Probability and Statistics
Tutorial 11

Siyi Wang

Southern University of Science and Technology

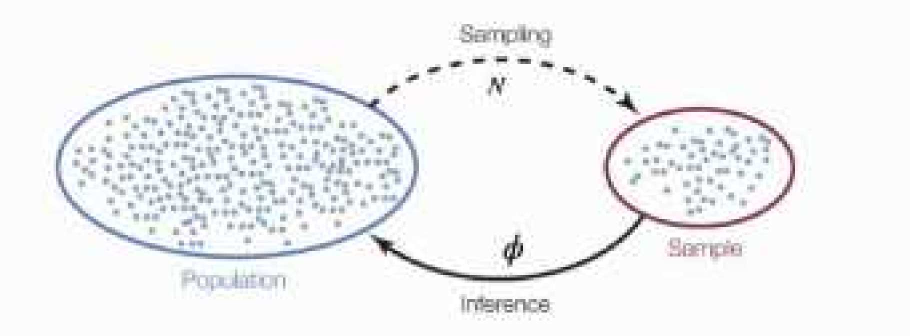
11951002@mail.sustech.edu.cn

December 5, 2020

Outline

Review

1. Population and Samples



(1). Samples $X_1, .., X_N$ are i.i.d random variables. (2). Sample Size: N.

2. Some Statistics

与均值和方差有什么不同?

样本均值 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 为什么不是 $\frac{1}{n}$（后续说明）

样本方差 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$

样本标准差 $S = \sqrt{S^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$

与第4章介绍的矩有什么不同?

样本k阶矩 $A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \quad (k = 1, 2, \cdots)$

样本k阶中心矩 $B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k \quad (k = 1, 2, \cdots)$

将样本 $X_1, X_2, \ldots, X_n$ 由小到大排列为 $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$

顺序统计量 $(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$

极小值 $X_{(1)} = \min\{X_1, X_2, \cdots, X_n\}$

极大值 $X_{(n)} = \max\{X_1, X_2, \cdots, X_n\}$

3. Sample Mean and Sample Variance

> 设总体 X 的均值和方差
>
> $$E(X)\triangleq\mu,\ D(X)\triangleq\sigma^2$$
>
> 都存在. $X_1,X_2,\cdots,X_n$是来自总体 X 的样本，则
>
> $$E(\bar{X})=\mu,\quad D(\bar{X})=\frac{\sigma^2}{n},\quad E(S^2)=\sigma^2$$

4. Distributions Derived From Normal Distribution

统计量的构造	抽样分布密度函数	期望	方差
$\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2$	$p(y) = \dfrac{1}{\Gamma\left(\frac{n}{2}\right) 2^{n/2}} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} \quad (y>0)$	n	$2n$
$F = \dfrac{(Y_1^2 + \cdots + Y_m^2)/m}{(X_1^2 + \cdots + X_n^2)/n}$	$p(y) = \dfrac{\Gamma\left(\frac{m+n}{2}\right)\left(\frac{m}{n}\right)^{m/2}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} y^{\frac{m}{2}-1}\left(1+\frac{m}{n}y\right)^{-\frac{m+n}{2}}$	$\dfrac{n}{n-2}$ $(n>2)$	$\dfrac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ $(n>4)$
$t = \dfrac{X_0}{\sqrt{(X_1^2 + \cdots + X_n^2)/n}}$	$p(y) = \dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{y^2}{n}\right)^{-\frac{n+1}{2}}$ $(-\infty < y < +\infty)$	0 $(n>1)$	$\dfrac{n}{n-2}$ $(n>2)$

4. Distributions Derived From Normal Distribution

(1). $\chi^2(n)$

χ^2-分布的可加性

设 $\chi_1^2 \sim \chi^2(n_1), \chi_2^2 \sim \chi^2(n_2)$, 且 χ_1^2, χ_2^2 相互独立，则

$$\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$$

χ^2-分布的数字特征

设 $\chi^2 \sim \chi^2(n)$, 则 $E(\chi^2) = n$, $D(\chi^2) = 2n$

$\chi^2(n)$ 的 α 分位点记为 $\chi_\alpha^2(n)$

Review

4. Distributions Derived From Normal Distribution

(2). $t(n)$

$t-$ 分布的密度函数及图形

$$f(x)=\frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\,\Gamma(n/2)}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2},\quad -\infty<x<\infty$$

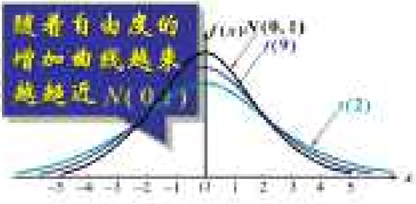
易知：$f(-x)=f(x)$

$f'(x)>0\ (x<0)$

$f'(x)<0\ (x>0)$

$\lim\limits_{x\to-\infty} f(x)=0$

$\lim\limits_{x\to+\infty} f(x)=0$



$t(n)$ 的 α 分位点记为 $t_\alpha(n)$

4. Distributions Derived From Normal Distribution

(3). $F(n_1, n_2)$

F – 分布的重要性质

若 $F \sim F(n_1, n_2)$，则 $\dfrac{1}{F} \sim F(n_2, n_1)$

若 $F \sim F(n_1, n_2)$ 则 $F^{-1} \sim F(n_2, n_1)$

故 $F_\alpha(n_1, n_2) = \dfrac{1}{F_{1-\alpha}(n_2, n_1)}$

"三反"公式

$F(n_1, n_2)$ 的 α 分位点记为 $F_\alpha(n_1, n_2)$

4. Distributions Derived From Normal Distribution
(4). Five Important Theorems

定理一 设$X_1, X_2, \cdots, X_n$是来自总体$X \sim N(\mu, \sigma^2)$的样本，则

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

4. Distributions Derived From Normal Distribution
(4). Five Important Theorems

定理二 设 $X_1, X_2, \cdots, X_n$ 是总体 $X \sim N(\mu, \sigma^2)$ 的样本，$\bar{X}, S^2$ 分别为样本均值和样本方差，则有

① $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

② $\bar{X}, S^2$ 相互独立

4. Distributions Derived From Normal Distribution
(4). Five Important Theorems

定理三 设 $X_1, X_2, \cdots, X_n$ 是总体 $X \sim N(\mu, \sigma^2)$ 的样本，$\bar{X}, S^2$ 分别为样本均值和样本方差，则有

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

4. Distributions Derived From Normal Distribution
(4). Five Important Theorems

定理四 设$X_1, X_2, \cdots, X_{n_1}$是总体$X \sim N(\mu_1, \sigma_1^2)$的样本；$Y_1, Y_2, \cdots, Y_{n_2}$是总体$Y \sim N(\mu_2, \sigma_2^2)$ 的样本，且两样本相互独立，两样本均值和样本方差分别为$\bar{X}, \bar{Y}, S_1^2, S_2^2$.则

$$\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$$

4. Distributions Derived From Normal Distribution

(4). Five Important Theorems

定理五 设$X_1, X_2, \cdots, X_{n_1}$是总体$X \sim N(\mu_1, \sigma^2)$的样本；$Y_1, Y_2, \cdots, Y_{n_2}$是总体$Y \sim N(\mu_2, \sigma^2)$的样本，且两样本相互独立，两样本均值和样本方差分别为$\bar{X}, \bar{Y}, S_1^2, S_2^2$，则

