# 概率论与数理统计B 第三次习题课

**主要内容: 区间估计, 假设检验, 13, 14周作业**

**2020.06.03　朱心远　PB17000015　xyzhu2001@gmail.com**

---

**统计推断:**

**1. 估计  1.1. 点估计 1.1.1. 矩估计 1.1.2. MLE 1.1.3. 点估计的优良准则**

**1.2 区间估计**

**2. 假设检验**

## 1. 区间估计

### *1.1 基本概念*

**定义.** 设总体分布 $F(x,\theta)$ 含有一个或多个未知的参数 $\theta,\theta\in\Theta$, 对给定的值 $\alpha,(0<\alpha<1)$, 若由样本 $X_1,\cdots,X_n$ 确定的两个统计量 $\bar{\theta}=\bar{\theta}\left(X_1,\cdots,X_n\right)$ 和 $\underline{\theta}=\underline{\theta}\left(X_1,\cdots,X_n\right)$, 满足

$$P_\theta(\underline{\theta}\le\theta\le\bar{\theta})=1-\alpha\quad\forall\theta\in\Theta$$

称$1-\alpha$为置信系数或置信水平, 而称$[\underline{\theta},\bar{\theta}]$ 为 $\theta$ 的置信水平为$1-\alpha$ 的置信区间.

值得指出的是, 这里的$\theta$虽然未知但是是常数, 而$\underline{\theta},\bar{\theta}$则是样本的函数, 是随机变量, 因此$(\underline{\theta},\bar{\theta})$是一个随机区间, 置信区间的更准确的讲法是随机区间$(\underline{\theta},\bar{\theta})$将以概率$1-\alpha$覆盖参数$\theta$.

**枢轴变量法**. 设待估参数为 $g(\theta)$,

- 找一个与待估参数 $g(\theta)$ 有关的统计量$T$，一般是其一个良好的点估计(多数是通过极大似然估计构造);
- 设法找出$T$与 $g(\theta)$ 的某一函数 $S(T,g(\theta))$ 的分布，其分布$F$要与参数 $\theta$ 无关 ($S$ 即为枢轴变量);
- 对任何常数 $a<b$, 不等式 $a\le S(T,g(\theta))\le b$ 要能表示成等价的形式 $A\le g(\theta)\le B$, 其中 $A,B$ 只与 $T,a,b$ 有关而与参数无关;
- 取分布F的上$\alpha/2$分位数$\omega_a/2$和上$(1-\alpha/2)$分位数 $\omega_{1-\alpha/2}$, 有 $F\left(\omega_{\alpha/2}\right)-F\left(\omega_{1-\alpha/2}\right)=1-\alpha$. 因此

$$P\left(\omega_{1-\alpha/2}\le S(T,g(\theta))\le\omega_{\alpha/2}\right)=1-\alpha$$

**区间长度与置信度.**

- 估计精度: 置信区间的区间长度 $\hat{\theta}_2-\hat{\theta}_1$ 是区间估计的估计精度, 区间长度越短, 精度越高.
- 置信度: 置信度的大小反映了这个区间估计的可靠程度, 即随机区间包含待估参数的概率的大小.
- 精度与置信度的关系: 在样本容量 $n$ 一定的前提下, 精度与置信度是彼此矛盾的. 当置信度增大时, 区间长度也增大, 精度减小. 当置信度减小时, 区间长度缩短, 精度增高. 只有增大样本容量 $n$, 才能同时提高区间估计的置信度与精度.

---

## *1.2 正态总体的区间估计*

### 1.2.1 单总体

| 待估参数 | 其他参数 | 枢轴量 | 分布 | 双侧置信区间上,下限 |
|---|---|---|---|---|
| $\mu$ | $\sigma^2$ 已知 | $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ | $N(0,1)$ | $\overline{X}\pm z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}$ |
| | $\sigma^2$ 未知 | $\frac{\overline{X}-\mu}{S/\sqrt{n}}$ | $t(n-1)$ | $\overline{X}\pm t_{\alpha/2}(n-1)\cdot\frac{S}{\sqrt{n}}$ |
| $\sigma^2$ | $\mu$ 已知 | $\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}$ | $\chi^2(n)$ | $\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}$, $\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}$ |
| | $\mu$ 未知 | $\frac{(n-1)S^2}{\sigma^2}$ | $\chi^2(n-1)$ | $\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}$, $\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}$ |

