概率论与数理统计 (Fall 2024)
习题课讲义

2024 年 12 月 16 日

聂嵘琢, 祝国隽

B01GB006Y-02/Note14

---

1. 记 X_1, X_2, X_3, X_4 iid $\sim$ Poisson (λ). 考察假设检验问题

$$H_0 : \lambda = \frac{1}{2} \leftrightarrow H_1 : \lambda = 2.$$

取拒绝域 $W = \left\{(X_1, X_2, X_3, X_4) \mid \sum_{i=1}^4 X_i \geqslant 3\right\}$. 求犯第 II 类错误概率和功效函数.

Ans: 解: 因为 X_i 服从 Poisson 分布, 由 Poisson 分布的可加性知 $\sum_{i=1}^n X_i \sim$ Poiss$(n\lambda)$. 所以功效函数为

$$\begin{aligned} g(\lambda) &= P_\lambda\left\{\sum_{i=1}^n X_i \geq 3\right\} = \sum_{k=3}^{+\infty} \frac{(n\lambda)^k}{k!} \mathrm{e}^{-n\lambda} \\ &= 1 - \sum_{k=0}^{2} \frac{(n\lambda)^k}{k!} \mathrm{e}^{-n\lambda} = 1 - \left(1 + n\lambda + \frac{n^2\lambda^2}{2}\right) \mathrm{e}^{-n\lambda}, \quad \lambda > 0 \end{aligned}$$

当 $\lambda = \frac{1}{2}$ 时, $g\left(\frac{1}{2}\right)$ 为犯第一类错误的概率, 即

$$g\left(\frac{1}{2}\right) = 1 - \left(1 + \frac{n}{2} + \frac{n^2}{8}\right) \mathrm{e}^{-\frac{n}{2}}$$

当 $\lambda = 2, g(2)$ 为功效, 即 $g(2) = 1 - (1 + 2n + 2n^2)\,\mathrm{e}^{-2n}$.

犯第二类错误的概率为 $1 - g(2) = (1 + 2n + 2n^2)\,\mathrm{e}^{-2n}$.

由 Poisson 分布的再生性, 则 $\sum_{i=1}^4 X_i \sim \text{Poisson}(4\lambda)$.

犯第 II 类错误概率为:

$$\eta=P_{H_1}\left(\sum_{i=1}^{4}X_i<3\right)=\sum_{k=0}^{2}\frac{(4\lambda)^k}{k!}e^{-4\lambda}\Bigg|_{\lambda=2}=\sum_{k=0}^{2}\frac{8^k}{k!}e^{-8}\approx 0.01375$$

功效函数为:

$$\beta_\phi(\lambda)=P\left(\sum_{i=1}^{4}X_i\geq 3\right)=1-P\left(\sum_{i=1}^{4}X_i<3\right)=1-\sum_{k=0}^{2}\frac{(4\lambda)^k}{k!}e^{-4\lambda}$$

2. 记 $X_1,X_2,\cdots,X_n$ iid $\sim N\left(\mu,\sigma_0^2\right)$, σ_0^2 已知. 设 $\mu_0,a_0>0$ 为已知常数, 取显著性水平为 $\alpha,0<\alpha<1$. 考察假设检验问题

$$H_0:\mu=\mu_0\longleftrightarrow H_1:\mu\geqslant\mu_0+a_0\sigma_0$$

(1) 构造检验统计量，并给出拒绝域使得犯第 I 类错误概率不超过 α;

(2) 求犯第 II 类错误的概率;

(3) 如果犯第 II 类错误的概率不超过 $1-\beta$, 问样本量 n 至少需取多大?

Ans: 令 $\bar{X}=\dfrac{1}{n}\sum_{i=1}^{n}X_i$. 取检验统计量 $T=\dfrac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{\sigma_0}$. 拒绝域为 $W=\{T\geq u_{1-a}\}$. 当 H_1 成立时, $\dfrac{\sqrt{n}(\bar{X}-\mu)}{\sigma_0}\sim N(0,1)$. 因而犯第二类错误的概率为

$$\begin{aligned}P_\mu\left(T<u_{1-a}\right)&=P_\mu\left(\frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{\sigma_0}<u_{1-a}\right)\\&=P_\mu\left(\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma_0}<u_{1-a}-\frac{\sqrt{n}\left(\mu-\mu_0\right)}{\sigma_0}\right)\\&=\Phi\left(u_{1-a}-\frac{\sqrt{n}\left(\mu-\mu_0\right)}{\sigma_0}\right).\end{aligned}$$

由不等式 $\mu\geq\mu_0+a_0\sigma_0$ 可得 $\dfrac{\mu-\mu_0}{\sigma_0}\geq a_0$, 进而

$$P_\mu\left(T<u_{1-a}\right)=\Phi\left(u_{1-\alpha}-\frac{\sqrt{n}\left(\mu-\mu_0\right)}{\sigma_0}\right)\leq\Phi\left(u_{1-a}-\sqrt{n}a_0\right).$$

记 $\Phi\left(u_{1-a}-\sqrt{n}a_0\right)\leq\beta$, 则 $u_{1-\alpha}-\sqrt{n}a_0\leq u_\beta$. 求解得

$$n\geq\frac{\left(u_{1-\alpha}-u_\beta\right)^2}{a_0^2}$$

故当 $\frac{\left(u_{1-\alpha}-u_\beta\right)^2}{a_0^2}$ 为整数时, 可取 $n=\frac{\left(u_{1-\alpha}-u_\beta\right)^2}{a_0^2}$; 当 $\frac{\left(u_{1-\alpha}-u_\beta\right)^2}{a_0^2}$ 不是整数时, 可取 $n=\left\lfloor\frac{\left(u_{1-\alpha}-u_\beta\right)^2}{a_0^2}\right\rfloor+1$, 其中 $\lfloor\cdot\rfloor$ 表示向下取整, 可使得当 $\mu\geq\mu_0+a_0\sigma_0$ 时, z 检验犯第二类错误的概率 $P_\mu\left(z<z_{1-a}\right)$ 不超过 β.

