§3.4 一致最小方差无偏估计—优化准则

均方误差准则

设$\mathcal{F}=\{p(x;\theta):\theta\in\Theta\}$是参数分布族, $g(\theta)$是待估参数.
$\widetilde{X}=(X_1,X_2,\cdots,X_n)$是取自$p(x;\theta)$的一个样本.

Definition

定义3.4.1 设$\widehat{g}=\widehat{g}(\widetilde{X})$是$g(\theta)$的一个估计量, 称

$$\mathsf{E}_\theta(\widehat{g}-g(\theta))^2$$

为$\widehat{g}$的均方误差(mean squared error).

若有$g(\theta)$的两个估计量, $\widehat{g}_1$和$\widehat{g}_2$, 满足

$$\mathsf{E}_\theta(\widehat{g}_1 - g(\theta))^2 \le \mathsf{E}_\theta(\widehat{g}_2 - g(\theta))^2, \quad \forall \theta \in \Theta,$$

则称在均方误差(MSE)准则下, $\widehat{g}_2$不优于$\widehat{g}_1$(或称$\widehat{g}_1$不劣于$\widehat{g}_2$).
若上式中的不等号至少对某一个$\theta \in \Theta$成立,则称$\widehat{g}_1$优于$\widehat{g}_2$.

Example

设$X_1, X_2, \cdots, X_n$是取自正态总体$N(\mu, \sigma^2)$的样本, $n \geq 2$.
$S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$ 和$S_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2$都是$\sigma^2$的估计. 前者是无偏估计. 计算它们的均方误差.

因为$(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$, 由$\chi^2(n)$的定义知,若$Y_i$为i.i.d.正态$N(0,1)$变量$(i = 1, 2, \cdots, n-1)$,则$\sum_{i=1}^{n-1} Y_i^2 \sim \chi^2(n-1)$. 所以

$$\begin{aligned}\mathsf{Var}\{S^2\} =& \mathsf{Var}\{\frac{\sigma^2}{n-1}\frac{(n-1)S^2}{\sigma^2}\} = \frac{\sigma^4}{(n-1)^2}\mathsf{Var}\{\sum_{i=1}^{n-1} Y_i^2\} \\ =& \frac{\sigma^4}{n-1}\mathsf{Var}(Y_1^2) = \frac{\sigma^4}{n-1}[\mathsf{E}Y_1^4 - (\mathsf{E}Y_1^2)^2] = \frac{2\sigma^4}{n-1}.\end{aligned}$$

所以S^2的均方误差为

$$\mathsf{E}\{S^2 - \sigma^2\}^2 = \mathsf{Var}\{S^2\} = \frac{2\sigma^4}{n-1}.$$

而$S_n^2=\frac{n-1}{n}S^2$, 所以

$$\mathsf{E}S_n^2=\frac{n-1}{n}\mathsf{E}S^2=\frac{n-1}{n}\sigma^2.$$

S_n^2的均方误差为

$$\begin{aligned}\mathsf{E}\{S_n^2-\sigma^2\}^2 =&\mathsf{Var}\{S_n^2\}+(\mathsf{E}S_n^2-\sigma^2)^2\\ =&(\frac{n-1}{n})^2\mathsf{Var}\{S^2\}+(-\frac{1}{n}\sigma^2)^2\\ =&(\frac{n-1}{n})^2\frac{2\sigma^4}{n-1}+\frac{\sigma^4}{n^2}=\frac{2n-1}{n^2}\sigma^4.\end{aligned}$$

$$\mathsf{E}\{S_n^2-\sigma^2\}^2<\mathsf{E}\{S^2-\sigma^2\}^2.$$

在均方误差意义下, S_n^2优于S^2.

问题是：是否存在一个$g(\theta)$的估计量$\widehat{g}$，使得在均方误差意义下它不劣于其它任意一个估计量？

即，是否存在：在均方误差意义下的最优估计？ 答案是否定的.

事实上，如果这样的统计量$\widehat{g}$存在，那么对任意给定的$\theta^* \in \Theta$，$\widehat{g}_2 \equiv g(\theta^*)$也是一个估计量，从而

$$\mathsf{E}_\theta\{\widehat{g} - g(\theta)\}^2 \le \mathsf{E}_\theta\{g(\theta^*) - g(\theta)\}^2, \quad \forall \theta \in \Theta.$$

特别时当$\theta = \theta^*$时有

$$\mathsf{E}_{\theta^*}\{\widehat{g} - g(\theta^*)\}^2 = 0.$$

由此得$\mathsf{P}_{\theta^*}\{\widehat{g} = g(\theta^*)\} = 1$. 由$\theta^*$的任意性，

$$\mathsf{P}_\theta\{\widehat{g} = g(\theta)\} = 1, \quad \forall \theta \in \Theta, \quad 即\ \widehat{g} = g(\theta)\ \ a.s.$$

最优无偏估计

在均方误差意义下找不到最优估计是因为, 所寻找的范围—-所有估计量组成的类—-太大.

我们将寻找范围缩小到由无偏估计组成的类:

$$U_g = \{\widehat{g} = \widehat{g}(\widetilde{X}) : \mathsf{E}_\theta \widehat{g} = g(\theta), \forall \theta \in \Theta\}.$$

称为$g(\theta)$的无偏估计类. U_g中统计量的均方误差就是它的方差.

U_g可能是空集. 当它是非空时, 我们称参数$g(\theta)$是可估参数.

如: 总体的均值, 若存在, 那它是可估参数.

注意：不是每个参数都是可估参数的.

Example

设容量为n的样本$\widetilde{X}=(X_1,X_2,\ldots,X_n)$取自$B(1,p)$, p为未知参数, $0<p<1$. 此时$g(p)=1/p$就是不可估参数.

事实上，对任一个统计量$T(\widetilde{X})$，其均值为

$$\mathsf{E}_p[T(\widetilde{X})]=\sum_{x_i=0,1;i=1,\ldots,n}T(\widetilde{x})p^{\sum_i x_i}(1-p)^{n-\sum_i x_i},\quad \forall p\in(0,1).$$

上式右端是一个p的n次多项式，它必是有界的,不可能在$(0,1)$上恒等于无界函数$1/p$. 所以$g(p)=1/p$的无偏估计不存在.

对于一个参数$g(\theta)$，要找其无偏估计往往也不容易,通常会从其充分统计量或一些已有的估计量中去挑选或修改. 例如方差的无偏估计S^2是从矩法估计S_n^2(或写为$m_{n,2}$)修改得到的.

Example

设$X_1,\ldots,X_n$为取自均匀分布$U(0,\theta)$, $\theta>0$的样本，问：$1/\theta$的无偏估计是否存在？如存在，请给出一个；如不存在,请说明理由.

解：考虑基于充分统计量$X_{(n)}=\max_{1\le i\le n}X_i$的估计量$T(X_{(n)})$，其均值为

$$\mathsf{E}_\theta[T(X_{(n)})]=\int_0^\theta T(x)\frac{nx^{n-1}}{\theta^n}dx.$$

若$T(X_{(n)})$是$1/\theta$的无偏估计，则有

$$\frac{1}{\theta}=\int_0^\theta T(x)\frac{nx^{n-1}}{\theta^n}dx,\ \ \forall\theta>0.$$

即

$$\theta^{n-1} = \int_0^\theta T(x) n x^{n-1} dx, \quad \forall \theta > 0.$$

再在两端对θ求导可解得$(n-1)\theta^{n-2} = T(\theta) n \theta^{n-1}$，所以

$$T(\theta) = \frac{n-1}{n\theta}.$$

即此时$T(x)$满足

$$T(x) = \frac{n-1}{nx}.$$

因此当$n \geq 2$时，$T(X_{(n)}) = \frac{n-1}{nX_{(n)}}$是$1/\theta$的无偏估计.
而当$n=1$时，上述$T(X_{(n)}) \equiv 0$，它不满足$E_\theta[T(X_{(n)})] = 1/\theta$.可以验证此时$1/\theta$的无偏估计不存在.