### 1.2.2 两总体

| 待估参数 | 其他参数 | 枢轴量 | 分布 | 双侧置信区间上,下限 |
|---|---|---|---|---|
| $\mu_1-\mu_2$ | $\sigma_1^2,\sigma_2^2$ 均已知 | $\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}-\frac{\sigma_2^2}{n_2}}}$ | $N(0,1)$ | $\overline{X}-\overline{Y}\pm z_{\alpha/2}\cdot\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$ |
| | $\sigma_1^2=\sigma_2^2=\sigma^2$ 但未知 | $\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ | $t(n_1+n_2-2)$ | $\overline{X}-\overline{Y}\pm t_{\alpha/2}(n_1+n_2-2)\cdot S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ |
| $\frac{\sigma_1^2}{\sigma_2^2}$ | $\mu_1,\mu_2$ 均已知 | $\frac{\sum_{i=1}^{n_1}(X_i-\mu_1)^2/n_1\sigma_1^2}{\sum_{i=1}^{n_2}(Y_i-\mu_2)^2/n_2\sigma_2^2}$ | $F(n_1,n_2)$ | $\frac{1}{f_{\alpha/2}(n_1,n_2)}\cdot\frac{n_2\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{n_1\sum_{i=1}^{n_2}(Y_i-\mu_2)^2}$, $\frac{1}{f_{1-\alpha/2}(n_1,n_2)}\cdot\frac{n_2\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{n_1\sum_{i=1}^{n_2}(Y_i-\mu_2)^2}$ |
| | $\mu_1,\mu_2$ 均未知 | $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$ | $F(n_1-1,n_2-1)$ | $\frac{1}{f_{\alpha/2}(n_1-1,n_2-1)}\cdot\frac{S_1^2}{S_2^2}$, $\frac{1}{f_{1-\alpha/2}(n_1-1,n_2-1)}\cdot\frac{S_1^2}{S_2^2}$ |

其中 $S_w=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$.

---

## ** 1.3 非正态总体的区间估计(大样本)*

### 1.3.1 正态逼近法

### 1.3.2 比例数$p$的估计

### 1.3.3 样本量的确定

> 阅读 Slides "参数估计(下)" 67-85 页

---

# 2.区间估计作业

**7-63.** 随机从一批钉子中抽取9枚, 测得其长度(cm)为:

$$2.15, 2.13, 2.10, 2.14, 2.15, 2.16, 2.12, 2.11, 2.13$$

假设钉子长度服从正态分布, 分别在下面两种情况下, 求出总体均值的90%置信区间:

(1) $\sigma = 0.01$　　　(2) $\sigma$ 未知

解:

首先, $\bar{X} = 2.132, \quad S = 0.0199, \quad z_{0.05} = 1.645, \quad t_{0.05}(8) = 1.860, \quad n = 9, \quad \alpha = 0.1$

(1) 在$\sigma$已知的情况下, $\mu$的区间估计是

$$\left[\bar{X} - \frac{z_{\alpha/2}\sigma}{\sqrt{n}}, \bar{X} + \frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right]$$

代入对应参数即可得到置信区间为

$$[2.127, 2.137]$$

(2) 在$\sigma$未知的情况下, $\mu$的区间估计是

$$\left[\bar{X} - \frac{t_{\alpha/2}(n-1)S}{\sqrt{n}}, \bar{X} + \frac{t_{\alpha/2}(n-1)S}{\sqrt{n}}\right]$$

代入对应参数即可得到置信区间为

$$[2.120, 2.144]$$

---

**7-72.** 假设用机器包装精盐的重量服从正态分布, 现从生产线上随机地抽取10袋, 测得其重量为(单位: 克):

$501.5, 500.7, 492.0, 504.7, 483.0, 512.8, 504.0, 490.3, 486.0, 520.0$

试在下列条件下分别求总体方差的95%和 90% 置信区间.

(1) $\mu = 500g$　　　(2) $\mu$ 未知

解:

(1)

$$\sum_{i=1}^{10} (x_i - \mu)^2 = 1247.76$$
$$\chi^2_{0.05}(10) = 18.31, \quad \chi^2_{0.95}(10) = 3.94$$
$$\chi^2_{0.025}(10) = 20.48, \quad \chi^2_{0.975}(10) = 3.25$$

$\mu$已知时, $\sigma^2$的置信区间是

$$\left[\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}, \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right]$$

令$\alpha=0.05$, 可以得到置信区间为:

$$[60.93, 380.93]$$

令$\alpha=0.1$, 可以得到置信区间为:

$$[68.15, 316.69]$$

(2)

$\mu$未知时, $\sigma^2$的置信区间是

$$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right]$$

而

$$\begin{gathered} S^2=138.362 \\ \chi^2_{0.05}(9)=16.92, \quad \chi^2_{0.95}(9)=3.33 \\ \chi^2_{0.025}(9)=19.02, \quad \chi^2_{0.975}(9)=2.70 \end{gathered}$$

令$\alpha=0.05$, 可以得到置信区间为:

$$[65.47, 461.21]$$

令$\alpha=0.1$, 可以得到置信区间为:

$$[73.60, 373.95]$$

---

**7-73.** 随机抽取16发子弹做试验, 测得子弹速度的样本标准差为 $S=12$ , 假设子弹速度服从正态分布$N(\mu,\sigma^2)$, 分别求 $\sigma$ 和 $\sigma^2$ 的95%置信区间.

解:

$\mu$未知时, $\sigma^2$的置信区间是

$$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right]$$

而

$$S^2=144, \quad \chi^2_{0.025}(15)=27.49, \quad \chi^2_{0.975}(15)=6.26$$

代入即得

$$[78.57, 345.05]$$

而$\sigma$的置信区间就是

$$[8.86, 18.58]$$

---

**7-86.** 为了了解一批灯泡的使用寿命, 共测试了16个灯泡的寿命, 得到寿命的平均为1600h, 样本标准差为15 h. 假设寿命服从正态分布$N(\mu,\sigma^2)$.

(1) 求$\mu$的95%置信下限

(2) 求$\sigma^2$的95%置信上限

解:

**这里我们计算单侧置信下限**

(1) 未知$\sigma^2$时, $\mu$的$1-\alpha$单侧置信下限是

$$\bar{X}-\frac{t_\alpha(n-1)S}{\sqrt{n}}$$

代入

$$t_{0.05}(15)=1.753, S=15, n=16, \bar{X}=1600$$

得到单侧置信下限为$1593.43$.

(2) 未知$\mu$时, $\sigma^2$的单侧置信上限是

$$\frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}$$

代入

$$\chi^2_{0.95}(15)=7.26, S=15, n=16$$

得到单侧置信上限为$464.88$.

---

# 3. 假设检验

## *3.1 基本概念和问题提法*

### 3.1.1 零假设, 对立假设, 两类错误, 拒绝域, 显著性水平, 检验统计量

- 原假设和备择假设: 当总体类型已知时, 对分布的一个或几个未知参数的值作出假设, 或者对总体分布的类型或某些特征提出某种假设, 这种假设称为原假设或零假设, 通常用$H_0$表示, 在抛弃原假设后可供选择的假设称为备择假设, 记为 $H_1$, $H_0$ 与 $H_1$ 是互不相容的.
- 在总体类型已知, 仅有某个或某几个参数未知的情况下, 对未知参数作出假设称为参数假设. 如对总体的某些特征提出假设, 则称为非参数假设.
- 拒绝域, 接受域和临界值: 检验统计量把样本空间分成两个区域, 使原假设$H_0$被拒绝的样本观察值 $x_1, x_2, \cdots, x_n$ 所组成的区域称为拒绝域. 而保留原假设$H_0$ 的样本观察值所组成的区域称为接受域. 拒绝域与接受域都是样本空间的子集, 并且是互不相容的, 而其区域之和为样本空间. 拒绝域与接受域的分界线处于一个特殊的地位, 当样本越过这个分界线时, 结论就发生了根本性的改变, 因此把分界线的值称为临界值.
- 第I类错误: 弃真; 第II类错误: 取伪

| | $H_0$ 成 立 | $H_1$ 成 立 |
|---|---|---|
| 接 受 $H_0$ | 不 犯 错 | 第 II 类 错 误 |
| 拒 绝 $H_0$ | 第 I 类 错 误 | 不 犯 错 |

- 解决两类错误之间的矛盾: 在控制第I类错误的基础上, 尽量少犯第II类错误. 这样的选取对原假设有保护作用.

- 显著性水平$\alpha$: 犯第I类错误的概率小于$\alpha$. 如果$\alpha$是显著性水平, 那么任意大于$\alpha$的数都是显著性水平, 通常指最小的那一个.

### 3.1.2 如何提出假设检验问题？

**原则一: 将受保护的对象置为零假设.** 如我国按照以前的司法制度, 公安机关抓到嫌疑犯后, 很多情况下要犯人自己证明无罪(有罪推断)，这对嫌疑犯很不利, 从而容易导致冤案. 现在的司法制度则总假定嫌疑犯是无罪的, 要司法部门证明其有罪(无罪推断), 这样做大大地有利于保护公民的利益, 如果要将真正的嫌疑犯绳之以法, 则司法部门必须有充分的证据, 这样做可以有效保护公民的权益, 对司法部门要求也变高了。又比如药厂生产出一种新药, 在上市前要通过食品与药品监管局的检验. 显然使用药品的病人是应该受保护的对象, 这时应该设定一个有利于病人的命题作为零假设, 这个命题就是“新药不比安慰剂效果好”, 以尽量避免病人用无效甚至有副作用的新药. 当然, 对立假设就是“新药比安慰剂效果好，将检验的显著性水平 $\alpha$ 设定得较小, 以证零假设不被轻易推翻. 在实际问题中, 如果根据某个合理的检验方法发现零假设被推翻, 则有充分的理由认为零假设不成立而对立假设成立, 这是因为万一零假设成立而被误据的概率不会超过 $\alpha$; 另一方面, 如果发现零假设未被拒绝, 并不表明有充分理由接受零假设, 而是因为零假设被保护得较严密以至于未被拒绝.

**原则二: 如果你希望“证明”某个命题, 就取相反结论或者其中一部分作为零假设(类似于反证法).** 这种提法往往是在两个假设命题中不太清楚哪个应受保护, 此时可以借用司法制度里的“谁主张, 谁举证”, 即若想用统计方法向人“证明”一个命题, 则将那个命题置为对立假设. 注意这里的证明不是数学上的严格证明, 而是允许犯错的一种统计推断方法. 用统计方法证明一个命题不是一件容易的事情, 所以如果没有足够把握, 人们应该避免用统计方法去证明一个命题.

