§2.3　最小方差无偏估计与充分统计量

一、最小方差无偏估计

对于一个无偏估计, 方差越小越有效, 方差最小者最有效. 我们有下列定义:

定义1　设$\hat{\theta}$是θ的一个无偏估计量,若对于θ的任一方差存在的无偏估计量$\tilde{\theta}$，都有

$$D(\hat{\theta}) \le D(\tilde{\theta}).$$

则称$\hat{\theta}$是θ的最小方差无偏估计(量),缩写为$MVUE$.

最小方差无偏估计存在的情况并不多，关于MVUE有如下一个判断准则.

定理2.3.1 设$X=(X_1, X_2, \cdots, X_n)$是来自某总体的一个样本, $\hat{\theta}=\hat{\theta}(X)$是$\theta$的一个无偏估计, $D(\hat{\theta})<\infty$. 如果对任意一个满足$E(L(X))=0$的$L(X)$，都有

$$\mathrm{Cov}_\theta(\hat{\theta}, L)=0, \ (\text{或者} E(\hat{\theta} L)=0)$$

则$\hat{\theta}$是θ的**最小方差无偏估计**(即MVUE).

证明 $\hat{\theta}_1(X)$是θ的任一无偏估计,记

$$L(X)=\hat{\theta}_1(X)-\hat{\theta}(X),$$

则$L(X)$的数学期望为0，由于

$$D[\hat{\theta}_1(X)]=D[\hat{\theta}(X)+L(X)]$$

$$=D[\hat{\theta}(X)]+D[L(X)]+2\mathrm{Cov}(\hat{\theta}, L)\geq D[\hat{\theta}(X)].$$

故$\hat{\theta}$ 是θ 的MVUE。

推论　参数θ的最小方差无偏估计如果存在，那么它是唯一的(在概率1的意义下).

例1 设$X=(X_1,X_2,\cdots,X_n)$是来自正态总体$N(\mu,\sigma^2)$的一个样本，已知$\overline{X}$和S^2分别是μ和σ^2的无偏估计，证明$\overline{X}$和S^2分别μ是和σ^2的MVUE.

证明 设$L(X)$满足$EL(X)=0$，则有

$$\frac{1}{(\sqrt{2\pi}\sigma)^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}L\cdot\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}dx=0 \qquad (1)$$

上式关于μ求导，并利用（1）整理得

$$\frac{1}{(\sqrt{2\pi}\sigma)^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}L\cdot(\sum_{i=1}^{n}x_i)\cdot\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}dx=0 \qquad (2)$$

故有 $E\{(\sum_{i=1}^{n}X_i)\cdot L(X)\}=0.$

因而 $E\{\overline{X}\cdot L(X)\}=0,$ $\overline{X}$是μ 的MVUE.

(2）式关于μ 求导，并利用(1), (2)得

$$\frac{1}{(\sqrt{2\pi}\sigma)^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}(\sum_{i=1}^{n}x_i)^2\cdot L\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}dx=0 \qquad (3)$$

(1）式左边关于σ^2求导，得

$$\frac{1}{(\sqrt{2\pi}\sigma)^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\sum_{i=1}^{n}(x_i-\mu)^2\cdot L\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}dx=0 \qquad (4)$$

利用 $\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}(x_i-\mu)^2-n\bar{x}^2+2n\mu\bar{x}-n\mu^2$ 得

$$\frac{1}{(\sqrt{2\pi}\sigma)^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\sum_{i=1}^{n}(x_i-\bar{x})^2\cdot L\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}dx=0 \qquad (5)$$

故有 $E\left\{S^2\cdot L(X)\right\}=0$

所以S^2是参数σ^2的MVUE.

为了寻找最小方差无偏估计，引进充分统计量和完备统计量的概念. 充分统计量是数理统计学的重要概念之一，因为它完整的简缩了样本提供的全部信息.

二、充分统计量和完备统计量

1 充分统计量

1922年英国统计学家Fisher提出了描述总体信息是否被完全提炼的概念 — 充分统计量.

样本$(X_1,\cdots,X_n)$有一个分布$F(x_1,\cdots,x_n;\theta)$，这个分布包含了样本中一切有关θ 的信息, 这是我们

一切统计推断的基础.

粗略的说, 一个统计量是充分的，如果有关估计参数的所有信息都集中在它的值中. 更确切地，当给定统计量$T(X_1, X_2, \cdots, X_n)$的值，即$T = t$，样本$(X_1, X_2, \cdots, X_n)$的条件分布$F(x_1, x_2, \cdots, x_n; \theta | t)$与$\theta$无关. 则称$T$是充分统计量.

