第 15 周讲稿 1

2 参数的假设检验

2.1 显著性检验的基本思想与检验的两类错误

★ 两类错误

零假设通常受到保护，而备选假设是当零假设被拒绝后才能被接受。

检验规则：构造一个统计量 $T(X_1,X_2,...,X_3)$，当 H_0 服从某一分布，当 H_0 不成立时，T 的偏大偏小特征。据此，构造拒绝域 W

第一类错误：“弃真”， H_0 **为真时，拒绝** H_0；$P\{T\in W\mid H_0\text{为真}\}$

第二类错误：“存伪”， H_0 **为不真时，接受** H_0. $P\{T\notin W\mid \mathrm{H}_0\text{为假}\}$

势函数：$\beta(\theta)=E_\theta(\delta(X))=P_\theta\{X\in W\}$ $\delta(X)=\begin{cases}1, & X\in W.\\ 0, & X\notin W.\end{cases}$

当 $\theta\in\Theta_0$ 时，$\beta(\theta)$ 为犯第一类错误的概率

当 $\theta\in\Theta_1$ 时，$1-\beta(\theta)$ 为犯第二类错误的概率

★ 假设检验的步骤：

（1） **根据实际问题中所关心的问题，建立原假设** H_0 **和备择假设** H_1；

（2） **选择一个合适的统计量 V（要求** H_0 **为真时，V 的分布已知），并根据其取值特点确定一个合适的拒绝域形式；**

（3） **由给定的检验水平** α，**利用关系式**

$$P_{H_0}(\{V<V_{\frac{\alpha}{2}}\}\bigcup\{V>V_{1-\frac{\alpha}{2}}\})=\alpha$$

$$P_{H_0}(\{V>V_{1-\alpha}\})=\alpha$$

$$P_{H_0}(\{V<V_{\alpha}\})=\alpha$$

中的某一个，求出水平为 α **的检验拒绝域.**

（4） **根据样本观察值，算出 V 的观察值，并根据它作出接受还是拒绝** H_0.

一个总体的情况： $X\sim N(\mu,\sigma^2)$

σ^2 **已知，检验** $\mathrm{H}_0{:}\mu=\mu_0\leftrightarrow H_1:\mu\neq\mu_0$：$U=\dfrac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}}\sim N(0,1)$

σ^2 未知，检验 $\mathrm{H}_0:\mu=\mu_0 \leftrightarrow H_1:\mu\neq\mu_0$： $T=\dfrac{\overline{X}-\mu_0}{S_n^*\big/\sqrt{n}}\sim t(n-1)$

μ 已知，检验 $\mathrm{H}_0:\sigma^2=\sigma_0^2 \leftrightarrow H_1:\sigma^2\neq\sigma_0^2$： $\chi^2=\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}\sim\chi^2(n)$

μ 未知，检验 $\mathrm{H}_0:\sigma^2=\sigma_0^2 \leftrightarrow H_1:\sigma^2\neq\sigma_0^2$： $\chi^2=\dfrac{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}{\sigma^2}\sim\chi^2(n-1)$

两个总体的情况： $X\sim N(\mu_1,\sigma_1^2)$， $Y\sim N(\mu_2,\sigma_2^2)$

$\sigma_1^2=\sigma_2^2=\sigma^2$ 未知时，检验 $\mathrm{H}_0:\mu_1=\mu_2 \leftrightarrow H_1:\mu_1\neq\mu_2$：

$$T=\frac{\overline{X}-\overline{Y}}{\sqrt{(n_1-1)S_{1n_1}^{*2}+(n_2-1)S_{2n_2}^{*2}}}\sqrt{\frac{n_1n_2(n_1+n_2-2)}{n_1+n_2}}\sim t(n_1+n_2-2)$$

μ_1,μ_2 未知时，检验 $\mathrm{H}_0:\sigma_1^2=\sigma_2^2 \leftrightarrow H_1:\sigma_1^2\neq\sigma_2^2$： $F=\dfrac{S_{1n_1}^{*2}}{S_{2n_2}^{*2}}\sim F(n_1-1,n_2-1)$