3. 记 $X_1,X_2,\cdots,X_n$ iid $\sim N\left(\mu_1,\sigma^2\right),Y_1,Y_2,\cdots,Y_m$ iid $\sim N\left(\mu_2,k\sigma^2\right),k>0$ 已知, σ^2 未知, 两组样本相互独立. 设 $a\neq0,b\neq0,c\in\mathcal{R}$ 是三个已知常数. 试给出下面两个假设检验问题

$$H_0:a\mu_1+b\mu_2=c\leftrightarrow H_1:a\mu_1+b\mu_2\neq c,$$
$$H_0:a\mu_1+b\mu_2\leqslant c\leftrightarrow H_1:a\mu_1+b\mu_2>c$$

的显著性水平为 α 的检检, $0<\alpha<1$.

Ans: 记

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\bar{Y}=\frac{1}{m}\sum_{j=1}^{m}Y_j$$

$$S_0^2=\frac{1}{m+n-2}\left[\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2+\frac{1}{k}\sum_{j=1}^{m}\left(Y_j-\bar{Y}\right)^2\right].$$

$$a\bar{X}+b\bar{Y}\sim N\left(a\mu_1+b\mu_{2,}\left(\frac{a^2}{n}+\frac{kb^2}{m}\right)\sigma^2\right),\quad\frac{(m+n-2)S_0^2}{\sigma^2}\sim\chi_{m+n-2}^2,$$

知

$$\frac{a\bar{X}+b\bar{Y}-\left(a\mu_1+b\mu_2\right)}{S_0\sqrt{\frac{a^2}{n}+\frac{kb^2}{m}}}\sim t_{m+n-2}$$

检验统计量为

$$T=\frac{a\bar{X}+b\bar{Y}-c}{S_0\sqrt{\frac{a^2}{n}+\frac{kb^2}{m}}}.$$

当 $a\mu_1+b\mu_2=c$ 时,

$$T\sim t_{m+n-2}.$$

对于第一小问, 拒绝域为 $|T|>t_{n+m-2}(1-\alpha/2)$, 对于第二问, 拒绝域为 $T>t_{n+m-2}(1-\alpha)$

4. 记 $X_1,X_2,\cdots,X_n$ iid $\sim \mathrm{Exp}(\lambda_1)$, $Y_1,Y_2,\cdots,Y_m$ iid $\sim \mathrm{Exp}(\lambda_2)$, 两组样本相互独立. 设 k 为已知的正数. 试给出下面两个假设检验问题

$$H_0:\lambda_1=k\lambda_2\leftrightarrow H_1:\lambda_1\neq k\lambda_2$$
$$H_0:\lambda_1\leqslant k\lambda_2\leftrightarrow H_1:\lambda_1>k\lambda_2$$

的显著性水平为 α 的检验, $0<\alpha<1$.

Ans: 由 χ^2 分布的性质 6 , 有

$$2n\lambda_1\bar{X}\sim\chi^2_{2n},2m\lambda_2\bar{Y}\sim\chi^2_{2m},\text{其中}\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\bar{Y}=\frac{1}{m}\sum_{j=1}^{m}Y_j$$

则

$$\frac{\frac{2n\lambda_1\bar{X}}{2n}}{\frac{2m\lambda_2\bar{Y}}{2m}}=\frac{\lambda_1}{k\lambda_2}\cdot\frac{k\bar{X}}{\bar{Y}}\sim F^2_{2n,2m}.$$

检验统计量为

$$T=\frac{k\bar{X}}{\bar{Y}}.$$

当 $\lambda_1>k\lambda_2$ 时, 随着两者差距越大, T 值就越小. 当 $\lambda_1<k\lambda_2$ 时, 随着两者差距越大, T 值就越大.

(1) 针对 $H_0:\lambda_1=k\lambda_2\leftrightarrow H_1:\lambda_1\neq k\lambda_2$.

T 值过大或者过小, 都倾向于拒绝 H_0, 故显著性水平为 α 的拒绝域为

$$W_1=\left\{T>F_{2n,2m}\left(1-\frac{\alpha}{2}\right),T<F_{2n,2m}\left(\frac{\alpha}{2}\right)\right\}.$$

(2) 针对 $H_0:\lambda_1\leq k\lambda_2\leftrightarrow H_1:\lambda_1>k\lambda_2$.

T 值越小, 倾向于拒绝 H_0, 故显著性水平为 α 的拒绝域为

$$W_2 = \{T < F_{2n,2m}(\alpha)\}$$

最后再说明它们均可以控制犯第 I 类错误概率.

5. 记 $X \sim B(n,p), p_0 \in (0,1)$ 已知. 试给出假设检验问题

$$H_0 : p = p_0 \leftrightarrow H_1 : p \neq p_0$$

的一个显著性水平为 α 的大样本检验, $0 < \alpha < 1$.