定义2　设$X_1, X_2, \cdots, X_n$是来自总体X具有分布函数$F(x, \theta)$的一个样本, $T = T(X_1, X_2, \cdots, X_n)$为一个(一维或多维)统计量，当给定$T = t$时，若样本$(X_1, X_2, \cdots, X_n)^T$的条件分布与参数$\theta$无关，则称$T$为$\theta$的充分统计量.

充分统计量T中包含了关于θ的全部信息，如果知道了T的观察值以后, 样本的条件分布与θ无关, 因此要做关于θ的统计推断, 只需用统计量T就足够了.

例1 设总体X服从两点分布$B(1, p)$，即

$$P\{X = x\} = p^x (1-p)^{1-x}, x = 1, 0,$$

$(X_1, X_2, \cdots, X_n)^T$是来自总体$X$的一个样本. 试证

(1) $\bar{X}$是参数p的充分统计量.

(2) $S = X_1 + X_2 (n > 2)$不是参数p的充分统计量.

证　　利用定义证明$\bar{X}$是充分统计量

$$P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n \mid \bar{X} = k/n\}$$

$$= \frac{P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n, \bar{X} = k/n\}}{P\{\bar{X} = k/n\}}$$

$$= \frac{P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n, n\bar{X} = k\}}{P\{n\bar{X} = k\}}$$

$$= \frac{P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n, \sum X_i = k\}}{P\{\sum X_i = k\}}$$

$$= \frac{P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\}}{P\{\sum X_i = k\}}.$$

$$=\frac{P\{X_1=x_1\}P\{X_2=x_2\}\cdots P\{X_n=x_n\}}{P\{\sum X_i=k\}}$$

$$=\begin{cases}\dfrac{p^{\sum x_i}(1-p)^{n-\sum x_i}}{C_n^k p^k(1-p)^{n-k}}, & \sum x_i=k,\\ 0, & \text{其他}.\end{cases}=\begin{cases}\dfrac{1}{C_n^k}, & \sum x_i=k,\\ 0, & \text{其他}.\end{cases}$$

显然该条件分布与p无关，因而$\bar{X}$是p的充分统计量.

对$S=X_1+X_2(n>2)$. 由于它只用了前面两个样本观测值，显然没有包含样本中所有关于p的信息, 在给定S的取值s后, 对任意的一组$x_1,\cdots,x_n(x_1+x_2=s)$.有

$$P(X_1 = x_1, \cdots, X_n = x_n \mid S = s)$$

$$= \frac{P(X_1 = x_1, X_2 = s - x_1, X_3 = x_3, \cdots, X_n = x_n)}{P(x_1 + x_2 = s)}$$

$$= \frac{1}{C_2^s p^s (1-p)^{2-s}} p^{s + \sum_{i=3}^{n} x_i} (1-p)^{n-s-\sum_{i=3}^{n} x_i}$$

$$= \frac{1}{C_2^s} p^{\sum_{i=3}^{n} x_i} (1-p)^{n-2-\sum_{i=3}^{n} x_i}$$

这个分布依赖于未知参数p，这说明样本中关于p 的信息没有完全包含在统计量S 中．因而$S = X_1 + X_2 (n > 2)$不是参数p 的充分统计量.

注：对例1而言

$$T_1=(X_1,X_2,X_3,\cdots,X_n),\quad T_2=(X_1+X_2,X_3,\cdots,X_n)$$

$$\cdots\cdots\cdots\cdots$$

$$T_{n-1}=(\sum_{i=1}^{n-1}X_i,X_n),\qquad T_n=\sum_{i=1}^{n}X_i.$$

都是参数p的充分统计量．显然T_n最好(维数最低).

利用定义判别充分统计量比较麻烦，因而需要需求更好的判别准则.

2. 充分统计量的判别准则

定理2.3.2 （因子分解定理）(Fisher-Neyman准则)

设总体概率函数为$f(x;\theta)$，$X_1,\cdots,X_n$为样本，则$T=T(X_1,\cdots,X_n)$为充分统计量的充分必要条件是：存在两个函数$g(t,\theta)$和$h(x_1,\cdots,x_n)$，使得对任意的θ和任意一组观测值$x_1,\cdots,x_n$，有

$$f(x_1,\cdots,x_n;\theta)=g\left[T(x_1,\cdots,x_n),\theta\right]h(x_1,\cdots,x_n)$$

其中$g(T,\theta)$是通过统计量T的取值而依赖于样本$x_1,x_2,\cdots,x_n$.