Definition

定义 设$\mathcal{F}=\{p(x;\theta):\theta\in\Theta\}$是一个参数分布族，$\widetilde{X}$是取自于$p(x;\theta)$的一个样本. 设$g(\theta)$是一个可估参数，$U_g$是$g(\theta)$的无偏估计类. 假如$\widehat{g}^*=\widehat{g}^*(\widetilde{X})$是这样一个无偏估计，对一切$\widehat{g}=\widehat{g}(\widetilde{X})\in U_g$，有

$$\mathsf{Var}_\theta\{\widehat{g}^*\}\le\mathsf{Var}_\theta\{\widehat{g}\},\quad\forall\theta\in\Theta.$$

则称$\widehat{g}^*$是$g(\theta)$的最优无偏估计，也称一致最小方差无偏估计(uniform minimum variance unbiased estimator)，记为UMVUE.

直接寻找一个UMVUE或证明一个估计量是UMVUE并非易事，下面的Rao-Blackwell定理给出了一个寻找UMVUE的途径.

Theorem

定理 *(Rao-Blackwell)* 设$T=T(\widetilde{X})$是参数$\theta\in\Theta$ 的充分统计量，$\varphi(\widetilde{X})$是参数$g(\theta)$的一个无偏估计．则

$$\widehat{g}(T)=E\{\varphi(\widetilde{X})|T\}$$

也是$g(\theta)$的无偏估计,且

$$Var_\theta\{\widehat{g}(T)\}\le Var\{\varphi(\widetilde{X})\},\quad \forall\theta\in\Theta.$$

其中等号成立的充分必要条件是

$$P_\theta\{\varphi(\widetilde{X})=\widehat{g}(T)\}=1,\quad \forall\theta\in\Theta.$$

定理告诉我们:

- 改善无偏估计方差的方法——现有的无偏估计对充分统计量求条件期望;
- UMVUE一定是充分统计量的函数.

定理的证明: 因为T是充分统计量,在给定T的条件下, $\widetilde{X}$的条件分布与参数θ无关, 所以条件期望$\widehat{g}(T)=\mathsf{E}\{\varphi(\widetilde{X})|T\}$与$\theta$无关, 从而是统计量. 且

$$\mathsf{E}_\theta\{\widehat{g}(T)\}=\mathsf{E}_\theta[\mathsf{E}\{\varphi(\widetilde{X})|T\}]=\mathsf{E}_\theta\{\varphi(\widetilde{X})\}=g(\theta).$$

因此$\widehat{g}(T)$是$g(\theta)$的无偏估计量.

下证第二部分结论. 写$\varphi = \varphi(\widetilde{X})$.

$$\begin{aligned}
\mathsf{Var}_\theta\{\varphi\} =& \mathsf{E}_\theta\{\varphi - \mathsf{E}_\theta(\varphi)\}^2 \\
=& \mathsf{E}_\theta\{\varphi - \mathsf{E}_\theta(\varphi|T) + \mathsf{E}_\theta(\varphi|T) - \mathsf{E}_\theta(\varphi)\}^2 \\
=& \mathsf{E}_\theta\{\varphi - \mathsf{E}_\theta(\varphi|T)\}^2 + \mathsf{E}_\theta\{\mathsf{E}_\theta(\varphi|T) - \mathsf{E}_\theta(\varphi)\}^2 \\
& + 2\mathsf{E}_\theta\left\{\left[\varphi - \mathsf{E}_\theta(\varphi|T)\right]\left[\mathsf{E}_\theta(\varphi|T) - \mathsf{E}_\theta(\varphi)\right]\right\} \\
=& \mathsf{E}_\theta\{\varphi - \widehat{g}(T)\}^2 + \mathsf{Var}_\theta\{\widehat{g}(T)\} \\
& + 2\mathsf{E}_\theta\left\{\left[\varphi - \mathsf{E}_\theta(\varphi|T)\right]\left[\mathsf{E}_\theta(\varphi|T) - \mathsf{E}_\theta(\varphi)\right]\right\}.
\end{aligned}$$

由于

$$
\begin{aligned}
&\mathsf{E}_\theta\left\{\left[\varphi-\mathsf{E}_\theta(\varphi|T)\right]\left[\mathsf{E}_\theta(\varphi|T)-\mathsf{E}_\theta(\varphi)\right]\right\}\\
=&\mathsf{E}_\theta\left(\mathsf{E}_\theta\left\{\left[\varphi-\mathsf{E}_\theta(\varphi|T)\right]\left[\mathsf{E}_\theta(\varphi|T)-\mathsf{E}_\theta(\varphi)\right]\Big|T\right\}\right)\\
=&\mathsf{E}_\theta\left(\mathsf{E}_\theta\left\{\left[\varphi-\mathsf{E}_\theta(\varphi|T)\right]\Big|T\right\}\cdot\left[\mathsf{E}_\theta(\varphi|T)-\mathsf{E}_\theta(\varphi)\right]\right)\\
=&\mathsf{E}_\theta\left(\left[\mathsf{E}_\theta(\varphi|T)-\mathsf{E}_\theta(\varphi|T)\right]\cdot\left[\mathsf{E}_\theta(\varphi|T)-\mathsf{E}_\theta(\varphi)\right]\right)=0.
\end{aligned}
$$

故得

$$
\begin{aligned}
\mathsf{Var}_\theta\{\varphi\}=&\mathsf{E}_\theta\{\varphi-\widehat{g}(T)\}^2+\mathsf{Var}_\theta\{\widehat{g}(T)\}\\
\geq&\mathsf{Var}_\theta\{\widehat{g}(T)\},\quad\forall\theta\in\Theta.
\end{aligned}
$$

等号成立当且仅当$\mathsf{E}_\theta\{\varphi-\widehat{g}(T)\}^2=0$, 即$\mathsf{P}_\theta\{\varphi=\widehat{g}(T)\}=1$.

Theorem

设$T=T(\widetilde{X})$是参数$\theta\in\Theta$ 的充分统计量，$\boldsymbol{\varphi}(\widetilde{X})=(\varphi_1(\widetilde{X}),\ldots,\varphi_k(\widetilde{X}))$是取值在$\mathbf{R}^k$上的参数$\mathbf{g}(\theta)$的一个无偏估计. 则

1. $\widehat{\mathbf{g}}(T)=\mathsf{E}\{\boldsymbol{\varphi}(\widetilde{X})|T\}$ 也是$\mathbf{g}(\theta)$的无偏估计，
2. 若记$\widehat{\mathbf{g}}(T)$ 和$\boldsymbol{\varphi}(\widetilde{X})$的协方差矩阵分别为

$$\mathbf{V}(\theta)=\{\mathsf{Cov}_\theta[\widehat{g}_i(T),\widehat{g}_j(T)]\}_{1\le i,j\le k},$$

$$\mathbf{U}(\theta)=\left\{\mathsf{Cov}_\theta[\varphi_i(\widetilde{X}),\varphi_j(\widetilde{X})]\right\}_{1\le i,j\le k},$$

则

$$\mathbf{U}(\theta)-\mathbf{V}(\theta)\ \text{非负定},\quad \forall\theta\in\Theta,$$

且$\mathbf{U}(\theta)-\mathbf{V}(\theta)$为零矩阵的充分必要条件是

$$\mathsf{P}_\theta\{\boldsymbol{\varphi}(\widetilde{X})=\widehat{\mathbf{g}}(T)\}=1,\quad \forall\theta\in\Theta.$$

Example

设$X_1, X_2, \cdots, X_n$是来自二点分布$B(1,p)$的一个样本，讨论p的无偏估计.