Ans: $X \sim B(n,p), EX = np, \operatorname{var}(X) = np(1-p)$.

$$\frac{X - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1), n \to \infty.$$

设检验统计量为

$$T = \frac{X - np_0}{\sqrt{np_0\,(1-p_0)}}.$$

显著性水平为 α 的拒绝域为 $\left\{|T| > U_{1-\frac{\alpha}{2}}\right\}$, 其中 $U_{1-\frac{\alpha}{2}}$ 是 $N(0,1)$ 的 $1 - \dfrac{\alpha}{2}$ 分位数.

补充材料参考张伟平老师 (USTC) 的讲义, 正态总体的假设检验比较经典, 大家应该也没什么疑惑了, 补充一下非正态总体的材料.

Lec10: 假设检验(二)

张伟平

2011 年 4 月 18 日

1 非正态总体下参数的检验

1.1 指数分布参数的检验

设$X_1, \cdots, X_n$为从期望是$1/\lambda$的指数分布总体中抽取的简单样本，考虑如下三种形式的假设检验问题：

1. $H_0: \lambda \geq \lambda_0 \leftrightarrow H_1: \lambda < \lambda_0$
2. $H_0': \lambda \leq \lambda_0 \leftrightarrow H_1': \lambda > \lambda_0$
3. $H_0'': \lambda = \lambda_0 \leftrightarrow H_1'': \lambda \neq \lambda_0$

注意到$\sum_{i=1}^n X_i$是参数λ的充分统计量，及参数$1/\lambda$的无偏估计为$\bar{X}$而且$2\lambda\sum_{i=1}^n X_i \sim \chi^2_{2n}$。因此，对检验假设问题1，一个合理的检验为

$$\phi: \text{当}\sum_{i=1}^n X_i > c\text{时，拒绝}H_0,\text{不然就接受}$$

其功效函数为

$$\beta_\phi(\lambda) = P(\sum_{i=1}^n X_i > c) = P(2\lambda\sum_{i=1}^n X_i > 2\lambda c)$$

为λ的减函数。故欲使上式小于等于α，只需$\beta_\phi(\lambda_0) = \alpha$，从而$c = \frac{1}{2\lambda_0}\chi^2_\alpha(2n)$。因而所求的检验为

$$\phi: \text{当}\sum_{i=1}^n X_i > \frac{1}{2\lambda_0}\chi^2_\alpha(2n)\text{时，拒绝}H_0,\text{不然就接受}$$

类似可以得到2，3的检验为

$$\phi': \text{当}\sum_{i=1}^n X_i < \frac{1}{2\lambda_0}\chi^2_{1-\alpha}(2n)\text{时，拒绝}H_0',\text{不然就接受}$$

和

$$\phi: \quad \text{当}\sum_{i=1}^n X_i < \frac{1}{2\lambda_0}\chi^2_{1-\alpha/2}(2n)\text{ 或}$$

$$\sum_{i=1}^{n} X_i > \frac{1}{2\lambda_0}\chi^2_{\alpha/2}(2n)\text{时，拒绝}H_0'',\text{不然就接受}$$

在指数总体中，我们更感兴趣的是如下两种类型的截尾：

(a) 定数截尾

以实例说明。假设某种电子元件的寿命服从指数分布，抽取n个元件测其寿命。试验前定下一个自然数$r < n$，试验进行到有r个元件失效时为止，记这r个元件失效的时刻为

$$0 \le t_1 \le t_2 \le \cdots \le t_r$$

其中t_i表示第i个失效元件的失效时刻。也就是说，我们得到的样本观察值是$0 \le t_1 \le t_2 \le \cdots \le t_r$. 此时，似然函数为

$$L(\lambda) = [\prod_{i=1}^{r} \lambda e^{-\lambda t_i}]e^{-(n-r)\lambda t_r}$$

因此得到$1/\lambda$的似然估计为$T/r = \frac{1}{r}[\sum\limits_{i=1}^{r} t_i + (n-r)t_r]$。

用样本$X_1, \cdots, X_n$表示抽取的样本，则上述叙述表明我们得到如下信息: "$X_{(1)}, \cdots, X_{(r)}$的观察值$t_1, \cdots, t_r$"。或者等价的说，有r个元件的寿命是$t_1, \cdots, t_r$，还有$n-r$个元件的寿命大于t_r。因此在这种情况下，对假设检验问题1，一个合理的检验是

$$\phi: \text{当}\sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)} > c\text{时，拒绝}H_0,\text{不然就接受}$$

为确定常数c，需要知道统计量$T = \sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)}$的分布。作变换

$$Y_1 = nX_{(1)}, Y_i = (n-i+1)(X_{(i)} - X_{(i-1)}), i = 2, \cdots, n.$$

则由$X_{(1)}, \cdots, X_{(n)}$的联合pdf

$$f(x;\lambda) = n!\lambda^n e^{-\lambda\sum_{i=1}^{n} x_{(i)}} I(0 < x_{(1)} \le \cdots \le x_{(n)})$$

容易得到$Y_1, \cdots, Y_n$的联合pdf

$$g(y;\lambda) = \lambda^n e^{-\lambda\sum_{i=1}^{n} y_i} I(y_i > 0, i = 1, \cdots, n)$$

即$Y_1, \cdots, Y_n$相互独立且服从都服从指数分布。而$T = \sum_{i=1}^{r} Y_i$，因此有$2\lambda T \sim \chi^2_{2r}$。从而类似于前面的处理方法可以得到$c = \chi^2_\alpha(2r)/2\lambda_0$，故检验$\phi$为

$$\phi: \text{当}\sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)} > \frac{1}{2\lambda_0}\chi^2_\alpha(2r)\text{时，拒绝}H_0,\text{不然就接受}$$