证明：给出离散型随机变量下的证明．此时

$$f(x_1,\cdots,x_n;\theta)=P(X_1=x_1,\cdots,X_n=x_n;\theta)$$

必要性　设T是充分统计量，则

$$P(X_1=x_1,\cdots,X_n=x_n\,|\,T=t)$$

与θ无关，记为$h(x_1,\cdots,x_n)$，则在$T=t$下，令

$$A(t)=\left\{(x_1,\cdots,x_n):T(x_1,\cdots,x_n)=t\right\},$$

当$(x_1,\cdots,x_n)\in A(t)$时，有

$$\left\{X_1=x_1,\cdots,X_n=x_n\right\}\subset\left\{T=t\right\}$$

$$P(X_1=x_1,\cdots,X_n=x_n;\theta)$$

$$=P(X_1=x_1,\cdots,X_n=x_n,T=t;\theta)$$

$$= P(X_1 = x_1, \cdots, X_n = x_n \mid T = t)P(T = t;\theta)$$

$$= h(x_1, \cdots, x_n)g(t, \theta)$$

其中$g(t,\theta) = P(T = t;\theta)$，而

$$h(x_1, \cdots, x_n) = P(X_1 = x_1, \cdots, X_n = x_n \mid T = t)$$

与θ无关.　　必要性得证.

充分性，由于

$$P(T = t;\theta) = \sum_{\{(x_1,\cdots,x_n):T=t\}} P(X_1 = x_1, \cdots, X_n = x_n;\theta)$$

$$= \sum_{\{(x_1,\cdots,x_n):T=t\}} g(t,\theta)h(x_1, \cdots, x_n)$$

对任给$(x_1, \cdots, x_n)$和 t 满足$(x_1, \cdots, x_n) \in A(t)$，有

$$P(X_1 = x_1, \cdots, X_n = x_n \mid T = t)$$

$$= \frac{P(X_1 = x_1, \cdots, X_n = x_n, T = t; \theta)}{P(T = t; \theta)}$$

$$= \frac{P(X_1 = x_1, \cdots, X_n = x_n, T = t; \theta)}{P(T = t; \theta)}$$

$$= \frac{g(t, \theta) h(x_1, \cdots, x_n)}{g(t, \theta) \sum_{\{(y_1, \cdots, y_n): T(y_1, \cdots, y_n) = t\}} h(y_1, \cdots, y_n)}$$

$$= \frac{h(x_1, \cdots, x_n)}{\sum_{\{(y_1, \cdots, y_n): T(y_1, \cdots, y_n) = t\}} h(y_1, \cdots, y_n)}$$

该分布与θ无关，充分性得证.

说明：如果参数θ为向量时，统计量T也是随机向量．例如，$\theta=(\mu,\sigma^2)$，则相应的统计量可以为$T=(\bar{X},S_n^2)$.

在统计学中有一个基本原则：在充分统计量存在场合，任何统计推断都可以基于充分统计量进行，这可以简化统计推断的程序，通常将该原则称为充分性原则.

以下将通过几个例子来说明判别法则的应用.

例2 根据因子分解定理证明例1.

证明： $P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\}$

$$= p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i} = (1-p)^n\left[p(1-p)^{-1}\right]^{\sum_{i=1}^{n} x_i}$$

$$= (1-p)^n\left[p(1-p)^{-1}\right]^{n\bar{X}} = g(T;p)h(x_1, x_2, \cdots, x_n).$$

其中 $T = \bar{X}$, $g(T,p) = (1-p)^n\left[p(1-p)^{-1}\right]^{nT}$,

$h(x_1, x_2, \cdots, x_n) = 1$,

因而$\bar{X}$是充分统计量.

例3 设$(X_1, X_2, \cdots, X_n)^T$是来自泊松分布$P(\lambda)$的一个样本，试证$\bar{X}$是参数$\lambda$的充分统计量.

证明: $P\{X_1=x_1, X_2=x_2, \cdots, X_n=x_n\}$

$$=\left(\prod_{i=1}^{n} x_i\,!\right)^{-1} \lambda^{\sum\limits_{i=1}^{n} x_i} \mathrm{e}^{-n\lambda} = \left(\prod_{i=1}^{n} x_i\,!\right)^{-1} \lambda^{n\bar{X}} \mathrm{e}^{-n\lambda}.$$

其中 $T(x_1, x_2, \cdots, x_n)=\bar{X}, \quad g(T, \lambda)=\lambda^{nT}\mathrm{e}^{-n\lambda},$

$$h(x_1, x_2, \cdots, x_n)=\left(\prod_{i=1}^{n} x_i\,!\right)^{-1},$$

因而$\bar{X}$是充分统计量.