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t(n_1+n_2-2)$$

5. Point Estimation
(1). Method of Moment (MoM)

8.4 矩方法

概率律的 k 阶矩定义为

$$\mu_k = E[X^k]$$

其中 X 是服从概率律的随机变量（当然，仅当期望存在时才有定义）。如果 $X_1, X_2, \cdots, X_n$ 是取自总体的 i.i.d. 随机变量，k 阶样本矩(sample moment) 定义为

$$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

我们将 $\hat{\mu}_k$ 视为 μ_k 的估计。矩方法首先利用最低阶矩表示估计参数，然后将样本矩代入表达式，最后得到参数的估计量。

Review

5. Point Estimation

(1). Method of Moment (MoM)

例如，假设我们希望估计两个参数 θ_1 和 θ_2. 如果 θ_1 和 θ_2 可以用前两阶矩表示成

$$\theta_1 = f_1(\mu_1, \mu_2)$$
$$\theta_2 = f_2(\mu_1, \mu_2)$$

那么矩估计方法是

$$\hat{\theta}_1 = f_1(\hat{\mu}_1, \hat{\mu}_2)$$
$$\hat{\theta}_2 = f_2(\hat{\mu}_1, \hat{\mu}_2)$$

矩估计方法包括三个基本步骤：

1. 计算低阶矩，找出利用参数表示的矩表达式. 通常，需要的低阶矩个数等同于参数个数.
2. 求解上一步的表达式，得到由矩表示的参数表达式.
3. 将样本矩代入第二步的表达式，得到基于样本矩的参数估计

5. Point Estimation

(1). Method of Moment (MoM)

例 8.4.2（正态分布） 正态分布的一阶矩和二阶矩是

$$\mu_1 = E(X) = \mu$$
$$\mu_2 = E(X^2) = \mu^2 + \sigma^2$$

因此，

$$\mu = \mu_1$$
$$\sigma^2 = \mu_2 - \mu_1^2$$

由样本矩得到的 μ 和 σ^2 的相应估计是

$$\hat{\mu} = \overline{X}$$
$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

5. Point Estimation

(1). Method of Moment (MoM)

Example 7.2.2 (Binomial method of moments) Let $X_1, \dots, X_n$ be iid binomial(k,p), that is,

$$P(X_i = x|k,p) = \binom{k}{x} p^x (1-p)^{k-x}, \quad x = 0, 1, \dots, k.$$

Here we assume that both k and p are unknown and we desire point estimators for both parameters. (This somewhat unusual application of the binomial model has been used to estimate crime rates for crimes that are known to have many unreported occurrences. For such a crime, both the true reporting rate, p, and the total number of occurrences, k, are unknown.)

Equating the first two sample moments to those of the population yields the system of equations

$$\bar{X} = kp,$$
$$\frac{1}{n}\sum X_i^2 = kp(1-p) + k^2p^2,$$

which now must be solved for k and p. After a little algebra, we obtain the method of moments estimators

$$\tilde{k} = \frac{\bar{X}^2}{\bar{X} - (1/n)\sum (X_i - \bar{X})^2}$$

and

$$\tilde{p} = \frac{\bar{X}}{\tilde{k}}.$$

Review

5. Point Estimation

(2). Maximum Likelihood Estimate (MLE)

a. Likelihood Function and Log Likelihood Function

假设随机变量 $X_1, \cdots, X_n$ 具有联合密度或频率函数 $f(x_1, x_2, \cdots, x_n|\theta)$，给定观测值 $X_i = x_i$，其中 $i = 1, \cdots, n$，作为 $x_1, x_2, \cdots, x_n$ 的函数，θ 的似然定义为

$$\mathrm{lik}(\theta) = f(x_1, x_2, \cdots, x_n|\theta)$$

注意，我们考虑作为 θ 函数的联合密度，而不是 x_i 的函数. 如果分布是离散的，f 是频率函数，似然函数给出了观测到给定数据的概率，它是参数 θ 的函数. θ 的最大似然估计(maximum likelihood estimate, mle) 是使得似然达到最大的 θ 值 —— 也就是说，观测数据 "最有可能" 出现.

如果 X_i 假设成 i.i.d 的，它们的联合密度是边际密度的乘积，似然是

$$\mathrm{lik}(\theta) = \prod_{i=1}^{n} f(X_i|\theta)$$

有时候更容易最大化似然的自然对数，而不是最大化似然本身 (由于对数是单调函数，因此二者等价)，对于 i.i.d 样本，对数似然(log likelihood) 是

$$l(\theta) = \sum_{i=1}^{n} \log[f(X_i|\theta)]$$

(在本书中，"log" 总是表示自然对数.)

Review

5. Point Estimation

(2). Maximum Likelihood Estimate (MLE)

例 8.5.2 (正态分布) 如果 $X_1, X_2, \cdots, X_n$ 是 i.i.d 的 $N(\mu, \sigma^2)$，那么联合密度是边际密度的乘积：

图 8.5 石棉数据 λ 的对数似然函数图

$$f(x_1, x_2, \cdots, x_n | \mu, \sigma) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\frac{x_i - \mu}{\sigma}\right]^2\right)$$

将其视为 μ 和 σ 的函数，这是似然函数。因此，对数似然是

$$l(\mu, \sigma) = -n \log \sigma - \frac{n}{2} \log 2\pi - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2$$

关于 μ 和 σ 求偏导

$$\frac{\partial l}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu)$$

$$\frac{\partial l}{\partial \sigma} = -\frac{n}{\sigma} + \sigma^{-3} \sum_{i=1}^{n} (X_i - \mu)^2$$

令第一个偏导等于零，求解最大似然估计，得到

$$\hat{\mu} = \overline{X}$$

令第二个偏导等于零，代入 μ 的最大似然估计，我们得到 σ 的最大似然估计

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2}$$

同样，这里的估计和抽样分布与矩方法得到的相同。

Review

5. Point Estimation
(2). Maximum Likelihood Estimate (MLE)

1. 最大似然估计　利用"最大似然原理"获得的估计，它只能在总体概率函数形式已知的情况下使用，具体步骤如下：设总体的概率函数为 $p(x;\theta)$，$\theta\in\Theta$，$x_1,\cdots,x_n$ 是来自该总体的样本。

- 写出似然函数 $L(\theta)=L(\theta;x_1,\cdots,x_n)=p(x_1;\theta)\cdot p(x_2;\theta)\cdot\cdots\cdot p(x_n;\theta)$；
- 使似然函数 $L(\theta)$ 达到最大的统计量 $\hat\theta=\hat\theta(x_1,\cdots,x_n)$ 称为 θ 的最大似然估计，简称 MLE，即 $\hat\theta$ 满足 $L(\hat\theta)=\max_{\theta\in\Theta}L(\theta)$.

