概率论第七章答案

1. 设随机变量 X 与 Y 相互独立且分别服从正态分布 $N(\mu,\sigma^2)$ 与 $N(\mu,2\sigma^2)$，其中 σ 是未知参数且 $\sigma>0$. 设 $Z=X-Y$.

   (1) 求 Z 的概率密度 $f(z;\sigma^2)$;

   (2) 设 $Z_1,Z_2,\ldots,Z_n$ 为来自总体 Z 的简单随机样本，求 σ^2 的极大似然估计量 $\hat{\sigma}^2$;

   (3) 证明：$\hat{\sigma}^2$ 是 σ^2 的无偏估计量.

解 (1) 根据题意，$Z=X-Y$ 服从正态分布，且

$$E(Z)=E(X-Y)=E(X)-E(Y)=\mu-\mu=0$$

$$D(Z)=D(X-Y)=D(X)+D(Y)=\sigma^2+2\sigma^2=3\sigma^2$$

所以 Z 的概率密度为

$$f(z;\sigma^2)=\frac{1}{\sqrt{6\pi}\sigma}e^{-\frac{z^2}{6\sigma^2}},\quad -\infty<z<+\infty.$$

(2) 似然函数为

$$L(\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{6\pi}\sigma}e^{-\frac{z_i^2}{6\sigma^2}}=\frac{1}{(6\pi)^{\frac{n}{2}}}\exp\left(-\frac{1}{6\sigma^2}\sum_{i=1}^{n}z_i^2\right),$$

对数似然函数为

$$\ln L(\sigma^2)=\ln\frac{1}{(6\pi)^{\frac{n}{2}}}-\frac{n}{2}\ln\sigma^2-\frac{1}{6\sigma^2}\sum_{i=0}^{n}z_i^2.$$

似然方程为 $0=\dfrac{\mathrm{d}[\ln L(\sigma^2)]}{\mathrm{d}\sigma^2}$，则 $-\dfrac{n}{2\sigma^2}+\dfrac{1}{6(\sigma^2)^2}\displaystyle\sum_{i=1}^{n}z_i^2=0$，由此解得 σ^2 的极大似然估计量为

$$\hat{\sigma}^2=\frac{1}{3n}\sum_{i=1}^{n}Z_i^2.$$

(3) 由于

$$\begin{aligned}E(\hat{\sigma}^2)&=E\left(\frac{1}{3n}\sum_{i=1}^{n}{Z_i}^2\right)=\frac{1}{3n}E\left(\sum_{i=1}^{n}Z_i^2\right)=\frac{1}{3}E(Z^2)\\&=\frac{1}{3}\{D(Z)+[E(Z)]^2\}=\frac{1}{3}(3\sigma^2+0)=\sigma^2.\end{aligned}$$

因此 $\hat{\sigma}^2$ 是 σ^2 的无偏估计量.

2. 设 $\hat{\theta}_1$，$\hat{\theta}_2$ 是参数 θ 的两个互相独立的无偏估计量，且 $D(\hat{\theta}_1)=2D(\hat{\theta}_2)$. 试求常数 k_1，k_2 满足什么条件才能使 $k_1\hat{\theta}_1+k_2\hat{\theta}_2$ 是 θ 的无偏估计量，并且求常数 k_1，k_2 使它在所有这种形式的估计量中方差达到最小.

解 为使 $k_1\hat{\theta}_1+k_2\hat{\theta}_2$ 是 θ 的无偏估计量，应有

$$\begin{aligned}E(k_1\hat{\theta}_1+k_2\hat{\theta}_2)&=k_1E(\hat{\theta}_1)+k_2E(\hat{\theta}_2)\\&=k_1\theta_1+k_2\theta_2=\theta.\end{aligned}$$

即有 $k_1+k_2=1$，又因 $\hat{\sigma_1}$，$\hat{\sigma_2}$ 是相互独立的，所以

$$\begin{aligned}D(k_1\hat{\theta}_1+k_2\hat{\sigma}_2)&=k_1^2D(\hat{\theta}_1)+k_2^2D(\hat{\theta}_2)\\&=k_1^2[2D(\hat{\theta}_2)]+k_2^2D(\hat{\theta}_2)\\&=(2k_1^2+k_2^2)D(\hat{\theta}_2)\\&=(3k_1^2-2k_1+1)D(\hat{\theta}_2).\end{aligned}$$

为使方差达到最小，即使 $f(k_1)=3k_1^2-2k_1+1$ 达到最小。利用极值原理，可得唯一驻点 $k_1=\dfrac{1}{3}$，使 $f(k_1)$ 达到最小，因此当 $k_1=\dfrac{1}{3}$，$k_2=\dfrac{2}{3}$ 时，所有这种形式的估计量中方差达到最小。

3. 设 $X_1,X_2,\cdots,X_n$ 为总体的一个样本, 求下列各总体的密度函数或分布律中的未知参数的矩估计量和最大似然估计量.