X_1是p的无偏估计($\because \mathsf{E}X_1 = p$)，而$T = X_1 + \cdots + X_n$是参数$p$的充分统计量.
我们用求条件期望的方法改进无偏估计.

$$\mathsf{E}(X_1|T) = \mathsf{E}(X_2|T) = \cdots = \mathsf{E}(X_n|T).$$

所以

$$\begin{aligned}\mathsf{E}(X_1|T) =& \frac{\mathsf{E}(X_1|T) + \mathsf{E}(X_2|T) \cdots + \mathsf{E}(X_n|T)}{n} \\ =& \frac{\mathsf{E}(X_1 + X_2 \cdots + X_n|T)}{n} \\ =& \frac{\mathsf{E}(T|T)}{n} = \frac{T}{n} = \overline{X}.\end{aligned}$$

零无偏估计法

Theorem

定理3.4.1 设$\widehat{g}=\widehat{g}(\widetilde{X})$是$g(\theta)$的一个无偏估计. $Var_\theta(\widehat{g})<\infty$(对$\forall\theta\in\Theta$).记零无偏估计集

$$\mathcal{L}=\{l=l(\widetilde{X}):E_\theta(l(\widetilde{X}))=0,\forall\theta\in\Theta\}.$$

若对$\forall l\in\mathcal{L}$, 有

$$Cov_\theta(\widehat{g},l)=E_\theta(\widehat{g}\cdot l)=0,\quad\forall\theta\in\Theta,\tag{3.4.2}$$

则$\widehat{g}$必为$g(\theta)$的$UMVUE$.

证明: 设$\widehat{g}_1 = \widehat{g}_1(\widetilde{X})$是$g(\theta)$的任一无偏估计. 记$l = l(\widetilde{X}) = \widehat{g}_1 - \widehat{g}$, 则$\mathsf{E}(l(\widetilde{X})) = 0$, 故$l \in \mathcal{L}$. 利用(3.4.2), 可得

$$\begin{aligned}\mathsf{Var}_\theta(\widehat{g}_1) &= \mathsf{Var}_\theta(\widehat{g} + l) = \mathsf{Var}_\theta(\widehat{g}) + \mathsf{Var}(l) + 2\mathsf{Cov}_\theta(\widehat{g}, l) \\ &= \mathsf{Var}_\theta(\widehat{g}) + \mathsf{Var}(l) \geq \mathsf{Var}_\theta(\widehat{g}), \quad \forall \theta \in \Theta.\end{aligned}$$

这就意味着$\widehat{g}$为$g(\theta)$的UMVUE.

Example

设$X_1, X_2, \cdots, X_n$是来自二点分布$B(1,p)$的一个样本, $0<p<1$, 证明$\overline{X}$为p的UMVUE.

证明: 由于$\mathsf{E}_p(\overline{X})=p$且$\mathsf{Var}_p(\overline{X})=p(1-p)/n<\infty$. 故由定理3.4.1知, 要证$\overline{X}$为$p$的UMVUE, 只需证明$\overline{X}$满足(3.4.2)即可. 现设$l=l(\tilde{X})$为零无偏估计, 则

$$0=\mathsf{E}_p(l)=\sum_{x_i=0,1;i=1,\ldots,n} l(\tilde{x})p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}, \quad \forall p\in(0,1).$$

令$\theta=\frac{p}{1-p}$, 则有

$$0=\sum_{x_i=0,1;i=1,\ldots,n} l(\tilde{x})\theta^{\sum_{i=1}^n x_i}, \quad \forall \theta>0 \qquad (✠1).$$

对(✠1)式的两边关于θ求导,得

$$0=\sum_{x_i=0,1;i=1,\ldots,n}\left(\sum_{i=1}^n x_i\right)\cdot l(\widetilde{x})\theta^{\sum_{i=1}^n x_i-1},\quad \forall\theta>0,$$

故

$$0=\sum_{x_i=0,1;i=1,\ldots,n}\left(\sum_{i=1}^n x_i\right)\cdot l(\widetilde{x})\theta^{\sum_{i=1}^n x_i},\quad \forall\theta>0,$$

从而$\mathsf{Cov}_p(\overline{X},l(\widetilde{X}))=\mathsf{E}_p(\overline{X}\cdot l(\widetilde{X}))=0$, 即(3.4.2)成立. 所以$\overline{X}$为$p$的UMVUE.

Corollary

推论3.4.1 设$T=T(\tilde{X})$为θ的充分统计量，$h(T)$是$g(\theta)$的一个无偏估计量，且$Var_\theta(h(T))<\infty$ (对$\forall\theta\in\Theta$). 记

$$\mathcal{L}_T=\{\delta=\delta(T): E_\theta(\delta(T))=0,\forall\theta\in\Theta\}.$$

若对$\forall\delta\in\mathcal{L}_T$，有

$$Cov_\theta(h(T),\delta(T))=E_\theta(h(T)\cdot\delta(T))=0,\quad\forall\theta\in\Theta,\qquad(3.4.3)$$

则$h(T)$必为$g(\theta)$的$UMVUE$.

Example

设$X_1, X_2, \cdots, X_n$是来自二点分布$B(1,p)$的一个样本, 证明$\overline{X}$为p的UMVUE.

证明:已知$T=\sum_{i=1}^n X_i$为p的充分统计量, 并且$\overline{X}=T/n$, $\mathsf{E}_p(\overline{X})=p$, $\mathsf{Var}_p(\overline{X})=p(1-p)/n<\infty$. 由推论3.4.1知,要证$\overline{X}$ 为p 的UMVUE, 只需证明$\overline{X}$满足(3.4.3)即可. 现设$\delta(T)\in\mathcal{L}_T$, 即$\mathsf{E}_p(\delta(T))=0$. 注意到$T\sim B(n,p)$, 则

$$0=\mathsf{E}_p(\delta(T))=\sum_{k=0}^n \delta(k)\binom{n}{k}p^k(1-p)^{n-k}, \quad \forall p\in(0,1).$$

令$\theta=\frac{p}{1-p}$, 并约去非零项, 则有

$$0=\sum_{k=0}^n \delta(k)\binom{n}{k}\theta^k, \quad \forall \theta>0 \qquad (\maltese 2).$$

对(✠2)式的两边关于θ求导,得

$$0=\sum_{k=0}^{n} k\cdot\delta(k)\tbinom{n}{k}\theta^{k-1},\quad \forall\theta>0,$$

故

$$0=\sum_{k=0}^{n} k\cdot\delta(k)\tbinom{n}{k}\theta^{k},\quad \forall\theta>0,$$

从而$\mathsf{Cov}_p(\overline{X},\delta(T))=\mathsf{E}_p(\overline{X}\cdot\delta(T))=\frac{1}{n}\mathsf{E}_p(T\cdot\delta(T))=0$, 即(3.4.3)成立. 据推论3.4.1, 可知$\overline{X}$为p的UMVUE.

思考:

由前面的讨论中, 我们得知:

设$h(T)$是$g(\theta)$的一个无偏估计, 其中T为θ的充分统计量. 如果$h(T)$与任意一个T的函数类中的零无偏估计的乘积的期望为0(此时称两个随机变量正交), 那么可以推出$h(T)$是$g(\theta)$的UMVUE.

