# 第三章 假设检验

- **主讲内容**
- 第一节 检验的求法
  1.1 似然比检验
  1.2 贝叶斯检验
- 第二节 检验的评价方法
  2.1 错误概率和功效函数
  2.2 最大功效检验（UMPT）
  *2.3 p-值

- **定义0.1** （【1】定义8.1.1） **假设**(Hypothesis)就是关于总体参数的一个陈述。
- 设$\theta$ 是一个未知总体参数，参数空间$\Theta$. 令$\Theta_0 \subset \Theta$，考虑

$$H_0 : \theta \in \Theta_0 \quad \Leftrightarrow \quad H_1 : \theta \in \Theta_1 = \Theta_0^c.$$

- **定义0.2** （【1】定义8.1.2） 一个假设检验问题中两个互补的假设$H_0$和$H_1$分别称为**原假设**（或零假设）（Null Hypothesis）和**备择假设** (Alternative Hypothesis)。
-
- **定义0.3** （【1】定义8.1.3） 一个**假设检验**就是一个规则，它明确描述:
  1. 对于哪些样本值应该决定接受$H_0$为真;
  2. 对于哪些样本值应该拒绝$H_0$而接受$H_1$为真。
- 那些由拒绝$H_0$的样本构成的样本空间子集就称为**拒绝域**（Rejection Region）或者**临界区域**（Critical Region），而拒绝域的补集称为**接受域**（Acceptance Region）。

注 假设检验的目的就是依靠样本来决定关于总体的两个互补的假设哪个为真

## Example (0.1)

（【0】例5.1.1） 某工厂生产的一大批产品要卖给商店，按规定次品率$p$不得超过0.01。今在其中随机抽取100件，经检验有3件次品，问这批产品可否出厂？

分析 真实的$p$值有两种可能性（i.e. 假设）

$$H_0: 0 < p \leq 0.01 \quad \Leftrightarrow \quad H_1: 0.01 < p < 1$$

问题等价于通过这次抽样来决定假设$H_0$是否为真。

## Example (0.2)

有两个硬币，第一个硬币是均匀的，第二个硬币抛到正面的概率为0.7. 现随机选取一个硬币，抛10次有6次为正面，问选的是第一枚硬币吗？

分析 这里$p$有两种可能性（i.e.假设）

$$H_0: p = 0.5 \quad \Leftrightarrow \quad H_1: p = 0.7$$

问题等价于通过观测到的这个样本来判定假设$H_0$是否成立。

- 设$X_1, \ldots, X_n$ $i.i.d. \sim X$，总体$X$的$p.d.f./p.m.f.$为$f(x|\vec{\theta})$，从而似然函数
$$\mathcal{L}(\theta|\vec{x}) = \prod_{i=1}^{n} f(x_i|\vec{\theta}),\ \vec{\theta} \in \Theta.$$
- **定义1.1**（【0】定义5.3.1，【1】定义8.2.1）设$\Theta_0$为参数空间$\Theta$的一个真子集，关于检验
$$\begin{aligned} H_0 &: \vec{\theta} \in \Theta_0, \\ H_1 &: \vec{\theta} \in \Theta_0^c \end{aligned}$$
的**似然比检验**（Likelihood Ratio Test，简记LRT）**统计量**为
$$\lambda(\vec{x}) = \frac{\sup_{\vec{\theta} \in \Theta_0} \mathcal{L}(\vec{\theta}|\vec{x})}{\sup_{\vec{\theta} \in \Theta} \mathcal{L}(\vec{\theta}|\vec{x})}$$
任一给出拒绝区域$\{\vec{x} : \lambda(\vec{x}) \le c\}$的检验都叫做**似然比检验**，这里$0 \le c < 1$ 为常数。

注1 在离散情形下，$\lambda(\vec{x})$的分子是参数$\vec{\theta}$取遍$\Theta_0$时，样本观测值$\vec{x}$出现的最大概率，而分母则是参数$\vec{\theta}$取遍全空间$\Theta$时，样本观测值$\vec{x}$出现的最大概率。