(1) $f(x)=\begin{cases}\theta c^\theta x^{-(\theta+1)}, & x>c\\ 0, & \text{其他}\end{cases}$，其中 $c>0$ 为已知 ,$\theta>1,\theta$ 为未知参数.

(2) $f(x)=\begin{cases}\sqrt{\theta}x^{\sqrt{\theta}-1}, & 0\leqslant x\leqslant 1\\ 0, & \text{其他}\end{cases}$，其中 $\theta>0,\theta$ 为未知参数.

(3) $P\{X=x\}=C_m^xp^x(1-p)^{m-x},x=0,1,2,\cdots,m,0<p<1,p$ 为未知参数.

解 (1)$E(X)=\int_c^\infty x\theta c^\theta x^{-(\theta+1)}\mathrm{d}x=\int_c^\infty \theta c^\theta x^{-\theta}\mathrm{d}x=\theta c^\theta\int_c^\theta x^{-\theta}\mathrm{d}x=\theta c^\theta\frac{x^{-\theta+1}}{-\theta+1}\Big|_c^{+\infty}$

$$=\frac{\theta c}{\theta-1}$$

令 $\frac{\theta c}{\theta-1}=\bar{X}$， 解得 $\hat{\theta}=\frac{\bar{X}}{\bar{X}-c}$, 即为 θ 的矩估计量. 似然函数为

$$L(\theta)=\prod_{i=1}^n\theta\cdot c^\theta x_i^{-(\theta+1)}=\theta^nc^{n\theta}\prod_{i=1}^nx_i^{-(\theta+1)}$$

而

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

令

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=\frac{n}{\theta}+n\ln c-\sum_{i=1}^n\ln x_i=0$$

解得 θ 的最大似然估计量 $\hat{\theta}=\frac{n}{\sum_{i=1}^n\ln X_i-n\ln c}$

(2) $EX=\int_0^1x\sqrt{\theta}x^{\sqrt{\theta}-1}\,\mathrm{d}x=\int_0^1\sqrt{\theta}x^{\sqrt{\theta}}\mathrm{d}x=\frac{\sqrt{\theta}}{\sqrt{\theta}+1}x^{\sqrt{\theta}+1}\Big|_0^1=\frac{\sqrt{\theta}}{\sqrt{\theta}+1}$ 令 $\frac{\sqrt{\theta}}{\sqrt{\theta}+1}=\bar{X}$, 解得 $\hat{\theta}=\left(\frac{\bar{X}}{\bar{X}-1}\right)^2$, 即为 θ 的矩估计量. 似然函数为:

$$L(\theta)=\prod_{i=1}^n\sqrt{\theta}x_i^{\sqrt{0}-1}=\theta^{\frac{n}{2}}\prod_{i=1}^nx_i^{\sqrt{\theta}-1}$$

对数似然函数为:

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

对数似然方程为：

$$\frac{\mathrm{d}\ln L(\theta)}{d\theta}=\frac{n}{2\theta}+\frac{\sum_{i=1}^{n}\ln x_i}{2\sqrt{\theta}}=0$$

其最大似然估计值为 $\hat{\theta}=\frac{n^2}{\left(\sum_{i=1}^{n}\ln x_i\right)^2}$， $\hat{\theta}=\frac{n^2}{\left(\sum_{i=1}^{n}\ln X_i\right)^2}$ 即为θ 的最大似然估计量

(3) 因为 $X\sim B(m,p)$, 所以

$$E(X)=\sum_{i=1}^{n}xC_m^rp^x(1-p)^{m-x}=mp$$

所以 $mp=\bar{X},\hat{p}=\frac{\bar{X}}{m}$ 为 p 的矩估计.

似然函数为:

$$L(p)=\prod_{i=1}^{n}C_m^xp^{x_i}(1-p)^{m-x_i}=p_{i=1}^{n}x_i(1-p)^{nm-\sum_{i=1}^{n}x_i}\prod_{i=1}^{n}C_m^{x_i}$$

对数似然函数为：

$$\ln L(p)=\sum_{i=1}^{n}x_i\ln p+\left(nm-\sum_{i=1}^{n}x_i\right)\ln(1-p)+\sum_{i=1}^{\infty}\ln C$$

对数似然方程为：

$$\frac{\mathrm{d}\ln L(p)}{\mathrm{d}p}=\frac{\sum_{i=1}^{n}x_i}{p}+\frac{nm-\sum_{i=1}^{*}x_i}{1-p}(-1)=0$$

则 $\hat{p}=\frac{\bar{x}}{m}$ 为 p 的最大似然估计量，$\hat{p}=\frac{\bar{X}}{m}$ 为 p 的最大似然估计量.

