ (简单假设)
- 备择假设 $H_1 : \mu \neq 100$ (双侧假设)

对于正态分布族来说, \bar{X} 或 $\sum_{i=1}^n X_i$ 是参数 μ 的充分统计量, 因此可以基于 \bar{X} 进行推断.

(计算概率 $P_{H_0}\{T(X) \text{ is more extreme than the observed } T(x)\}$, 作为 $P_{H_0}\{\text{observing } T(x)\}$ 的上界)

我们知道 $\frac{\sqrt{9}(\bar{X}-\mu)}{1.15} \stackrel{d}{=} N(0, 1)$

在原假设 $H_0 : \mu = 100$ 成立的前提条件下, 我们有 $Z = \frac{\sqrt{9}(\bar{X}-100)}{1.15} \stackrel{d}{=} N(0, 1)$

经计算可知 \bar{X} 的本次观测为:

$$\bar{x} = \frac{1}{9}(99.3 + 98.7 + 100.5 + 101.2 + 98.3 + 99.7 + 99.5 + 102.1 + 99.8) = 99.9$$

$$\text{于是 } z = \frac{\sqrt{9}(\bar{x}-100)}{1.15} \approx -0.261$$

$$P_{H_0}\{|Z| \geq |z| = 0.261\} = 2(1 - \Phi(0.261)) \approx 0.7942$$

说明本次观测 $\bar{x} = 99.9$ 虽然与 100 有偏差, 但这一事件发生的概率并不是很小, 因此拒绝零假设 $H_0 : \mu = 100$ 的证据不足.

下面我们考虑显著水平 $\alpha = 0.05$ 时 $\bar{X} = \frac{1}{9} \sum_{i=1}^9 X_i$ 的拒绝范围, 或者更一般地, 标准化后的 $T(X) = Z = \frac{\sqrt{9}(\bar{X}-100)}{1.15}$ 的拒绝范围:

- 求解 $P_{H_0}\{|Z| \geq \delta\} = 2(1 - \Phi(\delta)) = \alpha = 0.05$
 可知 $\delta = \Phi^{-1}(1 - \frac{1}{2}\alpha) = \Phi^{-1}(0.975) = 1.96$
 表明 $P_{H_0}\{|Z| \geq 1.96\} = \alpha = 0.05$
 因此我们有:

- 当 z 落入 $\{z : |z| \geq 1.96\}$ 时我们拒绝零假设 $H_0 : \mu = 100$
- 其余情况拒绝零假设 $H_0 : \mu = 100$ 的证据不足

因此 $T(X) = Z$ 的拒绝范围为 $\mathcal{T}_{\text{reject}} = \{z : |z| \geq 1.96\}$

于是拒绝域为 $\mathcal{X}_{\text{reject}} = \{x = (x_1, \dots, x_9) : |z| = \frac{\sqrt{9}|\frac{1}{9}\sum_{i=1}^9 x_i - 100|}{1.15} \geq 1.96\}$

检验函数 $\phi(X) = \begin{cases} 1 & X \in \mathcal{X}_{\text{reject}} \\ 0 & X \notin \mathcal{X}_{\text{reject}} \end{cases} = \begin{cases} 1 & T(X) \in \mathcal{T}_{\text{reject}} \\ 0 & T(X) \notin \mathcal{T}_{\text{reject}} \end{cases}$

请务必将 P_{H_0} 和 P_{μ} 的情况区分开来:

- 零假设 $H_0 : \mu = 100$ 成立的前提条件下 $Z = \frac{\sqrt{9}(\bar{X}-100)}{1.15} \stackrel{d}{=} N(0, 1)$
- 对于一般的 $\mu > 0$, 我们有 $Z = \frac{\sqrt{9}(\bar{X}-100)}{1.15} \stackrel{d}{=} N(\frac{3(\mu-100)}{1.15}, 1)$

功效函数为 (请务必将 P_{H_0} 和 P_{μ} 的情况区分开来):

$$\begin{aligned} \gamma_{\phi}(\mu) &= P_{\mu}\{|Z| \geq 1.96\} \\ &= P_{\mu}\left\{\left|N\left(\frac{3(\mu-100)}{1.15}, 1\right)\right| \geq 1.96\right\} \\ &= 1 - \Phi\left(1.96 - \frac{3(\mu-100)}{1.15}\right) + \Phi\left(-1.96 - \frac{3(\mu-100)}{1.15}\right) \end{aligned}$$

功效函数的图像如下:

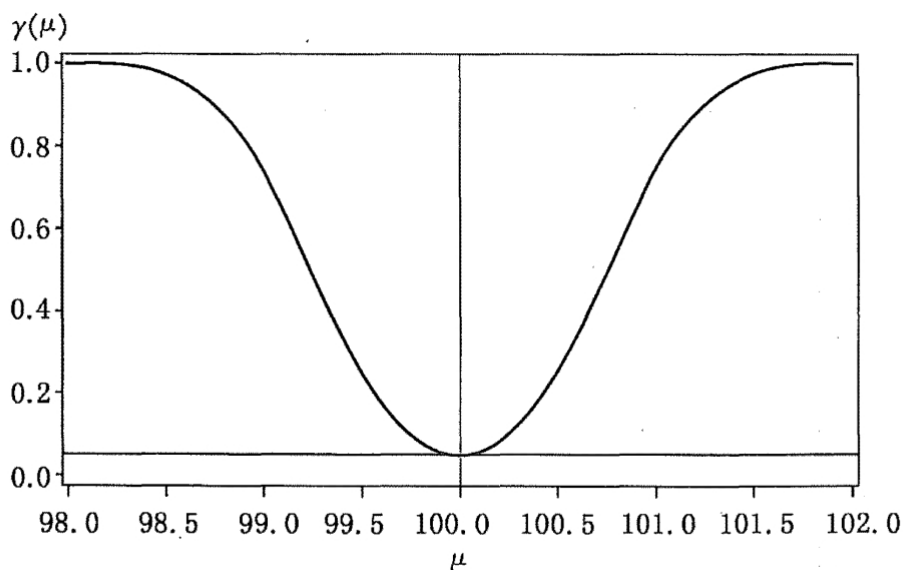


图 3.1-3 正态均值检验的功效函数图

- 当 $\mu \in \Theta_0$ (i.e. $\mu = 100$) 时, 犯第一类错误 (认为 $\mu \neq 100$) 的概率为:

$$\begin{aligned} P_{100}\{\text{error of type one}\} &= \gamma_{\phi}(100) \\ &= (1 - \Phi(1.96)) + \Phi(-1.96) \\ &= (1 - 0.975) + 0.025 \\ &= 0.05 \end{aligned}$$

- 当 $\mu \in \Theta_1$ (i.e. $\mu \neq 100$) 时, 犯第二类错误 (认为 $\mu = 100$) 的概率为:

$$\begin{aligned} P_{\mu}\{\text{error of type two}\} &= 1 - \gamma_{\phi}(\mu) \\ &= \Phi\left(1.96 - \frac{3(\mu - 100)}{1.15}\right) - \Phi\left(-1.96 - \frac{3(\mu - 100)}{1.15}\right) \end{aligned}$$

我们注意到这个双侧检验问题仍然有 "单调性" 在里面,
 $\gamma_{\phi}(\mu)$ 随着 μ 距离 $\mu_0 = 100$ 越来越远, 而越来越大, 即关于 $|\mu - \mu_0|$ 是单调递增的.
 (这在 3.2.2 节我们会更深入地讨论)

Neyman-Pearson 原则 (数理统计讲义 3.1.16)

对于假设检验问题 $H_0: \theta \in \Theta_0 \leftrightarrow H_1: \theta \in \Theta_1$

一个显著水平为 α 的检验法 ϕ 首先需要控制其犯第一类错误的概率,

即对于任意 $\theta \in \Theta_0$ 都必须满足 $\gamma_{\phi}(\theta) = \mathbb{E}_{\theta}[\phi(X)] \leq \alpha$,

然后再让功效函数 $\gamma_{\phi}(\theta)$ 在 $\theta \in \Theta_1$ 时尽可能大,

也就是使发生第二类错误的概率尽可能小.

在这一原则下, 零假设 H_0 和备择假设 H_1 的地位是不对称的, 前者是受保护的.

以 Φ_{α} 表示检验问题 $H_0: \theta \in \Theta_0 \leftrightarrow H_1: \theta \in \Theta_1$ 显著水平 α 的检验法全体.

- **Neyman-Pearson 原则**可以翻译为如下的优化问题:

$$\begin{aligned} \min_{\phi} \quad & \{1 - \inf_{\theta \in \Theta_1} \gamma_{\phi}(\theta)\} \\ \text{s.t.} \quad & \sup_{\theta \in \Theta_0} \gamma_{\phi}(\theta) \leq \alpha \quad (\text{i.e. } \phi \in \Phi_{\alpha}) \end{aligned}$$

其中约束 $\sup_{\theta \in \Theta_0} \gamma_{\phi}(\theta) \leq \alpha$ (即 $\phi \in \Phi_{\alpha}$) 保证了 ϕ 第一类错误概率的上确界被显著水平 α 所控制.

而我们的优化目标 $\{1 - \inf_{\theta \in \Theta_1} \gamma_{\phi}(\theta)\} = \sup_{\theta \in \Theta_1} \{1 - \gamma_{\phi}(\theta)\}$

则代表 $\phi \in \Phi_{\alpha}$ 第二类错误概率的上确界, 它需要尽可能小.

- 更强地, 若存在 $\phi^* \in \Phi_{\alpha}$, 对于任意 $\theta \in \Theta_1$ 都有 $\gamma_{\phi^*}(\theta) = \sup_{\phi \in \Phi_{\alpha}} \gamma_{\phi}(\theta)$,
 则我们称 ϕ^* 为显著水平 α 的**一致最有效检验法** (uniformly most powerful test, **UMP**).
 (所谓 "一致" 就是该检验法对于所有 $\theta \in \Theta_1$ 都是最优的)

3.2 N-P 引理及其应用

3.2.1 N-P 引理

设 p_0, p_1 为可能分布族中 (连续分布) 的密度或 (离散分布的) 概率.