显然零属于T的函数类, 也是一个零无偏估计, 而且零与任何随机变量正交.

如果T的函数类中的“零无偏估计只有零”, 那么这个无偏估计$h(T)$一定与之正交, 那个它也就是$g(\theta)$的UMVUE.

什么叫T的函数类中的“零无偏估计只有零”

$$\mathsf{E}_\theta \varphi(T) = 0, \quad \forall \theta \in \Theta$$

$$\Downarrow$$

$$\mathsf{P}_\theta\{\varphi(T) = 0\} = 1, \quad \forall \theta \in \Theta.$$

完备性(或称完全性)

Definition

定义 设X的分布所在的分布族为$\mathcal{F}=\{p(x;\theta):\theta\in\Theta\}$. 假如对任一个实函数$\varphi(x)$，由

$$\mathsf{E}_\theta\varphi(X)=0,\quad \forall\theta\in\Theta$$

可推出

$$\mathsf{P}_\theta\{\varphi(X)=0\}=1,\quad \forall\theta\in\Theta.$$

则称此分布族$\mathcal{F}$是完备的(或称为完全的).

Example

例 二项分布族$\{B(n,p):0<p<1\}$是完备的.

证：假如$\varphi(x)$满足

$$\mathsf{E}_p\varphi(X)=\sum_{x=0}^{n}\varphi(x)\binom{n}{x}p^x(1-p)^{n-x}=0,\quad \forall p\in(0,1).$$

令$\theta=p/(1-p)$,则上式可改写为

$$\sum_{x=0}^{n}\binom{n}{x}\varphi(x)\theta^x=0,\quad \forall\theta>0.$$

上式为θ的多项式，而$\binom{n}{x}\varphi(x)$是多项式的系数，所以$\varphi(x)=0$ $(x=0,1,\cdots,n)$. 又$\mathsf{P}_p(X=0,1,\cdots,n)=1,\forall p\in(0,1)$, 故$\mathsf{P}_p\{\varphi(X)=0\}=1,\forall p\in(0,1)$.所以二项分布族是完备的.

Example

正态分布族$\{N(0,\sigma^2):\sigma>0\}$是不完备的.

$$\varphi(x)=x.$$

Example

Gamma分布族$\{\Gamma(\alpha,\lambda):\lambda>0\}$是完备的(其中$\alpha>0$已知).

证明: 设实函数$\varphi(t)$ 满足$\mathsf{E}_\lambda\varphi(\Gamma)=0$, 其中$\Gamma\sim\Gamma(\alpha,\lambda)$, 即

$$\int_0^\infty \varphi(t)t^{\alpha-1}e^{-\lambda t}dt=0,\quad \forall\lambda>0.$$

上式左边是$\varphi(t)t^{\alpha-1}$的Laplace变换, 由Laplace变换的唯一性知:

$$\varphi(t)t^{\alpha-1}=0\quad a.e.,\quad \forall t>0,\ \text{故}\ \varphi(t)=0\quad a.e.\ \ \forall t>0.$$

注意到$\mathsf{P}_\lambda(\Gamma>0)=1,\forall\lambda>0$, 所以

$$\mathsf{P}_\lambda\{\varphi(\Gamma)=0\}=1,\quad \forall\lambda>0.$$

Example

正态分布族$\{N(\mu,\sigma^2): -\infty<\mu<\infty\}$是完备的,其中$\sigma$已知.

证明:设实函数$\varphi(x)$ 满足$\mathsf{E}_\mu\varphi(X)=0$, 其中$X\sim N(\mu,\sigma^2)$, 即

$$\int_{-\infty}^{\infty}\varphi(x)\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=0\ \ \forall\mu\in\mathbb{R},$$

$$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{\mu^2}{2\sigma^2}}\int_{-\infty}^{\infty}\varphi(x)e^{-\frac{x^2}{2\sigma^2}}e^{\frac{\mu}{\sigma^2}x}dx=0\ \ \forall\mu\in\mathbb{R}.$$

也即

$$\int_{-\infty}^{\infty}\left[\varphi(x)e^{-\frac{x^2}{2\sigma^2}}\right]e^{-\lambda x}dx=0\ \ \forall\lambda\in\mathbb{R}.$$

上式左边是$\varphi(x)e^{-\frac{x^2}{2\sigma^2}}$的Laplace变换，从而

$$\varphi(x)e^{-\frac{x^2}{2\sigma^2}}=0,\quad a.e.$$

即$\varphi(x)=0$ a.e. 所以

$$\mathsf{P}_\mu\{\varphi(X)=0\}=1.$$

Definition

定义设有参数分布族$\mathcal{F}=\{p(x;\theta):\theta\in\Theta\}$. T 是一个统计量. 假如对任一实函数$\varphi(t)$, 由

$$\mathsf{E}_\theta\varphi(T)=0,\quad \forall\theta\in\Theta$$

可推出

$$\mathsf{P}_\theta\{\varphi(T)=0\}=1,\forall\theta\in\Theta,$$

也就是说, 由T诱导出的分布族$\mathcal{F}^T=\{p_T(t;\theta):\theta\in\Theta\}$是完备的, 则称$T$是完备统计量.

$\{N(0,\sigma^2):\sigma>0\}$不完备, 但统计量$T=X_1^2+\cdots+X_n^2\sim\Gamma(\frac{n}{2},\frac{1}{2\sigma^2})$是完备统计量.

Example

设$X_1, X_2, \cdots, X_n$是来自正态总体$N(\mu, \sigma^2)$ 的样本. 证明$(\overline{X}, S^2)$ 是完备统计量.

证明: 设有$\varphi(x,t)$使得$E_{\mu,\sigma^2}\varphi(\overline{X}, S^2) = 0$. 由于$\overline{X} \sim N(\mu, \sigma^2/n)$, $(n-1)S^2/\sigma^2 \sim \chi^2(n-1) = \Gamma(\frac{n-1}{2}, \frac{1}{2})$, 所以$S^2 \sim \Gamma(\frac{n-1}{2}, \frac{n-1}{2\sigma^2})$.且$\overline{X}$ 和S^2相互独立, 从而

$$\int_{-\infty}^{\infty}\left[\int_0^{\infty}\varphi(x,t)f_{\sigma^2}(t)dt\right]\frac{1}{\sqrt{2\pi\sigma^2/n}}e^{-\frac{(x-\mu)^2}{2\sigma^2/n}}dx = 0, \quad \forall\sigma > 0, \mu \in \mathbb{R},$$

其中$f_{\sigma^2}(t)$为$\Gamma(\frac{n-1}{2}, \frac{n-1}{2\sigma^2})$的密度函数. 先固定$\sigma^2$, 上式对$\forall\mu \in \mathbb{R}$ 成立, 由正态分布族$\{N(\mu, \sigma^2/n), -\infty < \mu < \infty\}$的完备性知

$$\int_0^\infty \varphi(x,t) f_{\sigma^2}(t) dt = 0, a.e., \quad \forall \sigma > 0.$$

再由gamma分布族的完备性知道

$$\varphi(x,t) = 0, a.e..$$

所以$(\overline{X}, S^2)$是完备统计量.

Example

设$X_1, X_2, \cdots, X_n$是来自均匀分布总体$U(\theta-1/2, \theta+1/2)$的简单随机样本, $n \geq 1$, 其中$\theta \in \mathbb{R}$. 记$T=(X_{(1)}, X_{(n)})$, 证明T是θ的充分统计量, 但不是完备统计量.