例4 设$X_1, X_2, \cdots, X_n$是来自正态总体$N(\mu, 1)$的一个样本，试证$\bar{X}$是参数μ的充分统计量.

证明

$$L(\mu)=\frac{1}{(\sqrt{2\pi})^n}\mathrm{e}^{-\{\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\}}$$

$$=\frac{1}{(\sqrt{2\pi})^n}\exp\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x}+\bar{x}-\mu)^2$$

$$=\frac{1}{(\sqrt{2\pi})^n}\exp\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})^2-\frac{n}{2}(\bar{x}-\mu)^2\}$$

$$=\frac{1}{(\sqrt{2\pi})^n}\exp\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})^2\}\exp\{-\frac{n}{2}(\bar{x}-\mu)^2\}$$

$$=h(x_1, x_2, \cdots, x_n)g(T, \mu).$$

其中$h(x_1, x_2, \cdots, x_n) = \exp\{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \bar{x})^2\}$,

$$T(x_1, x_2, \cdots, x_n) = \bar{x},$$

$$g(T, \mu) = \frac{1}{(\sqrt{2\pi})^n}\exp\{-\frac{n}{2}(T - \mu)^2\},$$

因而, $\bar{X}$是充分统计量.

练习题

设$X_1, X_2, \cdots, X_n$是来自正态总体$N(\mu, \sigma^2)$的一个样本, 试证$T(X_1, X_2, \cdots, X_n) = (\bar{X}, \sum_{i=1}^{n} X_i^2)^T$是参数$\theta=(\mu, \sigma^2)^T$的(联合)充分统计量.

答案: $$L(\theta)=L(\mu,\sigma^2)=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\{\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\}}$$

$$=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2+\frac{n\mu}{\sigma^2}\bar{x}-\frac{n\mu^2}{2\sigma^2}\}$$

$$=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2+\frac{n\mu}{\sigma^2}\bar{x}-\frac{n\mu^2}{2\sigma^2}\},$$

其中$T(x_1,x_2,\cdots,x_n)=(\bar{x},\sum_{i=1}^{n}x_i^2)^T$, $h(x_1,x_2,\cdots,x_n)=1$,

$g(T(x_1,x_2,\cdots,x_n),\theta)$

$$=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2+\frac{n\mu}{\sigma^2}\bar{x}-\frac{n\mu^2}{2\sigma^2}\},$$

因而，$T(X_1,X_2,\cdots,X_n)=(\bar{X},\sum_{i=1}^{n}X_i^2)^T$是充分统计量.

为了从充分统计量出发，去寻找最小方差无偏估计，我们先给出几个定理.

定理2.3.3 (Rao – Blackwell定理)

设X和Y是两个随机变量，$EX = \mu$，$D(X) > 0$.定义

$$\varphi(y) = E(X \mid Y = y),$$

则有

$$E[\varphi(Y)] = \mu, \quad D[\varphi(Y)] \le D(X),$$

其中等号成立的充要条件是X和$\varphi(Y)$概率 1 相等.

证明： 不妨设X和Y都是连续型随机变量，设$p(x, y)$, $p_Y(y)$, $h(x \mid y)$分别是X和Y的联合密度函数Y的边际密度函数和给定$Y = y$下X的条件密度函数.

于是条件期望

$$\varphi(y)=E(X\,|\,Y=y)=\int_{-\infty}^{+\infty}xh(x\,|\,y)dx=\frac{\int_{-\infty}^{+\infty}xp(x,y)dx}{p_Y(y)},$$

$$E\left[\varphi(Y)\right]=\int_{-\infty}^{+\infty}\varphi(y)p_Y(y)dy=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xp(x,y)dxdy=EX=\mu,$$

这证明了第一个结论,下证第二个结论:

$$D(X)=E(X-\mu)^2$$

$$=E\left[\left(\varphi(Y)-\mu\right)+\left(X-\varphi(Y)\right)\right]^2$$

$$=D[\varphi(Y)]+E\left(X-\varphi(Y)\right)^2$$

$$+2E\left[\left(\varphi(Y)-\mu\right)\cdot\left(X-\varphi(Y)\right)\right]$$

由于

$$\int_{-\infty}^{+\infty}\left[x-\varphi(y)\right]h(x\,|\,y)dx = E(X\,|\,Y=y)-\varphi(y)=0,$$

故上式右端第三项为

$$E\left[\left(\varphi(Y)-\mu\right)\left(X-\varphi(Y)\right)\right]$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[\varphi(y)-\mu\right]\left[x-\varphi(y)\right]p(x,y)dxdy$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[\varphi(y)-\mu\right]\left[x-\varphi(y)\right]p_Y(y)h(x\,|\,y)dxdy$$

$$=\int_{-\infty}^{+\infty}\left[\varphi(y)-\mu\right]dx\left\{\int_{-\infty}^{+\infty}\left[x-\varphi(y)\right]h(x\,|\,y)dx\right\}p_Y(y)dy=0,$$

因而 $D(X)=D\left(\varphi(Y)\right)+E\left(X-\varphi(Y)\right)^2\geq D\left(\varphi(Y)\right)$

且等号成立的充要条件是 $P\left(X-\varphi(Y)=0\right)=1$

即X与$\varphi(Y)$概率1相等.