若存在 k 使得检验函数 ϕ 可以表示为
$$\phi(x) = \begin{cases} 1, & \text{if } \frac{p_1(x)}{p_0(x)} > k \\ 0, & \text{if } \frac{p_1(x)}{p_0(x)} < k \end{cases}$$

则相应的检验法称为**概率比检验** (Probability Ratio Test)

- 其中 $\frac{p_1(x)}{p_0(x)} = k$ 情况下 $\phi(x)$ 可以取 0 或 1,
 或者 (对于离散情况) 随机化处理至 $[0, 1]$ 区间上取值以满足某些条件,
 例如**定理 3.2.1** 中的 $\gamma_{\phi}(\theta_0) = \mathbb{E}_{\theta_0}[\phi(X)] = \int \phi(x) p_{\theta_0}(x) dx = \alpha$
 (其中 θ_0 为零假设 H_0 对应的参数)

定理 3.2.1: (Neyman-Pearson 引理, 数理统计讲义 命题 3.2.2)

设参数空间为 $\Theta = \{\theta_0, \theta_1\}$

样本 X 的分布具有分布密度 (或离散的概率) $p_0(\cdot) := p_{\theta_0}(\cdot)$ 和 $p_1(\cdot) := p_{\theta_1}(\cdot)$

对于假设检验问题 $H_0: \theta = \theta_0 \leftrightarrow H_1: \theta = \theta_1$ 和显著水平 $\alpha \in (0, 1)$

- **存在性:**

$$\text{存在非负常数 } k \text{ 以及概率比检验 } \phi_0(x) = \begin{cases} 1, & \text{if } \frac{p_1(x)}{p_0(x)} > k \\ 0, & \text{if } \frac{p_1(x)}{p_0(x)} < k \end{cases}$$

$$\text{满足 } \gamma_{\phi_0}(\theta_0) = \mathbb{E}_{\theta_0}[\phi_0(X)] = \int \phi_0(x)p_{\theta_0}(x)dx = \alpha$$

- **充分性:**

存在性中描述的 ϕ_0 是显著水平 α 的**最有效检验法** (MP, most powerful)

即对于显著水平 α 的任意检验函数 $\phi \in \Phi_\alpha$ 都有:

$$\gamma_{\phi_0}(\theta_1) = \mathbb{E}_{\theta_1}[\phi_0(X)] \geq \mathbb{E}_{\theta_1}[\phi(X)] = \gamma_\phi(\theta_1)$$

- **必要性:**

若 ϕ^* 为显著水平 α 的**最有效检验法**, 则 ϕ^* 必定是概率比检验.

定理 3.2.1 的注解:

- 概率比检验的临界值 k_α 为概率比统计量 $\frac{p_1(X)}{p_0(X)}$ 在 $\theta = \theta_0$ 假设下分布的 $1 - \alpha$ 分位数.
- 注意到这个检验问题的备择假设 $H_1: \theta \in \Theta_1 = \{\theta_1\}$ 是一个简单假设, 因此显著水平 α 的**最有效检验法** (MP) 就是**一致最有效检验法** (UMP) 所谓 "一致" 就是该检验法对于所有 $\theta \in \Theta_1$ 都是最优的, 而这里 $\Theta_1 = \{\theta_1\}$ 是单元素集, 因此没必要说明 "一致".
- 对于样本分布族的**充分统计量**, 根据因子化定理, 我们有:

$$\frac{p_1(x)}{p_2(x)} = \frac{g_1(T(x), \theta_1)h(x)}{g_2(T(x), \theta_2)h(x)} = \frac{g_1(T(x), \theta_1)}{g_2(T(x), \theta_2)} = f_{\theta_1, \theta_2}(T(x))$$

此时我们可以等价地使用 $T(X)$ 构造检验法.

证明定理 3.2.1 所基于 (即需要验证) 的事实:

- **存在性:**

- **事实 (1-a):**

$$\text{令 } G(k) = P_{\theta_0}\left\{\frac{p_1(x)}{p_0(x)} > k\right\}$$

注意到 $G(k)$ 关于 k 单调递减, 且 $\begin{cases} G(-\infty) = 1 \\ G(+\infty) = 0 \end{cases}$

而 $1 - G(k) = P_{\theta_0}\left\{\frac{p_1(x)}{p_0(x)} \leq k\right\}$ 正是 $\frac{p_1(X)}{p_2(X)}$ 在 $\theta = \theta_0$ 假设下的累计分布函数 (CDF).

- **事实 (1-b):**

对于任意 $\alpha \in (0, 1)$, 都存在 k_α 使得 $G(k_\alpha) \leq \alpha \leq G(k_\alpha^-)$ 成立.

即 $\frac{p_1(X)}{p_2(X)}$ 在 $\theta = \theta_0$ 假设下分布 $1 - G(k)$ 的 $1 - \alpha$ 分位数 k_α 对于任意 $\alpha \in (0, 1)$ 都是存在的.