证明: 由于$(X_1, X_2, \cdots, X_n)$的联合概率密度函数为

$$\begin{aligned} p(\tilde{x};\theta) &= \prod_{i=1}^{n} 1 \cdot I_{\{\theta-1/2<x_i<\theta+1/2, i=1,2,\cdots,n\}} \\ &= I_{\{\theta-1/2<x_{(1)}\leq x_{(n)}<\theta+1/2\}}, \end{aligned}$$

其中$x_{(1)}=\min\{x_i, i=1,\ldots,n\}$, $x_{(n)}=\max\{x_i, i=1,\ldots,n\}$. 令$h(\tilde{x})=1$, 根据因子分解定理, 可知$T=(X_{(1)}, X_{(n)})$是$\theta$的充分统计量.

为证T是不完备的，只需找到一个在概率意义下非0的$\varphi(T)=\varphi(X_{(1)},X_{(n)})$，且满足$\mathsf{E}_\theta(\varphi(T))=\mathsf{E}_\theta(\varphi(X_{(1)},X_{(n)}))=0$. 令$Y_i=\frac{X_i-(\theta-1/2)}{1}$，则$Y_1,Y_2,\cdots,Y_n$相互独立，且均服从$U(0,1)$，并记$Y_{(1)}=\min\{Y_i,i=1,...,n\}$，$Y_{(n)}=\max\{Y_i,i=1,...,n\}$，则

$$X_{(n)}-X_{(1)}=Y_{(n)}-Y_{(1)},$$

其分布与参数θ无关. 又注意到

$$\mathsf{E}_\theta(X_{(n)}-X_{(1)})=\mathsf{E}(Y_{(n)}-Y_{(1)})=c_n=\frac{n-1}{n+1}.$$

故可取$\varphi(T)=X_{(n)}-X_{(1)}-c_n$，此时有

$$\mathsf{E}_\theta(\varphi(T))=0,\quad\forall\theta\in\mathbb{R}.$$

但

$$\mathsf{P}_\theta(X_{(n)}-X_{(1)}-c_n=0)=0\neq1,\quad\forall\theta\in\mathbb{R}.$$

Theorem

定理2.8.2 (Basu定理) 设总体X的分布所在的分布族为$\mathcal{F}=\{p(x;\theta):\theta\in\Theta\}$，$\widetilde{X}=(X_1,X_2,\cdots,X_n)$是从该总体中抽取的简单随机样本．设$T=T(\widetilde{X})$是一完备统计量，且为参数$\theta$的充分统计量．若随机变量$V=V(\widetilde{X})$的分布与参数$\theta$无关，则对于任何$\theta\in\Theta$，$T$与$V$独立．

证明： 要证明T与V独立，只需证：对于任意的Borel集A，B，有

$$\mathsf{P}_\theta(T\in A,V\in B)=\mathsf{P}_\theta(T\in A)\cdot\mathsf{P}_\theta(V\in B),\quad \forall\theta\in\Theta.$$

由于V的分布与参数θ无关, 则可令

$$b = \mathsf{P}_\theta(V \in B),$$

b为与参数θ无关的常数. 并记

$$\psi(T) = \mathsf{E}(I_{\{V\in B\}}|T).$$

注意到T为参数θ的充分统计量, 因此$\psi(T)$与参数θ无关, 仅为T的函数. 又

$$\mathsf{E}_\theta(\psi(T) - b) = \mathsf{E}_\theta(\mathsf{E}(I_{\{V\in B\}}|T)) - b = \mathsf{P}(V \in B) - \mathsf{P}(V \in B) = 0, \quad \forall\theta \in \Theta.$$

由于T的完备性, 可知

$$\psi(T) - b = 0, \quad a.s. \quad \mathsf{P}_\theta, \quad \forall\theta \in \Theta,$$

即

$$\psi(T) = b, \quad a.s. \quad \mathsf{P}_\theta, \quad \forall\theta \in \Theta.$$

因此

$$
\begin{aligned}
&\mathsf{P}_\theta(T\in A, V\in B)=\mathsf{E}_\theta(I_{\{T\in A, V\in B\}})\\
=&\mathsf{E}_\theta(I_{\{T\in A\}}\cdot I_{\{V\in B\}})=\mathsf{E}_\theta(\mathsf{E}_\theta(I_{\{T\in A\}}\cdot I_{\{V\in B\}}|T))\\
=&\mathsf{E}_\theta(I_{\{T\in A\}}\cdot \mathsf{E}_\theta(I_{\{V\in B\}}|T))=\mathsf{E}_\theta(I_{\{T\in A\}}\cdot\psi(T))\\
=&\mathsf{E}_\theta(I_{\{T\in A\}}\cdot b)=\mathsf{E}_\theta(I_{\{T\in A\}})\cdot b\\
=&\mathsf{P}_\theta(T\in A)\cdot b=\mathsf{P}_\theta(T\in A)\cdot\mathsf{P}_\theta(V\in B).
\end{aligned}
$$

Example

设$X_1, X_2, \cdots, X_n$是来自正态总体$N(\mu, \sigma^2)$的样本，利用Busa定理，可得对于任意的参数$\mu \in \mathbb{R}, \sigma > 0$，样本偏度

$$\frac{\sqrt{n}\sum_{i=1}^n (X_i - \overline{X})^3}{\left(\sum_{i=1}^n (X_i - \overline{X})^2\right)^{3/2}} = \frac{m_{n,3}}{m_{n,2}^{2/3}}$$

与$(\overline{X}, S^2)$ 独立.

因为
$$\frac{m_{n,3}}{m_{n,2}^{2/3}} = \frac{\sqrt{n}\sum_{i=1}^n (Y_i - \overline{Y})^3}{\left(\sum_{i=1}^n (Y_i - \overline{Y})^2\right)^{3/2}},$$
其中$Y_i = \frac{X_i - \mu}{\sigma}$, i.i.d.,服从标准正态，即样本偏度的分布与参数μ, σ无关. 根据Basu定理，可知对于任意的参数$\mu \in \mathbb{R}, \sigma > 0$, $\frac{m_{n,3}}{m_{n,2}^{2/3}}$与参数$(\mu, \sigma)$的充分统计量$(\overline{X}, S^2)$ 独立.

Rao-Blackwell定理说明，如果T是θ的充分统计量，则可估参数$g(\theta)$的UMVUE可以在T的函数类中寻找.

由零无偏估计法(推论)可知，可估参数$g(\theta)$一个无偏估计$h(T)$，其中T为θ的充分统计量，如果满足$h(T)$与任意一个T的函数类中的零无偏估计的乘积的期望为0(此时称两个随机变量正交)，那么可以推出$h(T)$是$g(\theta)$的UMVUE.

而完备性告诉我们，如果T是完备统计量，则在T的函数类中,只有零是零的无偏估计，即对于任何满足$\mathsf{E}_\theta l(T) = 0, \quad \forall \theta \in \Theta$的$l(T)$必有$\mathsf{P}(l(T) = 0) = 1$.

从而，如果T既是充分的又是完备的，那么只要$\varphi(\widetilde{X})$是$g(\theta)$的无偏估计,那么$E[\varphi(\widetilde{X})|T]$就是$g(\theta)$的UMVUE.

充分完备统计量法

Theorem

定理3.4.2 ($Lehmann-Scheff\grave{e}$, 简称LS定理) 设$S=S(\tilde{X})$是参数θ的充分完备统计量(*complete sufficient statistic*), 则可估参数$g(\theta)$的$UMVUE$存在且唯一.
若$\varphi(\tilde{X})$是$g(\theta)$的一个无偏估计, 那么$g(\theta)$唯一的$UMVUE$就是

$$\mathsf{E}[\varphi(\tilde{X})|S].$$

上面的定理给出了我们找UMVUE的方法:

1. 先找一个充分完备统计量S, 再找一个可估参数的无偏估计φ, 计算条件期望$\mathsf{E}[\varphi|S]$就得UMVUE.