将定理2.3.3应用在参数估计中可得：

定理2.3.4 设总体的概率函数为$p(x;\theta)$，设$X_1,\cdots,X_n$是样本，$T=T(X_1,\cdots,X_n)$是θ的充分统计量，对θ的任一无偏估计$\hat{\theta}=\hat{\theta}(X_1,\ldots,X_n)$. 令$\tilde{\theta}=E(\hat{\theta}\,|\,T)$，则$\tilde{\theta}$也是$\theta$的无偏估计，且

$$D(\tilde{\theta})\le D(\hat{\theta}).$$

证明： 由于$T=T(x_1,\cdots,x_n)$是充分统计量，故$\tilde{\theta}=E(\hat{\theta}\,|\,T)$与$\theta$无关，因此它也是一个估计（统计量），只要在定理2中取$X=\hat{\theta},Y=T$ 即可.

定理4说明：如果无偏估计不是充分统计量的函数，则求其对充分统计量的条件期望，可以得到一个新的无偏估计，该新无偏估计的方差比原估计的方差要小，从而降低了无偏估计的方差.

只要最小方差无偏估计存在.它一定是充分统计量的函数.一般地,若依赖于充分统计量的无偏估计只有一个,它一定是**MVUE**.（见完备统计量）

考虑θ的估计问题只需要在基于充分统计量的函数中进行即可，该说法对所有的统计推断问题都是正确的，这便是所谓的充分性原则.

例6　设$X_1, X_2, \cdots, X_n$ 是来自$B(1, p)$的样本，则$T=\sum_{i=1}^{n} X_i$是p的充分统计量. 为估计$\theta = p^2$，可令

$$\hat{\theta}_1 = \begin{cases} 1, & x_1 = 1, x_2 = 1 \\ 0, & \text{其它} \end{cases}$$

由于$E(\hat{\theta}_1) = P(x_1 = 1, x_2 = 1) = p^2 = \theta$. 所以$\hat{\theta}_1$是$\theta$的无偏估计.

$$D(\hat{\theta}_1) = p^2(1 - p^2)$$

这个只使用了两个观测值的估计并不好，下面我们用Rao－Blackwell定理对之加以改进：

求$\hat{\theta}_1$关于充分统计量$T=\sum_{i=1}^{n}X_i$的的条件期望，得

$$\hat{\theta}=E(\hat{\theta}_1\mid T=t)$$

$$=1\times P(\hat{\theta}_1=1\mid T=t)+0\times P(\hat{\theta}_1=0\mid T=t)$$

$$=P(\hat{\theta}_1=1\mid T=t)=C_{n-2}^{t-2}\Big/C_n^t=\frac{t(t-1)}{n(n-1)},$$

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

T的函数

$$E(\hat{\theta})=E\left(\frac{T(T-1)}{n(n-1)}\right)=\frac{E(T^2)-E(T)}{n(n-1)}$$

$$=\frac{(np)^2+np(1-p)-np}{n(n-1)}=p^2$$

无偏估计

$T \sim B(n, p)$，根据二项分布的特征函数

$$f(t) = (pe^{it} + q)^n$$

可求得

$$f'(t) = i \cdot npe^{it}(pe^{it} + q)^{n-1}$$

$$f''(t) = if'(t) + i^2 n(n-1) p^2 e^{2it}(pe^{it} + q)^{n-2}$$

$$f'''(t) = 3if''(t) + 2f'(t) +$$

$$i^3 n(n-1)(n-2) p^3 e^{3it}(pe^{it} + q)^{n-3}$$

$$f^{(4)}(t) = 3if'''(t) + 2f''(t) + i^3 n(n-1)(n-2) p^3$$

$$\left[3ie^{3it}(pe^{it} + q)^{n-3} + i(n-3) pe^{4it}(pe^{it} + q)^{n-4}\right]$$