单边检验：举例说明， σ^2 已知，检验 $\mathrm{H}_0:\mu\le\mu_0 \leftrightarrow H_1:\mu>\mu_0$：

构造 $U_1=\dfrac{\overline{X}-\mu}{\sigma_0\big/\sqrt{n}}\sim N(0,1)$，给定显著性水平 α，有 $P\{U_1>u_{1-\alpha}\}=\alpha$。当 **H₀** 成立时 $U_1=\dfrac{\overline{X}-\mu}{\sigma_0\big/\sqrt{n}}\ge\dfrac{\overline{X}-\mu_0}{\sigma_0\big/\sqrt{n}}\overset{def}{\Rightarrow}U$，因此 $P\{U>u_{1-\alpha}\}\le P\{U_1>u_{1-\alpha}\}=\alpha$。故拒绝域为 $W=\{U>u_{1-\alpha}\}$

例 7 某厂生产的零件的直径服从正态分布，以往经验知其标准差为 3.6，考虑假设

$$H_0:\mu=68,\quad H_1:\mu\neq 68$$

现按下列方式进行判断：当 $|\overline{X}-68|>1$ 时，拒绝原假设 H_0，否则就接受原假设 H_0。现在抽取 **64** 件零件进行检验，则犯第一类错误的概率 $\alpha=$ ________ .【$2(1-\Phi(2.2))$】

例 8 设某糖厂用自动包装机包装糖果，包装好的糖果的重量服从正态分布，且由以往经验知标准差为 0.015（kg）. 某日开工后在生产线上抽测 9 袋，得数据 0.497，0.506，0.518，0.524，0.488，0.511，0.510，0.515，0.512（kg），如果规定包装的每装糖重量为 0.5kg，

问由抽测数据能否有 90%的把握判断生产线上包装机工作是否正常？（取 $\alpha = 0.1$）.

【$H_0 = 0.5; H_1 \neq 0.5,\quad \left|\dfrac{\bar{x}-0.5}{0.015/\sqrt{9}}\right| = 1.8 > z_{\alpha/2} = 1.645$，**拒绝** H_0】

2.2 正态总体的参数假设检验

2.2.1 单正态总体参数的双侧检验

1） **H₀**：$\mu = \mu_0$ **(a)** 设 σ^2 已知：采用 ***u*** **检验.**

(b) 设方差 σ^2 未知：采用 ***t*** **检验.**

2） **H₀**：$\sigma^2 = \sigma_0^2$ **(a)** 设 μ 已知: 采用 χ^2 **检验.**

(b) 设 μ 未知: 采用 χ^2 **检验.**

2.2.2 两正态总体参数的差异性检验

1) 两个总体均值差 $\mu_1 - \mu_2$ 的检验 **(a)** 设 ${\sigma_1}^2, {\sigma_2}^2$ 均为已知，采用 ***u*** **检验.**

(b) 设 ${\sigma_1}^2 = {\sigma_2}^2 = \sigma^2$, 但 σ^2 为未知，采用 ***t*** **检验.**

(c) 设 ${\sigma_1}^2, {\sigma_2}^2$ 均为未知, 且不知它们是否相等.

2) 方差比的假设检验. H_0： $\sigma_1^2 / \sigma_2^2 = c$.

2.2.3 一个正态总体参数的单侧检验

〖单侧假设检验，左边假设检验，右边假设检验的对照〗

〖假设检验与区间估计的对照〗

例：设 $X_1, \cdots, X_n$ 是总体 $X \sim N(\mu, \sigma^2)$（σ^2 已知）的一个样本，检验

$$H_0: \mu = \mu_0 \leftrightarrow H_1: \mu \neq \mu_0$$

检验统计量为 $U = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$，其拒绝域为

$$W = \{(x_1, x_2, \cdots, x_n): |u| > u_{1-\frac{\alpha}{2}}\}$$

$$= \{(x_1, x_2, \cdots, x_n): |\overline{x} - \mu_0| > u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\}$$

故接受域为

$$W^c=\{(x_1,x_2,\cdots,x_n):|\bar{x}-\mu_0|\le u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\}$$

$$=\{(x_1,x_2,\cdots,x_n):\bar{x}-u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\le\mu_0\le\bar{x}+u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\}$$

我们发现，将小写改为大写，μ_0改为μ，那么接受域就为参数的置信度为$1-\alpha$的置信区间，即$[\bar{X}-u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\bar{X}+u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}]$

注：给定样本观测值我们就得到了区间估计的一个实现$[\bar{x}-u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\bar{x}+u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}]$，而根据区间估计的定义，我们有$1-\alpha$的把握，使得参数真值$\mu$落入这个区间中，但是我们现在要检验的$\mu_0$并没有落入这个区间，那么我们自然会认为$\mu\ne\mu_0$，即这组样本落入了拒绝域。