2. 先找一个充分完备统计量S, 再找一个S的函数$h(S)$使得它是可估参数的无偏估计, $h(S)$即为所求的UMVUE.

Example

设总体X来自两点分布族$\{B(1,p); 0<p<1\}$. $\widetilde{X}=(X_1,X_2,\cdots,X_n)$为从中抽取的简单随机样本. 我们已证 $T=\sum_{i=1}^n X_i$是参数p的充分完备统计量. 下求p的UMVUE.

由无偏估计X_1算得$\mathsf{E}[X_1|T]=T/n=\overline{X}$, $\overline{X}$就是p的UMVUE.

Example

下求p^2的UMVUE.

因为

$$\mathsf{E}T^2 = \sum_{i=1}^{n} \mathsf{E}X_i^2 + \sum_{i\neq j} \mathsf{E}X_i \mathsf{E}X_j = \mathsf{E}T + n(n-1)p^2.$$

所以

$$p^2 = \frac{\mathsf{E}(T^2 - T)}{n(n-1)} = \mathsf{E}\left(\frac{T^2 - T}{n(n-1)}\right).$$

故

$$\frac{T^2 - T}{n(n-1)} = \frac{1}{n-1}(n\overline{X}^2 - \overline{X})$$

就是p^2的UMVUE.

LS定理的证明:

存在性: 设$\varphi=\varphi(\widetilde{X})$是可估参数$g(\theta)$的一个无偏估计. 记

$$\widehat{g}=\widehat{g}(\widetilde{X})=\mathsf{E}[\varphi|S].$$

由定理(Rao-Blackwell), 知$\widehat{g}$也是$g(\theta)$的无偏估计. 下面我们证明上述$\widehat{g}$就是$g(\theta)$的UMVUE. 设$\widehat{f}=f(\widetilde{X})$是$g(\theta)$的其它的任意一个无偏估计. 由Rao-Blackwell定理, $\mathsf{E}[\widehat{f}|S]$也是$g(\theta)$的无偏估计.且

$$\mathsf{Var}_\theta\{\mathsf{E}[\widehat{f}|S]\}\le\mathsf{Var}_\theta\{\widehat{f}\}.$$

由于$\widehat{g}$和$\mathsf{E}[\widehat{f}|S]$都是$g(\theta)$的无偏估计, 即, $\mathsf{E}_\theta\{\widehat{g}-\mathsf{E}[f|S]\}=0,\ \forall\theta$. 且$\widehat{g}$和$\mathsf{E}[\widehat{f}|S]$都是$S$的函数, 故由$S$的完备性, 可知$\mathsf{P}_\theta\{\widehat{g}=\mathsf{E}[\widehat{f}|S]\}=1$. 从而

$$\mathsf{Var}_\theta\{\widehat{g}\}=\mathsf{Var}_\theta\{\mathsf{E}[\widehat{f}|S]\}\le\mathsf{Var}_\theta\{\widehat{f}\}.$$

所以$\widehat{g}$ 是$g(\theta)$的UMVUE.

唯一性：设$W=W(\tilde{X})$和$Y=Y(\tilde{X})$都是$g(\theta)$的UMVUE. 则由定义必有

$$\mathsf{E}_\theta W=\mathsf{E}_\theta Y=g(\theta),\quad \mathsf{Var}_\theta\{W\}=\mathsf{Var}_\theta\{Y\}\ \ \forall\theta.$$

注意到

$$\begin{aligned}\mathsf{E}_\theta\{W-Y\}^2=&\mathsf{Var}_\theta\{W-Y\}\\=&\mathsf{Var}_\theta\{W\}+\mathsf{Var}_\theta\{Y\}-2\mathsf{Cov}_\theta\{W,Y\}\\=&2\mathsf{Var}_\theta\{W\}-2\mathsf{Cov}_\theta\{W,Y\}.\end{aligned}$$

如果能够证明

$$\mathsf{Cov}_\theta\{W, Y\} = \mathsf{Var}_\theta\{W\} = \mathsf{Var}_\theta\{Y\}.$$

那么就有$\mathsf{E}_\theta\{W - Y\}^2 = 0$. 因而得证：

$$\mathsf{P}_\theta\{W = Y\} = 1, \quad \forall \theta.$$

为证$\mathsf{Cov}_\theta\{W,Y\}=\mathsf{Var}_\theta\{W\}$, 记$Z=(W+Y)/2$, 那么$Z$也是$g(\theta)$的无偏估计. 由于$W$是UMVUE, 从而

$$
\begin{aligned}
&\mathsf{Var}_\theta\{W\}\le \mathsf{Var}_\theta\{Z\}\\
=&\frac{1}{4}\mathsf{Var}_\theta\{W\}+\frac{1}{4}\mathsf{Var}_\theta\{Y\}+\frac{1}{2}\mathsf{Cov}_\theta\{W,Y\}\\
\le&\frac{1}{4}\mathsf{Var}_\theta\{W\}+\frac{1}{4}\mathsf{Var}_\theta\{Y\}+\frac{1}{2}(\mathsf{Var}_\theta\{W\})^{1/2}\cdot(\mathsf{Var}_\theta\{Y\})^{1/2}\\
&\qquad\qquad(\text{Cauchy-Schwarz不等式})\\
=&\mathsf{Var}_\theta\{W\}.
\end{aligned}
$$

所以

$$\mathsf{Cov}_\theta\{W,Y\}=\mathsf{Var}_\theta\{W\}=\mathsf{Var}_\theta\{Y\}.$$

Example

设$X_1, X_2, \cdots, X_n$是来自参数为λ的Poisson总体的一个样本，$\lambda > 0$未知．求:

(1) 参数λ的UMVUE;

(2) 概率

$$P_\lambda(k) = \frac{\lambda^k}{k!}e^{-\lambda}$$

的UMVUE$(k = 0, 1, \cdots)$.

解: 先求λ的充分完备统计量.

注意到, $X_1, X_2, \cdots, X_n$的联合分布列为

$$
\begin{aligned}
&\mathsf{P}(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) \\
=&e^{-\lambda n}\lambda^{x_1+x_2+\cdots+x_n}\frac{1}{x_1!x_2!\cdots x_n!}\cdot I\{x_i = 0, 1, \cdots; i = 1, 2, \cdots, n\}.
\end{aligned}
$$

取$g(T(\widetilde{x}); \lambda) = e^{-\lambda n}\lambda^{x_1+x_2+\cdots+x_n}$,

$h(\widetilde{x}) = \frac{1}{x_1!x_2!\cdots x_n!}\cdot I\{x_i = 0, 1, \cdots; i = 1, 2, \cdots, n\}$, 那么由因子分解定理知,

$T_n = \sum_{i=1}^n X_i$是充分统计量.

下证明它是完备统计量. 由于$T_n \sim P(n\lambda)$, 如果

$$\mathsf{E}_\lambda f(T_n) = e^{-n\lambda}\sum_{t=0}^{+\infty} f(t)\frac{(n\lambda)^t}{t!} = 0, \quad \forall\lambda > 0,$$

令$\delta = n\lambda$, 去掉非零项, 即

$$\sum_{t=0}^{+\infty} f(t)\frac{\delta^t}{t!} = 0, \quad \forall\delta > 0$$

成立, 那么对$t = 0, 1, 2, \cdots,$ 必有δ^t的系数都为零, 即

$$f(t)\frac{1}{t!} = 0, \quad t = 0, 1, 2, \cdots,$$

所以$f(t) = 0$, $t = 0, 1, 2, \cdots$. 又$\mathsf{P}_\lambda(T_n = 0, 1, 2, \cdots) = 1$, 从而$\mathsf{P}_\lambda\{f(T_n) = 0\} = 1$, 故$T_n$是完备的.