注2 如果备择假设$H_1$中存在$\vec{\theta}_1 \in \Theta_0^c$, $s.t.$

$$\mathcal{L}(\vec{\theta}_1|\vec{x}) \geq \frac{1}{c} \sup_{\vec{\theta} \in \Theta_0} \mathcal{L}(\vec{\theta}|\vec{x}),$$

则我们拒绝$H_0$而接受$H_1$。

注3 由注2可知，似然比检验等价于给出一个定义在样本空间$\mathcal{X}$上的函数

$$\varphi(\vec{x}) = \begin{cases} 1 & \lambda(\vec{x}) \leq c \\ 0 & \lambda(\vec{x}) > c \end{cases}.$$

其支撑集$\{\vec{x} : \varphi(\vec{x}) > 0\} = \{\vec{x} : \lambda(\vec{x}) \leq c\}$ 即为$H_0$的拒绝域，$i.e.\varphi$为$H_0$拒绝域的示性函数。

# 举例说明

注4 回忆第二章第一节关于"极大似然估计"（MLE）部分，若$\widehat{\theta}_{MLE}$存在，则$\sup_{\vec{\theta}\in\Theta}\mathcal{L}(\vec{\theta}|\vec{x})=\mathcal{L}(\widehat{\theta}_{MLE}|\vec{x})$。类似地，设$\widehat{\theta}_0=\arg\max_{\vec{\theta}\in\Theta_0}\mathcal{L}(\vec{\theta}|\vec{x})$，则LRT统计量为

$$\lambda(\vec{x})=\frac{\mathcal{L}(\widehat{\theta}_0|\vec{x})}{\mathcal{L}(\widehat{\theta}_{MLE}|\vec{x})}$$

## Example (1.1)

设$X_1,\ldots,X_n\ i.i.d.\sim N(\mu,\sigma^2)$，考虑检验

$$\begin{aligned} H_0 &: \mu=\mu_0, \\ H_1 &: \mu\neq\mu_0 \end{aligned}$$

分别在如下情形找出其$LRT$统计量：(1)$\sigma^2$已知，(2)$\sigma^2$未知。

**注** 例1.1参考【0】例5.3.1, 【1】例8.2.2.

**Example (1.2)**

设$X_1, \ldots, X_n$ $i.i.d. \sim X$，总体$X$的$p.d.f.$ 如下

$$f(x|\theta) = e^{-(x-\theta)}\mathrm{I}_{[\theta,\infty)}(x), -\infty < \theta < \infty.$$

考虑检验

$$\begin{aligned} H_0 &: \theta \le \theta_0, \\ H_1 &: \theta > \theta_0 \end{aligned}$$

找出其$LRT$统计量并给出对应或等价的拒绝域。

**注** 例1.2参考【1】例8.2.3. 自行练习【0】例5.3.2，5.3.4.

# 举例说明

- 上述两例中相关统计量$\widehat{\theta}_{MLE}$均为相应参数分布族的充分统计量。考虑利用充分统计量简化寻找LRT统计量的过程。

**Theorem (1.1)**

（【1】定理8.2.4） 设$T(\vec{x})$是关于$\vec{\theta}$的一个充分统计量，而$\lambda^*(\vec{t})$和$\lambda(\vec{x})$分别是依赖于$T$和$\vec{X}$的$LRT$统计量，则对于样本空间内的每一点$\vec{x}$，均有$\lambda^*(T(\vec{x})) = \lambda(\vec{x})$。

注 T作为单个样本，似然函数$L^*(\vec{\theta}) = f_T(\vec{t}\,|\,\vec{\theta})$, $T$的$p.d.f./p.m.f.$.

**Example (1.3)**

利用定理1.1再次找出例1.1（1）和例1.2中检验的LRT统计量并给出对应或等价的拒绝域。

注 练习【0】例5.3.5.