类似可以得到2，3的检验为

$$\phi': \text{当}\sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)} < \frac{1}{2\lambda_0}\chi^2_{1-\alpha}(2r)\text{时，拒绝}H_0',\text{不然就接受}$$

和

$$\phi'': \qquad \text{当}\sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)} < \frac{1}{2\lambda_0}\chi^2_{1-\alpha/2}(2r)\text{或}$$

$$\sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)} > \frac{1}{2\lambda_0}\chi^2_{\alpha/2}(2r)$$
时，拒绝H_0'',不然就接受

(b) 定时截尾

与定数截尾相对的就是定时截尾，即在试验前，事先确定一个时间T_0，当实验进行到t_0时刻就停止整个试验。把到这时为止全部n 个元件的寿命加起来记为T^*，算法为：当某个元件在时刻T_0之前的某个时刻t失效，则该元件的寿命就是t，若到了T_0时刻该元件还没有失效，则该元件的寿命就记为T_0。显然，平均寿命越大，则T^*越倾向于取较大的值。于是对假设检验问题1，一个合理的检验可以取为

$$\phi: \text{当}T^* > c\text{时，拒绝}H_0,\text{不然就接受}$$

可以证明，近似地有$2\lambda T^* \sim \chi^2_{2u+1}$，这里$u$是到时刻$T_0$为止时失效的个数。因此假设1-3的检验为

$$\phi: \text{当}T^* > \frac{1}{2\lambda_0}\chi^2_{\alpha}(2u+1)\text{时，拒绝}H_0,\text{不然就接受}$$

类似可以得到2，3的检验为

$$\phi': \text{当}T^* < \frac{1}{2\lambda_0}\chi^2_{1-\alpha}(2u+1)\text{时，拒绝}H_0',\text{不然就接受}$$

和

$$\phi'': \quad \text{当}T^* < \frac{1}{2\lambda_0}\chi^2_{1-\alpha/2}(2u+1)\text{ 或 }T^* > \frac{1}{2\lambda_0}\chi^2_{\alpha/2}(2u+1)$$
时，拒绝H_0'',不然就接受

1.2 二项分布参数p的检验

设某个事件在一次试验中发生的概率为p，p未知。作n次独立的试验，每次观察该事件是否发生。以X记该事件发生的总次数，则$X \sim B(n,p)$，根据X去检验如下的假设：

1. $H_0: p \le p_0 \leftrightarrow H_1: p > p_0$
2. $H_0': p \ge p_0 \leftrightarrow H_1': p < p_0$
3. $H_0'': p = p_0 \leftrightarrow H_1'': p \ne p_0$

显然当p愈小，则X愈倾向于取较小的整数。因此，对假设1，一个合理的检验为

$$\phi: \text{当}X > c\text{时，拒绝}H_0,\text{不然就接受}$$

其功效函数为

$$\beta_\phi(p) = P(X > c) = 1 - P(X \le c) = 1 - \sum_{i=0}^{c}\binom{n}{i}p^i q^{n-i}$$

注意到$P(X \le k) = \frac{n!}{k!(n-k-1)!}\int_0^{1-p} t^k(1-t)^{n-k-1}dt$，即$\beta_\phi(p)$为$p$的增函数。因此欲使上式小于等于$\alpha$，只需$\beta_\phi(p_0)=\alpha$。即

$$\sum_{i=0}^{c}\binom{n}{i}p_0^i q_0^{n-i} = 1-\alpha$$

此方程往往没有整数解，较常见的是存在c_0，使得

$$\sum_{i=0}^{c_0}\binom{n}{i}p_0^i q_0^{n-i} < 1-\alpha < \sum_{i=0}^{c_0+1}\binom{n}{i}p_0^i q_0^{n-i}$$

这时，一个经常采用的检验是

$\phi:$ 当$X \le c_0$时，接受H_0;

当$X > c_0+1$时，拒绝H_0;

当$X = c_0+1$时，需要协商

(即再作随机试验或者按照某种都同意的准则)。

类似的对假设2和3，可以得到一个检验为

$$\phi': \text{当}X \ge c\text{时，接受}H_0',\text{不然就拒绝}$$

其中c由

$$\sum_{i=0}^{c-1}\binom{n}{i}p_0^i q_0^{n-i} = \alpha$$

确定。

$$\phi'': \text{当}c_1 \le X \le c_2\text{时，接受}H_0'',\text{不然就拒绝}$$

其中c_1, c_2由

$$\sum_{i=0}^{c_1-1}\binom{n}{i}p_0^i q_0^{n-i} + \sum_{i=c_2+1}^{n}\binom{n}{i}p_0^i q_0^{n-i} = \alpha$$

确定。常常令

$$\sum_{i=0}^{c_1-1}\binom{n}{i}p_0^i q_0^{n-i} = \alpha/2$$
$$\sum_{i=c_2+1}^{n}\binom{n}{i}p_0^i q_0^{n-i} = \alpha/2$$

以定出c_1, c_2。

1.3 Poisson总体参数的检验

对Poisson总体参数的检验，完全类似于二项分布总体参数的检验。考虑如下三种形式的假设

1. $H_0: \lambda \le \lambda_0 \leftrightarrow H_1: \lambda > \lambda_0$

2. $H_0': \lambda \geq \lambda_0 \leftrightarrow H_1': \lambda < \lambda_0$

3. $H_0'': \lambda = \lambda_0 \leftrightarrow H_1'': \lambda \neq \lambda_0$

设X为从Poisson总体$P(\lambda)$中抽取的样本，我们要根据X来检验上述三个假设。由于X为λ的无偏估计，从而对假设1，一个合理的检验为

$$\phi: \text{当}X > c\text{时，拒绝}H_0, \text{不然就接受}$$

其功效函数为

$$\beta_\phi(\lambda) = P(X > c) = 1 - \sum_{i=1}^{c} \frac{\lambda^i}{i!} e^{-\lambda}$$