(1) 由于统计量$\overline{X} = \frac{T_n}{n}$为$T_n$的函数，而且$\mathsf{E}\overline{X} = \mathsf{E}X = \lambda$，即$\overline{X}$是$\lambda$的无偏估计，据LS定理可知$\overline{X}$是$\lambda$的UMVUE.

(2) 统计量

$$\varphi_k(\widetilde{X}) = \begin{cases} 1, & X_1 = k, \\ 0, & X_1 \neq k \end{cases}$$

是$P_\lambda(k)$的无偏估计. 那么对于$t \geq k$, k, t均为非负整数, 有

$$\begin{aligned} &\mathsf{E}_\lambda[\varphi_k(\widetilde{X})|T_n = t] = 1 \cdot \mathsf{P}(\varphi_k(\widetilde{X}) = 1|T_n = t) = \mathsf{P}(X_1 = k|T_n = t) \\ =&\frac{\mathsf{P}(X_1 = k, T_n = t)}{\mathsf{P}(T_n = t)} = \frac{\mathsf{P}(X_1 = k)\mathsf{P}(X_2 + \cdots + X_n = t - k)}{\mathsf{P}(T_n = t)} \\ =&\frac{\frac{\lambda^k}{k!}e^{-\lambda} \cdot \frac{[(n-1)\lambda]^{t-k}}{(t-k)!}e^{-(n-1)\lambda}}{\frac{(n\lambda)^t}{t!}e^{-n\lambda}} = \binom{t}{k}\left(\frac{1}{n}\right)^k\left(1 - \frac{1}{n}\right)^{t-k}. \end{aligned}$$

故$P_\lambda(k)$的UMVUE为

$$\widehat{P_\lambda(k)} = \binom{T_n}{k}\left(\frac{1}{n}\right)^k\left(1-\frac{1}{n}\right)^{T_n-k} I\{T_n \geq k\}.$$

特别地

$$\widehat{P_\lambda(0)} = \widehat{e^{-\lambda}} = \left(1-\frac{1}{n}\right)^{T_n}.$$

Example

例 设$X_1, X_2, \cdots, X_n$是取自均匀分布$U(0,\theta)$的一个样本，$\theta > 0$未知，$n \geq 2$. 求参数θ的UMVUE. 并比较它与矩法估计的有效性.

解：先找一个充分完备统计量. 已证明$T = X_{(n)}$是充分统计量. 下证它的完备性. $X_{(n)}$的密度函数为

$$p(t;\theta) = \frac{n}{\theta^n} t^{n-1}, \quad 0 < t < \theta.$$

如果

$$\mathsf{E}_\theta\varphi(X_{(n)}) = \int_0^\theta \varphi(t)\frac{n}{\theta^n}t^{n-1}dt = 0, \quad \forall\theta > 0,$$

即

$$\int_0^\theta \varphi(t)t^{n-1}dt = 0, \quad \forall\theta > 0.$$

对θ求导数得

$$\varphi(\theta)\theta^{n-1} = 0, \quad \forall\theta > 0.$$

故

$$\varphi(t) = 0, \quad \forall t > 0.$$

注意到$\mathsf{P}(X_{(n)} > 0) \geq \mathsf{P}(0 < X_{(n)} < \theta) = 1$, 故$\mathsf{P}(\varphi(X_{(n)}) = 0) = 1$, 即$X_{(n)}$的完备性得证.

由于

$$\mathsf{E}X_{(n)}=\int_0^\theta t\cdot\frac{n}{\theta^n}t^{n-1}dt=\frac{n}{n+1}\theta.$$

所以$\widehat{\theta}=(1+\frac{1}{n})X_{(n)}$是$\theta$的无偏估计. 因为它是充分完备统计量的函数. 所以它是θ的UMVUE.

下面求$\widehat{\theta}$的方差.

$$\mathsf{E}X_{(n)}^2=\int_0^\theta t^2\cdot\frac{n}{\theta^n}t^{n-1}dt=\frac{n}{n+2}\theta^2.$$

所以

$$\mathsf{Var}(\widehat{\theta})=\mathsf{E}\widehat{\theta}^2-\theta^2=(1+\frac{1}{n})^2\frac{n}{n+2}\theta^2-\theta^2=\frac{\theta^2}{n(n+2)}.$$

由于$\mathsf{E}X=\frac{\theta}{2}$, $\theta=2\mathsf{E}X$. 所以θ的矩法估计为

$$\widehat{\theta}_1=2\overline{X}.$$

$$\mathsf{Var}(\widehat{\theta}_1)=4\mathsf{Var}(\overline{X})=4\frac{\theta^2}{12n}=\frac{\theta^2}{3n}.$$

这里，当$n\geq 2$时，UMVUE 比矩法估计有效.

Theorem

设总体来自指数型分布族,从总体中抽取的样本的联合pdf或pmf为

$$p(\tilde{x};\theta) = c^*(\theta)\exp\left\{\sum_{j=1}^{k} Q_j(\theta)T_j^*(\tilde{x})\right\} h^*(\tilde{x}),$$

若$\mathbf{Q} = (Q_1(\theta),\cdots,Q_k(\theta))$的值域有非空的内部，则$T = (T_1^*,...,T_k^*)$为充分完备统计量.

Example

设$X_1, X_2, \cdots, X_n$是来自正态总体$N(\mu, \sigma^2)$ 的样本，$-\infty < \mu < \infty$, $\sigma > 0$未知．求μ和σ^2的UMVUE.

解：已知$\mathcal{F} = \{N(\mu, \sigma^2) : -\infty < \mu < \infty, \sigma > 0\}$为指数型分布族，且$\widetilde{X} = (X_1, X_2, \cdots, X_n)$的联合密度函数为

$$p(\widetilde{x}; \mu, \sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{n\mu^2}{2\sigma^2}\right\}$$

$$\cdot \exp\left\{\left[\sum_{i=1}^{n} x_i\right]\left[\frac{\mu}{\sigma^2}\right] + \left[\sum_{i=1}^{n} x_i^2\right]\left[-\frac{1}{2\sigma^2}\right]\right\}.$$

对于$-\infty < \mu < \infty$, $\sigma > 0$, 注意到$(\mu/\sigma^2, -1/(2\sigma^2))$的值域有非空的内部，故$(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$是充分完备统计量.

而它的一一变换$(\overline{X}, S^2)$ 是这个统计量的函数，且分别是μ和σ^2的无偏估计.故它们分别是μ和σ^2的UMVUE.

Example

考虑Gamma分布$\Gamma(\alpha,\lambda)$，设$\alpha>0$为已知的，$\lambda>0$未知，求λ的UMVUE.

解：样本$\widetilde{X}=(X_1,\ldots,X_n)$的联合密度函数为

$$p(\widetilde{x};\lambda)=\frac{\lambda^{n\alpha}}{(\Gamma(\alpha))^n}\prod_{i=1}^n x_i^{\alpha-1}\exp\left\{-\lambda\sum_{i=1}^n x_i\right\}$$
$$\cdot I\{x_i>0, i=1,\cdots,n\};$$

利用指数型分布族的性质可知，当α已知时，$T=\sum_{i=1}^n X_i$是分布族$\{\Gamma(\alpha,\lambda):\lambda>0\}$的充分完备统计量.