注意：使得对数似然函数 $\ln L(\theta)$ 最大的 $\hat\theta$ 也使似然函数 $L(\theta)$ 最大，寻找最大值可以从定义出发，也可以对 $l(\theta)=\ln L(\theta)$ 使用微分法，后者更为常用。

5. Point Estimation
(1). Method of Moment (MoM)

Example 7.2.7 (Bernoulli MLE) Let $X_1, \ldots, X_n$ be iid Bernoulli(p). Then the likelihood function is

$$L(p|\mathbf{x}) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^y(1-p)^{n-y},$$

where $y = \sum x_i$. While this function is not all that hard to differentiate, it is much easier to differentiate the log likelihood

$$\log L(p|\mathbf{x}) = y \log p + (n-y)\log(1-p).$$

If $0 < y < n$, differentiating $\log L(p|\mathbf{x})$ and setting the result equal to 0 give the solution, $\hat{p} = y/n$. It is also straightforward to verify that y/n is the global maximum in this case. If $y = 0$ or $y = n$, then

$$\log L(p|\mathbf{x}) = \begin{cases} n\log(1-p) & \text{if } y = 0 \\ n\log p & \text{if } y = n. \end{cases}$$

In either case $\log L(p|\mathbf{x})$ is a monotone function of p, and it is again straightforward to verify that $\hat{p} = y/n$ in each case. Thus, we have shown that $\sum X_i/n$ is the MLE of p. ‖

5. Point Estimation
(3). Least Square Estimate (LSE)
Further Reading.

Homework

3. 令 $\overline{X}$ 是来自于均值为 0、方差为 1 的 16 个独立正态随机变量的样本平均值. 确定 c, 使其满足

$$P\{|\overline{X}| < c\} = 0.5$$

Homework

Solution

3\. Solution. $\bar{X} \sim N(0, \frac{1}{16})$

$$\frac{1}{2} = \mathbb{P}(|\bar{X}| > c)$$

$$\frac{1}{4} = \mathbb{P}\left(\frac{\bar{X}}{\frac{1}{4}} > 4c\right)$$

$$= 1 - \Phi(4c)$$

$$\Phi(4c) = \frac{3}{4}, \quad c = \frac{1}{4}\Phi^{-1}\left(\frac{3}{4}\right) \approx 0.17$$

Homework

6. 证明：如果 $T \sim t_n$，那么 $T^2 \sim F_{1,n}$.

Solution

6. Proof. If $t \sim t(n)$, then $\exists\, X \sim N(0,1)$, $Y \sim \chi^2(n)$, X, Y independent such that $t = \dfrac{X}{\sqrt{Y/n}}$.

Then, $t^2 = \dfrac{X^2}{Y/n}$.

Since $X^2 \sim \chi^2(1)$, then $t^2 \sim F(1,n)$. $\square$

Homework

8、证明：如果 X 和 Y 是 $\lambda = 1$ 的独立指数随机变量，那么 X/Y 服从 F 分布，同时指出自由度。

Homework

Solution

Proof. First, we prove the following property (*).

(*) If $X, Y \sim N(0, \frac{1}{2})$ and independent, then

$$X^2+Y^2 \sim Exp(1)$$

Pf of (*): Let $\begin{cases} R = X^2+Y^2 \\ T = \arctan(\frac{Y}{X}) \end{cases}$

Then, $\begin{cases} X = \sqrt{R}\cos(T) \\ Y = \sqrt{R}\sin(T) \end{cases}$

$$f_{R,T}(r,t) = \frac{1}{\pi} \cdot \exp(-R) \cdot 1_{\{-\frac{\pi}{2} < t < \frac{\pi}{2}\}} \cdot 1_{\{R>0\}}$$

Hence, $f_R(r) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(r,t)\,dt = e^{-R} 1_{\{R>0\}}$

That is, $R = X^2+Y^2 \sim Exp(1)$. #

Solution

Based on (*), we have if $X, Y \sim Exp(2)$ and independent, then

$$2X \sim \chi^2(2), \quad 2Y \sim \chi^2(2)$$

Hence, $\dfrac{X}{Y} = \dfrac{\frac{2X}{2}}{\frac{2Y}{2}} \sim F(2,2)$. $\square$

Homework

1. 从总体 $N\left(240, 20^2\right)$ 中独立地进行两次抽样, 容量分别为 36 和 49, 那么这两个样本均值之差的绝对值不超过 10 的概率是多少?

Solution

1. Solution. $\bar{X} \sim \mathcal{N}(240, (\frac{10}{6})^2)$

$$\bar{Y} \sim \mathcal{N}(240, (\frac{20}{7})^2)$$

$$\bar{X} - \bar{Y} \sim \mathcal{N}(0, (\frac{20}{6})^2 + (\frac{20}{7})^2)$$

$$\mathbb{P}(|\bar{X} - \bar{Y}| \leq 10) = 2\mathbb{P}(\bar{X} - \bar{Y} \leq 10) - 1$$

$$= 2\mathbb{P}\left(\frac{(\bar{X} - \bar{Y})}{\sqrt{(\frac{20}{6})^2 + (\frac{20}{7})^2}} \leq \frac{10}{\sqrt{(\frac{20}{6})^2 + (\frac{20}{7})^2}}\right) - 1$$

$$= 2\Phi\left(\frac{21}{\sqrt{85}}\right) - 1 \approx 0.9774. \square$$

2. 设 $X_1, X_2, \cdots, X_{10}$ 为 $N(0, 0.3^2)$ 的样本，求 C 使 $P\left\{\sum_{i=1}^{10} X_i^2 \le C\right\} = 0.95$，

Solution

2. Solution.

$$0.95 = \mathbb{P}\left(\sum_{i=1}^{10} X_i^2 \le C \right) = \mathbb{P}\left(\sum_{i=1}^{10} \left(\frac{X_i}{0.3}\right)^2 \le \frac{C}{(0.3)^2} \right)$$

i.e. $C = \chi^2_{0.95}(10) \times (0.3)^2 \approx 1.6479$. ☐

3. 设 X_1, X_2 是来自 $N(0,\sigma^2)$ 的样本.

(1) 求 $\dfrac{(X_1-X_2)^2}{(X_1+X_2)^2}$ 的分布;

(2) 求常数 k, 使 $P\left\{\dfrac{(X_1+X_2)^2}{(X_1+X_2)^2+(X_1-X_2)^2}>k\right\}=0.10$。

Solution

3\. Solution $\begin{cases} U = X_1 + X_2 \\ V = X_1 - X_2. \end{cases}$

$$f_{U,V}(u,v) = \frac{1}{2\pi\sigma^2} \cdot \frac{1}{2} \cdot \exp\left(-\frac{1}{2\sigma^2}\left(\frac{1}{2}u^2 + \frac{1}{2}v^2\right)\right)$$

$$= \left(\frac{1}{\sqrt{4\pi\sigma^2}} \exp\left(-\frac{u^2}{4\sigma^2}\right)\right)\left(\frac{1}{\sqrt{4\pi\sigma^2}} \exp\left(-\frac{v^2}{4\sigma^2}\right)\right)$$

Hence, X_1+X_2 and X_1-X_2 are independent.

$X_1+X_2 \sim \mathcal{N}(0, 2\sigma^2)$, $X_1-X_2 \sim \mathcal{N}(0, 2\sigma^2)$.

