# 习题 6.5

1．设一页书上的错别字个数服从泊松分布 $P(\lambda)$，有两个可能取值：1.5 和 1.8，且先验分布为

$$P\{\lambda = 1.5\} = 0.45,\quad P\{\lambda = 1.8\} = 0.55,$$

现检查了一页，发现有 3 个错别字，试求$\lambda$ 的后验分布．

解：总体 $X$ 表示一页书上的错别字个数，$X \sim P(\lambda)$，样本为 $X_1 = 3$，有 $P\{X_1 = k\} = \dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$，$k = 0, 1, 2, \cdots$，

则 $P\{X_1 = 3\} = P\{\lambda = 1.5\}P\{X_1 = 3 \mid \lambda = 1.5\} + P\{\lambda = 1.8\}P\{X_1 = 3 \mid \lambda = 1.8\}$

$$= 0.45 \times \frac{1.5^3}{6} \cdot \mathrm{e}^{-1.5} + 0.55 \times \frac{1.8^3}{6} \cdot \mathrm{e}^{-1.8} = 0.0565 + 0.0884 = 0.1449,$$

故参数$\lambda$ 的后验分布为 $P\{\lambda = 1.5 \mid X_1 = 3\} = \dfrac{P\{\lambda = 1.5\}P\{X_1 = 3 \mid \lambda = 1.5\}}{P\{X_1 = 3\}} = \dfrac{0.0565}{0.1449} = 0.3899$，

$$P\{\lambda = 1.8 \mid X_1 = 3\} = \frac{P\{\lambda = 1.8\}P\{X_1 = 3 \mid \lambda = 1.8\}}{P\{X_1 = 3\}} = \frac{0.0884}{0.1449} = 0.6101.$$

2．设总体为均匀分布 $U(\theta, \theta+1)$，$\theta$ 的先验分布是均匀分布 $U(10, 16)$．现有三个观测值：11.7, 12.1, 12.0．求 $\theta$ 的后验分布．

解：参数$\theta$ 的先验分布为 $\pi(\theta) = \dfrac{1}{6}\mathrm{I}_{10<\theta<16}$，

总体 $X$ 的条件分布为 $p(x \mid \theta) = \mathrm{I}_{\theta < x < \theta+1}$，

有样本 $X_1, X_2, X_3$ 的联合条件分布为 $p(x_1, x_2, x_3 \mid \theta) = \mathrm{I}_{\theta<x_1, x_2, x_3<\theta+1}$，

则样本 $X_1, X_2, X_3$ 和参数$\theta$ 的联合分布为

$$h(x_1, x_2, x_3, \theta) = \frac{1}{6}\mathrm{I}_{\theta<x_1, x_2, x_3<\theta+1,\, 10<\theta<16} = \frac{1}{6}\mathrm{I}_{x_{(3)}-1<\theta<x_{(1)},\, 10<\theta<16},$$

可得样本 $X_1, X_2, X_3$ 的边际分布为 $m(x_1, x_2, x_3) = \displaystyle\int_{-\infty}^{+\infty} \frac{1}{6}\mathrm{I}_{x_{(3)}-1<\theta<x_{(1)},\, 10<\theta<16} d\theta = \int_{11.1}^{11.7} \frac{1}{6} d\theta = 0.1$，

故参数$\theta$ 的后验分布为 $\pi(\theta \mid x_1, x_2, x_3) = \dfrac{h(x_1, x_2, x_3, \theta)}{m(x_1, x_2, x_3)} = \dfrac{5}{3}\mathrm{I}_{11.1<\theta<11.7}$．

3．设 $X_1, \cdots, X_n$ 是来自几何分布的样本，总体分布列为

$$P\{X = k \mid \theta\} = \theta(1-\theta)^k,\quad k = 0, 1, 2, \cdots,$$

$\theta$ 的先验分布是均匀分布 $U(0, 1)$．

（1）求$\theta$ 的后验分布；

（2）若 4 次观测值为 4, 3, 1, 6，求$\theta$ 的贝叶斯估计．

解：（1）参数$\theta$ 的先验分布为 $\pi(\theta) = \mathrm{I}_{0<\theta<1}$，

因样本 $X_1, \cdots, X_n$ 的联合条件分布为

$$p(x_1, \cdots, x_n \mid \theta) = \prod_{i=1}^{n} \theta(1-\theta)^{x_i} = \theta^n (1-\theta)^{x_1+\cdots+x_n},\quad x_1, \cdots, x_n = 0, 1, 2, \cdots,$$

