§5.3 假设检验与区间估计

Example

设$\widetilde{X}=(X_1,\cdots,X_n)$为来自正态分布总体$N(\mu,\sigma^2)$的样本，其中$\sigma$ 已知. 考虑假设检验问题

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

取检验统计量为

$$U=U(\widetilde{X})=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}.$$

则显著性水平为α的检验的拒绝域为

$$D=\{\widetilde{x}:|u(\widetilde{x})|>u_{\alpha/2}\}=\{\widetilde{x}:|\overline{x}-\mu_0|>u_{\alpha/2}\frac{\sigma}{\sqrt{n}}\}.$$

接受域为

$$\overline{D} = \{\tilde{x} : \overline{x} - u_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu_0 \le \overline{x} + u_{\alpha/2}\frac{\sigma}{\sqrt{n}}\}.$$

而

$$[\overline{X} - u_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \quad \overline{X} + u_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$$

恰好是参数μ的置信水平为$1-\alpha$的最优置信区间.

Theorem

设$\widetilde{X}=(X_1,\cdots,X_n)$为来自总体$X\sim F(x;\theta)$的样本，其中$F(x;\theta)$为$X$的分布函数. 对每个$\theta_0\in\Theta$，记$A(\theta_0)$为假设检验问题$H_0:\theta=\theta_0\longleftrightarrow H_1:\theta\neq\theta_0$的显著性水平为$\alpha$的检验的接受域. 对每个$\widetilde{x}$，令

$$C(\widetilde{x})=\{\theta_0:\widetilde{x}\in A(\theta_0)\}.$$

则随机集$C(\widetilde{X})$是参数θ的置信水平为$1-\alpha$的置信集.

Theorem

反过来，设$C(\tilde{X})$是参数θ的置信水平为$1-\alpha$的置信集. 对任意的$\theta_0\in\Theta$，令

$$A(\theta_0)=\{\tilde{x}:\theta_0\in C(\tilde{x})\}.$$

则$A(\theta_0)$为假设检验问题$H_0:\theta=\theta_0\longleftrightarrow H_1:\theta\neq\theta_0$的显著性水平为$\alpha$的检验的接受域.

证: 第一部分结论. 因为$A(\theta_0)$为假设检验问题$H_0: \theta = \theta_0 \longleftrightarrow H_1: \theta \neq \theta_0$的显著性水平为$\alpha$的检验的接受域. 所以

$$\mathsf{P}_{\theta_0}\{\widetilde{X} \notin A(\theta_0)\} \le \alpha,$$

即

$$\mathsf{P}_{\theta_0}\{\widetilde{X} \in A(\theta_0)\} \ge 1 - \alpha.$$

由θ_0的任意性, 将θ_0改写为θ, 得

$$\mathsf{P}_{\theta}\{\theta \in C(\widetilde{X})\} = \mathsf{P}_{\theta}\{\widetilde{X} \in A(\theta)\} \ge 1 - \alpha.$$

故$C(\widetilde{X})$是参数θ的置信水平为$1-\alpha$的置信集.

第二部分结论. 因为

$$\mathsf{P}_{\theta_0}\{\widetilde{X} \notin A(\theta_0)\} = \mathsf{P}_{\theta_0}\{\theta_0 \notin C(\widetilde{X})\} \leq \alpha.$$

所以$A(\theta_0)$为假设检验问题$H_0 : \theta = \theta_0 \longleftrightarrow H_1 : \theta \neq \theta_0$的显著性水平为$\alpha$的检验的接受域.

Example

例 $X \sim U(0,\theta)$, $\widetilde{X}=(X_1,X_2,\cdots,X_n)$为来自该总体的样本. 在第四章的讨论中，已经得到了：对给定的α $(0<\alpha<1)$, θ的置信水平为$1-\alpha$的同等置信区间为$[X_{(n)}, X_{(n)}/\sqrt[n]{\alpha}]$.