注意到$P(X \leq k) = \int_\lambda^\infty \frac{t^k}{k!} e^{-t} dt$，所以$\beta_\phi(\lambda)$为$\lambda$的增函数。从而欲使$\beta_\phi(\lambda) \leq \alpha$对任意的$\lambda \leq \lambda_0$成立，只需$\beta_\phi(\lambda_0) = \alpha$,即

$$\sum_{i=1}^{c} \frac{\lambda_0^i}{i!} e^{-\lambda_0} = 1 - \alpha$$

若方程有整数解，则可以定出c，检验ϕ也就完全确定了。但是此方程也是往往没有整数解，常见的是存在整数c_0，满足

$$\sum_{i=1}^{c_0} \frac{\lambda_0^i}{i!} e^{-\lambda_0} < 1 - \alpha < \sum_{i=1}^{c_0+1} \frac{\lambda_0^i}{i!} e^{-\lambda_0}$$

这时，一个经常采用的检验是

$$\begin{aligned} \phi: \quad & \text{当}X \leq c_0\text{时，接受}H_0; \\ & \text{当}X > c_0 + 1\text{时，拒绝}H_0; \\ & \text{当}X = c_0 + 1\text{时，需要协商} \\ & \text{(即再作随机试验或者按照某种都同意的准则)。} \end{aligned}$$

类似的对假设2和3，可以得到一个检验为

$$\phi': \text{当}X < c\text{时，拒绝}H_0', \text{不然就接受}$$

其中c由

$$\sum_{i=0}^{c-1} \frac{\lambda_0^i}{i!} e^{-\lambda_0} = \alpha$$

确定。

$$\phi'': \text{当}X < c_1 \text{ 或 } X > c_2\text{时，拒绝}H_0'', \text{不然就接受}$$

其中c_1, c_2由

$$\sum_{i=0}^{c_1-1} \frac{\lambda_0^i}{i!} e^{-\lambda_0} + \sum_{i=c_2+1}^{n} \frac{\lambda_0^i}{i!} e^{-\lambda_0} = \alpha$$

确定。常常令

$$\sum_{i=0}^{c_1-1} \frac{\lambda_0^i}{i!} e^{-\lambda_0} = \alpha/2$$

$$\sum_{i=c_2+1}^{n}\frac{\lambda_0^i}{i!}e^{-\lambda_0}=\alpha/2$$

以定出c_1, c_2。

1.4 极限分布为正态分布的检验*

当样本容量n计较大，这个时候我们可以根据中心极限定理，对参数进行检验。这种方法称为“大样本检验方法”。我们举几个例子说明。

1. Behrens-Fisher 问题

例1. **[Behrens-Fisher Problem]** 设$X_1, \cdots, X_n$和$Y_1, \cdots, Y_m$为分别来自正态总体$N(\theta_1, \sigma_1^2)$和$N(\theta_2, \sigma_2^2)$，且两组样本独立。$\theta_1, \theta_2, \sigma_1^2, \sigma_2^2$都未知。要检验假设$H_0: \theta_1=\theta_2 \leftrightarrow H_1: \theta_1 \neq \theta_2$。

解：由于

$$\frac{\bar{X}-\bar{Y}-(\theta_1-\theta_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}} \sim N(0,1)$$

中含有未知参数σ_1^2, σ_2^2，故不能从上式确定临界值。于是以S_X^2来估计σ_1^2和S_Y^2来估计σ_2^2，根据大数律，当n, m较大时，有

$$\frac{\bar{X}-\bar{Y}-(\theta_1-\theta_2)}{\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}} \sim AN(0,1)$$

于是，我们得到假设H_0的拒绝域为

$$|\bar{X}-\bar{Y}|\Big/\sqrt{S_X^2/n+S_Y^2/m} > u_{\alpha/2}.$$

例2. 考虑二项分布参数p的检验: $H_0: p=p_0$，当n很大时，c_1, c_2无法从二项分布表上查到。但根据中心极限定理，当原假设$p=p_0$成立且n足够大时，有$(X-np_0)/\sqrt{np_0q_0} \sim AN(0,1)$，因此可以提出如下的检验：当

$$|X-np_0|/\sqrt{np_0q_0} > u_{\alpha/2}$$

时拒绝H_0，不然就接受。这等于用上述不等式的两端的值作为c_1, c_2的值。这两个值比c_1, c_2的确切值要容易计算的多。

大样本检验方法是不得已而用的方法。

从本质上讲，这里用的大样本方法仍然是需要从直观上给出检验形式的。直接从数学推导上得到检验的方法还有Bayes方法和似然比检验方法等，此处我们简单介绍一下似然检验比方法。

2. 二项分布和Poisson分布参数的大样本检验

设$X_1, \ldots, X_n$ i.i.d. $\sim b(1,p)$, 显见$T=\sum_{i=1}^{n} X_i \sim b(n,p)$, 考虑下列检验问题:

$$H_0: p=p_0 \longleftrightarrow H_1: p \neq p_0 \tag{1.1}$$

此处p_0和检验水平α给定.

