1.数理统计（哈工大）

33 设总体 $X \sim N(\mu,1$，参数 $\mu \sim N(0,1$， $X_1,\ldots,X_n$ 是来自于总体 X 的样本，并且 $L(\mu,d)=(\mu-d)^2$，求参数 μ 的贝叶斯估计量 $\hat{\mu}$.

解 设 $x=(x_1,\ldots,x_n), X=(X_1,\ldots,X_n)$，先验分布密度 $h(y)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$，

当 $\mu=y$ 时，样本的概率密度分布为

$$f(x|y)=\prod_{i=1}^{n} f(x_i;y)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}}e^{-\frac{(x_i-y)^2}{2}}=(\frac{1}{\sqrt{2\pi}})^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-y)^2}$$

关于参数 μ 的后验分布为

$$\pi(y|x)=\frac{h(y)f(x\mid y)}{g(x)}\propto h(y)f(x\mid y)\propto e^{-\frac{y^2}{2}}e^{-\frac{(x_i-y)^2}{2}}\propto e^{-\frac{n+1}{2}(y-\frac{n\bar{x}}{1+n})^2}$$

μ 的后验分部为 $\mu|x \sim N(\dfrac{n\bar{x}}{1+n},\dfrac{1}{1+n})$，所以关于 μ 的 Bayes 估计量 $\hat{\mu}=\dfrac{n\overline{X}}{1+n}$.

2.数理统计（哈工大）

35 设总体 X 服从几何分布：$P(X=k)=p(1-p)^{k-1}$， $k=1,2,\cdots$，并且参数 $p\sim\beta(\alpha_1,\alpha_2)$，

其中 α_1,α_2 为已知参数 .在平方差损失下，求参数 p 的贝叶斯估计量 T.

解

设 $x=(x_1,\ldots,x_n), X=(X_1,\ldots,X_n)$，

先验分布密度$h(y)=\dfrac{1}{\mathrm{B}(\alpha_1,\alpha_2)}y^{\alpha_1-1}(1-y)^{\alpha_1-1}$

当$p=y$时，样本的概率密度分布为：

$$f(x|y)=\prod_{i=1}^{n}f(x_i;y)=\prod_{i=1}^{n}y(1-y)^{x_i-1}=y^n(1-y)^{n\bar{x}-n}$$

关于参数p的后验分部为

$$\pi(y|x)=\frac{h(y)f(x\mid y)}{g(x)}\propto h(y)f(x\mid y)\propto y^{\alpha_1-1}(1-y)^{\alpha_2-1}y^n(1-y)^{n\bar{x}-n}$$

p的后验分部为 $p|x\square\ \beta(\alpha_1+n,n\bar{X}+\alpha_2-n)$

关于p的 Bayes 估计量$\hat{p}=\dfrac{n+\alpha_1}{\alpha_1+\alpha_2+n\bar{X}}$.

3.数理统计（哈工大）

36 设$X_1,\dots,X_n$为总体$X\sim B(10,p)$的样本，$0<p<1$

1）求参数p是有效估计量T_1与相应的信息量$I(p)$；

2）如果$p\sim U[0,1]$，在平方差损失下，求参数p的贝叶斯估计量T_2.

3）试比较两个估计量T_1和T_2.

解 1）因为似然函数为：

$$L(p;x_1,\dots,x_n)=\prod_{i=1}^{n}f(x_i;p)=\prod_{i=1}^{n}C_{10}^{x_i}p^{x_i}(1-p)^{n-x_i}=p^{\sum_{i=1}^{n}x_i}(1-p)^{10n-\sum_{i=1}^{n}x_i}\prod_{i=1}^{n}C_{10}^{x_i}$$

所以

$$\frac{d}{dp}\ln L(p;x_1,\dots,x_n)=\frac{n\bar{x}}{p}-\frac{10n-n\bar{x}}{1-p}=\frac{n(\bar{x}-10p)}{p(1-p)}=\frac{10n(\frac{\bar{x}}{n}-p)}{p(1-p)}$$