那么对假设检验问题$H_0:\theta=\theta_0 \longleftrightarrow H_1:\theta\neq\theta_0$. 可得其显著性水平为$\alpha$的检验的拒绝域为:

$$
\begin{aligned}
D &= \{\widetilde{x}:\theta_0\notin[x_{(n)}, x_{(n)}/\sqrt[n]{\alpha}]\} \\
&= \{\widetilde{x}: x_{(n)}>\theta_0 \ \text{ 或 } \ x_{(n)}<\theta_0\sqrt[n]{\alpha}\}.
\end{aligned}
$$

§单参数指数型分布族的假设检验

一、单参数指数型分布族的性质

Theorem

定理 设总体X服从单参数指数型分布，$\widetilde{X}=(X_1,X_2,\cdots,X_n)$是取自该总体的样本，其联合密度函数(或分布列)为:

$$p(\widetilde{x};\theta)=C^*(\theta)\cdot\exp\{Q(\theta)\cdot U(\widetilde{x})\}\cdot h^*(\widetilde{x}),$$

其中，$Q(\theta)$是θ的严格增函数．若$\psi(U(\widetilde{x}))$是$U(\widetilde{x})$的一个非降函数，则$E_\theta\psi(U(\widetilde{X}))$也是$\theta$的一个非降函数．

Corollary

推论 设总体X服从单参数指数型分布，$\widetilde{X}=(X_1,X_2,\cdots,X_n)$是取自该总体的样本，其联合密度函数(或分布列)为：

$$p(\widetilde{x};\theta)=C^*(\theta)\cdot\exp\{Q(\theta)\cdot U(\widetilde{x})\}\cdot h^*(\widetilde{x}),$$

其中，$Q(\theta)$是θ的严格增函数．则对任意给定的常数C，$P_\theta\{U(\widetilde{X})>C\}$和$P_\theta\{U(\widetilde{X})<C\}$分别是$\theta$的一个非降和非增函数．

定理的证明: 不妨设$h^*(\widetilde{x}) > 0$, 即$p(\widetilde{x};\theta) > 0$.

设$\theta_1 < \theta_2$, 记

$$A = \{\widetilde{x} : p(\widetilde{x};\theta_1) < p(\widetilde{x};\theta_2)\} = \{\widetilde{x} : p(\widetilde{x};\theta_2)/p(\widetilde{x};\theta_1) > 1\}$$

$$B = \{\widetilde{x} : p(\widetilde{x};\theta_1) > p(\widetilde{x};\theta_2)\} = \{\widetilde{x} : p(\widetilde{x};\theta_2)/p(\widetilde{x};\theta_1) < 1\}$$

由于
$$\frac{p(\widetilde{x};\theta_2)}{p(\widetilde{x};\theta_1)}=\frac{C^*(\theta_2)}{C^*(\theta_1)}\exp\left\{[Q(\theta_2)-Q(\theta_1)]\cdot U(\widetilde{x})\right\},$$
和$Q(\theta_2)-Q(\theta_1)>0$, 所以$p(\widetilde{x};\theta_2)/p(\widetilde{x};\theta_1)$是$U(\widetilde{x})$ 的严格增函数. 所以

$$U(\widetilde{x})>U(\widetilde{y}),\quad \forall \widetilde{x}\in A,\widetilde{y}\in B,$$

若记
$$a=\inf_{\widetilde{x}\in A}\psi\big(U(\widetilde{x})\big),\ \ b=\sup_{\widetilde{y}\in B}\psi\big(U(\widetilde{y})\big).$$
注意到$\psi(U(\widetilde{x}))$是$U(\widetilde{x})$的一个非降函数, 因此有

$$a\geq b.$$

故

$$
\begin{aligned}
&\mathsf{E}_{\theta_2}\psi\big(U(\widetilde{X})\big)-\mathsf{E}_{\theta_1}\psi\big(U(\widetilde{X})\big)\\
=&\int \psi\big(U(\widetilde{x})\big)\cdot[p(\widetilde{x};\theta_2)-p(\widetilde{x};\theta_1)]\,d\widetilde{x}\\
=&\left(\int_A+\int_B\right)\psi\big(U(\widetilde{x})\big)\cdot[p(\widetilde{x};\theta_2)-p(\widetilde{x};\theta_1)]\,d\widetilde{x}\\
\geq& a\int_A[p(\widetilde{x};\theta_2)-p(\widetilde{x};\theta_1)]\,d\widetilde{x}+b\int_B[p(\widetilde{x};\theta_2)-p(\widetilde{x};\theta_1)]\,d\widetilde{x}\\
=&(a-b)\int_A[p(\widetilde{x};\theta_2)-p(\widetilde{x};\theta_1)]\,d\widetilde{x}+b\int[p(\widetilde{x};\theta_2)-p(\widetilde{x};\theta_1)]\,d\widetilde{x}\\
\geq& 0.
\end{aligned}
$$

定理得证.