则样本 $X_1, \cdots, X_n$ 和参数$\theta$ 的联合分布为

$$h(x_1, \cdots, x_n, \theta) = \theta^n (1-\theta)^{x_1+\cdots+x_n}\mathrm{I}_{0<\theta<1},\quad x_1, \cdots, x_n = 0, 1, 2, \cdots,$$

样本 $X_1, \cdots, X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^1\theta^n(1-\theta)^{x_1+\cdots+x_n}d\theta=\frac{\Gamma(n+1)\Gamma(x_1+\cdots+x_n+1)}{\Gamma(n+x_1+\cdots+x_n+2)},\quad x_1,\cdots,x_n=0,1,2,\cdots,$$

故参数$\theta$的后验分布为

$$\pi(\theta\,|\,x_1,\cdots,x_n)=\frac{h(x_1,\cdots,x_n,\theta)}{m(x_1,\cdots,x_n)}=\frac{\Gamma(n+x_1+\cdots+x_n+2)}{\Gamma(n+1)\Gamma(x_1+\cdots+x_n+1)}\theta^n(1-\theta)^{x_1+\cdots+x_n}\mathrm{I}_{0<\theta<1};$$

（2）因 $E(\theta\,|\,x_1,\cdots,x_n)=\int_0^1\theta\cdot\pi(\theta\,|\,x_1,\cdots,x_n)d\theta=\dfrac{\Gamma(n+x_1+\cdots+x_n+2)}{\Gamma(n+1)\Gamma(x_1+\cdots+x_n+1)}\int_0^1\theta^{n+1}(1-\theta)^{x_1+\cdots+x_n}d\theta$

$$=\frac{\Gamma(n+x_1+\cdots+x_n+2)}{\Gamma(n+1)\Gamma(x_1+\cdots+x_n+1)}\cdot\frac{\Gamma(n+2)\Gamma(x_1+\cdots+x_n+1)}{\Gamma(n+x_1+\cdots+x_n+3)}=\frac{n+1}{n+x_1+\cdots+x_n+2},$$

则贝叶斯估计 $\hat{\theta}_B=E(\theta\,|\,X_1,\cdots,X_n)=\dfrac{n+1}{n+X_1+\cdots+X_n+2}$，

因样本观测值为 4, 3, 1, 6，即 $x_1+\cdots+x_n=15$，$n=4$，

故 $\hat{\theta}_B=\dfrac{4+1}{4+14+2}=\dfrac{1}{4}$.

4. 验证：泊松分布的均值$\lambda$的共轭先验分布是伽玛分布.

证：设参数$\lambda$的先验分布是伽玛分布 $Ga(\alpha,\beta)$，密度函数为 $\pi(\lambda)=\dfrac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}\mathrm{e}^{-\beta\lambda}\mathrm{I}_{\lambda>0}$，

因样本 $X_1,\cdots,X_n$ 的联合条件分布为

$$p(x_1,\cdots,x_n\,|\,\lambda)=\prod_{i=1}^n\frac{\lambda^{x_i}}{x_i!}\mathrm{e}^{-\lambda}=\frac{\lambda^{x_1+\cdots+x_n}}{x_1!\cdots x_n!}\mathrm{e}^{-n\lambda},\quad x_1,\cdots,x_n=0,1,2,\cdots,$$

则样本 $X_1,\cdots,X_n$ 和参数$\lambda$的联合分布为

$$h(x_1,\cdots,x_n,\lambda)=\frac{\beta^\alpha\lambda^{x_1+\cdots+x_n+\alpha-1}}{\Gamma(\alpha)x_1!\cdots x_n!}\mathrm{e}^{-(n+\beta)\lambda}\mathrm{I}_{\lambda>0},\quad x_1,\cdots,x_n=0,1,2,\cdots,$$

样本 $X_1,\cdots,X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^{+\infty}\frac{\beta^\alpha\lambda^{x_1+\cdots+x_n+\alpha-1}}{\Gamma(\alpha)x_1!\cdots x_n!}\mathrm{e}^{-(n+\beta)\lambda}d\lambda=\frac{\beta^\alpha}{\Gamma(\alpha)x_1!\cdots x_n!}\int_0^{+\infty}\lambda^{x_1+\cdots+x_n+\alpha-1}\mathrm{e}^{-(n+\beta)\lambda}d\lambda$$

$$=\frac{\beta^\alpha}{\Gamma(\alpha)x_1!\cdots x_n!}\cdot\frac{\Gamma(x_1+\cdots+x_n+\alpha)}{(n+\beta)^{x_1+\cdots+x_n+\alpha}},\quad x_1,\cdots,x_n=0,1,2,\cdots,$$

