概率论与数理统计
第七章 参数估计

金玲飞

复旦大学计算机学院
Email: lfjin@fudan.edu.cn

2018.12.18

不变性

- 若$\hat{\theta}$是总体分布中未知参数θ的最大似然估计，则$g(\hat{\theta})$是否仍是$g(\theta)$的最大似然估计？

不变性

- 若$\hat{\theta}$是总体分布中未知参数θ的最大似然估计，则$g(\hat{\theta})$是否仍是$g(\theta)$的最大似然估计？

性质 (7.1.4 最大似然估计的不变性)

设$g(\theta)$是单射，若$\hat{\theta}$是总体分布中未知参数θ的最大似然估计，则$g(\hat{\theta})$是$g(\theta)$的最大似然估计。

表 7.1 常见分布中参数的估计量

分布	矩估计	最大似然估计	性质
两点分布 $B(1,p)$	$\hat{p}=\overline{X}$	$\hat{p}=\overline{X}$	有效估计，相合估计
二项分布 $B(N,p)$	$\hat{p}=\frac{\overline{X}}{N}$	$\hat{p}=\frac{\overline{X}}{N}$	无偏估计，相合估计
泊松分布（参数λ）	$\hat{\lambda}=\overline{X}$	$\hat{\lambda}=\overline{X}$	有效估计，相合估计
正态分布 $N(\mu,\sigma^2)$	$\hat{\mu}=\overline{X}$ $\hat{\sigma}^2=\frac{n-1}{n}S^2$ $\hat{\sigma}=\sqrt{\frac{n-1}{n}}S$	$\hat{\mu}=\overline{X}$ $\hat{\sigma}^2=\frac{n-1}{n}S^2$ $\hat{\sigma}=\sqrt{\frac{n-1}{n}}S$	$\hat{\mu}=\overline{X}$是μ的有效、相合估计；S^2为σ^2的渐近有效估计，一致最小方差无偏估计，相合估计；S为σ的相合估计，有偏估计
正态分布$N(\mu,\sigma_0^2)$（σ_0^2已知）	$\hat{\mu}=\overline{X}$	$\hat{\mu}=\overline{X}$	有效估计，相合估计
正态分布$N(\mu_0,\sigma^2)$（μ_0已知）	$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2$	$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2$	有效估计，相合估计
指数分布(参数λ)（记$\theta=\frac{1}{\lambda}$）	$\hat{\lambda}=\frac{1}{\overline{X}}$ $\hat{\theta}=\overline{X}$	$\hat{\lambda}=\frac{1}{\overline{X}}$ $\hat{\theta}=\overline{X}$	$\hat{\lambda}=\frac{1}{\overline{X}}$是$\lambda$的相合估计，$\frac{1}{\overline{X}}-\frac{1}{n\overline{X}}$是$\lambda$的一致最小方差无偏估计，相合估计；$\hat{\theta}=\overline{X}$是$\theta$的有效、相合估计
均匀分布 $U(0,\theta)$	$\hat{\theta}=2\overline{X}$	$\hat{\theta}=\max_{i\leq n}X_i$	$2\overline{X}$是θ的无偏、相合估计；$\frac{n+1}{n}\max_{i\leq n}X_i$是$\theta$的一致最小方差无偏估计，相合估计
均匀分布 $U(a,b)$	$\hat{a}=\overline{X}-\sqrt{3}\sqrt{\frac{n-1}{n}}S$ $\hat{b}=\overline{X}+\sqrt{3}\sqrt{\frac{n-1}{n}}S$	$\hat{a}=\min_{i\leq n}X_i$ $\hat{b}=\max_{i\leq n}X_i$	相合估计，有偏估计

7.2 估计量的优良性准则

估计量的优良性准则

一个“好”的估计量应该具有如下的条件：

- 无偏性：$E(\hat{\theta}) = \theta$
- 有效性：无偏估计量的方差尽可能小。
- 相合性：当样本容量越来越大时，$\hat{\theta}$应越来越接近θ的真值。