- **事实 (1-c):**

$$\text{存在 } \phi_0(x) = \begin{cases} 1, & \frac{p_1(x)}{p_0(x)} > k_\alpha \\ c_\alpha, & \frac{p_1(x)}{p_0(x)} = k_\alpha \text{ 满足 } \mathbb{E}_{\theta_0}[\phi_0(X)] = \alpha \\ 0, & \frac{p_1(x)}{p_0(x)} < k_\alpha \end{cases}$$

我们可以根据 $\mathbb{E}_{\theta_0}[\phi_0(X)] = \alpha$ 推出随机化处理的系数 $c_\alpha = \frac{\alpha - G(k_\alpha)}{G(k_\alpha^-) - G(k_\alpha)}$

- **充分性:**

任意给定 $\alpha \in (0, 1)$

记 Φ_α 为检验问题 $H_0: \theta = \theta_0 \leftrightarrow H_1: \theta = \theta_1$ 显著水平 α 的检验法全体.

- **事实 (2-a):**

事实 (1-c) 中的 ϕ_0 满足:

对于任意 $\begin{cases} \phi \in \Phi_\alpha \\ x \in \mathcal{X} \end{cases}$ 都有 $(\phi_0(x) - \phi(x))(p_1(x) - k_\alpha p_0(x)) \geq 0$ 成立

◦ **事实 (2-b) :**

对于任意 $\phi \in \Phi_\alpha$, 根据事实 (2-a) 有:

$$\begin{aligned} 0 &\leq \int (\phi_0(x) - \phi(x))(p_1(x) - k_\alpha p_0(x)) dx \quad (\text{拆开四项求积分}) \\ &\leq \mathbb{E}_{\theta_1}[\phi_0(X)] - \mathbb{E}_{\theta_1}[\phi(X)] \end{aligned}$$

因此有 $\mathbb{E}_{\theta_1}[\phi_0(X)] \geq \mathbb{E}_{\theta_1}[\phi(X)]$ ($\forall \phi \in \Phi_\alpha$)

• **充分性:**

若 ϕ' 是最有效检验法 (自然有 $\phi' \in \phi_\alpha$)

◦ **事实 (3-a):**

根据定义有 $\begin{cases} \mathbb{E}_{\theta_1}[\phi_0(X)] = \mathbb{E}_{\theta_1}[\phi'(X)] \\ \mathbb{E}_{\theta_0}[\phi_0(X)] \end{cases}$

◦ **事实 (3-b):**

几乎处处成立 $(\phi_0(x) - \phi'(x))(p_1(x) - k_\alpha p_2(x)) = 0$

于是得到 $\phi'(x) = \begin{cases} 1, & \phi_0(x) = 1 \\ 0, & \phi_0(x) = 0 \end{cases}$

即 $\phi' = \phi_0$

(数理统计讲义 例 3.2.3)

设 p_0, p_1 分别为正态分布 $N(\mu_0, 1)$ 和 $N(\mu_1, 1)$ 的分布密度, $\mu_1 > \mu_0$

对于检验问题 $H_0: \mu = \mu_0 \leftrightarrow H_1: \mu = \mu_1$

基于样本 $X = (X_1, \dots, X_n)$, 概率比统计量为:

$$\frac{p_1(x)}{p_0(x)} = \frac{(\frac{1}{\sqrt{2\pi}})^n \exp\{-\frac{1}{2} \sum_{i=1}^n (x_i - \mu_1)^2\}}{(\frac{1}{\sqrt{2\pi}})^n \exp\{-\frac{1}{2} \sum_{i=1}^n (x_i - \mu_0)^2\}} = \exp\left\{(\mu_1 - \mu_0) \sum_{i=1}^n x_i - n \frac{\mu_1^2 - \mu_0^2}{2}\right\}$$

由于上述概率比是 $\sum_{i=1}^n x_i = n\bar{x}$ 的单调递增函数,

故 $\frac{p_1(x)}{p_2(x)}$ 和常数 k 的比较等价于 \bar{x} 与适当常数 c 的比较,

因此概率比检验具有形式 $\phi(x) = \begin{cases} 1 & \bar{x} > c \\ 0 & \bar{x} < c \end{cases}$

下面我们通过令 $\gamma_\phi(\mu_0) = \mathbb{E}_{\mu_0}[\phi(X)] = \alpha$ 来确定**最有效检验法**:

• 在 $H_0: \mu = \mu_0$ 的前提条件下, 我们有 $\sqrt{n}(\bar{X} - \mu_0) \stackrel{d}{=} N(0, 1)$

因此概率比检验具有等价形式 $\phi(x) = \begin{cases} 1 & \sqrt{n}(\bar{x} - \mu_0) > c' \\ 0 & \sqrt{n}(\bar{x} - \mu_0) < c' \end{cases}$

• 我们有:

$$\begin{aligned} \alpha &= \mathbb{E}_{\theta_0}[\phi(X)] \\ &= P_{\theta_0}\{\sqrt{n}(\bar{X} - \mu_0) > c'\} \\ &= P_{\theta_0}\{N(0, 1) > c'\} \\ &= 1 - \Phi(c') \end{aligned}$$