二、单参数指数型分布族的假设检验问题

设总体X服从单参数指数型分布, 参数为θ. $\widetilde{X} = (X_1, X_2, \cdots, X_n)$是取自该总体的样本, 其联合密度函数(或分布列)为:

$$p(\widetilde{x}; \theta) = C^*(\theta) \cdot \exp\{Q(\theta) \cdot U(\widetilde{x})\} \cdot h^*(\widetilde{x}),$$

其中, $Q(\theta)$是θ的严格增函数.

θ_0为一已知常数, 检验的显著性水平为$\alpha(0 < \alpha < 1)$.

假设检验问题:

$$H_0 : \theta \geq \theta_0 \longleftrightarrow H_1 : \theta < \theta_0, \tag{1}$$

$$H_0 : \theta \leq \theta_0 \longleftrightarrow H_1 : \theta > \theta_0, \tag{2}$$

$$H_0 : \theta = \theta_0 \longleftrightarrow H_1 : \theta \neq \theta_0, \tag{3}$$

(1) 和(2)是单边假设检验问题, (3)是双边假设检验问题.

常取检验统计量为$U(\widetilde{X})$. 当$E_\theta U(\widetilde{X})$是$\theta$的单调非降函数时, 上述三个假设检验问题的拒绝域形式分别为

$$\{U(\widetilde{x}) < C\},$$

$$\{U(\widetilde{x}) > C\},$$

$$\{U(\widetilde{x}) < C_1 \text{ or } U(\widetilde{x}) > C_2\}.$$

又

$$\sup_{\theta \geq \theta_0} \mathsf{P}_\theta\{U(\widetilde{X}) < C\} = \mathsf{P}_{\theta_0}\{U(\widetilde{X}) < C\},$$

$$\sup_{\theta \leq \theta_0} \mathsf{P}_\theta\{U(\widetilde{X}) > C\} = \mathsf{P}_{\theta_0}\{U(\widetilde{X}) > C\}.$$

拒绝域的临界值可以由$\theta = \theta_0$时，$U(\widetilde{X})$的分布求得.

三、指数分布下参数假设检验问题

总体X服从参数为θ^{-1}的指数分布$(\theta > 0)$，其密度函数为

$$p(x;\theta) = \theta^{-1}\exp\{-x/\theta\}, \quad x > 0.$$

$\widetilde{X} = (X_1, X_2, \cdots, X_n)$是取自该总体的样本.其联合密度函数为

$$p(\widetilde{x};\theta) = \theta^{-n}\exp\{-\frac{1}{\theta}\cdot\sum_{i=1}^{n} x_i\}, \quad x_i > 0, i = 1, 2, \cdots, n.$$

考虑假设检验问题

$$H_0 : \theta \geq \theta_0 \longleftrightarrow H_1 : \theta < \theta_0. \tag{1}$$

$\overline{X}$是θ的UMVUE，取它作检验统计量(或取$T = \sum_{i=1}^{n} X_i = n\overline{X}$)，则拒绝域为$\{\widetilde{x} : \overline{x} < C\}$，