无偏性

反映估计量相对待估参数有无系统偏差。

定义 (7.2.1 无偏性)

设$\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$是未知参数$\theta$的估计量。若

$$E(\hat{\theta}) = \theta, \forall \theta \in \Theta$$

则称$\hat{\theta}$是θ的无偏估计量 *unbiased estimator*。若

$$\lim_{n \to \infty} E(\hat{\theta}) = \theta, \forall \theta \in \Theta$$

则称$\hat{\theta}$是θ的渐近无偏估计量。

无偏性

反映估计量相对待估参数有无系统偏差。

定义 (7.2.1 无偏性)

设$\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$是未知参数$\theta$的估计量。若

$$E(\hat{\theta}) = \theta, \forall \theta \in \Theta$$

则称$\hat{\theta}$是θ的无偏估计量 *unbiased estimator*。若

$$\lim_{n \to \infty} E(\hat{\theta}) = \theta, \forall \theta \in \Theta$$

则称$\hat{\theta}$是θ的渐近无偏估计量。

称$E(\hat{\theta}) - \theta$为系统偏差。

- 样本均值是总体均值的无偏估计
- 样本矩是总体矩的无偏估计
- 样本方差是总体方差的无偏估计。

- 样本均值是总体均值的无偏估计
- 样本矩是总体矩的无偏估计
- 样本方差是总体方差的无偏估计。

- 样本标准差S是总体标准差σ的无偏估计吗？

例子 (7.2.1)

设$X_1, \ldots, X_n$是来自总体$X \sim N(\mu, 1)$的样本，$x_1, \ldots, x_n$是样本观察值，求μ的最大似然估计量$\hat{\mu}$，并说明$\hat{\mu}$是无偏估计量。

例子 (7.2.1)

设$X_1, \ldots, X_n$是来自总体$X \sim N(\mu, 1)$的样本，$x_1, \ldots, x_n$是样本观察值，求μ的最大似然估计量$\hat{\mu}$，并说明$\hat{\mu}$是无偏估计量。

$$\widehat{\mu} = \overline{X}, E(\widehat{\mu}) = E(\overline{X}) = \mu$$

例子 (7.2.1)

设$X_1,\dots,X_n$是来自总体$X\sim N(\mu,1)$的样本，$x_1,\dots,x_n$是样本观察值，求μ的最大似然估计量$\hat{\mu}$，并说明$\hat{\mu}$是无偏估计量。

$$\widehat{\mu}=\overline{X}, E(\widehat{\mu})=E(\overline{X})=\mu$$

例子 (7.2.2)

设总体X服从$[0,\theta]$上的均匀分布，试证：θ的矩估计量$\hat{\theta}_1=2\overline{X}$是$\theta$的无偏估计量，$\theta$的最大似然估计量$\hat{\theta}_2=X_{(n)}$是$\theta$的渐近无偏估计量。

(1) 无偏估计一般不唯一；如均匀分布

(1) 无偏估计一般不唯一；如均匀分布

(2) 无偏估计没有不变性（除非$g(\theta)$是θ的线性函数）；

(1) 无偏估计一般不唯一；如均匀分布

(2) 无偏估计没有不变性（除非$g(\theta)$是θ的线性函数）；

(3) 只有大量重复使用同一估计量时，无偏性才有意义；

(4) 无偏估计不一定存在；

(5) 无偏性的要求可能导致不合理的要求；

例子

设$x_1, \cdots, x_n$是来自二点分布$b(1, p)$的一个样本，

(1) 寻求$p(1-p)$的无偏估计；

有效性

在无偏估计中，方差小的估计量较好。

定义 (7.2.2 有效性)