而由Gamma分布的可加性可得$T \sim \Gamma(n\alpha, \lambda)$, 所以

$$\mathsf{E}_\lambda T = \frac{n\alpha}{\lambda},$$

即

$$\lambda = \frac{n\alpha}{\mathsf{E}_\lambda T}.$$

??????

$$
\begin{aligned}
\mathsf{E}_\lambda\left[\frac{1}{T}\right] &= \int_0^\infty \frac{1}{t}p(t;\lambda)dt = \frac{\lambda^{n\alpha}}{\Gamma(n\alpha)}\int_0^\infty t^{(n\alpha-1)-1}e^{-\lambda t}dt \\
&= \frac{\lambda\Gamma(n\alpha-1)}{\Gamma(n\alpha)} = \frac{\lambda}{n\alpha-1}.
\end{aligned}
$$

因此

$$
\mathsf{E}_\lambda\left[\frac{n\alpha-1}{T}\right] = \lambda.
$$

根据定理3.4.2,

$$
\widehat{\lambda} = \frac{n\alpha-1}{\sum_{i=1}^n X_i} = \frac{n\alpha-1}{n\overline{X}}
$$

是λ 的UMVUE. 当$\alpha=1$时, $(n-1)/(n\overline{X})$为指数分布$E(\lambda)$的参数λ的UMVUE.

另解: 我们要在T的函数$h(T)$中找无偏估计, 即

$$\mathsf{E}_\lambda h(T) = \frac{\lambda^{n\alpha}}{\Gamma(n\alpha)}\int_0^{+\infty} h(t)t^{n\alpha-1}e^{-\lambda t}dt = \lambda, \quad \forall\lambda > 0.$$

即

$$\frac{1}{\Gamma(n\alpha)}\int_0^{+\infty} h(t)t^{n\alpha-1}e^{-\lambda t}dt = \lambda^{-n\alpha+1}, \quad \forall\lambda > 0.$$

所以

$$\begin{aligned}
&\frac{1}{\Gamma(n\alpha)}\int_0^{+\infty} h(t)t^{n\alpha-1}e^{-\lambda t}dt \\
=&\frac{1}{\Gamma(n\alpha-1)}\int_0^{+\infty} t^{(n\alpha-1)-1}e^{-\lambda t}dt \\
=&\frac{n\alpha-1}{\Gamma(n\alpha)}\int_0^{+\infty} t^{(n\alpha-1)-1}e^{-\lambda t}dt, \quad \forall\lambda > 0.
\end{aligned}$$

从而

$$\frac{1}{\Gamma(n\alpha)}\int_0^\infty \left((h(t)-\frac{n\alpha-1}{t})t^{n\alpha-1}\right)e^{-\lambda t}dt=0,\quad \forall\lambda>0.$$

由Laplace变换的唯一性知

$$h(t)=\frac{n\alpha-1}{t},\quad a.e.,\quad \forall t>0.$$

故

$$\widehat{\lambda}=\frac{n\alpha-1}{\sum_{i=1}^n X_i}=\frac{n\alpha-1}{n\overline{X}}$$

是λ 的UMVUE.

Example

总体X的分布来自指数分布族$\mathcal{F}=\{E(\lambda),\lambda>0\}$, 其分布函数为$F(x;\lambda)$, 概率密度函数为$p(x;\lambda)$. 设$\widetilde{X}=\{X_1,X_2,\cdots,X_n\}$是来自该总体的简单随机样本. 对某$x_0>0$, 求$F(x_0;\lambda)$和$p(x_0;\lambda)$的UMVUE.

解: 由于X的概率密度函数为$p(x;\lambda)=\lambda\cdot\exp\{-\lambda x\}I\{x>0\}$, 故为指数型分布族. 此时样本$\widetilde{X}$的联合概率密度函数为

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

其中$Q(\lambda)=-\lambda$, 其值域为$(-\infty,0)$具有非空的内部, 故由指数型分布族的性质可知, $T=\sum_{i=1}^n X_i$是该分布族的充分完备统计量.
对于某$x_0>0$, 记$\varphi(\widetilde{X})=I\{X_1<x_0\}$, 则

$$\mathsf{E}(\varphi(\widetilde{X}))=\mathsf{P}(X_1<x_0)=F(x_0;\lambda),$$

即$\varphi(\widetilde{X})$为$F(x_0;\lambda)$的一个无偏估计. 那么由LS定理可知, $F(x_0;\lambda)$的UMVUE为$\mathsf{E}(\varphi(\widetilde{X})|T)$.

注意到对于$t>0$, $\mathsf{E}(\varphi(\widetilde{X})|T=t)=\mathsf{P}(X_1<x_0|T=t)$即为在$\{T=t\}$的条件下, X_1的条件分布函数在x_0处的函数值$F_{X_1|T}(x_0|t)$. 当$0<x_0\le t$时,

$$\begin{aligned}\mathsf{P}(X_1<x_0|T=t)&=\mathsf{P}(\frac{X_1}{T}<\frac{x_0}{t}|T=t)\\&=\mathsf{P}(\frac{X_1}{X_1+\sum_{i=2}^n X_i}<\frac{x_0}{t}|T=t).\end{aligned}$$

而$X_1\sim\Gamma(1,\lambda)$, $\sum_{i=2}^n X_i\sim\Gamma(n-1,\lambda)$且与$X_1$独立,
故$\frac{X_1}{T}=\frac{X_1}{X_1+\sum_{i=2}^n X_i}\sim\beta(1,n-1)$, 其分布与参数$\lambda$无关. 而$T$为参数$\lambda$的充分统计量, 且为完备统计量, 根据Basu定理, 知$\frac{X_1}{T}$与T独立.

因此

$$\mathsf{P}(X_1 < x_0 | T = t) = \mathsf{P}(\frac{X_1}{T} < \frac{x_0}{t} | T = t) = \mathsf{P}(\frac{X_1}{T} < \frac{x_0}{t})$$
$$= \int_0^{\frac{x_0}{t}} \frac{\Gamma(1 + (n-1))}{\Gamma(1)\Gamma(n-1)} y^{1-1}(1-y)^{(n-1)-1} dy$$
$$= \int_0^{\frac{x_0}{t}} (n-1)(1-y)^{n-2} dy = 1 - (1 - \frac{x_0}{t})^{n-1}.$$

当$x_0 > t$时，$\mathsf{P}(X_1 < x_0 | T = t) = 1$.

所以$F(x_0; \lambda)$的UMVUE为

$$\widehat{F(x_0; \lambda)} = 1 - (1 - \frac{x_0}{T})^{n-1} I\{0 < x_0 \le T\}.$$

对于$t > 0$, 记$p(x|t) := p_{X_1|T}(x|t)$为在$\{T = t\}$的条件下, X_1的条件概率密度函数. 并记$p(x_0, t)$为X_1与T的联合密度函数. 则

$$\begin{aligned} \mathsf{E}(p(x_0|T)) &= \int_0^\infty p(x_0|t)p_T(t)dt \\ &= \int_0^\infty \frac{p(x_0, t)}{p_T(t)} \cdot p_T(t)dt = p_{X_1}(x_0; \lambda). \end{aligned}$$

即$p(x_0|T)$为$p(x_0; \lambda)$的无偏估计, 注意到它还是充分完备统计量T的函数, 因此$p(x_0|T)$为$p(x_0; \lambda)$的UMVUE. 根据条件密度函数与条件分布函数的关系, 可知

$$\widehat{p(x_0; \lambda)} = p(x_0|T) = (n-1)(1 - \frac{x_0}{T})^{n-2} \cdot \frac{1}{T} I\{0 < x_0 \le T\}.$$