4. 设随机变量 X 的分布函数为

$F(x;\alpha,\beta)=\begin{cases}1-\left(\frac{\alpha}{x}\right)^{\beta}, & x>\alpha\\ 0, & x\leqslant\alpha\end{cases}$

其中参数 $\alpha>0,\beta>1$, 设 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本.

(1) 当 $\alpha=1$ 时, 求未知参数 β 的矩估计量;

(2) 当 $\alpha=1$ 时, 求未知参数 β 的最大似然估计量;

(3) 当 $\beta=2$ 时, 求未知参数 α 的最大似然估计量.

解（1）由已知 X 的分布函数可得其概率密度为 $f(x;\alpha,\beta)=\begin{cases}\frac{\beta\alpha^{\beta}}{x^{\beta+1}}, & x>\alpha\\ 0, & x\leqslant\alpha\end{cases}$ 当 $\alpha=1$ 时, X 的概率密度为

$$f(x;\beta)=\begin{cases}\frac{\beta}{x^{\beta+1}}, & x>1\\ 0, & x\leqslant 1\end{cases}$$

$$EX=\int_{-\infty}^{+\infty}xf(x;\beta)\mathrm{d}x=\int_{1}^{+\infty}\frac{\beta}{x^{\beta}}\mathrm{d}x=\frac{\beta}{\beta-1}\tag{1}$$

令$\frac{\beta}{\beta-1}=\bar{X}$，解得$\beta=\frac{\bar{X}}{\bar{X}-1}$，所以$\beta$ 的矩估计量为$\hat{\beta}=\frac{\bar{X}}{\bar{X}-1}$，其中$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$.

(2) 对于总体 X 的样本值 $x_1,x_2,\cdots,x_n$, 似然函数为

$$L(\beta)=\begin{cases}\frac{\beta^n}{(x_1x_2\cdots x_n)^{\beta+1}}, & x_i>1(i=1,2,\cdots,n)\\ 0, & \text{其他}\end{cases}$$

当 $x_i > 1$ 时, 两边取对数得

$$\ln L(\beta) = n\ln\beta - (\beta+1)\sum_{i=1}^{n}\ln x_i, \quad \frac{\mathrm{d}\ln L(\beta)}{\mathrm{d}\beta} = \frac{n}{\beta} - \sum_{i=1}^{n}\ln x_i$$

令 $\frac{\mathrm{d}\ln L(\beta)}{\mathrm{d}\beta} = 0$, 解之得 $\beta = \frac{n}{\sum_{i=1}^{n}\ln x_i}$ 故 β 的最大似然估计量为 $\hat{\beta} = \frac{n}{\sum_{i=1}^{n}\ln X_i}$.

(3) 当 $\beta = 2$ 时, X 的概率密度为

$$f(x;\alpha) = \begin{cases} \frac{2\alpha^2}{x^3}, & x > \alpha \\ 0, & x \leqslant \alpha \end{cases}$$

对于总体 X 的样本值 $x_1, x_2, \cdots, x_n$, 其似然函数为

$$L(\alpha) = \begin{cases} \frac{2^n\alpha^{2n}}{(x_1x_2\cdots x_n)^3}, & x_i > \alpha (i=1,2,\cdots,n) \\ 0, & \text{其他} \end{cases}$$

$$\ln L(\alpha) = n\ln 2 + 2n\ln\alpha - 3\sum_{i=1}^{n}\ln x_i, \quad \frac{\mathrm{d}\ln L(\alpha)}{\mathrm{d}\alpha} = \frac{2n}{\alpha} > 0$$

所以 $L(\alpha)$ 单调递增. 当 $x_i > \alpha(i=1,2,\cdots,n)$ 时 ,α 越大, $L(\alpha)$ 就越大, 因此 α 的最大似然估计值为

$$\hat{\alpha} = \min(x_1, x_2, \cdots, x_n)$$

则 α 的最大似然估计量为 $\hat{\alpha} = \min(x_1, x_2, \cdots, x_n)$

5. 设某种清漆的 9 个样品其干燥时间（以 h 计）分布为

6.0　5.7　5.8　6.5　7.0　6.3　5.6　6.1　5.0.

设干燥时间服从正态分布 $N(\mu, \sigma^2)$，求总体均值 μ 的置信水平为 0.95 的置信区间.