Homework

Solution

(1) $$\frac{(X_1-X_2)^2}{(X_1+X_2)^2} = \frac{\left(\frac{X_1-X_2}{\sqrt{2\sigma^2}}\right)^2}{\left(\frac{X_1+X_2}{\sqrt{2\sigma^2}}\right)^2} \sim F(1,1).$$

(2) $$0.1 = \mathbb{P}\left(\frac{U^2}{U^2+V^2} > k\right) = \mathbb{P}\left(\frac{V^2}{U^2} < \frac{1}{k} - 1\right)$$

Then, $\frac{1}{k} - 1 = F_{0.1}(1,1) = \frac{1}{F_{0.9}(1,1)} \approx 0.9755.$ $\square$

Homework

4、设 $X_1, X_2, \ldots, X_n$ 是来自 $N(\mu, 16)$的样本，n 多大时才能使得 $P(|\overline{X}-\mu|<1) \geq 0.95$ 成立。

Homework

Solution

4 Solution. $\bar{X} - \mu \sim N(0, \frac{16}{n})$

$$\mathbb{P}(|\bar{X} - \mu| < 1) = 2\,\mathbb{P}(\bar{X} - \mu < 1) - 1$$

$$= 2\,\mathbb{P}\left(\frac{\bar{X} - \mu}{\frac{4}{\sqrt{n}}} < \frac{\sqrt{n}}{4}\right) - 1 = 2\Phi\left(\frac{\sqrt{n}}{4}\right) - 1$$

$$2\Phi\left(\frac{\sqrt{n}}{4}\right) - 1 \geq 0.95, \quad \frac{\sqrt{n}}{4} \geq \Phi^{-1}(0.975)$$

$$n_{min} = 62. \quad \square$$

Homework

5. 设 $X_1,X_2,\dots,X_n,X_{n+1}$ 为 $X\sim N(\mu,\sigma^2)$ 的样本，$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$，$S_n^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2$，试求常数 c 使得 $t_c=c\frac{X_{n+1}-\overline{X}_n}{S_n}$ 服从 t 分布，并指出分布的自由度。

Solution

5. Solution Since S_n is independent of $\bar{X}_n$ and X_{n+1}, then S_n is independent of $X_{n+1}-\bar{X}_n$.

Also, we have $\frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$,

At the same time, we have

$$X_{n+1}-\bar{X}_n \sim N(0, \frac{n+1}{n}\sigma^2)$$

Then, $c \cdot \frac{\sqrt{\frac{n+1}{n}\sigma^2}}{\sqrt{\sigma^2}} = 1$, i.e. $c = \sqrt{\frac{n}{n+1}}$

$$t_c \sim t(n-1). \square$$

Homework

8. 假设 X 是离散随机变量，具有 $P(X=1)=\theta$ 和 $P(X=2)=1-\theta$. 取 X 的三个独立观测，$x_1=1, x_2=2, x_3=2$.
   a. 计算 θ 的矩方法估计.
   b. 似然函数是什么.
   c. θ 的最大似然估计是什么？
   d. 如果 Θ 的先验分布是 $[0,1]$ 上的均匀分布，后验密度是什么？

Solution

5. Solution. (a) $\mathbb{E}X = \theta + 2(1-\theta) = 2 - \theta$.

$$2 - \hat{\theta} = \frac{1}{3}(1+2+2), \text{ i.e. } \hat{\theta} = \frac{1}{3}.$$

(b) $L(\theta; x_1, x_2, x_3) = \mathbb{P}(X_1 = x_1, X_2 = x_2, X_3 = x_3)$

$$= \prod_{i=1}^{3} \theta^{2-x_i}(1-\theta)^{x_i-1} = \theta^{6-\sum_{i=1}^{3} x_i}(1-\theta)^{\sum_{i=1}^{3} x_i - 3}$$

(c) $$\hat{\theta} = \arg\max_{\theta \in [0,1]} \mathbb{P}_\theta(X_1 = 1, X_2 = 2, X_3 = 2)$$

$$= \arg\max_{\theta \in [0,1]} \theta(1-\theta)^2$$

$$= \frac{1}{3}.$$

Homework

1. 设总体 X 具有密度函数

$$f(x;\theta)=\begin{cases}\dfrac{2}{\theta^2}(\theta-x), & 0<x<\theta\\ 0, & \text{其它}\end{cases}$$

$X_1,X_2,\cdots,X_n$ 是其样本，求 θ 的矩估计。

Homework

Solution

1. Solution. $\mathbb{E}X = \int_0^\theta \frac{2}{\theta^2}(\theta - x)x\,dx = \frac{\theta}{3}$

$\frac{\hat{\theta}}{3} = \overline{X}_n$, i.e. $\hat{\theta} = 3\overline{X}_n$. $\square$

Homework

2. 设总体 X 的密度函数为 $f(x;\theta)$, $X_1, X_2, \cdots, X_n$ 为其样本，求下列情况下 θ 的最大似然估计.

(1) $f(x;\theta)=\begin{cases} \dfrac{\theta^x}{x!}e^{-\theta}, & x=0,1,2\cdots \\ 0, & \text{其它} \end{cases}$ $(\theta>0)$

(2) $f(x;\theta)=\begin{cases} \theta\alpha x^{\alpha-1}e^{-\theta x^{\alpha}}, & x>0 \\ 0, & \text{其它} \end{cases}$ $(\alpha\text{已知})$

Homework

Solution

2. Solution. (a)

$$L(\theta;\vec{x}) = \frac{1}{\prod_{i=1}^{n}(x_i!)} e^{-n\theta}\, \theta^{\sum_{i=1}^{n} x_i} \cdot 1_{\{x_{(1)} > 0\}}.$$

$$\ell(\theta) = \ln L(\theta) = \ln\left(\prod_{i=1}^{n}(x_i!)\right) - n\theta + n\bar{x}\ln\theta, \quad x_{(1)} > 0.$$

$$\hat{\theta} = \underset{\theta>0}{\operatorname{argmax}}\ \ell(\theta) = \underset{\theta>0}{\operatorname{argmax}}\ n(\bar{x}\ln\theta - \theta)$$

$$= \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Solution

(b) $L(\theta;\vec{x}) = \alpha^n \left(\prod_{i=1}^n x_i\right)^{\alpha-1} \cdot \theta^n \cdot e^{-\theta \sum_{i=1}^n x_i^\alpha} \cdot \mathbb{1}_{\{x_{(1)}>0\}}$.

$\ell(\theta) = \ln L(\theta) = n\ln\alpha + (\alpha-1)\sum_{i=1}^n \ln x_i + n\ln\theta - \theta\sum_{i=1}^n x_i^\alpha$.

$\hat{\theta} = \underset{\theta>0}{\operatorname{argmax}}\, \ell(\theta) = \dfrac{n}{\sum_{i=1}^n x_i^\alpha}$. $\square$

Homework

3. 设总体X具有密度函数

$$f(x;\theta)=\begin{cases}\theta(1-x)^{\theta-1}, & 0\le x\le 1\\ 0, & 其它\end{cases}$$

$X_1,X_2,\cdots,X_n$是其样本，求θ的矩估计及最大似然估计.

Solution

3. Solution.

① moment estimate

$$\mathbb{E}X = \int_0^1 \theta x(1-x)^{\theta-1}dx$$

$$= (-(1-x)^{\theta}\cdot x)\Big|_0^1 + \int_0^1 (1-x)^{\theta}dx = \frac{1}{\theta+1}$$

Then, $\overline{X}_n = \dfrac{1}{\hat{\theta}+1}$, i.e. $\hat{\theta} = \dfrac{1}{\overline{X}_n} - 1$.