同样若样本观测值确定，改变原假设中的μ_0，当μ_0取值不同时，我们对原假设的判断也会发生变化。而区间$[\bar{x}-u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\bar{x}+u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}]$的含义，就是所有能使原假设被接受的的$\mu_0$集合。

一般有：

结论：设$A(\theta_0)$是$H_0:\theta=\theta_0\leftrightarrow H_1:\theta\ne\theta_0$的接受域，令

$$C(x_1,x_2,\cdots,x_n)=\{\theta:(x_1,x_2,\cdots,x_n)\in A(\theta_0)\}$$

则$C(X_1,X_2,\cdots,X_n)$为参数θ的置信度为$1-\alpha$的置信区间。

反之，若$C(X_1,X_2,\cdots,X_n)$为参数θ的置信度为$1-\alpha$的置信区间，令

$$A(\theta_0)=\{(x_1,x_2,\cdots,x_n):\theta_0\in C(x_1,x_2,\cdots,x_n)\}$$

则$A(\theta_0)$是$H_0:\theta=\theta_0\leftrightarrow H_1:\theta\ne\theta_0$的置信水平为$\alpha$的接受域。

单参数指数型分布族假设检验问题

对于单参数指数型分布族，若样本 $X_1,\cdots,X_n$ 的联合密度函数为

$$\prod_{i=1}^{n} f(x_i;\theta)=C(\theta)h(x_1,x_2,...,x_n)\exp\{Q(\theta)T(x_1,x_2,...,x_n)\}$$

其中 $Q(\theta)$ 关于 θ 严格增，我们一般取检验统计量为 $T(X_1,X_2,\cdots,X_n)$，考察三类检验问题

$$\begin{aligned}
&H_0:\theta\ge\theta_0 \leftrightarrow H_1:\theta<\theta_0;\\
&H_0:\theta\le\theta_0 \leftrightarrow H_1:\theta>\theta_0;\\
&H_0:\theta=\theta_0 \leftrightarrow H_1:\theta\ne\theta_0;
\end{aligned}$$

在确定拒绝域的形式时，我们需要先求检验统计量的期望 $E_\theta T(X_1,X_2,\cdots,X_n)$，如果它是 θ 的单增函数，那么对于第一个假设而言，当 $T(X_1,X_2,\cdots,X_n)$ 比较小的时候我们会拒绝原假设；类似地可以得到上述三种检验问题的拒绝域形式分别为

$$\begin{aligned}
&\{T(x_1,x_2,\cdots,x_n)<c\};\\
&\{T(x_1,x_2,\cdots,x_n)>c\};\\
&\{T(x_1,x_2,\cdots,x_n)<c_1\}\cup\{T(x_1,x_2,\cdots,x_n)>c_2\};
\end{aligned}$$

计算临界值时，我们要使得犯第一类错误的概率小于显著性水平 α，即原假设成立的条件下样本落入拒绝域的概率的上确界小于显著性水平。概率的上确界就是 $\theta=\theta_0$ 时的概率，即

$$\sup_{\theta\ge\theta_0} P_\theta\{T(X_1,X_2,\cdots,X_n)<c\}=P_{\theta_0}\{T(X_1,X_2,\cdots,X_n)<c\}$$

（以第一个为例）

若$\theta=\theta_0$时分布完全确定，令上式为α即可求出临界值。

例　设样本$X_1,\cdots,X_n$是来自$E(\frac{1}{\theta})$，检验假设

$$H_0:\theta\ge\theta_0\leftrightarrow H_1:\theta<\theta_0$$

解：样本的分布为

$$f(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}f(x_i;\theta)=(\frac{1}{\theta})^n e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i}1_{\{x_1,x_2,\cdots,x_n>0\}}$$

取检验统计量为$T(X_1,X_2,\cdots,X_n)=\sum_{i=1}^{n}X_i$。其期望为

$E_\theta T(X_1,X_2,\cdots,X_n)=n\theta$是$\theta$的单增函数，$T=\sum_{i=1}^{n}X_i\sim\Gamma(n,\frac{1}{\theta})$