即参数$\lambda$的后验分布为

$$\pi(\lambda\,|\,x_1,\cdots,x_n)=\frac{h(x_1,\cdots,x_n,\lambda)}{m(x_1,\cdots,x_n)}=\frac{(n+\beta)^{x_1+\cdots+x_n+\alpha}}{\Gamma(x_1+\cdots+x_n+\alpha)}\lambda^{x_1+\cdots+x_n+\alpha-1}\mathrm{e}^{-(n+\beta)\lambda}\mathrm{I}_{\lambda>0},$$

后验分布仍为伽玛分布 $Ga(x_1+\cdots+x_n+\alpha,\,n+\beta)$，
故伽玛分布是泊松分布的均值$\lambda$的共轭先验分布.

5. 验证：正态总体方差（均值已知）的共轭先验分布是倒伽玛分布.

证：设参数$\sigma^2$的先验分布是倒伽玛分布$IGa(\alpha,\lambda)$，密度函数为$\pi(\sigma^2)=\dfrac{\lambda^\alpha}{\Gamma(\alpha)}\left(\dfrac{1}{\sigma^2}\right)^{\alpha+1}e^{-\frac{\lambda}{\sigma^2}}$，

又设总体分布为$N(\mu_0,\sigma^2)$，其中$\mu_0$已知，密度函数为$p(x\,|\,\sigma^2)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu_0)^2}{2\sigma^2}}$，

有样本$X_1,\cdots,X_n$的联合条件分布为

$$p(x_1,\cdots,x_n\,|\,\sigma^2)=\prod_{i=1}^n\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\mu_0)^2}{2\sigma^2}}=\frac{1}{(\sqrt{2\pi})^n\sigma^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu_0)^2},$$

则样本$X_1,\cdots,X_n$和参数$\sigma^2$的联合分布为

$$h(x_1,\cdots,x_n,\sigma^2)=\frac{\lambda^\alpha}{(\sqrt{2\pi})^n\Gamma(\alpha)}\cdot\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+\alpha+1}e^{-\frac{1}{\sigma^2}\left[\lambda+\frac{1}{2}\sum_{i=1}^n(x_i-\mu_0)^2\right]},$$

样本$X_1,\cdots,X_n$的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^{+\infty}\frac{\lambda^\alpha}{(\sqrt{2\pi})^n\Gamma(\alpha)}\cdot\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+\alpha+1}e^{-\frac{1}{\sigma^2}\left[\lambda+\frac{1}{2}\sum_{i=1}^n(x_i-\mu_0)^2\right]}d(\sigma^2)$$

$$=\frac{\lambda^\alpha}{(\sqrt{2\pi})^n\Gamma(\alpha)}\cdot\int_{+\infty}^0 t^{\frac{n}{2}+\alpha+1}e^{-t\left[\lambda+\frac{1}{2}\sum_{i=1}^n(x_i-\mu_0)^2\right]}\left(-\frac{1}{t^2}\right)dt$$

$$=\frac{\lambda^\alpha}{(\sqrt{2\pi})^n\Gamma(\alpha)}\cdot\int_0^{+\infty} t^{\frac{n}{2}+\alpha-1}e^{-t\left[\lambda+\frac{1}{2}\sum_{i=1}^n(x_i-\mu_0)^2\right]}dt=\frac{\lambda^\alpha}{(\sqrt{2\pi})^n\Gamma(\alpha)}\cdot\frac{\Gamma\left(\frac{n}{2}+\alpha\right)}{\left[\lambda+\frac{1}{2}\sum_{i=1}^n(x_i-\mu_0)^2\right]^{\frac{n}{2}+\alpha}},$$

即参数$\sigma^2$的后验分布为

$$\pi(\sigma^2\,|\,x_1,\cdots,x_n)=\frac{\left[\lambda+\frac{1}{2}\sum_{i=1}^n(x_i-\mu_0)^2\right]^{\frac{n}{2}+\alpha}}{\Gamma\left(\frac{n}{2}+\alpha\right)}\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+\alpha+1}e^{-\frac{1}{\sigma^2}\left[\lambda+\frac{1}{2}\sum_{i=1}^n(x_i-\mu_0)^2\right]},$$