上述两原则是统一的: 一般不应该让受保护对象去证明一个命题。

---

## *3.2 常见的参数检验问题:*

| 检验对象 | 检验统计量 | 分布 | 拒绝域† |
|---|---|---|---|
| $\mu$ ($\sigma^2$已知) | $U=\sqrt{n}(\bar{X}-\mu_0)/\sigma$ | $N(0,1)$ | $\begin{cases} \lvert U\rvert>u_{\alpha/2} \\ U>u_\alpha \\ U<-u_\alpha \end{cases}$ |
| $\mu$ ($\sigma^2$ 未知) | $T=\sqrt{n}(\bar{X}-\mu_0)/S$ | $t_{n-1}$ | $\begin{cases} \lvert T\rvert>t_{n-1}(\alpha/2) \\ T>t_{n-1}(\alpha) \\ T<-t_{n-1}(\alpha) \end{cases}$ |
| $\sigma^2$ ($\mu$ 已知) | $\chi^2=\frac{1}{\sigma_0^2}\sum_{i=1}^n(X_i-\mu)^2$ | $\chi_n^2$ | $\begin{cases} \chi^2>\chi_n^2(\alpha/2)\text{或者}\chi^2<\chi_n^2(1-\alpha/2) \\ \chi^2>\chi_n^2(\alpha) \\ \chi^2<\chi_n^2(1-\alpha) \end{cases}$ |
| $\sigma^2$ ($\mu$ 未知) | $\chi^2=\frac{1}{\sigma_0^2}\sum_{i=1}^n(X_i-\bar{X})^2$ | $\chi_{n-1}^2$ | $\begin{cases} \chi^2>\chi_{n-1}^2(\alpha/2)\text{或者}\chi^2<\chi_{n-1}^2(1-\alpha/2) \\ \chi^2>\chi_{n-1}^2(\alpha) \\ \chi^2<\chi_{n-1}^2(1-\alpha) \end{cases}$ |

†有关均值的检验：对立假设分别为$\mu\neq\mu_0$，$\mu>\mu_0$ 和$\mu<\mu_0$．有关方差的检验：对立假设分别为$\sigma^2\neq\sigma_0^2$, $\sigma^2>\sigma_0^2$ 和$\sigma^2<\sigma_0^2$.

| 检验对象 | 检验统计量 | 分布 | 拒绝域† |
|---|---|---|---|
| 均值(方差已知) | $U=\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}$ | $N(0,1)$ | $\begin{cases} \lvert U\rvert>u(\alpha/2) \\ U>u(\alpha) \\ U<-u(\alpha) \end{cases}$ |
| 均值(方差未知)‡ | $T=\frac{\bar{X}-\bar{Y}}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}}$ | $t_{m+n-2}$ | $\begin{cases} \lvert T\rvert>t_{m+n-2}(\alpha/2) \\ T>t_{m+n-2}(\alpha) \\ T<-t_{m+n-2}(\alpha) \end{cases}$ |
| 方差(均值已知) | $F=\frac{\sum_{i=1}^m(X_i-\mu_1)^2/m}{\sum_{i=1}^n(X_i-\mu_2)^2/n}$ | $F_{m,n}$ | $\begin{cases} F>F_{m,n}(\alpha/2)\text{或}F<\frac{1}{F_{n,m}(\alpha/2)} \\ F>F_{m,n}(\alpha) \\ F<\frac{1}{F_{n,m}(\alpha)} \end{cases}$ |
| 方差(均值未知) | $F=\frac{S_1^2}{S_2^2}$ | $F_{m-1,n-1}$ | $\begin{cases} F>F_{m-1,n-1}(\alpha/2)\text{或}F<\frac{1}{F_{n-1,m-1}(\alpha/2)} \\ F>F_{m-1,n-1}(\alpha) \\ F<\frac{1}{F_{n-1,m-1}(\alpha)} \end{cases}$ |

†有关均值的检验：对立假设分别为$\mu_1\neq\mu_2$，$\mu_1>\mu_2$ 和$\mu_1<\mu_2$．有关方差的检验：对立假设分别为$\sigma_1^2\neq\sigma_2^2$, $\sigma_1^2>\sigma_2^2$ 和$\sigma_1^2<\sigma_2^2$.
‡假定方差相等