其中C由下式确定

$$\sup_{\theta\geq\theta_0} \mathsf{P}_\theta(\overline{X} < C) = \alpha.$$

由于$X/\theta \sim E(1) = \Gamma(1,1)$，所以$n\overline{X}/\theta \sim \Gamma(n,1)$，从而$2n\overline{X}/\theta \sim \Gamma(n,1/2) = \chi^2(2n)$. 故

$$\sup_{\theta\geq\theta_0} \mathsf{P}_\theta(\overline{X} < C) = \sup_{\theta\geq\theta_0} \mathsf{P}_\theta(2n\overline{X}/\theta < C\cdot 2n/\theta) = \sup_{\theta\geq\theta_0} \mathsf{P}_\theta(\chi^2(2n) < C\cdot 2n/\theta)$$
$$= \mathsf{P}_{\theta_0}(\chi^2(2n) < C\cdot 2n/\theta) = \mathsf{P}(\chi^2(2n) < C\cdot 2n/\theta_0)\Big(= \mathsf{P}_{\theta_0}(\overline{X} < C)\Big).$$

令上式等于α，则有$C\cdot 2n/\theta_0 = \chi^2_{1-\alpha}(2n)$，得假设检验问题(1)的显著性水平为$\alpha$的检验的拒绝域为

$$\{\tilde{x} : 2n\overline{x}/\theta_0 < \chi^2_{1-\alpha}(2n)\}.$$

或: 由于指数分布族属于单参数指数型分布族，且$u(\tilde{x}) = \sum_{i=1}^n x_i$，故取检验统计量为$T = u(\tilde{X}) = \sum_{i=1}^n X_i = n\overline{X}$. 注意到$E_\theta U(\tilde{X}) = E(T) = n\theta$是$\theta$的单调增函数，因此拒绝域形为$\{\tilde{x} : n\overline{x} < C'\}$,而

$$\sup_{\theta \geq \theta_0} \mathsf{P}_\theta(n\overline{X} < C') = \mathsf{P}_{\theta_0}(n\overline{X} < C'), \quad \text{令其为}\alpha.$$

则$C' = \frac{\theta_0}{2}\chi^2_{1-\alpha}(2n)$，得假设检验问题(1)的显著性水平为$\alpha$的检验的拒绝域为

$$\{\tilde{x} : n\overline{x} < \frac{\theta_0}{2}\chi^2_{1-\alpha}(2n)\} = \{\tilde{x} : 2n\overline{x}/\theta_0 < \chi^2_{1-\alpha}(2n)\}.$$

或: 联想到区间估计中的枢轴量, 取检验统计量为$2n\overline{X}/\theta_0$.

假设检验问题

$$H_0 : \theta \le \theta_0 \longleftrightarrow H_1 : \theta > \theta_0. \tag{2}$$

的显著性水平为α的检验的拒绝域为

$$\{\widetilde{x} : 2n\overline{x}/\theta_0 > \chi^2_\alpha(2n)\}.$$

假设检验问题

$$H_0 : \theta = \theta_0 \longleftrightarrow H_1 : \theta \ne \theta_0. \tag{3}$$

的显著性水平为α的检验的拒绝域为

$$\{\widetilde{x} : 2n\overline{x}/\theta_0 < \chi^2_{1-\alpha/2}(2n) \ \text{或}\ 2n\overline{x}/\theta_0 > \chi^2_{\alpha/2}(2n)\}.$$

四、两点分布下参数假设检验问题

设总体X服从二点分布$B(1,p)$,$0<p<1$. $\widetilde{X}=(X_1,X_2,\cdots,X_n)$是取自该总体的样本, 其联合分布列为:

$$p(\widetilde{x};p)=(1-p)^n\cdot\exp\{\sum_{i=1}^n x_i\cdot\ln\frac{p}{1-p}\},\quad x_i=0,1,i=1,2,\cdots,n.$$

而$\ln\frac{p}{1-p}$关于p严格递增. $U(\widetilde{x})=\sum_{i=1}^n x_i$.