(1) 若由以往经验知总体标准差 $\sigma = 0.6$;

(2) 若 σ 未知.

解 样本容量为 $n = 9$，根据所给的样本可算得：$\bar{x} = 6$，$s^2 = 0.33$，且 $1-\alpha = 0.95$，$\frac{\alpha}{2} = 0.025$.

(1) 又因为 σ 已知，查表知 $z_{0.025} = 1.96$，于是得置信区间得上下限分别为

$$\bar{x} - \frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}} = 5.608, \quad \bar{x} - \frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}} = 6.392.$$

所以，此时 μ 的置信水平为 0.95 的置信区间为 (5.608, 6.392).

(2) 又因 σ 未知，查表知 $t_{0.025}(8) = 2.306$，于是得置信区间得上下限分别为

$$\bar{x} - \frac{s}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1) = 5.558, \quad \bar{x} + \frac{s}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1) = 6.442.$$

所以，此时 μ 得置信水平为 0.95 得置信区间为 (5.558, 6.442).

6. 某车间生产钢丝，设钢丝折断力服从正态分布，现随机地抽取 10 根，检查折断力，得数据如下（单位：N）:

578　572　570　568　572　570　570　572　596　584.

试求钢丝折断力方差的置信区间和置信上限（置信度为 0.95）.

解 (1) 这是一个正态总体，期望未知，对方差作双侧置信限的估计问题，应选统计量

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

则 σ^2 的 $1-\alpha$ 的置信区间是

$$\left(\frac{(n-1)S^2}{\chi^2_{n/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-n/2}(n-1)} \right).$$

由所给样本得

$$\bar{x} = 575.2, \quad (n-1)s^2 = \sum_{i=1}^{10} (x_i - \bar{x})^2 = 681.6.$$

根据给定的置信度 $1-\alpha = 0.95$，查自由度为 10-1 = 9 的 χ^2 分布表，得双侧临界值

$$\chi^2_{\alpha/2}(n-1) = 19.0$$
$$\chi^2_{1-\alpha/2}(n-1) = 2.7$$

代入上式得 σ^2 得 0.95 得置信区间为 (35.87, 232.44).

(2) 这是一个正态总体期望未知时，σ^2 得单侧区间估计问题，则 σ^2 的置信度为 0.95 的单侧置信上限为

$$\frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)} = \frac{\sum\limits_{i=1}^{10} (x_i - \bar{x})^2}{\chi^2_{1-\alpha}(n-1)},$$

已算得 $(n-1)S^2 = \sum\limits_{i=1}^{10} (x_i - \bar{x})^2 = 681.6$，根据置信度 $1-\alpha = 0.95$，查自由度为 9 的 χ^2 分布表得单侧临界值

$$\chi^2_{1-\alpha}(n-1) = 3.325$$

代入上式便得 σ^2 的 0.95 的置信上限为 205.

7. 某车间有两台自动机床加工一类套筒，假设套筒直径服从正态分布。现在从两个班次的产品中分别检查 5 个和 6 个套筒，得其直径（单位：cm）数据如下：

甲班：	5.06	5.08	5.03	5.00	5.07	
乙班：	4.98	5.03	4.97	4.99	5.02	4.95

试求两班加工套筒直径的方差比 $\sigma^2_{\text{甲}}/\sigma^2_{\text{乙}}$ 的 0.95 置信区间。

解 此处，$m=5, n=6$, 若取 $1-\alpha = 0.95$，查表则知：

$$F_{0.025}(4,5) = \frac{1}{F_{0.975}(5,4)} = \frac{1}{9.36},$$
$$F_{0.975}(4,5) = 7.39.$$

由数据算得 $s^2_{\text{甲}} = 0.00107, s^2_{\text{乙}} = 0.00092$，故置信区间的两端分别为：

$$\begin{aligned} \frac{s^2_{\text{甲}}}{s^2_{\text{乙}}} \cdot \frac{1}{F_{0.975}(4,5)} &= \frac{0.00107}{0.00092} \cdot \frac{1}{7.39} = 0.1574, \\ \frac{s^2_{\text{甲}}}{s^2_{\text{乙}}} \cdot \frac{1}{F_{0.025}(4,5)} &= \frac{0.00107}{0.00092} \cdot 9.36 = 10.8899, \end{aligned} \tag{2}$$

由此可知，$\sigma^2_{甲}/\sigma^2_{乙}$ 的 0.95 置信区间为 $[0.1574, 108899]$。

8. 设随机地从 A 批导线中抽 4 根，又从 B 批导线中抽 5 根，测得电阻 Ω 为：

A 批导线：0.143　0.142　0.143　0.137

B 批导线：0.140　0.142　0.136　0.138　0.140

设测定数据分别来自分布 $N(\mu_1,\sigma^2)$, $N(\mu_2,\sigma^2)$，且两样本相互独立，又 μ_1，μ_2，σ^2 均未知。试求 $\mu_1-\mu_2$ 的置信水平为 0.95 的置信区间.