再由$f^{(k)}(0) = i^k ET^k$　得

$$E(T^2) = np + n(n-1) p^2$$

$$E(T^3)=np+3n(n-1)p^2+n(n-1)(n-2)p^3$$

$$E(T^4)=np+7n(n-1)p^2+6n(n-1)(n-2)p^3$$

$$+n(n-1)(n-2)(n-3)p^4$$

$$E(T^2-T)^2=E(T^4-2T^3+T^2)$$

$$=2n(n-1)p^2+4n(n-1)(n-2)p^3$$

$$+n(n-1)(n-2)(n-3)p^4$$

$$D(\hat{\theta})=D\left(\frac{T(T-1)}{n(n-1)}\right)=E\left(\frac{T(T-1)}{n(n-1)}\right)^2-p^4$$

$$=p^2\times\left(\frac{2+4(n-2)p+(n-2)(n-3)p^2}{n(n-1)}-p^2\right)$$

方差变小

$$<p^2\times\left(\frac{2+4(n-2)+(n-2)(n-3)}{n(n-1)}-p^2\right)=p^2(1-p^2)$$

3 完备统计量

由上面可知:只要最小方差无偏估计存在,它一定是充分统计量的函数.如果依赖于充分统计量的无偏估计只有一个,它一定是最小方差无偏估计.那么何时它是唯一的?为此引入完备性的概念.

定义3 设总体的分布函数为$F(x,\theta)(\theta\in\Theta)$,若对于任意一个满足

$$E_\theta[g(X)]=0,\text{对一切}\theta\in\Theta,$$

的随机变量$g(X)$,总有

$$P_\theta\{g(X)=0\}=1,\text{对一切}\theta\in\Theta,$$

则称$\{F(x,\theta),\theta\in\Theta\}$是完备的分布函数族.

设$X_1, X_2, \cdots, X_n$是来自总体X具有分布函数$F(x,\theta)$的一个样本, $T = T(X_1, X_2, \cdots, X_n)$的分布函数是完备的，则称$T = T(X_1, X_2, \cdots, X_n)$为完备统计量.

完备统计量：若T的实值函数$\mathrm{g}(T)$满足：

$$E(\mathrm{g}(T)) = 0,$$

对一切$\theta \in \Theta$成立时，总有

$$P\left\{\mathrm{g}(T) = 0\right\} = 1.$$

则称T是完备统计量.

即：如果$\mathrm{g}(T)$数学期望为0，则它必概率1取值为0.

例7 设总体X服从两点分布$B(1, p)$，即

$$P\{X=x\}=p^x(1-p)^{1-x}, x=1,0.$$

$X_1, X_2, \cdots, X_n$是来自总体X的一个样本，试证$\bar{X}$是参数p的完备统计量.

证明：由于$P\{\bar{X}=\frac{k}{n}\}=P\{n\bar{X}=k\}=C_n^k p^k(1-p)^{n-k}$,

如果$E_p(g(\bar{X}))=\sum_{k=0}^{n} g(\frac{k}{n})C_n^k p^k(1-p)^{n-k}=0$，则

$$(1-p)^n\sum_{k=0}^{n} g(\frac{k}{n})C_n^k(\frac{p}{1-p})^k=0,$$

得$g(\frac{k}{n})=0$, $\quad g(\bar{X})=0.$

充分完备统计量

如果一个统计量既是充分的，又是完备的，则称为充分完备统计量.

定理2.3.5：设总体X的分布函数为$F(x,\theta),\theta\in\Theta$. $X_1,X_2,\cdots,X_n$是总体样本, $T=T(X_1,X_2,\cdots,X_n)$是充分完备统计量. 如果θ的无偏估计存在，记为$\hat{\theta}$. 则$\tilde{\theta}=E(\hat{\theta}\,|\,T)$是唯一的最小方差无偏估计.

证明：练习题.

判断一个统计量$T(X_1,X_2,\cdots,X_n)$是否为充分完备统计量比较复杂，为此介绍一类分布族，其参数的充分完备统计量容易寻找.

4、指数型分布族的概念

定义1.7 设总体X的概率函数为$f(x,\theta)$，其中θ未知且$\theta\in\Theta$，如果存在定义在参数空间上的实值函数$c(\theta)$，$d(\theta)$和实值函数$T(x),S(x)$，使得

$$f(x,\theta)=\exp\{C(\theta)T(x)+d(\theta)+h(x)\}$$

则称$\{f(x,\theta)\mid\theta\in\Theta\}$为单参数指数分布族.

例如:(1)正态分布$N(\mu_0,\sigma^2)$,已知μ_0,关于参数σ;

(2)正态分布$N(\mu,\sigma_0^2)$,已知σ_0，关于参数μ;

(3)泊松分布$P(\lambda)$关于参数λ.