设$\hat{\theta}_1 = \hat{\theta}_1(X_1, \ldots, X_n)$和$\hat{\theta}_2 = \hat{\theta}_2(X_1, \ldots, X_n)$都是未知参数$\theta$的无偏估计量，若

$$D(\hat{\theta}_1) \leq D(\hat{\theta}_2), \quad \forall \theta \in \Theta$$

且至少存在一个$\theta \in \Theta$，使得上述不等号严格成立，则称$\hat{\theta}_1$比$\hat{\theta}_2$有效。

例子 (7.2.3)

根据例子*7.2.2*，如果总体X服从$[0,\theta]$上的均匀分布，则我们有θ的两个无偏估计量

$$\hat{\theta}_1 = 2\overline{X}$$

$$\hat{\theta}_3 = \frac{n+1}{n} X_{(n)}$$

试比较它们的有效性。

一致最小方差无偏估计

定义 (7.2.3 Uniformly minimum variance unbiased estimate)

设$\hat{\theta}(X_1,\ldots,X_n)$是未知参数$\theta$的无偏估计量，若对另外任意一个$\theta$的无偏估计$\tilde{\theta}$，都有

$$D(\hat{\theta}) \leq D(\tilde{\theta}),$$

则称$\hat{\theta}$是θ的一致最小方差无偏估计*UMVUE*。

方差能小到什么程度，有没有下界？

相合性

当样本容量$n \to \infty$

定义 (7.2.4 Consistent estimator)

设$\hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n)$是未知参数$\theta$的估计量，若对任意正数$\epsilon$，有

$$\lim_{n \to \infty} P\{|\hat{\theta}(X_1, X_2, \ldots, X_n) - \theta| \geq \epsilon\} = 0, \forall \theta \in \Theta$$

则称$\hat{\theta}$是θ的相合估计量。

相合性

当样本容量$n \to \infty$

定义 (7.2.4 Consistent estimator)

设$\hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n)$是未知参数$\theta$的估计量，若对任意正数$\epsilon$，有

$$\lim_{n\to\infty} P\{|\hat{\theta}(X_1, X_2, \ldots, X_n) - \theta| \geq \epsilon\} = 0, \forall \theta \in \Theta$$

则称$\hat{\theta}$是θ的相合估计量。

一个估计量必须具有相合性。

- 样本均值是总体均值的相合估计
- 样本矩是总体矩的相合估计
- 样本方差是总体方差的相合估计。

- 样本均值是总体均值的相合估计
- 样本矩是总体矩的相合估计
- 样本方差是总体方差的相合估计。

相合估计量不变性：若$\hat{\theta}$是θ的相合估计，对于θ的某个连续函数$g(\theta)$，则$g(\hat{\theta})$ 是$g(\theta)$ 的相合估计。

估计的优良性

例子

$X_1, X_2, \ldots, X_n$是来自总体X的样本，记μ为总体均值，$\alpha_i, i = 1, \cdots, n$为常数，且$\sum_{i=1}^{n} \alpha_i = 1$。证明

- $\sum_{i=1}^{n} \alpha_i X_i$是$\mu$的无偏估计；
- 说明在μ的所有形如$\sum_{i=1}^{n} \alpha_i X_i$的无偏估计中，$\overline{X}$最有效。

估计的优良性

例子

$X_1, X_2, \ldots, X_n$是来自总体X的样本，记μ为总体均值，$\alpha_i, i = 1, \cdots, n$为常数，且$\sum_{i=1}^n \alpha_i = 1$。证明

- $\sum_{i=1}^n \alpha_i X_i$是$\mu$的无偏估计；
- 说明在μ的所有形如$\sum_{i=1}^n \alpha_i X_i$的无偏估计中，$\overline{X}$最有效。

- $\frac{X_1+X_2}{2}$也是μ的无偏估计。

估计的优良性

例子

$X_1, X_2, \ldots, X_n$是来自总体X的样本，记μ为总体均值，$\alpha_i, i=1,\cdots,n$为常数，且$\sum_{i=1}^n \alpha_i = 1$。证明