Homework

Solution

② MLE.

$$L(\theta;\vec{x}) = \theta^n \prod_{i=1}^{n} (1-x_i)^{\theta-1} \mathbb{1}_{\{0 \le x_{(1)} \le x_{(n)} \le 1\}}$$

$$\ell(\theta) = n \ln\theta + (\theta-1)\sum_{i=1}^{n} \ln(1-x_i)$$

$$\hat{\theta} = \underset{\theta>0}{\arg\max}\ n\ln\theta + (\theta-1)\sum_{i=1}^{n}\ln(1-x_i)$$

$$= -\frac{n}{\sum_{i=1}^{n}\ln(1-x_i)}. \quad \square$$

Supplement Exercises

5. 某厂生产的电容器的使用寿命服从指数分布.为了解其平均寿命,从中抽出 n 件产品测其实际使用寿命,试说明什么是总体,什么是样本,并指出样本的分布.

Solution

解　总体是该厂生产的电容器的寿命全体，或者可以说总体是指数分布，其分布为 $Exp(\lambda)$；

样本是该厂中抽出的 n 个电容器的寿命；

记第 i 个电容器的寿命为 x_i，则 $x_i \sim Exp(\lambda)$，$i=1,2,\cdots,n$，样本 $(x_1,\cdots,x_n)$ 的分布为 $\prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda t}$，其中 $t = x_1 + \cdots + x_n$。

Supplement Exercises

7. 设有 N 个产品，其中有 M 个次品. 进行放回抽样. 定义 x_i 如下：

$$x_i = \begin{cases} 1, & \text{第 } i \text{ 次取得次品}, \\ 0, & \text{第 } i \text{ 次取得正品}. \end{cases}$$

求样本 $x_1, x_2, \cdots, x_n$ 的联合分布.

Supplement Exercises

Solution

解 总体的分布列为

$$P(X=1)=\frac{M}{N}, P(X=0)=1-\frac{M}{N},$$

也可以写成

$$P(X=x)=\left(\frac{M}{N}\right)^{x}\left(1-\frac{M}{N}\right)^{1-x}, \quad x=0,1.$$

因此样本 $x_1,x_2,\cdots,x_n$ 的联合分布列为

$$P(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}\left(\frac{M}{N}\right)^{x_i}\left(1-\frac{M}{N}\right)^{1-x_i}=\left(\frac{M}{N}\right)^{t}\left(1-\frac{M}{N}\right)^{n-t}, \quad x_i=0,1,$$

其中 $t=x_1+\cdots+x_n$.

Supplement Exercises

2. 证明:对任意常数 c,d,有

$$\sum_{i=1}^{n}(x_i-c)(y_i-d)=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})+n(\bar{x}-c)(\bar{y}-d).$$

Supplement Exercises

Solution

证 $$\sum_{i=1}^{n}(x_i-c)(y_i-d)=\sum_{i=1}^{n}(x_i-\bar{x}+\bar{x}-c)(y_i-\bar{y}+\bar{y}-d)$$
$$=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})+\sum_{i=1}^{n}(\bar{x}-c)(y_i-\bar{y})+$$
$$\sum_{i=1}^{n}(x_i-\bar{x})(\bar{y}-d)+\sum_{i=1}^{n}(\bar{x}-c)(\bar{y}-d),$$

由 $\sum_{i=1}^{n}(x_i-\bar{x})=0$, $\sum_{i=1}^{n}(y_i-\bar{y})=0$，得

$$\sum_{i=1}^{n}(x_i-c)(y_i-d)=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})+n(\bar{x}-c)(\bar{y}-d),$$

因而结论成立.

5. 从同一总体中抽取两个容量分别为 n, m 的样本，样本均值分别为 $\bar{x}_1, \bar{x}_2$，样本方差分别为 s_1^2, s_2^2，将两组样本合并，其均值、方差分别为 $\bar{x}, s^2$，证明：

$$\bar{x} = \frac{n\bar{x}_1 + m\bar{x}_2}{n+m},$$

$$s^2 = \frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-1} + \frac{nm(\bar{x}_1 - \bar{x}_2)^2}{(n+m)(n+m-1)}.$$

Supplement Exercises

14. 利用切比雪夫不等式求抛均匀硬币多少次才能使正面朝上的频率落在(0.4,0.6)间的概率至少为0.9. 如何才能更精确地计算这个次数? 是多少?

Supplement Exercises

Solution

解 均匀硬币正面朝上的概率 $p=0.5$，设 x_n 为 n 次抛硬币中正面朝上的次数，则有 $x_n \sim b(n,p)$. 据题意选取次数 n 应满足

$$P\left(0.4 < \frac{x_n}{n} < 0.6\right) \geqslant 0.9,$$

此式等价于 $P(|x_n - 0.5n| \geqslant 0.1n) < 0.1$，利用切比雪夫不等式估计上式左端概率的上界

Solution

$$P\{|x_n - 0.5n| \geqslant 0.1n\} \leqslant \frac{n \times 0.5(1-0.5)}{(0.1n)^2} = \frac{25}{n},$$

再由不等式$\frac{25}{n} \leqslant 0.1$可得粗糙的估计 $n \geqslant 250$. 即抛均匀硬币 250 次后可满足要求.

讨论:利用 $\bar{x}$ 的渐近正态性可以得到更精确的结论. 由中心极限定理知,样本均值 $\bar{x} = \frac{x_n}{n}$,$\sqrt{n}(\bar{x}-0.5)/\sqrt{0.5 \times 0.5} \dot{\sim} N(0,1)$,故

$$P(0.4 < \bar{x} < 0.6) = P(\sqrt{n}|\bar{x}-0.5|/0.5 < \sqrt{n}/5)$$
$$= 2\Phi(\sqrt{n}/5) - 1 \geqslant 0.9,$$

即 $\Phi(\sqrt{n}/5) \geqslant 0.95$,故$\sqrt{n}/5 \geqslant 1.645$,这就给出较精确的上界 $n \geqslant (5 \times 1.645)^2 = 67.65$,这表明只需抛均匀硬币 68 次就可满足要求. 两个结果差异很大,说明切比雪夫不等式是一个较为粗糙的不等式,在能够使用大样本结果的情况下应尽量使用中心极限定理.

Supplement Exercises

4. 由正态总体 $N(\mu,\sigma^2)$ 抽取容量为 20 的样本，试求 $P(10\sigma^2 \le \sum_{i=1}^{20}(x_i - \mu)^2 \le 30\sigma^2)$.

Solution

解 因为 $x_i \sim N(\mu,\sigma^2)$，所以 $\frac{x_i-\mu}{\sigma} \sim N(0,1)$，$\sum_{i=1}^{20} \frac{(x_i-\mu)^2}{\sigma^2} \sim \chi^2(20)$

用 $k_{20}(x)$ 表示服从 $\chi^2(20)$ 的随机变量的分布函数值，则

$$P\left(10\sigma^2 \leqslant \sum_{i=1}^{20}(x_i-\mu)^2 \leqslant 30\sigma^2\right) = P\left(10 \leqslant \frac{\sum_{i=1}^{20}(x_i-\mu)^2}{\sigma^2} \leqslant 30\right)$$

$$= k_{20}(30) - k_{20}(10).$$

7. 设随机变量 $X \sim F(n,n)$，证明 $P(X<1)=0.5$.