故拒绝域的形式为$\{T(x_1,x_2,\cdots,x_n)<c\}=\{n\bar{x}<c\}$，犯第一类错误的概率的

最大值为$\sup\limits_{\theta\ge\theta_0}P_\theta\{n\bar{X}<c\}=P_{\theta_0}\{n\bar{X}<c\}=\alpha$

由于$T=\sum_{i=1}^{n}X_i\sim\Gamma(n,\frac{1}{\theta})$，故$\frac{2n\bar{X}}{\theta}\sim\Gamma(n,\frac{1}{2})=\chi^2(2n)$，

因此临界值$c=\frac{\theta_0}{2}\chi^2{}_\alpha(2n)$。

拒绝域为

$$\{(x_1,x_2,\cdots,x_n):n\bar{x}<\frac{\theta_0}{2}\chi^2{}_\alpha(2n)\}$$

$$=\{(x_1,x_2,\cdots,x_n):\frac{2n\bar{x}}{\theta_0}<\chi^2{}_\alpha(2n)\}$$

例 9 正常生产条件下，某产品的生产指标 **X**～$N(\mu_0,\sigma_0^2)$，其中 σ_0 **= 0.23.** 现在改变了生产工艺，产品的生产指标变为 **X′** ～$N(\mu,\sigma^2)$. 从新工艺产品中任意抽取 **10** 件，测得均方差为 **0.33.** 　试在显著性水平α **= 0.05** 下检验　**1)** σ^2无明显变化；　**2)** σ^2明显增大.

例 10　某种柴油发动机，每升柴油的运转时间服从正态分布，现测试 **6** 台柴油机，每升柴

油的运转时间为 28，27，31，29，30，27（分钟），按设计要求每升柴油的运转时间平均应在 30 分钟以上，问在显著性水平 α =0.05 下，这种柴油机是否符合设计要求？

【$H_0: \mu \geq 30; H_1: \mu < 30,$ **拒绝域为** $\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}} < -t_{1-\alpha/2}(n-2)$，

$\overline{x} = 28.67, s^2 = 1.633^2; \dfrac{\overline{x}-\mu_0}{S/\sqrt{n}} = -2.0 > -t_{1-\alpha/2}(n-2) = -2.015$，**接受** H_0】

例 11 测得两批电子器件的样品的电阻（欧姆）为

样品 A(x)	0.140	0.138	0.143	0.142	0.144	0.137
样品 B (y)	0.135	0.140	0.142	0.136	0.138	0.140

设这两批器件的电阻值总体分别服从 $N(\mu_1,\sigma_1^2)$ $N(\mu_2,\sigma_2^2)$，且两样本独立.

（1） 检验假设（α =0.05）

$$H_0: \ \sigma_1^2 = \sigma_2^2; \quad H_1: \sigma_1^2 \neq \sigma_2^2$$

（2） 在（1）的基础上检验（α =0.05）

$$H_0': \ \mu_1 = \mu_2; \quad H_1': \mu_1 \neq \mu_2$$

【（1） $F = \dfrac{S_1^2}{S_2^2}$，**拒绝域为** $F \geq F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)$ 或 $F \leq F_{\frac{\alpha}{2}}(n_1-1, n_2-1)$，

接受 H_0

（2）**拒绝域为，** $\dfrac{|\overline{X}-\overline{Y}|}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \geq t_{1-\frac{\alpha}{2}}(n_1+n_2-2)$ ，**接受** H_0'

$\overline{x} = 0.14067, \overline{y} = 0.1385; s_1^2 = 7.866\times 10^{-6}, s_1^2 = 7.1\times 10^{-6}; s_w = 0.002732$

$F_{0.025}(5,5) = 7.15, F_{0.975}(5,5) = 0.13986; t_{1-\alpha/2}(10) = 2.2281$】

例 12 一药厂生产一种新的止痛片，厂方希望验证服用新药片后至开始起作用的时间间隔较原有止痛片至少缩短一半，因此厂方提出需检验假设

$$H_0: \ \mu_1 = 2\mu_2; \quad H_1: \mu_1 > 2\mu_2$$

此处，μ_1, μ_2 分别是服用原有止痛片和服用新止痛片后起作用的时间间隔的总体的均值。

设两总体均为正态，且方差分别为 σ_1^2，σ_2^2，现分别在两总体中任取一样本 $x_1, x_2, \cdots, x_{n_1}$ 和

$y_1, y_2, \cdots, y_{n_2}$，设两样本独立，试给出上述假设 H_0 的拒绝域（显著性水平为 α）

【检验统计量： $\dfrac{\overline{X}-2\overline{Y}}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{4\sigma_2^2}{n_2}}} \sim N(0,1)$**，拒绝域：** $\dfrac{\overline{X}-2\overline{Y}}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{4\sigma_2^2}{n_2}}} \geq u_{1-\alpha}$ **】**