由独立同分布场合的中心极限定理可知: $(T-np)/\sqrt{np(1-p)} \xrightarrow{\mathscr{L}} N(0,1)$, 当$n \longrightarrow \infty$时, 故当$H_0$成立, 即$p=p_0$时有

$$U=\frac{T-np_0}{\sqrt{np_0(1-p_0)}} \xrightarrow{\mathscr{L}} N(0,1), \text{ 当}n \longrightarrow \infty\text{时}$$

因此取U作为检验统计量. 当n较大时, U可以近似认为服从$N(0,1)$分布. 由U检验法可知检验问题(1.1)水平近似为α的否定域为

$$D=\left\{(X_1,\ldots,X_n):\ |T-np_0|/\sqrt{np_0(1-p_0)}>u_{\alpha/2}\right\}$$

类似可求两个单边检验问题

$$H_0':\ p\le p_0 \longleftrightarrow H_1':\ p>p_0$$
$$H_0'':\ p\ge p_0 \longleftrightarrow H_1'':\ p<p_0$$

的大样本检验.

再考虑Poisson分布的大样本检验问题. 设$X_1,\ldots,X_n$为自Poisson总体$p(\lambda)$中抽取的随机样本, 考虑检验问题

$$H_0:\lambda=\lambda_0 \longleftrightarrow H_1:\lambda\ne\lambda_0 \tag{1.2}$$

此处λ_0和检验水平α给定.

由于$T=\sum_{i=1}^n X_i$服从参数为$n\lambda$的Poisson分布$P(n\lambda)$.由中心极限定理可知: $(T-n\lambda)/\sqrt{n\lambda} \xrightarrow{\mathscr{L}} N_(0,1)$,当$n \longrightarrow \infty$时. 因此当$H_0$成立, 即$\lambda=\lambda_0$时有

$$U_0=\frac{T-n\lambda_0}{\sqrt{n\lambda_0}} \xrightarrow{\mathscr{L}} N(0,1), \text{ 当}n \longrightarrow \infty\text{时.}$$

因此可取U_0作为检验统计量. 当n较大时, U_0可以近似认为服从$N(0,1)$分布. 由U检验法可知双边检验问题(1.2)的水平近似为α的否定域为

$$D=\{(X_1,\ldots,X_n):\ |T-n\lambda_0|\,/\sqrt{n\lambda_0}>u_{\alpha/2}\}$$

类似方法可求关于λ的两个单边检验问题

$$H_0':\ \lambda\le\lambda_0 \longleftrightarrow H_1':\ \lambda>\lambda_0$$
$$H_0'':\ \lambda\ge\lambda_0 \longleftrightarrow H_1'':\ \lambda<\lambda_0$$

的大样本检验.

下面考虑两样本检验问题. 设$X_1,\ldots,X_m$ i.i.d. $\sim b(1,p_1)$, $Y_1,\ldots,Y_n$ i.i.d. $\sim b(1,p_2)$,且合样本$X_1,\ldots,X_m,Y_1,\ldots,Y_n$相互独立. 求下列检验问题:

$$H_0:p_2-p_1=0 \longleftrightarrow H_1:p_2-p_1\ne 0 \tag{1.3}$$

检验水平α给定.

记$\bar{X}$和$\bar{Y}$分别为两组样本的均值. 由中心极限定理可知

$$\frac{\bar{X}-\bar{Y}-(p_2-p_1)}{\sqrt{p_1(1-p_1)/m+p_2(1-p_2)/n}} \xrightarrow{\mathscr{L}} N(0,1), \text{ 当 } n,\ m \longrightarrow \infty \text{ 时.}$$

当H_0成立时, 即$p_1 = p_2 = p$时, 将p用合样本估计, 即取

$$\hat{p} = \frac{1}{m+n}\left(\sum_{i=1}^{m} X_i + \sum_{j=1}^{n} Y_j\right)$$

则有

$$U^* = \frac{\bar{X}-\bar{Y}}{\sqrt{\hat{p}(1-\hat{p})}}\sqrt{\frac{mn}{m+n}} \xrightarrow{\mathscr{L}} N(0,1), \text{ 当 } m,\ n \longrightarrow \infty \text{ 时}.$$

因此取U^*为检验统计量, 当m, n都较大时, U可认为近似服从$N(0,1)$分布. 由U检验法得到双边检验问题(1.3)的检验水平近似为α的否定域为

$$D = \{(X_1,\ldots,X_m,Y_1,\ldots,Y_n):\ |\,U^*| > u_{\alpha/2}\}.$$

还可以用类似方法讨论下列两个单边检验问题

$$H_0':\ p_2 \le p_1 \longleftrightarrow H_1':\ p_2 > p_1$$
$$H_0'':\ p_2 \ge p_1 \longleftrightarrow H_1'':\ p_2 < p_1$$

的大样本检验.

Poisson分布的两样本检验问题可用类似方法讨论, 检验统计量的选取和检验否定域的形式留给读者作为练习.

2 假设检验与区间估计

假设检验与区间估计这两个统计推断的形式表面上看好像完全不同, 而实际上两者之间有着非常密切的关系. 由单参数假设检验问题的水平为α 的检验, 往往可以得到该参数的置信系数为$1-\alpha$的置信区间,反之亦然. 具体说明如下.