解 将 A 批导线测定数据的均值、方差分别记为 $\bar{x}_1$，s_1^2；B 批的均值、方差分别记为 $\bar{x}_2$，s_2^2。则有：

$$
\begin{aligned}
&n_1=4, &&\bar{x}_1=0.14125, &&3s_1^2=0.00002475,\\
&n_2=5, &&\bar{x}_2=0.1392, &&4s_2^2=0.0000208,
\end{aligned}
$$

$$
s_W^2=\frac{3s_1^2+4s_2^2}{7}=(0.00255)^2, 1-\alpha=0.95, \alpha/2=0.025,
$$
$$
t_{\alpha/2}(n_1+n_2-2)=t_{0.025}(7)=2.3646.
$$

得 $\mu_1-\mu_2$ 的一个置信水平为 0.95 的置信区间：

$$
\left(\bar{x}_1-\bar{x}_2\pm t_{\alpha/2}(n_1+n_2-2)s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)=(0.002\pm0.004)=(-0.002,0.006)
$$

9. 假如 0.50,1.25,0.80,2.00 是来自总体 X 的简单随机样本值, 已知 $Y=\ln X$ 服从正态分布 $N(\mu,1)$.
(1) 求 X 的数学期望 EX(记 EX 为 b);

(2) 求 μ 的置信度为 0.95 的置信区间;

(3) 利用上述结果求 b 的置信度为 0.95 的置信区间.

解（1）由题意知 Y 的概率密度为

$$
f(y)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(y-\mu)^2}{2}}
$$

又由 $Y=\ln X$, 得 $X=\mathrm{e}^Y$, 故

$$
\begin{aligned}
b=EX=E\mathrm{e}^Y&=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^ye^{\frac{(y-\mu)^2}{2}}\ \mathrm{d}y\\
&=e^{\mu+\frac{1}{2}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}[y-(\mu+1)]^2}\ \mathrm{d}y=e^{\mu+\frac{1}{2}}
\end{aligned}
$$

(2) 经过分析, μ 的置信区间公式为 $\left(\bar{Y}-\frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}},\bar{Y}+\frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}}\right)$.

由 $1-\alpha=0.95$, 查表得 $u_{\frac{\alpha}{2}}=1.96$. 代入 $\sigma=1,n=4,\bar{y}=\frac{1}{4}(\ln0.5+\ln1.25+\ln0.8+\ln2)=0$

得 $\left(-\frac{1}{2}\times1.96,\frac{1}{2}\times1.96\right)$

故 μ 的置信度为 0.95 的置信区间为 (-0.98,0.98)

(3) 由（1）可知，b=EX=$e^{\mu+\frac{1}{2}}$, 又由 (2) 知，μ 的置信区间为 (-0.98,0.98) 因为 e^x 为严格增函数，所以 b 的置信区间为 $(e^{-0.98+\frac{1}{2}}\ e^{0.98+\frac{1}{2}})$, 即为 $(e^{-0.48},e^{1.48})$

10. 设样本 $(X_1, X_2, \cdots, X_n)$ 为总体 $X \sim N(\mu, \sigma^2)$ 的样本, 其中 μ, σ^2 为未知参数, 设随机变量 L 是关于 μ 的置信度为 $1-\alpha$ 的置信区间的长度, 求 $E(L^2)$.

解 取 $T = \frac{\bar{X}-\mu}{S}\sqrt{n}$, 则 $T \sim t(n-1)$ 由 $P\left\{|T| < t_{\frac{\alpha}{2}}(n-1)\right\} = 1-\alpha$ 可得 μ 的 $1-\alpha$ 置信区间为

$$\left(\bar{X} - t_{\frac{\alpha}{2}}(n-1) \cdot \frac{S}{\sqrt{n}}, \quad \bar{X} + t_{\frac{\alpha}{2}}(n-1) \cdot \frac{S}{\sqrt{n}}\right)$$

易见区间长度 $L = 2t_{\frac{\alpha}{2}}(n-1) \cdot \frac{S}{\sqrt{n}}$ 故 $E(L^2) = E\left[4t_{\frac{\alpha}{2}}^2(n-1) \cdot \frac{S^2}{n}\right] = \frac{4}{n}t_{\frac{\alpha}{2}}^2(n-1) \cdot \sigma^2$