因此 $c' = \Phi^{-1}(1 - \alpha) = z_{1-\alpha}$ (标准正态分布的 $(1 - \alpha)$ -分位数)

因此 $H_0: \mu = \mu_0 \leftrightarrow H_1: \mu = \mu_1$ 的**最有效检验法**就是

$$\phi(x) = \begin{cases} 1 & \sqrt{n}(\bar{x} - \mu_0) > z_{1-\alpha} \\ 0 & \sqrt{n}(\bar{x} - \mu_0) < z_{1-\alpha} \end{cases}$$

这个例子表明，在应用概率比检验时，

并不是直接使用形如 $\phi_0(x) = \begin{cases} 1 & \text{if } \frac{p_1(x)}{p_0(x)} > k \\ 0 & \text{if } \frac{p_1(x)}{p_0(x)} < k \end{cases}$ 的检验函数，

而是由它推导出一个（在零假设成立时）分布已知的检验统计量和拒绝接受范围，得到相应的检验函数。

(数理统计讲义 例 3.2.4)

设 p_0, p_1 分别为均匀分布 $\text{Uniform}(0, \theta_0)$ 和 $\text{Uniform}(0, \theta_1)$ 的分布密度, $\theta_1 > \theta_0$

对于检验问题 $H_0: \theta = \theta_0 \leftrightarrow H_1: \theta = \theta_1$

基于样本 $X = (X_1, \dots, X_n)$, 概率比统计量为:

$$\frac{p_1(x)}{p_0(x)} = \frac{\frac{1}{\theta_1^n} I(x_{(1)} > 0) I(x_{(n)} < \theta_1)}{\frac{1}{\theta_0^n} I(x_{(1)} > 0) I(x_{(n)} < \theta_0)} = \begin{cases} \left(\frac{\theta_0}{\theta_1}\right)^n & 0 \leq x_{(1)} < x_{(n)} < \theta_0 \\ +\infty & 0 < x_{(1)}, \theta_0 \leq x_{(n)} < \theta_1 \\ \text{undefined} & \text{otherwise} \end{cases}$$

注意到 $\frac{p_1(x)}{p_0(x)}$ 表达式未定的部分在 $H_0: \theta = \theta_0$ 和 $H_1: \theta = \theta_1$ 的假设下都是一个零概率事件，因此概率比可写为:

$$\frac{p_1(x)}{p_0(x)} = \left(\frac{\theta_0}{\theta_1}\right)^n I(x_{(n)} < \theta_0) + \infty I(x_{(n)} \geq \theta_0)$$

在 $H_0: \theta = \theta_0$ 的假设下, $\frac{p_1(X)}{p_0(X)}$ 以概率 1 (即几乎处处) 只取 $\left(\frac{\theta_0}{\theta_1}\right)^n$

即 $P_{\theta_0} \left\{ \frac{p_1(X)}{p_0(X)} \leq k \right\} = \begin{cases} 0, & k < \left(\frac{\theta_0}{\theta_1}\right)^n \\ 1, & k > \left(\frac{\theta_0}{\theta_1}\right)^n \end{cases}$ (这比较特殊, 是一个退化分布)

因此对于任意显著水平 $\alpha \in (0, 1)$,

统计量 $\frac{p_1(X)}{p_0(X)}$ 在 P_{θ_0} 下分布的 $1 - \alpha$ 分位数都是 $\left(\frac{\theta_0}{\theta_1}\right)^n$,

因此检验法的临界值 $k_\alpha = \left(\frac{\theta_0}{\theta_1}\right)^n$

取随机化的检验法 $\phi(X) = \alpha I(X_{(n)} < \theta_0) + 1 \cdot I(X_{(n)} \geq \theta_0)$ 便可满足

$$\gamma_\phi(\theta_0) = \mathbb{E}_{\theta_0}[\phi(X)] = \alpha$$

因此是检验问题 $H_0: \theta = \theta_0 \leftrightarrow H_1: \theta = \theta_1$ 在显著水平 α 下的最有效检验法。

值得注意的是:

若分布族存在充分统计量 $T(X)$, 则概率比检验必定是充分统计量的函数。

在上述两个例子中, \bar{X} 和 $X_{(n)}$ 分别是对应分布族的充分统计量,

最后的概率比检验函数都依赖于这些充分统计量。

3.2.2 单侧检验

设 $\mathcal{P}_X = \{p_\theta(x) : \theta \in (a, b)\}$ 为样本分布族。

若对于任意 $a < \theta_1 < \theta_2 < b$,

似然比 (Likelihood Ratio) $\frac{p_{\theta_2}(x)}{p_{\theta_1}(x)} = f_{\theta_1, \theta_2}(T(x))$ 都是关于 $T(x)$ 的严格递增 (递减) 函数,

则我们称分布族 \mathcal{P} 关于 T 有单调递增 (递减) 似然比。

(数理统计讲义 例 3.2.6)