***备择假设中一般不含有等号***.

---

## 3.3 拟合优度检验

设某总体X服从一个离散分布, 且根据经验得知总体落在类别 $a_1,\cdots,a_k$ 的理论频率分别为 $p_1,\cdots,p_k$, 现从该总体抽得一个样本量为n的样本, 其落在类别 $a_1,\cdots,a_k$ 的观测数分别为 $n_1,\cdots,n_k$. 感兴趣的问题是检验理论频率是否正确, 即下面假设是否正确:

$$H_0: P(X\in a_1)=p_1,\cdots,P(X\in a_k)=p_k.$$

这类问题只提零假设而不提对立假设, 相应的检验方法称为拟合优度检验. 显然, 在零假设下, 各类别的理论频数分别为$np_1,..,np_k$, 将理论频数和观测频数列于下表:

| 类 别 | $a_1$ | $a_2$ | $\cdots$ | $a_k$ |
|---|---|---|---|---|
| 理 论 频 数 | $np_1$ | $np_2$ | $\cdots$ | $np_k$ |
| 观 测 频 数 | $n_1$ | $n_2$ | $\cdots$ | $n_k$ |

由大数定律知, 在零假设成立时, $n_i/n$ 依概率收敛于 $p_i$, 故理论频数$np_i$与观测频数$n_i$接近. 而检验统计量取为

$$\chi^2=\sum_{i=1}^{k}\frac{(n_i-np_i)^2}{np_i}\sim\chi^2(k-1)$$

当理论总体总含有未知的参数时，理论频数$np_i$一般也与这些参数有关, 此时应该用适当的估计如极大似然估计代替这些参数以得到 $p_i$ 的估计 $\hat{p}_i$, 得到的统计量记为

$$\chi^2=\sum_{i=1}^{k}\frac{(n_i-n\hat{p}_i)^2}{n\hat{p}_i}$$

拟合优度检验的提出者Karl Pearson最初认为在零假设下, 检验统计量的 $\chi^2$ 的极限分布仍等于自由度为 $k-1$ 的 $\chi^2$ 分布, R.A. Fisher 发现自由度应该等于 $k-1$ 减去估计的独立参数的个数r, 即 $k-1-r$.

---

## *3.4 列联表的独立性检验, 齐一性检验*

列联表是一种按两个属性作双向分类的表. 例如肝癌病人可以按所在医院(属性A) 和是否最终死亡(属性B) 分类. 目的是看不同医院的疗效是否不同. 又如婴儿可按喂养方式(属性A, 分两个水平: 母乳喂养与人工喂养) 和小儿牙齿发育状况(属性B, 分两个水平: 正常与异常) 来分类. 这两个例子中两个属性都只有两个水平, 相应的列联表称为“四格表”, 一般地, 如果第一个属性有$a$个水平, 第二个属性有$b$个水平, 称为 $a\times b$ 表. 实际应用中, 常见的一个问题是考察两个属性是否独立。即零假设是$H_0$ : 属性A 与属性B 独立. 这是列联表的独立性检验问题。

| | $B_1$ | $B_2$ | $\cdots$ | $B_b$ | 合 计 |
|---|---|---|---|---|---|
| $A_1$ | $n_{11}$ | $n_{12}$ | $\cdots$ | $n_{1b}$ | $n_{1\bullet}$ |
| $A_2$ | $n_{21}$ | $n_{22}$ | $\cdots$ | $n_{2b}$ | $n_{2\bullet}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ | |
| $A_a$ | $n_{a1}$ | $n_{a2}$ | $\cdots$ | $n_{ab}$ | $n_{a\bullet}$ |
| 合 计 | $n_{\bullet 1}$ | $n_{\bullet 2}$ | $\cdots$ | $n_{\bullet b}$ | $n$ |

假设样本量为n, 第 $(i,j)$ 格的频数为 $n_{ij}$. 记 $p_{ij}=P($属性A,B分别处于水平 $i,j), p_i=P($属性A 有水平 $i), q_i=P($ 属性B 有水平 $j)$ . 则零假设就是 $p_{ij}=p_iq_j$. 将 $p_i$ 和 $q_j$ 看成参数, 则总的独立参数有 $a-1+b-1=a+b-2$ 个. 它们的极大似然估计为

$$\hat{p}_i=\frac{n_{i\bullet}}{n},\hat{q}_j=\frac{n_{\bullet j}}{n}$$

正好是它们的频率(证明参看教材) . 其中 $n_{i\bullet}=\sum_{j=1}^{b}n_{ij}, n_{\bullet j}=\sum_{i=1}^{a}n_{ij}$. 在 $H_0$ 下, 第 $(i,j)$ 格的理论频数为 $n\hat{p}_{ij}=n_{i\bullet}n_{\bullet j}/n$, 因此在 $H_0$ 下 , $\sum_{i=1}^{a}\sum_{j=1}^{b}\left(n_{ij}-n\hat{p}_{ij}\right)$ 应该较小。故取检验统计量为

$$\chi^2=\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{(n_{ij}-n_{i\bullet}n_{\bullet j}/n)^2}{(n_{i\bullet}n_{\bullet j}/n)}=n\left(\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{i\bullet}n_{\bullet j}}-1\right)$$

在零假设下 $\chi^2$ 的极限分布是有自由度为 $k-1-r=ab-1-(a+b-2)=(a-1)(b-1)$的 $\chi^2$ 分布. 对于四格表, 自由度为$1$.