Solution

证　若随机变量 $X \sim F(n,n)$，则 $Y = 1/X$ 也服从 $F(n,n)$，从而

$$P(X<1)=P(Y<1)=P(1/X<1)=P(X>1).$$

而

$$P(X<1)+P(X>1)=1,$$

这就证明了 $P(X<1)=0.5$.

11. 设 $x_1,\cdots,x_n$ 是来自 $N(\mu_1,\sigma^2)$ 的样本，$y_1,\cdots,y_m$ 是来自 $N(\mu_2,\sigma^2)$ 的样本，两总体独立. c,d 是任意两个不为 0 的常数，证明

$$t=\frac{c(\bar{x}-\mu_1)+d(\bar{y}-\mu_2)}{s_w\sqrt{\frac{c^2}{n}+\frac{d^2}{m}}}\sim t(n+m-2),$$

其中 $s_w^2=\frac{(n-1)s_x^2+(m-1)s_y^2}{n+m-2}$，$s_x^2$ 与 s_y^2 分别是两个样本方差.

Solution

证 由条件有

$$c(\bar{x}-\mu_1) \sim N\left(0, \frac{c^2\sigma^2}{n}\right), d(\bar{y}-\mu_2) \sim N\left(0, \frac{d^2\sigma^2}{m}\right),$$

$$\frac{(n-1)s_x^2}{\sigma^2} \sim \chi^2(n-1), \quad \frac{(m-1)s_y^2}{\sigma^2} \sim \chi^2(m-1),$$

且 $\bar{x}, \bar{y}, s_x^2, s_y^2$ 相互独立,故

$$c(\bar{x}-\mu_1) + d(\bar{y}-\mu_2) \sim N\left(0, \frac{c^2\sigma^2}{n} + \frac{d^2\sigma^2}{m}\right),$$

$$\frac{(n+m-2)s_w^2}{\sigma^2} = \frac{(n-1)s_x^2}{\sigma^2} + \frac{(m-1)s_y^2}{\sigma^2} \sim \chi^2(n+m-2).$$

于是

$$t = \frac{c(\bar{x}-\mu_1) + d(\bar{y}-\mu_2)}{s_w\sqrt{\frac{c^2}{n} + \frac{d^2}{m}}}$$

$$= \frac{[c(\bar{x}-\mu_1) + d(\bar{y}-\mu_2)] \Big/ \sqrt{\frac{c^2\sigma^2}{n} + \frac{d^2\sigma^2}{m}}}{\sqrt{\frac{(n+m-2)s_w^2}{\sigma^2} \Big/ (n+m-2)}} \sim t(n+m-2).$$

Supplement Exercises

19. 设 $x_1, x_2, \cdots, x_n$ 是来自某连续总体的一个样本. 该总体的分布函数 $F(x)$ 是连续严增函数，证明：统计量 $T = -2\sum_{i=1}^{n} \ln F(x_i)$ 服从 $\chi^2(2n)$.

Supplement Exercises

Solution

证　分几步进行：

(1) 若 $X \sim F(x)$，且 $F(x)$ 为连续严增函数，则 $Y=F(X) \sim U(0,1)$。这是因为 $F(x)$ 的反函数 F^{-1} 也存在。于是 $Y=F(X)$ 的分布函数为

$$F_Y(y)=P(F(X)\leqslant y)=P(X\leqslant F^{-1}(y))=F(F^{-1}(y))=y,$$

其中 $y\in(0,1)$，当 $y\leqslant 0$，$F_Y(y)=0$，当 $y\geqslant 1$，$F_Y(y)=1$，所以 $F(X)\sim U(0,1)$。

(2) 若 $Y\sim U(0,1)$，则 $Z=-\ln Y\sim\chi^2(2)$。这是由于 y 仅在 $(0,1)$ 上取值，故 $Z=-\ln Y$ 仅在 $(0,+\infty)$ 上取值，所以当 $z\leqslant 0$ 时，$F_Z(z)=0$；当 $z>0$ 时，有

$$F_Z(z)=P(-\ln Y\leqslant z)=P(Y\geqslant e^{-z})=1-e^{-z},$$

这是参数为 1 的指数分布函数，也是自由度为 2 的 χ^2 分布函数，即 $Z=-\ln Y\sim\chi^2(2)$。

(3) 由 $X_1,X_2,\cdots,X_n$ 的相互独立性可导致 $F(X_1),F(X_2),\cdots,F(X_n)$ 相互独立。由(1)与(2)可知 $u=-2\sum_{i=1}^{n}\ln F(x_i)\sim\chi^2(2n)$。

Supplement Exercises

4. 设总体 $X \sim N(\mu,\sigma^2)$, $x_1,\cdots,x_n$ 是来自该总体的一个样本. 试确定常数 c 使 $c\sum\limits_{i=1}^{n-1}(x_{i+1}-x_i)^2$ 为 σ^2 的无偏估计.

Supplement Exercises

Solution

解　由于总体 $X \sim N(\mu,\sigma^2)$,这给出

$$E(x_i^2)=\sigma^2+\mu^2,$$

$$E(x_i x_{i+1})=E(x_i)E(x_{i+1})=\mu^2,\quad i=1,2,\cdots,n.$$

于是

$$E\left(\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2\right)=E(x_1^2+2x_2^2+\cdots+2x_{n-1}^2+x_n^2-2x_1x_2-\cdots-2x_{n-1}x_n)$$

$$=[2(n-1)(\sigma^2+\mu^2)-2(n-1)\mu^2]=2(n-1)\sigma^2.$$

若要使 $c\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2$ 为 σ^2 的无偏估计,即 $cE\left(\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2\right)=\sigma^2$,这给出

$$c=\frac{1}{2(n-1)}.$$

Supplement Exercises

11. 设总体 X 服从正态分布 $N(\mu,\sigma^2)$，$x_1,x_2,\cdots,x_n$ 为来自总体 X 的样本，为了得到标准差 σ 的估计量，考虑统计量：

$$y_1=\frac{1}{n}\sum_{i=1}^{n}|x_i-\bar{x}|,\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,n\geq 2,$$

$$y_2=\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i-x_j|,n\geq 2,$$

求常数 C_1 与 C_2，使得 C_1y_1 与 C_2y_2 都是 σ 的无偏估计。

Solution

解 由期望的公式及对称性，我们只需要求出 $E|x_1-\bar{x}|$ 和 $E|x_1-x_2|$ 即可，注意到 $x_1-\bar{x}\sim N(0,\frac{n-1}{n}\sigma^2)$（为什么？）和 $x_1-x_2\sim N(0,2\sigma^2)$，我们只需要求出如下期望即可完成本题：设 $y\sim N(0,\sigma^2)$，则

$$E|y| = 2\int_0^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} y e^{-\frac{y^2}{2\sigma^2}}dy$$

$$= \int_0^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{z}{2\sigma^2}}dz = \int_0^{\infty} \frac{\sqrt{2}\sigma}{\sqrt{\pi}} e^{-t}dt = \frac{\sqrt{2}\sigma}{\sqrt{\pi}},$$

于是有 $E|x_1-\bar{x}| = \sqrt{\frac{n-1}{n}}\frac{\sqrt{2}\sigma}{\sqrt{\pi}}$ 和 $E|x_1-x_2| = \frac{2\sigma}{\sqrt{\pi}}$，从而给出 $C_1=\sqrt{\frac{n\pi}{2(n-1)}}$，$C_2=\frac{\sqrt{\pi}}{2}$.