但是$[0,\theta]$上的均匀分布不是参数θ的单参数分布族.

一般地，泊松分布族, 二项分布族, 正态分布族, 伽玛分布族等, 其中只要有一个参数是已知常数，则都属于单参数指数分布族.

注：定义中的函数c,d,h和T不唯一. 类似的有多参数的指数分布族的概念.

设$X_1, X_2,\cdots, X_n$是来自单参数指数分布族X的样本，则样本的联合概率函数为

$$\prod_{i=1}^{n} f(x_i,\theta)=\exp\{C(\theta)\sum_{i=1}^{n}T(x_i)+nd(\theta)+\sum_{i=1}^{n}h(x_i)\}$$

根据因子分解定理,$\sum_{i=1}^{n}T(x_i)$是θ 的充分统计量.

可以证明它也是完备统计量.

例8 设$(X_1, X_2, \cdots, X_n)^T$是来自泊松分布$P(\lambda)$的一个样本，试证$\bar{X} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$是参数$\lambda$的充分完备统计量.

证明:

$$P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\} = \lambda^{\sum\limits_{i=1}^{n} x_i} \mathrm{e}^{-n\lambda} \left(\prod_{i=1}^{n} x_i\,!\right)^{-1}$$

$$= \lambda^{n\bar{x}} \mathrm{e}^{-n\lambda} \left(\prod_{i=1}^{n} x_i\,!\right)^{-1} = \exp\left\{\bar{X} \times n \ln \lambda - n\lambda - \ln\left(\prod_{i=1}^{n} x_i\,!\right)\right\}$$

其中 $T = \bar{X}$, $C(\lambda) = n \ln \lambda$,

$$nd(\lambda) = -n\lambda, \quad \sum_{i=1}^{n} h(x_i) = \left(\prod_{i=1}^{n} x_i\,!\right)^{-1}.$$

因而$\bar{X}$是充分完备统计量

练习题　设$(X_1, X_2, \cdots, X_n)^T$是来自均匀分布$U(0,\theta)$的一个样本，$\theta > 0$。试求参数$\theta$的最小方差无偏估计.

(提示：从参数的极大似然估计出发去求解)

关于多维参数的情况有下列定理：

定理2.3.6：设总体X的分布密度为$f(x,\theta)$为指数型分布族，即样本的联合分布密度具有形式

$$\prod_{i=1}^{n} f(x_i,\theta)=\exp\{\sum_{j=1}^{m} b_j(\theta)T_j(x_1,x_2,\cdots,x_n)$$

$$+d(\theta)+h(x_1,x_2,\cdots,x_n)\}$$

其中$\theta=(\theta_1,\cdots,\theta_m)^T,\theta\in\Theta$. 如果$\Theta$包含一个$m$维矩形，且$B=(b_1(\theta),\cdots,b_n(\theta))^T$的值域包含一个$m$维开集, 则

$T=(T_1(X_1,\cdots,X_n),T_2(X_1,\cdots,X_n),\cdots T_m(X_1,\cdots,X_n))^T$

是参数$\theta=(\theta_1,\theta_2,\cdots,\theta_m)^T$的充分完备统计量.