| | $B_1$ | $B_2$ | 合 计 |
|---|---|---|---|
| $A_1$ | $a$ | $b$ | $a+b$ |
| $A_2$ | $c$ | $d$ | $c+d$ |
| 合 计 | $a+c$ | $b+d$ | $n=a+b+c+d$ |

$$\chi^2=\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(n_{ij}-n\hat{p}_{i\bullet}\hat{p}_{\bullet j}\right)^2}{n\hat{p}_{i\bullet}\hat{p}_{\bullet j}}=\frac{n(ad-bc)^2}{(a+b)(a+c)(b+d)(c+d)}$$

跟列联表有关的另一类重要的检验是齐一性检验，即检验某一个属性A 的各个水平对应的另一个属性B的分布全部相同，这种检验跟独立性检验有着本质的区别. 独立性问题中两属性都是随机的; 而齐一性问题中属性A是非随机的，这样涉及到的分布实际上是条件分布. 虽然如此, 所采用的检验方法跟独立性检验完全一样。

---

# 4. 假设检验作业

**8-1.** 假设$X_1, \cdots, X_{16}$服从正态分布$N(\mu, 0.16)$, 检验问题

$$H_0 : \mu = 0.5 \leftrightarrow H_1 : \mu > 0.5$$

显著水平为$0.05$

(1) 检验的拒绝域是什么?

(2) $\mu = 0.65$时犯第二类错误的概率是多少?

解:

(1) 拒绝域是

$$W = \{z > u_{0.05}\} = \{z > 1.645\}$$

而$z = \frac{\bar{X}-0.5}{0.4/\sqrt{16}}$, 即$\bar{X} = 0.1z + 0.5$, 则拒绝域对应$\{\bar{X} > 0.6645\}$.

(2)

犯第二类错误, 则要接受$H_0$, 即满足$\bar{X} > 0.6645$.

$$P(\bar{X} < 0.6645) = P(Z < \frac{0.6645 - 0.65}{0.1}) = \Phi(0.145) \approx 0.56.$$

---

**8-24.** 用传统工艺加工的某种水果罐头中每瓶维生素C的含量平均为19毫克, 现采用一种新的加工工艺, 试图减少在加工过程中对维生素C的破坏, 抽查了16瓶罐头, 测得维生素C的含量(单位: 毫克)为:

$$23, 20.5, 21, 20, 22.5, 19, 20, 23, 20.5, 18.8, 20, 19.5, 22, 18, 23, 22.$$

已知水果罐头中维生素C的含量服从正态分布. 在方差未知的情况下, 问新工艺下维生素的含量是否比旧工艺有所提高($\alpha = 0.01$)?

解:

$$H_0 : \mu = 19 \qquad H_1 : \mu > 19.$$

检验统计量$t = \frac{\bar{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1)$

计算和查表可以得到

$$\bar{X} = 20.8 \quad n = 16 \quad S \approx 1.617 \quad t_{0.01}(15) = 2.602$$

注意到

$$t = \frac{20.8 - 19}{1.617/4} \approx 4.453 > t_{0.01}(15) = 2.602$$

因此拒绝$H_0$, 接受$H_1$, 即认为新工艺下维生素的含量有所提高.

**8-29.** 某车间生产铜丝, 生产一向比较稳定. 今从产品中随机抽取10根检查其折断力, 得数据如下(单位: kg):

$$288.8, 294.7, 300.2, 286.6, 290.3, 280.1, 296.4, 295.4, 290.2, 289.2.$$

假设铜丝的折断力服从正态分布, 问是否可以相信该车间生产的铜丝的折断力的方差是$16(\alpha=0.05)$?

解:

$$H_0: \sigma^2 = 16 \qquad H_1: \sigma^2 \neq 16$$
$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

计算和查表可以得到:

$$S^2 = 32.652 \quad n = 10 \quad \chi^2_{0.975}(9) = 2.70 \quad \chi^2_{0.025}(9) = 19.02$$

注意到

$$\chi^2 = \frac{9 \times 32.652}{16} = 18.267 \in [\chi^2_{0.975}(9), \chi^2_{0.025}(9)] = [2.70, 19.02]$$

因此接受$H_0$, 相信该车间生产的铜丝的折断力的方差是16.

---

**8-35.** 一个以减肥为主要目的的健美俱乐部声称, 参加其训练班至少可以使肥胖者平均减少体重8kg以上. 为检验该宣传是否可信, 调查人员随机调查了9名参加者, 得到他们训练前后的体重数据如下(单位:kg):

| 训练前 | 104.5 | 94.0 | 104.7 | 96.4 | 91.6 | 90.9 | 92.0 | 99.9 | 109.7 |
|---|---|---|---|---|---|---|---|---|---|
| 训练后 | 94.2 | 86.6 | 97.5 | 91.7 | 82.6 | 83.8 | 81.3 | 92.2 | 101.0 |

现假设训练前后人的体重均服从正态分布. 问在0.05的显著水平下, 是否可以认为该俱乐部的宣传是可信的?