设 $X = (X_1, \dots, X_n)$ 为 Bernoulli 分布族 $\{B(1, p) : p \in (0, 1)\}$ 的简单随机样本,

则样本概率分布的概率比为 $\frac{p_{p_2}(x)}{p_{p_1}(x)} = \left(\frac{p_2}{p_1}\right)^{\sum_{i=1}^n x_i} \left(\frac{1-p_2}{1-p_1}\right)^{n-\sum_{i=1}^n x_i} = \left(\frac{p_2}{1-p_2} \frac{1-p_1}{p_1}\right)^{\sum_{i=1}^n x_i} \left(\frac{1-p_2}{1-p_1}\right)^n$

当 $p_1 < p_2$ 时, 我们有 $\left(\frac{p_2}{1-p_2} \frac{1-p_1}{p_1}\right) > 1$ 成立,

因此 $\frac{p_{p_2}(x)}{p_{p_1}(x)}$ 关于 $T(X) = \sum_{i=1}^n X_i$ 是严格递增函数,

于是样本分布族关于 $T(X) = \sum_{i=1}^n X_i$ 有单调递增似然比。

定理 3.2.2: (数理统计讲义 命题 3.2.7)

对于单参数的指数型分布族 $\{p_\theta(x) = C(\theta) \exp\{Q(\theta)T(x)\} : \theta \in (a, b)\}$
若 $Q(\theta)$ 严格单调递增 (递减), 则分布族关于 $T(X)$ 有单调递增 (递减) 似然比.

定理 3.2.3: (数理统计讲义 定理 3.2.8)

设 X 的分布族 $\{p_\theta(x) : \theta \in (a, b)\}$ 关于 $T(x)$ 有单调递增似然比,
对于检验问题 $H_0 : \theta \leq \theta_0 \leftrightarrow H_1 : \theta > \theta_0$ 和显著水平 $\alpha \in (0, 1)$ 我们有:

- 所有形如 $\phi(x) = \begin{cases} 1, & T(x) > c \\ 0, & T(x) < c \end{cases}$ 的 ϕ 的功效函数 $\gamma_\phi^*(\theta) = \mathbb{E}_\theta[\phi^*(X)]$ 都是 θ 的单增函数.
 - **(存在性)** 存在检验函数 $\phi^*(x) = \begin{cases} 1, & T(x) > c \\ 0, & T(x) < c \end{cases}$ 满足 $\gamma_{\phi^*}(\theta_0) = \mathbb{E}_{\theta_0}[\phi^*(X)] = \alpha$
 - 存在性所描述的 ϕ^* 都是检验问题 $H_0 : \theta \leq \theta_0 \leftrightarrow H_1 : \theta > \theta_0$ 显著水平为 α 的一致最有效检验.
- 即对于任意 $\theta \in \Theta$ 都有 $\gamma_{\phi^*}(\theta) = \sup_{\phi \in \Phi_\alpha} \gamma_\phi(\theta)$ 成立.
其中 Φ_α 代表检验问题 $H_0 : \theta \leq \theta_0 \leftrightarrow H_1 : \theta > \theta_0$ 显著水平 α 的检验法全体.

证明定理 3.2.3 的思路:

- 前两点 (待补充)
 - 至于第三点, 结合 $\gamma_{\phi^*}(\theta) = \mathbb{E}_\theta[\phi^*(X)]$ 关于 θ 单调递增的事实,
存在性中 $\gamma_{\phi^*}(\theta_0) = \alpha$ 的条件可以保证 $\gamma_{\phi^*}(\theta) \leq \gamma_{\phi^*}(\theta_0) = \alpha$ ($\forall \theta \leq \theta_0$) 成立,
表明存在性所描述的 ϕ^* 都是检验问题 $H_0 : \theta \leq \theta_0 \leftrightarrow H_1 : \theta > \theta_0$ 显著水平为 α 的概率比检验.
- 注意到在这里 $H_0 : \theta \leq \theta_0 \leftrightarrow H_1 : \theta > \theta_0$ 等价于 $H_0 : \theta = \theta_0 \leftrightarrow H_1 : \theta > \theta_0$
-

(数理统计讲义 例 3.2.9)

设 $X = (X_1, \dots, X_n)$ 为 Bernoulli 分布族 $\{B(1, p) : p \in (0, 1)\}$ 的简单随机样本.
考虑检验问题 $H_0 : p \leq p_0 \leftrightarrow H_1 : p > p_0$, 显著水平 $\alpha \in (0, 1)$

根据数理统计讲义 例 3.2.6 可知样本分布族关于 $T(X) = \sum_{i=1}^n X_i$ 有单调递增似然比.