一、如何由假设检验得到置信区间

设要求θ 的置信系数为$1-\alpha$ 的置信区间. 考虑双边检验问题

$$H_0:\ \theta = \theta_0 \longleftrightarrow H_1:\ \theta \ne \theta_0\,.$$

求出水平为α 的否定域D,故接受域为$\bar{D}$, 则必有

$$P(\bar{D}|H_0) = 1-\alpha, \tag{2.1}$$

由$\bar{D}$确定的不等式得到如下不等式: $\hat{\theta_1}(\mathbf{X}) \leqslant \theta_0 \leqslant \hat{\theta_2}(\mathbf{X})$, 由于(2.1)是在条件“$H_0: \theta = \theta_0$”下成立,改$\theta_0$ 为θ 得$\hat{\theta_1}(\mathbf{X}) \leqslant \theta \leqslant \hat{\theta_2}(\mathbf{X})$, 则$[\hat{\theta_1}(\mathbf{X}), \hat{\theta_2}(\mathbf{X})]$即为所求的置信系数为$1-\alpha$ 的置信区间.

若要求θ 的置信上、下限, 就需要考虑单边检验 $H_0: \theta \leqslant \theta_0 \longleftrightarrow H_1\theta > \theta_0$ 或$H_0: \theta \geqslant \theta_0 \longleftrightarrow H_1: \theta < \theta_0$ 的检验问题. 下面通过例子来说明.

例5.3.1 设$X_1,\ldots,X_n$ 为自正态总体$N(\mu,\sigma^2)$ 抽取的随机样本. μ, σ^2皆未知, 要分别求μ 和σ^2 的置信系数为$1-\alpha$ 的置信区间和置信上、下限.

先考虑μ 的置信区间和置信上、下限问题. 在§5.2 中已给出假设

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

的水平为α的检验的否定域

$$D = \{(X_1, \ldots, X_n): |T| \geqslant t_{n-1}(\alpha/2)\},$$

其中$T = \sqrt{n}(\bar{X} - \mu_0)/S$.记$\theta = (\mu, \sigma^2)$, 故有$P_\theta(|T| > t_{n-1}(\alpha/2) \,|\, H_0) = \alpha$. 等价地, 对接受域 $\bar{D}$ 有

$$P_\theta(|\sqrt{n}(\bar{X} - \mu_0)/S| \leqslant t_{n-1}(\alpha/2) \mid H_0) = 1 - \alpha. \tag{2.2}$$

由于上述等式是在条件H_0成立, 即$\mu = \mu_0$时获得的, 因此我们将下面出现的所有μ_0用μ代替是等价的. 解(2.2)括号中的不等式得H_0成立的条件下有$\mu = \mu_0$, μ

$$\bar{X} - \frac{S}{\sqrt{n}} t_{n-1}(\alpha/2) \leqslant \mu \leqslant \bar{X} + \frac{S}{\sqrt{n}} t_{n-1}(\alpha/2)$$

因此

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

即为μ的置信系数为$1 - \alpha$ 的置信区间.

若要求μ 的置信下限, 则考虑检验问题

$$H_0: \mu \leqslant \mu_0 \longleftrightarrow H_1: \mu > \mu_0.$$

在§5.2中已给出水平为α 的否定域$D = \{(X_1, \ldots, X_n):\ T > t_{n-1}(\alpha)\}$, 其接受域

$$\bar{D} = \{(X_1, \ldots, X_n):\ T \leqslant t_{n-1}(\alpha)\}.$$

因此有

$$P_\theta\big(\sqrt{n}(\bar{X} - \mu_0)/S \leqslant t_{n-1}(\alpha) \,|\, H_0\big) = 1 - \alpha$$

解括号中的不等式得

$$\bar{X} - \frac{S}{\sqrt{n}} t_{n-1}(\alpha) \leqslant \mu_0,$$

再改μ_0 为μ $\bar{X} - \frac{S}{\sqrt{n}} t_{n-1}(\alpha) \leq \mu < \infty$ 因此μ 的置信系数为$1 - \alpha$ 的置信下限为$\bar{X} - \frac{S}{\sqrt{n}} t_{n-1}(\alpha)$. 同理可求$\mu$ 的置信系数为$1 - \alpha$ 的置信上限为$\bar{X} + \frac{S}{\sqrt{n}} t_{n-1}(\alpha)$.

关于正态总体方差σ^2 的置信区间和置信上、下限留给读者作为练习.

例5.3.2 设$X_1, \ldots, X_n$ 和$Y_1, \ldots, Y_n$ 分别自正态总体$N(\mu_1, \sigma^2)$ 和$N(\mu_2, \sigma^2)$ 中抽取的简单随机样本. 且合样本$X_1, \ldots, X_n, Y_1, \ldots, Y_n$ 相互独立. 令$\mu = \mu_2 - \mu_1$, 求μ 的置信系数为$1 - \alpha$ 的置信区间和置信上、下限.

解 在§5.2中已找到了

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

的两样本t 检验的否定域:

$$D = \{(\mathbf{X}, \mathbf{Y}):\ |T| > t_{n+m-2}(\tfrac{\alpha}{2})\},$$

其中检验统计量为

$$T=\frac{\bar{Y}-\bar{X}-\mu_0}{S_w}\sqrt{\frac{mn}{m+n}},$$

此处$S_w^2=\frac{1}{n+m-2}[(m-1)S_1^2+(n-1)S_2^2]$, 而$S_1^2$和$S_2^2$分别为两组样本的样本方差. 若记$\theta=(\mu_1,\mu_2,\sigma^2)$,则有$P_\theta(|T|>t_{n+m-2}(\alpha/2)\,|\,H_0)=\alpha$. 类似于上例的讨论, 对接受域$\bar{D}$ 有

$$P_\theta\left(\left|\frac{\bar{Y}-\bar{X}-\mu_0}{S_w}\sqrt{\frac{mn}{m+n}}\right|\leqslant t_{m+n-2}(\alpha/2)\Big|H_0\right)=1-\alpha, \tag{2.3}$$