又因为$E\dfrac{1}{n}\bar{X}=p$

所以取$c(p)=\dfrac{10n}{p(p-1)}$，有定理 2.3.2 得 $T_1=\dfrac{1}{n}\bar{X}$是p的有效估计量

$$I(p)=\frac{c(p)g'(p)}{n}=\frac{\frac{10n}{p(1-p)}}{n}=\frac{10}{1-p}$$

2）设$x=(x_1,\dots,x_n),X=(X_1,\dots,X_n)$

先验分布密度$h(y)=1$ 当$p=y$时，样本的概率密度分布为

$$f(x|y)=\prod_{i=1}^{n} f(x_i;y)=\prod_{i=1}^{n} C_{10}^{x_i} y^{x_i}(1-y)^{10-x_i} =\prod_{i=1}^{n} C_{10}^{x_i} y^{n\bar{x}}(1-y)^{10n-n\bar{x}}$$

关于参数 μ 的后验分部为

$$\pi(y|x)=\frac{h(y)f(x\mid y)}{g(x)}\propto h(y)f(x\mid y)\propto y^{n\bar{x}}(1-y)^{10n-n\bar{x}} \quad 0<y<1$$

p 的后验分部为 $p|x$ □ $\beta(1+n\bar{X},10n-n\bar{X}+1)$，关于 p 的 Bayes 估计量

$$T_2=\frac{n\bar{X}+1}{10n+2}$$

（3）比较估计量 T_1,T_2 ,有：$D(T_1)=D(\frac{1}{n}\bar{X})=\frac{1}{n^2}D(\bar{X})$

$$D(T_2)=D(\frac{n\bar{X}+1}{10n+2})=(\frac{n}{10n+2})^2 D(\bar{X})>D(T_1)$$

所以，T_1 优于 T_2

4.高等数理统计（茆诗松）

5.34设随机变量X服从几何分布，即

$$P(X=k)=\theta(1-\theta)^k,k=0,1,2,\ldots$$

其中参数θ的先验分布为均匀分布$U(0,1)$.

(1)若只对X作一次观察，观察值为3，在平方损失函数下求θ的Bayes估计；

(2)若对X三次观察，观察值为2, 3, 5，在平方损失函数下求θ的Bayes估计.

习题解答

解：(1)由题可得$\pi(\theta)=1$，

θ的后验分布为：

$$\pi(\theta|X=k)=\frac{P(X=k|\theta)\pi(\theta)}{\int_0^1 P(X=k|\theta)\pi(\theta)\mathrm{d}\theta}=\frac{\theta(1-\theta)^k}{\int_0^1 \theta(1-\theta)^k\mathrm{d}\theta}=\frac{\Gamma(k+3)}{\Gamma(2)\Gamma(k+1)}\theta(1-\theta)^k$$

由于在平方损失函数下，θ的Bayes估计$\delta^\pi(x)=\mathrm{E}(\theta|x)$，所以

$$\mathrm{E}(\theta|X=k)=\int_0^1 \frac{\Gamma(k+3)}{\Gamma(2)\Gamma(k+1)}\theta^2(1-\theta)^k\mathrm{d}\theta=\frac{\Gamma(k+3)}{\Gamma(2)\Gamma(k+1)}\frac{\Gamma(3)\Gamma(k+1)}{\Gamma(k+4))}=\frac{2}{k+3}$$

将$k=3$代入得θ的Bayes估计为$\hat{\theta}=\frac{1}{3}$

(2)θ的后验分布为：

$$\pi(\theta|X_1=k_1,X_2=k_2,X_3=k_3)=\frac{P(X_1=k_1,X_2=k_2,X_3=k_3|\theta)\pi(\theta)}{\int_0^1 P(X_1=k_1,X_2=k_2,X_3=k_3|\theta)\pi(\theta)\mathrm{d}\theta}=\frac{P(X_1=k_1|\theta)P(X_2=k_2|\theta)P(X_3=k_3|\theta)}{\int_0^1 P(X_1=k_1|\theta)P(X_2=k_2|\theta)P(X_3=k_3|\theta)\mathrm{d}\theta}=\frac{\theta^3(1-\theta)^k}{\int_0^1 \theta^3(1-\theta)^i}$$

其中$a=k_1+k_2+k_3$.