附录 假设检验与区间估计的对照

以单正态总体 $X \sim N(\mu, \sigma^2)$，σ^2 未知为例，考察参数 μ 的区间估计和假设检验问题：

区间估计	假设检验（原假设为 $H_0: \mu = \mu_0$）
枢轴变量：$T = \dfrac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$	检验统计量： $T = \dfrac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}} \sim t(n-1)$
双侧区间：由 $\lvert t \rvert < t_{1-\frac{\alpha}{2}}(n-1)$ 解出 μ 即 $(\overline{X} - t_{1-\frac{\alpha}{2}}(n-1)\dfrac{\sigma}{\sqrt{n}},\ \overline{X} + t_{1-\frac{\alpha}{2}}(n-1)\dfrac{\sigma}{\sqrt{n}})$	双边检验拒绝域： $\lvert t \rvert > t_{1-\frac{\alpha}{2}}(n-1)$ 即 $t > t_{1-\frac{\alpha}{2}}(n-1)$ 或 $t < -t_{1-\frac{\alpha}{2}}(n-1)$
单侧置信上限：$\overline{X} + t_{1-\alpha}(n-1)\dfrac{\sigma}{\sqrt{n}}$ 单侧置信下限：$\overline{X} - t_{1-\alpha}(n-1)\dfrac{\sigma}{\sqrt{n}}$	右边检验 $(H_1: \mu > \mu_0)$ 拒绝域： $t > t_{1-\alpha}(n-1)$； 左边检验 $(H_1: \mu < \mu_0)$ 拒绝域： $t < -t_{1-\alpha}(n-1)$

- ***p*** **值**

p 和控制犯第一类错的概率 α 的关系：

Neyman-Pearson 派：α 是建立在 $H_0 \leftrightarrow H_1$ 体系下的，需要事先设定 H_0、H_1，通过事先给定 α，即犯第一类错误的概率来划定拒绝域的边界。
Neyman-Pearson 派不会刻意关注 p 的具体值，他们也不去计算 p 值，而是只看观察结果是否落在拒绝域之内。

Fisher 派：p 值，是在原假设正确的前提下，出现观察结果或比之更极端情形的概率，其计算根本不涉及备择假设。**Fisher** 派在计算出 p 值以后，根据其大小判断是否拒绝原假设。认为 p 值的门限可以取 0.05 或 0.01 等等比较小的数。把这个门限称为显著性水平。

- 定义假设检验的 ***p*** 值是：
-

1) 在原假设为真的前提下出现观察样本以及更极端情况的概率。
2) 利用样本数据能拒绝原假设的最小显著性水平。
3) 观察到的(实例样本数据的)显著性水平。
4) 表示对原假设的支持程度，是用于确定是否应该拒绝原假设的另一种方法。

- ***p*** **值的计算：**

设 T 表示假设检验的检验的统计量，当 **H0** 为真时，可由样本数据计算出该统计量的观测值 C，根据检验统计量 T 的具体分布，可求出 p 值。

（**1**）左侧检验的 p 值为：$p = P(T \le C)$；

（**2**）右侧检验的 p 值为：$p = P(T \ge C)$；

（**3**）双侧检验的 p 值为：检验统计量 T 落在样本统计值 ***C*** 为端点的尾部区域内的概率的 **2** 倍：即

$p = 2P(T \ge C)$ **(**当 **C** 位于分布曲线的右端时**)**；

$p = 2P(T \leq C)$ **(当 C 位于分布曲线的左端时)**

若 ***T*** 服从正态分布和 ***t*** 分布，其分布曲线是关于纵轴对称的，则 $p = P(|T| \geq C)$

- 如何利用 ***p*** 值的检验

  计算出 p 值后，将给定的显著性水平 α 与 p 值比较，就可作出检验的结论：

如果 p 值 $< \alpha$，则在显著性水平 α 下拒绝原假设。

如果 p 值 $\geq \alpha$，则在显著性水平 α 下不拒绝原假设。