因此上述检验问题的一致最有效检验具有形式 $\phi(X) = \begin{cases} 1, & \sum_{i=1}^n X_i > c \\ 0, & \sum_{i=1}^n X_i < c \end{cases}$

下面我们令 $\gamma_\phi(p_0) = \mathbb{E}_{p_0}[\phi(X)] = \alpha$ 来确定常数 c :

- 我们知道在 $p = p_0$ 条件下 $T(X) = \sum_{i=1}^n X_i \stackrel{d}{=} B(n, p_0)$
- 记 $\bar{B}(k; n, p_0) = P_{p_0}\{\sum_{i=1}^n X_i \geq k\} = P\{B(n, p_0) \geq k\} = \sum_{j=k}^n \binom{n}{j} p_0^j (1 - p_0)^{n-j}$
设整数 k_0 使得 $\bar{B}(k_0 + 1; n, p_0) \leq \alpha \leq \bar{B}(k_0; n, p_0)$ 成立,
则我们可以取 $\phi(X)$ 为:

$$\phi(X) = \frac{\alpha - \bar{B}(k_0 + 1; n, p_0)}{\bar{B}(k_0; n, p_0) - \bar{B}(k_0 + 1; n, p_0)} I(\sum_{i=1}^n X_i = k_0) + I(\sum_{i=1}^n X_i \geq k_0 + 1)$$

- 可以验证它满足 $\gamma_\phi(p_0) = \mathbb{E}_{p_0}[\phi(X)] = \alpha$:

$$\begin{aligned}
\gamma_\phi(p_0) &= \mathbb{E}_{p_0}[\phi(X)] \\
&= \frac{\alpha - \bar{B}(k_0 + 1; n, p_0)}{\bar{B}(k_0; n, p_0) - \bar{B}(k_0 + 1; n, p_0)} \mathbb{E}_{p_0}[I(\sum_{i=1}^n X_i = k_0)] + \mathbb{E}_{p_0}[I(\sum_{i=1}^n X_i \geq k_0 + 1)] \\
&= \frac{\alpha - \bar{B}(k_0 + 1; n, p_0)}{\bar{B}(k_0; n, p_0) - \bar{B}(k_0 + 1; n, p_0)} (\bar{B}(k_0; n, p_0) - \bar{B}(k_0 + 1; n, p_0)) + \bar{B}(k_0 + 1; n, p_0) \\
&= \alpha
\end{aligned}$$

因此它便是检验问题 $H_0: \theta \leq p_0 \leftrightarrow H_1: \theta > p_0$ 显著水平 $\alpha \in (0, 1)$ 的 UMP 检验.

(一致最有效检验法, uniformly most powerful test, UMP)

- 事实上, 我们就是取 c_α 为 $B(n, p_0)$ 的 $(1 - \alpha)$ 分位数, 但由于这是一个离散分布, 我们并不一定能取到准确的分位数, 故我们需要随机化处理.

重新考虑数理统计讲义例 3.1.14,

其检验问题为 $H_0: \pi \leq 0.06 \leftrightarrow H_1: \pi > 0.06$

设显著水平 $\alpha = 0.05$

根据 $P_{0.06}\{\sum_{i=1}^{50} X_i \geq 7\} = 0.02892 < 0.05 < 0.07764 = P_{0.06}\{\sum_{i=1}^{50} X_i \geq 6\}$

可知 $k_0 = 6$

因此显著水平 $\alpha = 0.05$ 的 UMP 检验法为:

$$\begin{aligned}
\phi(X) &= \frac{\alpha - \bar{B}(k_0 + 1; n, p_0)}{\bar{B}(k_0; n, p_0) - \bar{B}(k_0 + 1; n, p_0)} I(\sum_{i=1}^n X_i = k_0) + I(\sum_{i=1}^n X_i \geq k_0 + 1) \\
&= \frac{\alpha - P_{0.06}\{\sum_{i=1}^{50} X_i \geq 7\}}{P_{0.06}\{\sum_{i=1}^{50} X_i \geq 6\} - P_{0.06}\{\sum_{i=1}^{50} X_i \geq 7\}} I(\sum_{i=1}^n X_i = 6) + I(\sum_{i=1}^n X_i \geq 7) \\
&= \frac{0.05 - 0.02892}{0.07764 - 0.02892} I(\sum_{i=1}^n X_i = 6) + I(\sum_{i=1}^n X_i \geq 7) \\
&= 0.4327 \cdot I(\sum_{i=1}^n X_i = 6) + I(\sum_{i=1}^n X_i \geq 7)
\end{aligned}$$

(数理统计讲义例 3.2.10)

设 $X = (X_1, \dots, X_n)$ 为正态分布族 $\{N(\mu, \sigma_0^2) : \mu \in \mathbb{R}\}$ (σ_0^2 已知) 的简单随机样本.

考虑检验问题 $H_0: \mu \leq \mu_0 \leftrightarrow H_1: \mu > \mu_0$, 显著水平 $\alpha \in (0, 1)$

注意到样本 X 的分布密度为:

$$\begin{aligned}
p_\mu(x) &= \left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n \exp\left\{-\frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \mu)^2\right\} \\
&= \left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n \exp\left\{\frac{n\mu}{\sigma_0^2} \bar{x} - \frac{n\mu^2}{2\sigma_0^2} - \frac{1}{2\sigma_0^2} \sum_{i=1}^n x_i^2\right\} \\
&= \left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n \exp\left\{-\frac{n\mu^2}{2\sigma_0^2}\right\} \cdot \exp\left\{-\frac{n\mu}{\sigma_0^2} \cdot \bar{x}\right\} \cdot \exp\left\{-\frac{1}{2\sigma_0^2} \sum_{i=1}^n x_i^2\right\}
\end{aligned}$$

$$\text{令} \begin{cases} C(\mu) = \left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n \exp\left\{-\frac{n\mu^2}{2\sigma_0^2}\right\} \\ Q(\mu) = -\frac{n\mu}{\sigma_0^2} \\ T(x) = \bar{x} \\ h(x, \mu) = \exp\left\{-\frac{1}{2\sigma_0^2} \sum_{i=1}^n x_i^2\right\} \end{cases} \quad \text{可知它是指数量分布族.}$$

注意到 $Q(\mu) = -\frac{n\mu}{\sigma_0^2}$ 关于 μ 是单调递减的,

故根据定理 3.2.2 可知样本分布族关于 $T(X) = \bar{X}$ 有单调递减的似然比.