因此

$$\mathrm{E}(\theta|X_1=k_1,X_2=k_2,X_3=k_3)=\int_0^1 \frac{\Gamma(a+4)}{\Gamma(4)\Gamma(a+1)}\theta^4(1-\theta)^a\mathrm{d}\theta=\frac{\Gamma(a+4)}{\Gamma(4)\Gamma(a+1)}\frac{\Gamma(5)\Gamma(a+1)}{\Gamma(a+6)}=\frac{4}{a+5}$$

将$k_1=2,k_2=3,k_3=5$ 代入得θ的Bayes估计为$\hat{\theta}=\frac{4}{15}$

5.高等数理统计（茆诗松）

设为一顾客的服务时间（单位：分）服从指数分布$Exp(\lambda)$，其中λ未知，又设λ的先验分布是均值为0.2，标准差为1的$Gamma$分布。如今对20位顾客服务，平均时间是3.8分钟。在平方损失函数下，分别求λ和$\theta=\lambda^{-1}$的$Bayes$估计。

由题设知样本X的联合条件密度函数为

$$p(x|\lambda)=Exp(\lambda)=\lambda e^{-\lambda x}$$

又有λ的先验分布为

$$\pi(\lambda)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda},$$

且有其期望$\frac{\alpha}{\beta}=0.2$，方差$\frac{\alpha}{\beta^2}=1$，解之得$\alpha=0.04,\beta=0.2$。
则有λ的后验分布

$$\begin{aligned}\pi(\lambda|x)&=\frac{p(x|\lambda)\pi(\lambda)}{\int_{\Theta}p(x|\lambda)\pi(\lambda)d\lambda}\\&=\frac{\frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha}e^{-(\beta+x)\lambda}}{\int_0^{\infty}\frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha}e^{-(\beta+x)\lambda}d\lambda}\end{aligned}$$

其中

$$\begin{aligned}&\int_0^{\infty}\frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha}e^{-(\beta+x)\lambda}d\lambda\\=&\frac{\beta^{\alpha}}{\Gamma(\alpha)(\beta+x)^{\alpha+1}}\int_0^{\infty}(\beta+x)^{\alpha}\lambda^{\alpha}e^{-(\beta+x)\lambda}d(\beta+x)\lambda\\=&\frac{\beta^{\alpha}}{\Gamma(\alpha)(\beta+x)^{\alpha+1}}\Gamma(\alpha+1).\end{aligned}$$

所以λ的后验分布为

$$\pi(\lambda|x)=\frac{\lambda^{\alpha}}{\Gamma(\alpha+1)}(\beta+x)^{\alpha+1}e^{-(\beta+x)\lambda}.$$

由题设$x0=3.8$，则由定理5.6可知，λ的$Bayes$估计为

$$\begin{aligned}\delta^{\pi}(x0)&=\int_0^{\infty}\frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)}(\beta+x0)^{\alpha+1}e^{-(\beta+x0)\lambda}d\lambda\\&=\frac{1}{\Gamma(\alpha+1)(\beta+x0)}\int_0^{\infty}(\beta+x0)^{\alpha+1}\lambda^{\alpha+1}e^{-(\beta+x0)\lambda}d(\beta+x0)\lambda\\&=\frac{1}{\Gamma(\alpha+1)(\beta+x0)}\Gamma(\alpha+2)\\&=\frac{\alpha+1}{\beta+x0}\\&=\frac{0.04+1}{0.2+3.8}\\&=0.26.\end{aligned}$$

进一步有$\hat{\theta}=\hat{\lambda}^{-1}\approx3.85$。

6.高等数理统计（茆诗松）

设θ是不合格品率，它的先验分布为$Beta$分布$Be(5,10)$.