(1) $H_0 : p \geq p_0 \longleftrightarrow H_1 : p < p_0$
取检验统计量为$U(\widetilde{X}) = \sum_{i=1}^{n} X_i$，则拒绝域为

$$D = \{\widetilde{x} : \sum_{i=1}^{n} x_i < C^\star\},$$

其中,

$$\begin{aligned} C^\star &= \sup_C \{C : \mathsf{P}_{p_0}\{\sum_{i=1}^{n} X_i < C\} \leq \alpha\} \\ &= \sup_C \{C : \sum_{j=0}^{-[-C]-1} \binom{n}{j} p_0^j (1-p_0)^{n-j} \leq \alpha\}. \end{aligned}$$

另一种分析:

已知$\hat{p} = \overline{X}$为p的UMVUE. 在前提$p \geq p_0$下, $\{\sum_{i=1}^n X_i < C\}$是稀有事件. 若H_0为真, $\mathsf{P}\{\sum_{i=1}^n X_i < C\}$应该比较小. 在仅作一次观测的情况下, 事件$\{\sum_{i=1}^n X_i < C\}$几乎是不发生的. 若不然，其原因就是: H_0不对了!

故取拒绝域为$D = \{\tilde{x} : \sum_{i=1}^n x_i < C\}$.

(2) $H_0: p \le p_0 \longleftrightarrow H_1: p > p_0$

取检验统计量为$U(\widetilde{X}) = \sum_{i=1}^n X_i$,则拒绝域为

$$D = \{\widetilde{x} : \sum_{i=1}^n x_i > C^\star\},$$

其中,

$$\begin{aligned} C^\star &= \inf_C \{C : \mathsf{P}_{p_0}\{\sum_{i=1}^n X_i > C\} \le \alpha\} \\ &= \inf_C \{C : \sum_{j=[C]+1}^n \binom{n}{j} p_0^j (1-p_0)^{n-j} \le \alpha\}. \end{aligned}$$

(3) $H_0: p = p_0 \longleftrightarrow H_1: p \neq p_0$

取检验统计量为$U(\widetilde{X}) = \sum_{i=1}^n X_i$,则拒绝域为

$$D = \{\widetilde{x} : \sum_{i=1}^n x_i < C_1^\star\} \bigcup \{\widetilde{x} : \sum_{i=1}^n x_i > C_2^\star\},$$

其中,

$$C_1^\star = \sup_{C_1}\{C_1 : \sum_{j=0}^{-[-C_1]-1} \binom{n}{j} p_0^j (1-p_0)^{n-j} \leq \alpha/2\};$$

$$C_2^\star = \inf_{C_2}\{C_2 : \sum_{j=[C_2]+1}^{n} \binom{n}{j} p_0^j (1-p_0)^{n-j} \leq \alpha/2\}.$$

H_0	H_1	拒绝域
$p \geq p_0$	$p < p_0$	$\{\widetilde{x} : \sum_{i=1}^n x_i \leq C'\}$, $C' = \max\left\{C \in \mathcal{N} : \sum_{i=0}^C \binom{n}{i} p_0^i (1-p_0)^{n-i} \leq \alpha\right\}$
$p \leq p_0$	$p > p_0$	$\{\widetilde{x} : \sum_{i=1}^n x_i \geq C'\}$, $C' = \min\left\{C \in \mathcal{N} : \sum_{i=C}^n \binom{n}{i} p_0^i (1-p_0)^{n-i} \leq \alpha\right\}$
$p = p_0$	$p \neq p_0$	$\{\widetilde{x} : \sum_{i=1}^n x_i \leq C_1'$ 或 $\sum_{i=1}^n x_i \geq C_2'\}$, $C_1' = \max\left\{C \in \mathcal{N} : \sum_{i=0}^C \binom{n}{i} p_0^i (1-p_0)^{n-i} \leq \alpha/2\right\}$ $C_2' = \min\left\{C \in \mathcal{N} : \sum_{i=C}^n \binom{n}{i} p_0^i (1-p_0)^{n-i} \leq \alpha/2\right\}$

当n充分大时，当$p=p_0$时

$$\frac{\sum_{i=1}^n X_i - np_0}{\sqrt{np_0(1-p_0)}}$$

近似服从$N(0,1)$. 那么此时，取上面的统计量为检验统计量，然后可用U-检验法进行检验.