14. 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^2 的无偏估计；

(2) 寻求 $p(1-p)$ 的无偏估计；

(3) 证明 $\dfrac{1}{p}$ 的无偏估计不存在.

Supplement Exercises

Solution

解 (1) $\overline{x}^2$ 是 p^2 的一个直观估计，但不是 p^2 的无偏估计，这是因为

$$E(\overline{x}^2)=\mathrm{Var}(\overline{x})+[E(\overline{x})]^2=\frac{p(1-p)}{n}+p^2=\frac{p}{n}+\frac{n-1}{n}p^2\neq p^2,$$

由此可见 $\hat{p}^2=\frac{n}{n-1}\left[\overline{x}^2-\frac{\overline{x}}{n}\right]$ 是 p^2 的无偏估计.

(2) $\overline{x}(1-\overline{x})=\overline{x}-\overline{x}^2$ 是 $p(1-p)$ 的直观估计，但不是 $p(1-p)$ 的无偏估计，这是因为

$$E(\overline{x}-\overline{x}^2)=p-\left(\frac{p(1-p)}{n}+p^2\right)=\frac{n-1}{n}p(1-p)\neq p(1-p),$$

由此可见 $\frac{n}{n-1}\overline{x}(1-\overline{x})$ 是 $p(1-p)$ 的一个无偏估计.

Supplement Exercises

Solution

(3) 反证法，倘若 $g(x_1,\cdots,x_n)$ 是 $\frac{1}{p}$ 的无偏估计，则有

$$\sum_{x_1,\cdots,x_n} g(x_1,\cdots,x_n)p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}=\frac{1}{p},$$

或者

$$\sum_{x_1,\cdots,x_n} g(x_1,\cdots,x_n)p^{\sum_{i=1}^{n}x_i+1}(1-p)^{n-\sum_{i=1}^{n}x_i}-1=0,$$

上式是 p 的 $n+1$ 次方程，它最多有 $n+1$ 个实根，而 p 可在 $(0,1)$ 取无穷多个值，所以不论取什么形式都不能使上述方程在 $0<p<1$ 上成立，这表明 $\frac{1}{p}$ 的无偏估计不存在.

Supplement Exercises

4. 设总体密度函数如下，$x_1,\cdots,x_n$ 是样本，试求未知参数的矩估计.

(1) $p(x;\theta)=\frac{2}{\theta^2}(\theta-x)$, $0<x<\theta$, $\theta>0$;

(2) $p(x;\theta)=(\theta+1)x^{\theta}$, $0<x<1$, $\theta>0$;

(3) $p(x;\theta)=\sqrt{\theta}\,x^{\sqrt{\theta}-1}$, $0<x<1$, $\theta>0$;

(4) $p(x;\theta,\mu)=\frac{1}{\theta}e^{-\frac{x-\mu}{\theta}}$, $x>\mu$, $\theta>0$.

Solution

(4) 先计算总体均值与方差

$$E(X)=\int_{\mu}^{+\infty} x\frac{1}{\theta}e^{-\frac{x-\mu}{\theta}}dx=\int_{0}^{+\infty} t\frac{1}{\theta}e^{-\frac{t}{\theta}}dt+\int_{0}^{+\infty}\frac{1}{\theta}\mu e^{-\frac{t}{\theta}}dt$$
$$=\theta+\mu,$$

$$E(X^2)=\int_{\mu}^{+\infty} x^2\frac{1}{\theta}e^{-\frac{x-\mu}{\theta}}dx=\int_{0}^{+\infty}(t+\mu)^2\frac{1}{\theta}e^{-\frac{t}{\theta}}dt$$
$$=\int_{0}^{+\infty} t^2\frac{1}{\theta}e^{-\frac{t}{\theta}}dt+\int_{0}^{+\infty}2\mu t\frac{1}{\theta}e^{-\frac{t}{\theta}}dt+\int_{0}^{+\infty}\mu^2\frac{1}{\theta}e^{-\frac{t}{\theta}}dt$$
$$=2\theta^2+2\mu\theta+\mu^2,$$

$$\mathrm{Var}(X)=E(X^2)-(E(X))^2=\theta^2.$$

由此可以推出 $\theta=\sqrt{\mathrm{Var}(X)}$, $\mu=E(X)-\sqrt{\mathrm{Var}(X)}$, 从而参数 θ,μ 的矩估计为

$$\hat{\theta}=s,\qquad \hat{\mu}=\bar{x}-s.$$

8. 设 $x_1,\cdots,x_n$ 是来自对数级数分布

$$P(X=k)=-\frac{1}{\ln(1-p)}\cdot\frac{p^k}{k},0<p<1;k=1,2,\cdots$$

的一个样本,求参数 p 的矩估计.

Supplement Exercises

Solution

解　由于

$$EX=\sum_{k=1}^{+\infty}kP(X=k)=-\frac{1}{\ln(1-p)}\sum_{k=1}^{+\infty}p^k=-\frac{p}{(1-p)\ln(1-p)},$$

$$EX^2=\sum_{k=1}^{+\infty}k^2P(X=k)=-\frac{1}{\ln(1-p)}\sum_{k=1}^{+\infty}kp^k=-\frac{p}{(1-p)^2\ln(1-p)},$$

因此有 $1-p=\dfrac{EX}{EX^2}$，从而得到 p 的一个矩估计 $\hat{p}=1-\dfrac{\sum x_i}{\sum x_i^2}$.

Supplement Exercises

1. 设总体概率函数如下，$x_1,\cdots,x_n$ 是样本，试求未知参数的最大似然估计.

(1) $p(x;\theta)=\sqrt{\theta}\,x^{\sqrt{\theta}-1}$，$0<x<1$，$\theta>0$；

(2) $p(x;\theta)=\theta c^{\theta}x^{-(\theta+1)}$，$x>c$，$c>0$ 已知，$\theta>1$.

Solution

解　(1) 似然函数为 $L(\theta)=(\sqrt{\theta})^n(x_1\cdots x_n)^{\sqrt{\theta}-1}$，其对数似然函数为

$$\ln L(\theta)=\frac{n}{2}\ln\theta+(\sqrt{\theta}-1)(\ln x_1+\cdots+\ln x_n).$$

Supplement Exercises

Solution

将 $\ln L(\theta)$ 关于 θ 求导并令其为 0 即得到似然方程

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{n}{2\theta} + (\ln x_1 + \cdots + \ln x_n)\frac{1}{2\sqrt{\theta}} = 0.$$

解之得

$$\hat{\theta} = \left(\frac{1}{n}\sum_{i=1}^{n} \ln x_i\right)^{-2}.$$

由于

$$\left.\frac{\partial^2 \ln L(\theta)}{\partial \theta^2}\right|_{\hat{\theta}} = \left.\left(-\frac{n}{2\theta^2} - \frac{\sum \ln x_i}{4\theta^{3/2}}\right)\right|_{\hat{\theta}} = -\frac{3(\sum \ln x_i)^4}{4n^3} < 0,$$

所以 $\hat{\theta}$ 是 θ 的最大似然估计.