解(2.3)括号中的不等式得到

$$\bar{Y}-\bar{X}-S_w\,t_{n+m-2}\Big(\frac{\alpha}{2}\Big)\sqrt{\frac{1}{m}+\frac{1}{n}}\leqslant\mu\leqslant\bar{Y}-\bar{X}+S_w\,t_{n+m-2}\Big(\frac{\alpha}{2}\Big)\sqrt{\frac{1}{m}+\frac{1}{n}}$$

在H_0 成立的前提下, 可改上式中的μ_0 为μ, 因此$\mu=\mu_2-\mu_1$ 的置信系数为$1-\alpha$ 的置信区间为

$$\left[\bar{Y}-\bar{X}-S_w\,t_{n+m-2}\Big(\frac{\alpha}{2}\Big)\sqrt{\frac{1}{m}+\frac{1}{n}},\,\bar{Y}-\bar{X}+S_w\,t_{n+m-2}\Big(\frac{\alpha}{2}\Big)\sqrt{\frac{1}{m}+\frac{1}{n}}\right]$$

类似方法求得$\mu=\mu_2-\mu_1$ 的置信系数为$1-\alpha$ 的置信下、上限分别为$\bar{Y}-\bar{X}-S_w\,t_{n+m-2}(\alpha)\sqrt{\frac{1}{m}+\frac{1}{n}}$ 和$\bar{Y}-\bar{X}+S_w\,t_{n+m-2}(\alpha)\sqrt{\frac{1}{m}+\frac{1}{n}}$.

这里我们假定了两总体有相同的方差σ^2. 若去掉这一假设, 假定两总体的方差分别为σ_1^2和σ_2^2, 则就得到著名的*Behrens-Fisher*问题, 由§5.2中第五部分中给出的Behrens-Fisher 问题的大样本检验方法和一个小样本的近似方法, 用类似的方法也同样可以得到一个近似的*Behrens-Fisher*问题的区间估计形式(这已在§4.2中讨论过, 从略) .

两正态总体方差比的置信区间和置信上、下限如何通过假设检验方法的得到, 留给读者作练习.

二、如何由置信区间得到假设检验

若我们用某种方法建立了θ 的置信水平为$1-\alpha$ 的区间估计$[\widehat{\theta}_1,\widehat{\theta}_2]$, 对给定的$\theta_0$ 不难求出检验问题$H_0:\theta=\theta_0\longleftrightarrow H_1:\theta\neq\theta_0$ 的一个水平为α 的检验. 事实上, 一个简单方法就是若$\theta_0\in[\widehat{\theta}_1,\widehat{\theta}_2]$ 则接受H_0, 否则就拒绝H_0.

类似方法可由置信系数为$1-\alpha$ 的置信上、下限求出检验问题$H_0':\theta\geqslant\theta_0\longleftrightarrow H_1':\theta<\theta_0$ 和$H_0'':\theta\leqslant\theta_0\longleftrightarrow H_1'':\theta>\theta_0$ 的水平为α 的检验.

三、假设检验和区间估计的比较

与点估计和假设检验比较, 区间估计这一推断形式有一个显著的特点, 即它的精确度 (一般可用区间的长度刻画) 和可靠度(用其置信系数刻画) 一目了然. 点估计不具备这个特点, 才促使人们考虑区间估计. 而且区间估计可以在精确度、可靠度和样本大小n之间调整, 以达到预先指定的要求. 而假设检验提供的信息不如区间估计确切, 请看下例:

设从正态总体$N(\mu,\sigma^2)$ 中抽取一定大小的样本去检验假设$H_0:\mu=0\longleftrightarrow H_1:\mu\neq 0$. 结果假设被接受了. 如我们在§5.1中所述, 这并不意味着“证明”了$\mu=0$. 假如我们只知道$\mu=0$ 被

接受了, 我们甚至无法估量真正的μ 值与0 相差有多大. 但如果我们被告知μ 的具有95% 的置信系数的区间估计为$[-0.05, 0.07]$ 或者是$[-15, 20]$, 则在前一个场合, μ 与0相距最大不超过0.07,这么大小一个值在实用上可能无关紧要. 这时我们就有一定的把握(0.95)说μ "事实上"可以认为是0,而不止是接受"$\mu = 0$"了. 若在后一场合, 虽则$\mu = 0$ 这个假设也被接受(因为0这个点在区间$[-15, 20]$ 内) , 但因μ 的可能范围很大, 实际上我们只能说对μ "知之甚少".

反之, 若我们得到"$\mu = 0$ 被否定". 我们从这句话也只知道有比较显著的证据认为$\mu \neq 0$, 但还无法知道其实际意义如何. 但如果我们被告知: μ 的区间估计为$[0.01, 0.02]$ 或$[-40, -30]$. 在前一场合, 虽然$\mu = 0$ 被否定(因为0不在区间$[0.01, 0.02]$ 内) , 但μ 与0的最大差距不过0.02, 这么小一个值可能实际上与0无异. 因此, 虽然在统计上否定了$\mu = 0$, 但事实上可以认为$\mu = 0$. 在后一个场合μ 的值与0相距至少是30, 不仅要否定$\mu = 0$,从实际上看μ 也显著异于0.

这些分析说明, 区间估计所提供的信息比假设检验更为确切. 这也提醒我们: 1. 对假设检验结果的实际含义的解释要十分小心. 2.在得到假设检验结果时, 最好也将被检验参数的区间估计求出来作为参考.