- $\sum_{i=1}^n \alpha_i X_i$是$\mu$的无偏估计；
- 说明在μ的所有形如$\sum_{i=1}^n \alpha_i X_i$的无偏估计中，$\overline{X}$最有效。

- $\frac{X_1+X_2}{2}$也是μ的无偏估计。
- $\overline{X}$ 是 Best Linear Unbiased Estimator。

7.3 区间估计

区间估计 Interval estimate

- 对未知参数θ，除了求它的点估计外，我们希望估出一个范围，并知道这个范围包含未知参数θ的可靠程度。
- 区间估计：找两个统计量$\underline{\theta} = \underline{\theta}(X_1, \ldots, X_n)$ 和$\overline{\theta} = \overline{\theta}(X_1, \ldots, X_n)$来把未知参数$\theta$估计在$(\underline{\theta}, \overline{\theta})$内。由于样本的随机性，$(\underline{\theta}, \overline{\theta})$盖住$\theta$的可能性并不确定。一般的，要求概率尽可能大。

置信区间

定义 (7.3.1 confidence interval)

设总体X的分布函数为$F(x;\theta)$，未知参数$\theta \in \Theta \subset \mathbf{R}$，$X_1, X_2, \ldots, X_n$是来自$X$的样本，$\alpha$是给定值($0 < \alpha < 1$)。若两个统计量 $\underline{\theta} = \underline{\theta}(X_1, \ldots, X_n)$ 和 $\overline{\theta} = \overline{\theta}(X_1, \ldots, X_n)$ 满足

$$P\{\underline{\theta} < \theta < \overline{\theta}\} = 1 - \alpha, \forall \theta \in \Theta$$

则称随机区间$(\underline{\theta}, \overline{\theta})$是$\theta$的置信度为$1-\alpha$的置信区间，分别称$\overline{\theta}$ 和$\underline{\theta}$为置信上限 *confidence upper limit*和置信下限 *confidence lower limit*。

- 置信区间又称区间估计，置信度又称置信水平；
- 置信区间的长度刻画了区间估计的精度；
- 置信度表达了置信区间包含未知参数真值的可靠度。

- 置信区间又称区间估计，置信度又称置信水平；
- 置信区间的长度刻画了区间估计的精度；
- 置信度表达了置信区间包含未知参数真值的可靠度。

置信区间好坏的两个要素：

- 精度：区间的长度越短，精度越高；
- 置信度 $1-\alpha$越大越好。

- 置信区间又称区间估计，置信度又称置信水平；
- 置信区间的长度刻画了区间估计的精度；
- 置信度表达了置信区间包含未知参数真值的可靠度。

置信区间好坏的两个要素：

- 精度：区间的长度越短，精度越高；
- 置信度 $1-\alpha$越大越好。

一般的，我们先确定置信度，然后再寻找长度最短的一个置信区间。

如何选取最短的一个置信区间？

枢轴量法

(1) 构造一个样本$X_1, X_2, \ldots, X_n$和θ的函数 $Z = Z(X_1, X_2, \ldots, X_n; \theta)$。$Z$的分布已知，不依赖于$\theta$。一般称这样的函数为枢轴量 pivotal variable。

(2) 对于给定的置信度$1-\alpha$，适当的选取两个常数a, b使得

$$P\{a < Z(X_1, X_2, \ldots, X_n; \theta) < b\} = 1-\alpha, \quad \forall \theta \in \Theta$$

(3) 利用不等式变形，求得未知参数θ的置信区间。 若可以求得$P\{\underline{\theta} < \theta < \overline{\theta}\} = 1-\alpha$，则$(\underline{\theta}, \overline{\theta})$就是$\theta$的置信度为$1-\alpha$的置信区间。

单侧置信区间

有时只关心上限或下限。

定义 (7.3.2)