(1)假设随机检查20个产品，只发现1个不合格品。在平方损失函数下求θ的$Bayes$估计；

解(1)：

θ的先验分布为$Beta(\alpha,\beta)$，则：

$$x|\theta \sim Bin(n,\theta)$$

$\therefore$

$$\begin{aligned}h(x,\theta) &= \frac{n!}{x!(n-x)!}\theta^{x}(1-\theta)^{n-x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} \\ &= \frac{\Gamma(\alpha+\beta)n!}{\Gamma(\alpha)\Gamma(\beta)x!(n-x)!}\theta^{x+\alpha-1}(1-\theta)^{n+\beta-x-1} \\ &= \theta^{x+\alpha-1}(1-\theta)^{n+\beta-x-1}\end{aligned}$$

$\therefore$

$$\pi(\theta|x) = \frac{\Gamma(\alpha^*+\beta^*)}{\Gamma(\alpha^*)\Gamma(\beta^*)}\theta^{\alpha^*-1}(1-\theta)^{\beta^*-1}$$

其中$\alpha^* = x+\alpha,\qquad \beta^* = n+\beta-x$　则

$$E[\pi(\theta|x)] = \frac{\alpha^*}{\alpha^*+\beta^*}$$

在本题中，$\alpha=5,\beta=10,n=20,x=1$　而根据定理5.6易知，在平方损失下θ的贝叶斯估计：

$$\pi(x) = E[\pi(\theta|x)] = \frac{x+\alpha}{n+\alpha+\beta} = \frac{6}{35}$$

7.高等数理统计（茆诗松）

习题5.48

设有一批产品，其不合格率为p。若将每N件都装为一箱。现从一箱中随机抽检n件产品，得知不合格产品数是r。试求这箱的不合格率$q=\omega/N$的$Bayes$估计。其中ω是这箱中的不合格品数，假如取平方损失函数$L(\omega,\delta)=(\delta-\omega/N)^2$。

解：
该样本的联合密度函数为：

$$P(x|q) = \mathrm{C}_n^r q^r (1-q)^{n-r}$$

取$Be(a,b)$为先验分布，$a>0, b>0$ 为两个待定参数，即先验分布密度为：

$$\pi(q;a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} q^{a-1}(1-q)^{b-1}$$

故q的期望为$E(q) = \frac{a}{a+b}$，方差为：

$$var(q) = \frac{ab}{(a+b)^2(a+b+1)}$$

易见其后验分布为$Be(a+r, b+n-r)$。若取平方损失函数，则不合格率$q=\omega/N$ 的$Bayes$ 估计为后验分布的期望，即：

$$\hat{q} = \frac{a+r}{a+b+n}$$

先验信息只有不合格率为p，则可令$p=E(q)=\frac{a}{a+b}$ 。同时对p比较确信，故可试验不同的(a,b)组合，使$p=\frac{a}{a+b}$ ，且$a+b$较大，而$\frac{ab}{(a+b)^2(a+b+1)}$ 较小即可。设确定的a为$\hat{a}$，b为$\hat{b}$ ，且$\hat{a}$与$\hat{b}$都与p有关，则不合格率q的$Bayes$ 估计为：

$$\hat{q} = \frac{\hat{a}+r}{\hat{a}+\hat{b}+n}$$

8.高等数理统计（茆诗松）

习题5.51

对正态分布$N(\theta,1)$观察，获得三个独立观察值$X_1=2, X_2=3, X_3=4$，若θ的先验分布为$N(3,1)$，求θ的0.95可信区间。

解

例题中已经求出θ的后验分布为$N(\mu_1,\sigma_1^2)$，其中$\mu_1 = \frac{\sigma_0^{-2}\bar{x}+\tau^{-2}\mu}{\sigma_0^{-2}+\tau^{-2}}, \sigma_1^2 = (\sigma_0^{-2}+\tau^{-2})^{-1}, \sigma_0^2 = \sigma^2/n.$

在题目中，$X\sim N(\theta,1), \theta\sim N(3,1). n=3, \bar{x}=3.$

将σ^2,τ^2代入得，$\mu_1 = \frac{n\bar{x}+3}{n+1}, \sigma_1^2 = \frac{1}{n+2} = 1/4, \sigma_1 = 1/2.$

θ的$Bayes$估计为$\hat{\theta} = \mu_1 = 3.$

可信区间为$[\mu_1 - \sigma_1 u_{0.975}, \mu_1 + \sigma_1 u_{0.975}], u_{0.975} = 1.96,$**即$\theta$的0.95可信区间为**$[2.02, 3.98].$

9.高等数理统计（茆诗松）

题目

设x是从二项分布$b(n,\theta)$中抽取得一个样本，先考察如下两个假设

$$H_0:\theta=\frac{1}{2}\,,\quad H_1=\theta\neq\frac{1}{2}$$

若设在$\theta\neq\frac{1}{2}$ 上的先验密度$g_1(\theta)$为区间$(0,1)$上的均匀分布$U(0,1)$.