2. 设总体概率函数如下，$x_1, \cdots, x_n$ 是样本，试求未知参数的最大似然估计.

(1) $p(x;\theta) = c\theta^c x^{-(c+1)}, x > \theta, \theta > 0, c > 0$ 已知；

(2) $p(x;\theta,\mu) = \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}}, x > \mu, \theta > 0$；

(3) $p(x;\theta) = (k\theta)^{-1}, \theta < x < (k+1)\theta, \theta > 0$.

Solution

解 (1) 样本 $x_1,\cdots,x_n$ 的似然函数为

$$L(\theta)=c^n\theta^{nc}(x_1\cdots x_n)^{-(c+1)}I_{\{x_{(1)}>\theta\}}$$

要使 $L(\theta)$ 达到最大,首先示性函数应为1,其次是 θ^{nc} 尽可能大. 由于 $c>0$, 故 θ^{nc} 是 θ 的单调增函数,所以 θ 的取值应尽可能大,但示性函数的存在决定了 θ 的取值不能大于 $x_{(1)}$,由此给出 θ 的最大似然估计为 $x_{(1)}$.

Solution

(3) 设有样本 $x_1,\cdots,x_n$,其似然函数为

$$L(\theta)=\left(\frac{1}{k\theta}\right)^n I_{\{\theta\le x_{(1)}\le x_{(n)}\le(k+1)\theta\}}.$$

由于 $L(\theta)$ 的主体 $\left(\frac{1}{k\theta}\right)^n$ 是关于 θ 的单调递减函数,要使 $L(\theta)$ 达到最大,θ 应尽可能小,但由限制 $\theta\le x_{(1)}\le x_{(n)}\le(k+1)\theta$ 可以得到 $\frac{x_{(n)}}{k+1}\le\theta\le x_{(1)}$,这说明 θ 不能小于 $\frac{x_{(n)}}{k+1}$,因而 θ 的最大似然估计为 $\hat{\theta}=\frac{x_{(n)}}{k+1}$.

Supplement Exercises

3. 设总体概率函数如下，$x_1, \cdots, x_n$ 是样本，试求未知参数的最大似然估计.

(1) $p(x;\theta) = \dfrac{1}{2\theta} e^{-|x|/\theta}, \theta > 0$;

(2) $p(x;\theta) = 1, \theta - 1/2 < x < \theta + 1/2$;

(3) $p(x;\theta_1, \theta_2) = \dfrac{1}{\theta_2 - \theta_1}, \theta_1 < x < \theta_2$.

Supplement Exercises

Solution

解 (1) 不难写出似然函数为

$$L(\theta)=\left(\frac{1}{2\theta}\right)^n e^{-\frac{\sum_{i=1}^{n}|x_i|}{\theta}}$$

对数似然函数为

$$\ln L(\theta)=-n\ln 2\theta-\frac{\sum_{i=1}^{n}|x_i|}{\theta}$$

Supplement Exercises

Solution

（2）此处的似然函数为

$$L(\theta)=I_{\{\theta-\frac{1}{2}<x_{(1)}<x_{(n)}<\theta+\frac{1}{2}\}}\ .$$

它只有两个取值：0 和 1，为使得似然函数取 1，θ 的取值范围应是 $x_{(n)}-\frac{1}{2}<\theta<x_{(1)}+\frac{1}{2}$，因而 θ 的最大似然估计 $\hat{\theta}$ 可取 $\left(x_{(n)}-\frac{1}{2},x_{(1)}+\frac{1}{2}\right)$ 中的任意值. 这说明 MLE 可能不止一个.

Supplement Exercises

Solution

（3）由条件，似然函数为

$$L(\theta)=\frac{1}{(\theta_2-\theta_1)^n}I_{\{\theta_1<x_{(1)}<x_{(n)}<\theta_2\}}.$$

要使 $L(\theta)$ 尽量大，首先示性函数应为 1，这说明 $\theta_1<x_{(1)}<x_{(n)}<\theta_2$；其次 $\theta_2-\theta_1$ 要尽量小，综上可知，θ_1 的最大似然估计应为 $x_{(1)}$，θ_2 的最大似然估计应为 $x_{(n)}$。

4. 一地质学家为研究密歇根湖的湖滩地区的岩石成分，随机地自该地区取100个样品，每个样品有10块石子，记录了每个样品中属石灰石的石子数. 假设这100次观察相互独立，求这地区石子中石灰石的比例p的最大似然估计. 该地质学家所得的数据如下：

样本中的石子数	0	1	2	3	4	5	6	7	8	9	10
样品个数	0	1	6	7	23	26	21	12	3	1	0

Supplement Exercises

Solution

解 本题中，总体 X 为样品中石灰石的个数，且 X 服从参数为 $(10, p)$ 的二项分布，即

Supplement Exercises

Solution

$$p(X=x)=\binom{10}{x}p^x(1-p)^{10-x},$$

$x_1,x_2,\cdots,x_{100}$ 为样本,则其似然函数为(忽略常数)

$$L(p)=p^{\sum_{i=1}^{100}x_i}(1-p)^{10\times 100-\sum_{i=1}^{100}x_i},$$

对数似然函数为

$$\ln L(p)=\sum_{i=1}^{100}x_i\ln p+(10\times 100-\sum_{i=1}^{100}x_i)\ln(1-p).$$

将对数似然函数关于 p 求导并令其为 0 得到似然方程

$$\frac{\partial \ln L(p)}{\partial p}=\frac{\sum_{i=1}^{100}x_i}{p}-\frac{10\times 100-\sum_{i=1}^{100}x_i}{1-p}=0.$$

解之得

$$\hat{p}=\frac{\sum_{i=1}^{100}x_i}{1\ 000}$$

Solution

1 000

由于

$$\frac{\partial \ln^2 L(p)}{\partial p^2} = -\frac{\sum_{i=1}^{100} x_i}{p^2} - \frac{1\,000 - \sum_{i=1}^{100} x_i}{(1-p)^2} < 0$$

由二阶导数的性质知，p 的最大似然估计为

$$\hat{p} = \frac{\sum_{i=1}^{100} x_i}{1\,000} = \frac{499}{1\,000} = 0.499.$$

Supplement Exercises

11. 证明：对正态分布 $N(\mu,\sigma^2)$，若只有一个观测值，则 σ^2 的最大似然估计不存在.

Supplement Exercises

Solution

证　在只有一个观测值场合，对数似然函数为

$$l(\mu,\sigma^2;x)=-\ln(\sqrt{2\pi}\,\sigma)-\frac{(x-\mu)^2}{2\sigma^2}.$$

该函数在 $\sigma\to0$ 时趋于 ∞，这说明该函数没有最大值，或者说极大值无法实现，从而 σ^2 的最大似然估计不存在.

Further Reading

1. LSE

Reference. [Casella, Berger] Statistical Inference, Chapter 12, Regression Models

Thank you!