## Example (1.4)

(【0】例5.3.3)设$X_1, \ldots, X_n$ $i.i.d. \sim N(\mu, \sigma^2)$，$\mu \in \mathbb{R}, \sigma^2 > 0$ 均未知，考虑检验问题

$$\begin{aligned} H_0 &: \sigma^2 = \sigma_0^2, \\ H_1 &: \sigma^2 \neq \sigma_0^2 \end{aligned}$$

请找出其$LRT$统计量并给出对应或等价的拒绝域。

作业1 找出习题5 Ex. 30, 31, 32, 37中假设检验的LRT统计量并给出对应或等价的拒绝域。

作业2 利用定理1.1找出例1.1（2）中检验的LRT统计量并给出对应或等价的拒绝域。

# 1.2 贝叶斯检验

- 贝叶斯统计模型：给定先验分布$\vartheta \sim \pi(\theta)$ (p.d.f./p.m.f.) 和抽样$\vec{X}$的联合p.d.f./p.m.f. $f(\vec{x}|\theta)$，得到$\vartheta$的后验分布$\vartheta|\vec{x} \sim \pi(\theta|\vec{x})$.

注 所有关于$\vartheta$的统计推断都是基于后验分布进行。

**Definition (贝叶斯检验)**

对于检验问题 $\quad H_0 : \theta \in \Theta_0 \quad \leftrightarrow \quad H_1 : \theta \in \Theta_1,$

1. 分别计算$H_0$和$H_1$为真/成立的后验概率

$$\alpha_0 = \mathbb{P}(H_0|\vec{x}) = \mathbb{P}(\theta \in H_0|\vec{x})$$
$$\alpha_1 = \mathbb{P}(H_1|\vec{x}) = \mathbb{P}(\theta \in H_1|\vec{x})$$

2. 设定检验法则：对于固定常数$b > 0$,
当$\frac{\alpha_0}{\alpha_1} < b$时拒绝$H_0$，否则接受$H_0$.

为简化问题，在本课程中，我们约定取$b = 1$（参考【0】7.3.3(1)）。

注 与似然比检验方法相比：（1）它无需选择检验统计量；（2）无需给出检验水平（下一节讲述）以确定拒绝域；（3）容易推广到多重假设检验情形（分$H_1, H_2, \ldots, H_k$，详见【0】7.3.3（2）*）。

**Example (1.5)**

考虑对一个儿童作智力测验，设测验结果$X \sim N(100.5, \sigma^2)$，其中$\sigma^2$为测验中智商IQ的方差真实值，满足$\xi = \frac{1}{\sigma^2}$的先验分布为$Gamma(\frac{5}{2}, \frac{3}{8})$。现该儿童测验得分$x = 101$，求下列检验问题的贝叶斯检验结果，

$$H_0 : \sigma^2 \leq 0.29 \quad \leftrightarrow \quad H_1 : \sigma^2 > 0.29.$$

注 练习【0】例7.3.8，7.3.9。

**Example (1.6)**

（引言例0.2续）有两个硬币，第一个硬币是均匀的，第二个硬币抛到正面的概率为0.7。现随机选取一个硬币，抛10次有6次为正面。假设已知第一个硬币选到的概率为$\frac{1}{3}$，第二个硬币为$\frac{2}{3}$，问选的硬币均匀吗？

作业1 习题7, Ex. 18.

作业2 一个餐馆老板决定每盈利的总天数达到5天就休息一天，设每天盈利的概率$\mathbb{P}(X = 1) = \theta$，不盈利的概率$\mathbb{P}(X = 0) = 1 - \theta$，且盈利的情况$X_1, \ldots, X_n$ $i.i.d.$，$n \in \mathbb{N}^+$。假设参数$\theta$的先验分布为均匀分布$U(0, 1)$。现老板工作了6天后休息了1天，问如下关于盈利概率假设哪个为真？

$$H_0 : \theta > 0.5 \quad \leftrightarrow \quad H_1 : \theta \leq 0.5.$$