解:

计算可得样本差值为

$$10.3, 7.4, 7.2, 4.7, 9.0, 7.1, 10.7, 7.7, 8.7.$$

检验

$$H_0: \mu \leq 8 \qquad H_1: \mu > 8$$
$$t = \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}} \sim t(n-1)$$

计算和查表可以得到

$$\bar{X}_n = 8.089 \quad S = 1.830 \quad n = 9 \quad t_{0.05}(8) = 1.860$$

注意到

$$t = \frac{8.089 - 8}{1.830/3} = 0.146 < t_{0.05}(8) = 1.860$$

因此接受$H_0$, 不能认为宣传是可信的.

---

**8-40.** 为了解甲、乙两企业职工工资水平, 分别从两企业各随机抽取若干名职工调 查, 得如下数据(单位: 元):

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 甲公司 | 3750 | 5300 | 3750 | 9100 | 5700 | 5250 | 5000 | |
| 乙公司 | 5000 | 9500 | 4500 | 9000 | 6000 | 8500 | 9750 | 6000 |

设两企业职工工资分别服从正态分布, 而总体独立且均值方差均未知. 试根据以上数据判断: 两企业职工工资的方差是否相等? 甲企业职工平均工资是否低于乙企业职工平均工资($\alpha = 0.05$)?

解:

(1)

$$H_0 : \sigma_1^2 = \sigma_2^2 \qquad H_1 : \sigma_1^2 \neq \sigma_2^2$$
$$F = \frac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1)$$

计算和查表得到

$$S_1^2 = 3236190, S_2^2 = 4525670 \quad n_1 = 7, n_2 = 8$$
$$F_{0.025}(6,7) = 5.120, F_{0.975}(6,7) = 0.175$$

注意到

$$F = \frac{3236190}{4525670} = 0.715 \in [F_{0.975}(6,7), F_{0.025}(6,7)] = [0.175, 5.120]$$

因此接受$H_0$, 即认为两企业职工工资方差相等.

(2)

$$H_0 : \mu_1 \geqslant \mu_2 \qquad H_1 : \mu_1 < \mu_2$$
$$t = \frac{\bar{X}_n - \bar{Y}_n}{S_w\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2), S_w^2 = \frac{(n_1 - 1)\, S_1^2 + (n_2 - 1)\, S_2^2}{n_1 + n_2 - 2}$$

计算和查表得到

$$S_w^2 = 3930525, n_1 = 7, n_2 = 8, \bar{X}_n = 5407.1, \bar{Y}_n = 7281.3, t_{0.05}(13) = 1.771$$

注意到

$$t = \frac{5407.1 - 7281.3}{\sqrt{3930525(\frac{1}{7} + \frac{1}{8})}} = -1.827 < -t_{0.05}(13) = -1.771$$

因此拒绝$H_0$, 接受$H_1$, 即可以认为甲企业职工平均工资低于乙企业职工平均工资.

---

**8-62.** 某工厂为了了解白班和夜班的产品合格率是否有差异, 进行调查得到如下数据

| | 合格 | 不合格 |
|---|---|---|
| 白班 | 232 | 19 |
| 夜班 | 54 | 18 |

试据此判断, 产品合格率是否与班次有关?($\alpha = 0.05$)

解:

$$H_0 : \text{产品合格率与班次无关} \qquad H_1 : \text{产品合格率与班次有关}$$

$$V_n = \frac{(232 + 54 + 19 + 18)(232 \times 18 - 19 \times 54)^2}{(232 + 52)(232 + 19)(18 + 19)(18 + 54)} = 16.759 > \chi^2_{0.05}(1) = 3.84.$$

拒绝$H_0$, 可以认为产品合格率与班次有关.

---

**8-63.** 为了解男性和女性对三种类型的啤酒: 淡啤酒、普通啤酒和黑啤酒的偏好有没有差异, 分别调查了180位男士和120位女士的喜好, 得如下数据

| | 淡啤酒 | 普通啤酒 | 黑啤酒 |
|---|---|---|---|
| 男性 | 49 | 31 | 100 |
| 女性 | 51 | 20 | 49 |

请问男性和女性对这三种类型的啤酒的偏好有显著差异吗?($\alpha = 0.05$)

解:

$H_0$ : 男性和女性对这三种类型的啤酒的偏好没有显著差异　　$H_1$ : 男性和女性对这三种类型的啤酒的偏好有显著差异

$$\begin{aligned}V_n &= n\left(\sum_{i=1}^{2}\sum_{j=1}^{3}\frac{n_{ij}^2}{n_{i\bullet}n_{\bullet j}} - 1\right)\\ &= 300\times\left(\frac{49^2}{180\times 100}+\frac{31^2}{180\times 51}+\frac{100^2}{180\times 149}+\frac{51^2}{120\times 100}+\frac{20^2}{120\times 51}+\frac{49^2}{120\times 149}-1\right)\\ &= 8.197\\ &> \chi^2_{0.05}(2) = 5.99\end{aligned}$$

拒绝$H_0$, 可以认为男性和女性对这三种类型的啤酒的偏好有显著差异.