(1)求 Bayes 因子;

(2)求原假设$H_0:\theta=\theta_0=\frac{1}{2}$ 的后验概率;

(3)若取$\pi_0=\frac{1}{2}\,,n=5,x=3$ ，请做出判断.

解答

(1)

我们不妨取H_0 和H_1 的先验概率为π_0 和π_1

样本分布为$p(x|\theta)=\binom{n}{x}\theta^x(1-\theta)^{(n-x)}$，$\theta$的先验密度为$\pi(\theta)=\pi_0 I_{\theta=1/2}+\pi_1 g_1(\theta)$，其中$g_1(\theta)$为均匀分布$U(0,1)$，即$g_1(\theta)=1$

由于

$$m_1(x)=\int_{\theta\neq\theta_0}p(x|\theta)g_1(\theta)d\theta=\binom{n}{x}\int_0^1\theta^x(1-\theta)^{n-x}=\binom{n}{x}\frac{\Gamma(x+1)\Gamma(n-x+1)}{\Gamma(n+2)}$$

从而Bayes因子为

$$B^\pi(x)=\frac{p(x|\theta_0)}{m_1(x)}=\frac{0.5^x(1-0.5)^{n-x}}{\frac{\Gamma(x+1)\Gamma(n-x+1)}{\Gamma(n+2)}}$$

(2)

原假设H_0 的后验概率为

$$\pi(\Theta_0|x)=\frac{\pi_0 p(x|\theta_0)}{m(x)}=[1+\frac{(1-\pi_0)}{\pi_0 B^\pi(x)}]^{-1}$$

(3)

有题目数据得

$$B^{\pi}(x)=\frac{0.5^{x}(1-0.5)^{n-x}}{\frac{\Gamma(x+1)\Gamma(n-x+1)}{\Gamma(n+2)}}=\frac{15}{8}$$

则后验概率为

$$\frac{\alpha_0}{\alpha_1}=\frac{\pi_0}{\pi_1}B^{\pi}(x)=\frac{15}{23}$$

则$\frac{15}{23}<1$，故不拒绝原假设.

10.概率论与数理统计 （茆诗松）

例 10.15 设随机变量 X 服从几何分布，其分布列为

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

其中参数 θ 只能取 1/4，2/4 和 3/4 三个值，并以相同概率取这三个值。如今只获得一个观察值 $X=2$，要在 0—1 损失函数下寻求 θ 的贝叶斯估计。

解：在这个问题中，$X=2$ 的条件概率为

$$P(X=2|\theta)=\theta(1-\theta)^{2}。$$

θ 的先验分布为

$$P(\theta=i/4)=1/3,\quad i=1,2,3。$$

于是联合概率

$$P(X=2,\theta=i/4)=\frac{1}{3}\cdot\frac{i}{4}\left(1-\frac{i}{4}\right)^{2}。$$

$X=2$ 的无条件概率为

$$P(X=2)=\sum_{i=1}^{3}P(X=2,\theta=i/4)=\frac{1}{3}\left[\frac{1}{4}\left(\frac{3}{4}\right)^{2}+\frac{2}{4}\left(\frac{2}{4}\right)^{2}+\frac{3}{4}\left(\frac{1}{4}\right)^{2}\right]=\frac{5}{48}。$$

于是在 $X=2$ 条件下，θ 的后验分布为

$$P(\theta=i/4|x=2)=\frac{P(x=2,\theta=i/4)}{P(x=2)}=\frac{4i}{5}\left(1-\frac{i}{4}\right)^{2}。$$

它的分布列为

	1/4	2/4	3/4
$P(\theta=i/4\|x=2)$	$\frac{9}{20}$	$\frac{8}{20}$	$\frac{3}{20}$

根据定理 10.3 知，在 0—1 损失函数下，θ 的贝叶斯估计应是 $\hat{\theta}=1/4$，因为这是后验分布的众数。