对给定的$\alpha(0<\alpha<1)$，若统计量$\underline{\theta}=\underline{\theta}(X_1,\ldots,X_n)$满足

$$P\{\underline{\theta}<\theta\}=1-\alpha,\forall\theta\in\Theta$$

则称随机区间$(\underline{\theta},+\infty)$是$\theta$的置信度为$1-\alpha$的单侧置信区间，称$\underline{\theta}$ 为置信下限。 若统计量$\overline{\theta}=\overline{\theta}(X_1,\ldots,X_n)$满足

$$P\{\theta<\overline{\theta}\}=1-\alpha,\forall\theta\in\Theta$$

则称随机区间$(-\infty,\overline{\theta})$是$\theta$的置信度为$1-\alpha$的单侧置信区间，称$\underline{\theta}$ 为置信上限。

7.4 正态总体均值与方差的区间估计

单个正态总体均值的区间估计

1. 当σ^2已知时，μ的区间估计

设$X_1, X_2, \ldots, X_n$是来自正态总体$N(\mu, \sigma^2)$的样本。

单个正态总体均值的区间估计

1. 当σ^2已知时，μ的区间估计

设$X_1, X_2, \ldots, X_n$是来自正态总体$N(\mu, \sigma^2)$的样本。

1. 选取枢轴量$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$，已知$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0, 1)$。
2. 对给定的α，

$$P\left\{\left|\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right| < z_{\frac{\alpha}{2}}\right\} = 1-\alpha$$

3. 由此得到μ的置信度为$1-\alpha$的置信区间为

$$\left(\overline{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}, \overline{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right)$$

注：从$P\left\{-z_{\frac{\alpha}{3}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < z_{\frac{2\alpha}{3}}\right\} = 1-\alpha$ 也可得到置信区间，但这个置信区间比$\left(\overline{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}, \overline{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right)$ 长。

注：从$P\left\{-z_{\frac{\alpha}{3}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < z_{\frac{2\alpha}{3}}\right\} = 1-\alpha$ 也可得到置信区间，但这个置信区间比$\left(\overline{X}-\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}, \overline{X}+\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}\right)$ 长。

注：对于标准正态分布和t分布这种对称分布，取对称的分位点所得的区间最短。

2. 当σ^2未知时，μ的区间估计

设$X_1, X_2, \ldots, X_n$是来自正态总体$N(\mu, \sigma^2)$的样本。

单个正态总体均值的区间估计

2. 当σ^2未知时，μ的区间估计

设$X_1, X_2, \ldots, X_n$是来自正态总体$N(\mu, \sigma^2)$的样本。

1. 选取枢轴量$\frac{\overline{X}-\mu}{S/\sqrt{n}}$，已知$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$，有

$$P\left\{\left|\frac{\overline{X}-\mu}{S/\sqrt{n}}\right| < t_{\frac{\alpha}{2}}(n-1)\right\} = 1-\alpha$$

2. 由此得到μ的置信度为$1-\alpha$的置信区间为

$$\left(\overline{X} \pm \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)\right)$$

单个正态总体均值的区间估计——例子

例子 (7.4.1)

设某种清漆的9个样品，其干燥时间（单位：h）分别为

$$6.0 \quad 5.7 \quad 5.8 \quad 6.5 \quad 7.0 \quad 6.3 \quad 5.6 \quad 6.1 \quad 5.0$$

设干燥时间总体服从正态分布$N(\mu, \sigma^2)$。分别就下列两种情形求μ的置信度为0.95的置信区间：

1. $\sigma = 0.6(h)$
2. σ未知

单个正态总体均值的区间估计——例子

例子 (7.4.1)

设某种清漆的9个样品，其干燥时间（单位：h）分别为

$$6.0 \quad 5.7 \quad 5.8 \quad 6.5 \quad 7.0 \quad 6.3 \quad 5.6 \quad 6.1 \quad 5.0$$

设干燥时间总体服从正态分布$N(\mu, \sigma^2)$。分别就下列两种情形求μ的置信度为0.95的置信区间：

1. $\sigma = 0.6(h)$
2. σ未知

$$\overline{X} = 6, s = 0.577$$