---

**8-64.** 检查一本书的 150 页, 记录各页中印刷错误的个数, 其结果为

| 错误的个数 $f_i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | $\geq 7$ |
|---|---|---|---|---|---|---|---|---|
| 含 $f_i$ 个错误的页数 | 86 | 40 | 19 | 2 | 0 | 2 | 1 | 0 |

试在显著性水平0.05下检验假设$H_0$: 每页上的印刷错误个数服从泊松分布.

解:

首先估计$H_0$成立时, 泊松分布$P(\lambda)$的参数$\lambda$, 我们知道, $\lambda$的极大似然估计是

$$\hat{\lambda} = \bar{X} = \frac{1}{150}(40 + 19\times 2 + 2\times 3 + 5\times 2 + 6\times 1) = \frac{2}{3}.$$

如果进行如下分组

$$l_1 = [0,1),\quad l_2 = [1,2),\quad l_3 = [2,3),\quad l_4 = [3,4),\quad l_5 = [4,5),\quad l_6[5,6),\quad l_7 = [6,+\infty)$$

那么有

$$\begin{aligned}p_1 &= e^{-\hat{\lambda}} = 0.513\\ p_2 &= \hat{\lambda}e^{-\hat{\lambda}} = 0.342\\ p_3 &= \frac{\hat{\lambda}^2}{2!}e^{-\hat{\lambda}} = 0.114\\ p_4 &= \frac{\hat{\lambda}^3}{3!}e^{-\hat{\lambda}} = 0.025\\ p_5 &= \frac{\hat{\lambda}^4}{4!}e^{-\hat{\lambda}} = 0.0042\\ p_6 &= \frac{\hat{\lambda}^5}{5!}e^{-\hat{\lambda}} = 0.00056\\ p_7 &= 1 - \sum_{j=1}^{6}p_j = 0.0012\end{aligned}$$

为了保证分组满足$np_i \geq 5$, 我们将$l_3, l_4, l_5, l_6, l_7$合并为$l_3$, 并计算对应的$\hat{p}$.

$$p_1 = e^{-\hat{\lambda}} = 0.513, \qquad \hat{p}_1 = \frac{86}{150} = 0.573$$
$$p_2 = \hat{\lambda} e^{-\hat{\lambda}} = 0.342, \qquad \hat{p}_2 = \frac{40}{150} = 0.267$$
$$p_3 = 1 - 0.513 - 0.342 = 0.145, \qquad \hat{p}_3 = \frac{24}{150} = 0.160$$

进而

$$\begin{aligned} U &= \sum_{j=1}^{3} \frac{n}{p_j} \left(\hat{p}_j - p_j\right)^2 \\ &= \frac{150}{0.513}(0.573 - 0.513)^2 + \frac{150}{0.342}(0.267 - 0.342)^2 + \frac{150}{0.145}(0.160 - 0.145)^2 \\ &= 3.752 \\ &< \chi^2_{0.05}(1) = 3.84 \end{aligned}$$

所以可以认为每页上的印刷错误个数服从泊松分布.

---

**8-68.** 设某两家工厂生产同一种产品, 产品分 1, 2, 3三个等级(分别表示高, 中, 低三个等级), 现从两家工厂生产的产品中各随机抽取 110 和 100 个产品, 检查产品等级后结果如下

| 工厂 \ 等级 | 1 | 2 | 3 |
|---|---|---|---|
| 甲 | 58 | 40 | 12 |
| 乙 | 40 | 40 | 20 |

试分别在水平$\alpha = 0.1$和$\alpha = 0.05$下判断两家工厂的产品质量等级是否相同? 并解释你的结论.

解:

$$H_0\text{: 两家工厂的产品质量等级相同} \quad vs. \quad H_1\text{: 两家工厂的产品质量等级不同}$$

$$\begin{aligned} V_n &= n\left(\sum_{i=1}^{k}\sum_{j=1}^{l} \frac{(n_{ij})^2}{n_{i\bullet} n_{\bullet j}} - 1\right) \\ &= 210 \times \left(\frac{1}{110} \times \left(\frac{58^2}{98} + \frac{40^2}{80} + \frac{12^2}{32}\right) + \frac{1}{100} \times \left(\frac{40^2}{98} + \frac{40^2}{80} + \frac{20^2}{32}\right) - 1\right) \\ &= 4.841 \end{aligned}$$

而

$$\chi^2_{0.1}(2) = 4.61 < V_n$$
$$\chi^2_{0.05}(2) = 5.99 > V_n$$

结论:

在显著性水平$\alpha = 0.1$下, 不可以认为两家工厂的产品质量等级相同;

在显著性水平$\alpha = 0.05$下, 可以认为两家工厂的产品质量等级相同.

解释:

显著性水平$\alpha$越小, 那么检验的结果越倾向于扩大接受域, 也就是更倾向于接受原假设,