五、泊松分布下参数λ假设检验问题

$\widetilde{X}=(X_1,X_2,\cdots,X_n)$取自总体$X\sim P(\lambda)$, $\mathsf{E}X=\lambda$. 假设检验问题

$$H_0:\lambda\geq\lambda_0\longleftrightarrow H_1:\lambda<\lambda_0, \tag{1}$$

$$H_0:\lambda\leq\lambda_0\longleftrightarrow H_1:\lambda>\lambda_0, \tag{2}$$

$$H_0:\lambda=\lambda_0\longleftrightarrow H_1:\lambda\neq\lambda_0. \tag{3}$$

$\overline{X}$是λ的UMVUE, 取它, 或等价地$T=\sum_{i=1}^n X_i$, 作为检验统计量. 以问题(2)为例, 拒绝域的形式为

$$\{T>C\}.$$

由定理的推论知，

$$\sup_{\lambda\leq\lambda_0}\mathsf{P}_\lambda(T>C)=\mathsf{P}_{\lambda_0}(T>C).$$

所以拒绝域为

$$\{\sum_{i=1}^n x_i>C^\star\},$$

其中，

$$C^\star=\inf_C\{C:\mathsf{P}_{\lambda_0}\{\sum_{i=1}^n X_i>C\}\leq\alpha\}$$
$$=\inf_C\{C:\sum_{j=[C]+1}^{+\infty}\frac{(n\lambda_0)^j}{j!}e^{-n\lambda_0}\leq\alpha\}.$$

或写成: 拒绝域为

$$\{\sum_{i=1}^{n} x_i \geq C'\},$$

其中,

$$C' = \min_{C}\{C \in \mathcal{N} : \mathsf{P}_{\lambda_0}\{\sum_{i=1}^{n} X_i \geq C\} \leq \alpha\}$$
$$= \min_{C}\{C \in \mathcal{N} : \sum_{j=C}^{+\infty} \frac{(n\lambda_0)^j}{j!} e^{-n\lambda_0} \leq \alpha\},$$

即C'为满足$\sum_{j=C}^{+\infty} \frac{(n\lambda_0)^j}{j!} e^{-n\lambda_0} \leq \alpha$的最小正整数.

当C为正整数时, 检验犯第一类错误的概率为

$$\begin{aligned}\alpha(\lambda) =& \mathsf{P}_\lambda(T \ge C) = \sum_{t=C}^{\infty} \frac{(n\lambda)^t}{t!} e^{-n\lambda} \\ =& \Gamma(n\lambda; C, 1) = \Gamma(2n\lambda; C, 1/2) = \chi^2(2n\lambda; 2C), \qquad \lambda \le \lambda_0.\end{aligned}$$

从而

$$\sup_{\lambda \le \lambda_0} \alpha(\lambda) = \sup_{\lambda \le \lambda_0} \mathsf{P}_\lambda(T \ge C) = \mathsf{P}_{\lambda_0}(T \ge C) = \chi^2(2n\lambda_0; 2C).$$

取C'为满足

$$2n\lambda_0 \le \chi^2_{1-\alpha}(2C)$$

的最小正整数, 即拒绝域为$\{\sum_{i=1}^n x_i \ge C'\}$, 其中

$$C' = \min_C\{C \in \mathcal{N} : 2n\lambda_0 \le \chi^2_{1-\alpha}(2C)\}.$$

类似地，检验问题

$$H_0 : \lambda \geq \lambda_0 \longleftrightarrow H_1 : \lambda < \lambda_0, \tag{1}$$

的水平为α的检验的拒绝域为$\{T \leq C'\}$，其中C'为满足$2n\lambda_0 \geq \chi^2_\alpha(2(C+1))$的最大正整数.

检验问题

$$H_0 : \lambda = \lambda_0 \longleftrightarrow H_1 : \lambda \neq \lambda_0, \tag{3}$$

的水平为α的检验的拒绝域为$\{T \geq C_1'$，或$T \leq C_2'\}$，其中C_1'为满足$2n\lambda_0 \leq \chi^2_{1-\alpha/2}(2C)$的最小正整数，$C_2'$为满足$2n\lambda_0 \geq \chi^2_{\alpha/2}(2(C+1))$的最大正整数.

当n充分大时，当$\lambda=\lambda_0$时

$$\frac{T-n\lambda_0}{\sqrt{n\lambda_0}}$$

近似服从$N(0,1)$. 我们可用U-检验.