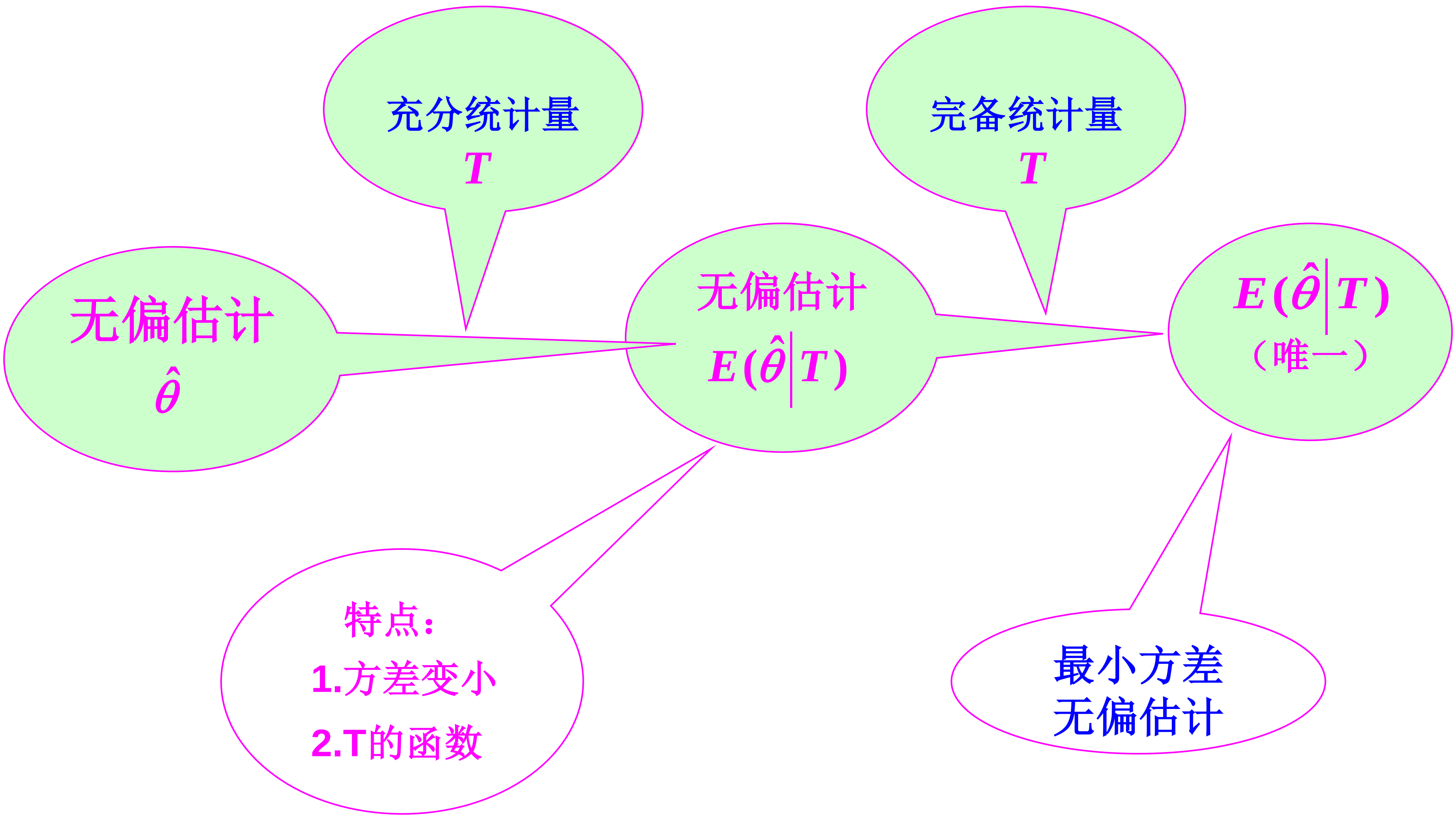
例9 设$(X_1,X_2,\cdots,X_n)^T$是来自正态总体$N(\mu,\sigma^2)$的一个样本，试证$T(X_1,X_2,\cdots,X_n)=(\bar{X},\frac{1}{n}\sum_{i=1}^{n}X_i^2)^T$是参数$\theta=(\mu,\sigma^2)^T$的联合充分完备统计量.

$(\bar{X},S^2)$也是$(\mu,\sigma^2)^T$的充分完备统计向量. 又$\bar{X},S^2$分别为μ和σ^2的无偏估计，且

$$\hat{\mu}=E(\overline{X}\mid T)=\overline{X},\quad \hat{\sigma}^2=E(S^2\mid T)=S^2.$$

由定理知：$\bar{X},S^2$分别为μ和σ^2的惟一的最小方差无偏估计.

单指数分布族具有充分完备统计量



练习题 设$(X_1, X_2, \cdots, X_n)^T$是来自均匀分布$U(0,\theta)$的一个样本，$\theta > 0$。试求参数$\theta$的最小方差无偏估计.

(提示：从参数的极大似然估计出发去求解)

证明：$T = X_{(n)}$是充分统计量，也是完备统计量，因为其密度函数为

$g(t)$

$$f_T(t) = n[F(t)]^{n-1} f(t) = \frac{n}{\theta^n} t^{n-1}. \quad 0 < t < \theta, \theta > 0.$$

如果对任一可测函数$g(t)$,有

$$E(\mathrm{g}(T)) = 0,$$

即 $$\int_0^{\theta} g(t)\frac{n}{\theta^n}t^{n-1}\mathrm{dt}=0, \qquad \theta>0.$$

所以 $$\int_0^{\theta} g(t)t^{n-1}\mathrm{dt}=0, \qquad \theta>0.$$

等式两边对θ求导数得

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

所以$\theta>0$时，$g(\theta)=0$，也就是

$$g(t)=0, \quad \theta>0.$$

所以$T=X_{(n)}$完备统计量.

由于$\dfrac{n+1}{n}X_{(n)}$是θ的无偏估计，且是$X_{(n)}$的函数。

因此是θ的最小方差无偏估计.