定理 3.2.2: (数理统计讲义 命题 3.2.7)

对于指数分布族 $\{p_\theta(x) = C(\theta) \exp\{Q(\theta)T(x)\} : \theta \in (a, b)\}$

若 $Q(\theta)$ 严格单调递增 (递减), 则分布族关于 $T(x)$ 有单调递增 (递减) 似然比.

因此检验问题 $H_0 : \mu \leq \mu_0 \leftrightarrow H_1 : \mu > \mu_0$ 的 UMP 检验法具有形式:

$$\phi(X) = \begin{cases} 1, & \sum_{i=1}^n X_i > c \\ 0, & \sum_{i=1}^n X_i < c \end{cases} = \begin{cases} 1, & \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0} > z_{1-\alpha} \\ 0, & \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0} < z_{1-\alpha} \end{cases}$$

其中 $z_{1-\alpha} = \Phi^{-1}(1 - \alpha)$ 代表 $N(0, 1)$ 的 $1 - \alpha$ 分位数.

可以验证它满足 $\gamma_\phi(\mu_0) = \mathbb{E}_{\mu_0}[\phi(X)] = \alpha$:

$$\begin{aligned} \gamma_\phi(\mu_0) &= \mathbb{E}_{\mu_0}[\phi(X)] \\ &= P_{\mu_0} \left\{ \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0} > z_{1-\alpha} \right\} \\ &= P\{N(0, 1) > z_{1-\alpha}\} \\ &= 1 - \Phi(z_{1-\alpha}) \\ &= 1 - (1 - \alpha) \\ &= \alpha \end{aligned}$$

(数理统计讲义 例 3.2.10)

设 $X = (X_1, \dots, X_n)$ 为正态分布族 $\{N(\mu_0, \sigma^2) : \sigma^2 > 0\}$ (μ_0 已知) 的简单随机样本.

考虑检验问题 $H_0 : \sigma^2 \leq \sigma_0^2 \leftrightarrow H_1 : \sigma^2 > \sigma_0^2$, 显著水平 $\alpha \in (0, 1)$

注意到样本 X 的分布密度为:

$$p_{\sigma^2}(x) = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_0)^2 \right\}$$

$$\text{令 } \begin{cases} C(\sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \\ Q(\sigma^2) = -\frac{1}{2\sigma^2} \\ T(x) = \sum_{i=1}^n (x_i - \mu_0)^2 \\ h(x, \sigma^2) \equiv 1 \end{cases} \text{ 可知它是指数量分布族.}$$

注意到 $Q(\sigma^2) = -\frac{1}{2\sigma^2}$ 关于 σ^2 是单调递增的,

故根据定理 3.2.2 可知样本分布族关于 $T(X) = \sum_{i=1}^n (X_i - \mu_0)^2$ 有单调递增的似然比.

因此检验问题 $H_0 : \mu \leq \mu_0 \leftrightarrow H_1 : \mu > \mu_0$ 的 UMP 检验法具有形式:

$$\phi(X) = \begin{cases} 1, & \sum_{i=1}^n (X_i - \mu_0)^2 > c \\ 0, & \sum_{i=1}^n (X_i - \mu_0)^2 < c \end{cases} = \begin{cases} 1, & \sum_{i=1}^n \left(\frac{X_i - \mu_0}{\sigma_0} \right)^2 > \chi_{(n), (1-\alpha)}^2 \\ 0, & \sum_{i=1}^n \left(\frac{X_i - \mu_0}{\sigma_0} \right)^2 < \chi_{(n), (1-\alpha)}^2 \end{cases}$$

其中 $\chi_{(n), (1-\alpha)}^2$ 代表自由度为 n 的卡方分布 $\chi_{(n)}^2$ 的 $1 - \alpha$ 分位数.

可以验证它满足 $\gamma_\phi(\sigma_0^2) = \mathbb{E}_{\sigma_0^2}[\phi(X)] = \alpha$:

$$\begin{aligned} \gamma_\phi(\sigma_0^2) &= \mathbb{E}_{\sigma_0^2}[\phi(X)] \\ &= P_{\sigma_0^2} \left\{ \sum_{i=1}^n \left(\frac{X_i - \mu_0}{\sigma_0} \right)^2 > \chi_{(n), (1-\alpha)}^2 \right\} \\ &= P\{\chi_{(n)}^2 > \chi_{(n), (1-\alpha)}^2\} \\ &= 1 - (1 - \alpha) \\ &= \alpha \end{aligned}$$