后验分布仍为倒伽玛分布$IGa\left(\dfrac{n}{2}+\alpha,\lambda+\dfrac{1}{2}\sum_{i=1}^n(x_i-\mu_0)^2\right)$，

故倒伽玛分布是参数$\sigma^2$的共轭先验分布．

6．设$X_1,\cdots,X_n$是来自如下总体的一个样本，

$$p(x\,|\,\theta)=\frac{2x}{\theta^2},\quad 0<x<\theta.$$

（1）若$\theta$的先验分布为均匀分布$U(0,1)$，求$\theta$的后验分布；

（2）若$\theta$的先验分布为$\pi(\theta)=3\theta^2$，$0<\theta<1$，求$\theta$的后验分布．

解：样本$X_1,\cdots,X_n$的联合条件分布为

$$p(x_1,\cdots,x_n\,|\,\theta)=\prod_{i=1}^n\frac{2x_i}{\theta^2}I_{0<x_i<\theta}=\frac{2^n x_1\cdots x_n}{\theta^{2n}}I_{0<x_1,\cdots,x_n<\theta},$$

（1）因参数$\theta$的先验分布为$\pi(\theta)=I_{0<\theta<1}$，

则样本 $X_1,\cdots,X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{2^n x_1\cdots x_n}{\theta^{2n}}\mathrm{I}_{0<x_1,\cdots,x_n<\theta<1}=\frac{2^n x_1\cdots x_n}{\theta^{2n}}\mathrm{I}_{x_1,\cdots,x_n>0,\,x_{(n)}<\theta<1},$$

样本 $X_1,\cdots,X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_{x_{(n)}}^1\frac{2^n x_1\cdots x_n}{\theta^{2n}}\mathrm{I}_{x_1,\cdots,x_n>0}d\theta=\frac{2^n x_1\cdots x_n}{2n-1}[x_{(n)}^{-(2n-1)}-1]\cdot\mathrm{I}_{x_1,\cdots,x_n>0},$$

故参数 $\theta$ 的后验分布为

$$\pi(\theta\,|\,x_1,\cdots,x_n)=\frac{h(x_1,\cdots,x_n,\theta)}{m(x_1,\cdots,x_n)}=\frac{2n-1}{\theta^{2n}[x_{(n)}^{-(2n-1)}-1]}\mathrm{I}_{x_{(n)}<\theta<1};$$

（2）因参数 $\theta$ 的先验分布为 $\pi(\theta)=3\theta^2\,\mathrm{I}_{0<\theta<1}$，

则样本 $X_1,\cdots,X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{3\cdot 2^n x_1\cdots x_n}{\theta^{2n-2}}\mathrm{I}_{0<x_1,\cdots,x_n<\theta<1}=\frac{3\cdot 2^n x_1\cdots x_n}{\theta^{2n-2}}\mathrm{I}_{x_1,\cdots,x_n>0,\,x_{(n)}<\theta<1},$$

样本 $X_1,\cdots,X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_{x_{(n)}}^1\frac{3\cdot 2^n x_1\cdots x_n}{\theta^{2n-2}}\mathrm{I}_{x_1,\cdots,x_n>0}d\theta=\frac{3\cdot 2^n x_1\cdots x_n}{2n-3}[x_{(n)}^{-(2n-3)}-1]\cdot\mathrm{I}_{x_1,\cdots,x_n>0},$$

故参数 $\theta$ 的后验分布为

$$\pi(\theta\,|\,x_1,\cdots,x_n)=\frac{h(x_1,\cdots,x_n,\theta)}{m(x_1,\cdots,x_n)}=\frac{2n-3}{\theta^{2n-2}[x_{(n)}^{-(2n-3)}-1]}\mathrm{I}_{x_{(n)}<\theta<1}.$$

7. 设 $X_1,\cdots,X_n$ 是来自如下总体的一个样本，

$$p(x\,|\,\theta)=\theta x^{\theta-1},\quad 0<x<1.$$

若取 $\theta$ 的先验分布为伽玛分布，即 $\theta\sim Ga(\alpha,\lambda)$，求 $\theta$ 的后验期望估计.

解：参数 $\theta$ 的先验分布为 $Ga(\alpha,\lambda)$，密度函数为 $\pi(\theta)=\dfrac{\lambda^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}\mathrm{e}^{-\lambda\theta}\,\mathrm{I}_{\theta>0}$，

因样本 $X_1,\cdots,X_n$ 的联合条件分布为

$$p(x_1,\cdots,x_n\,|\,\theta)=\prod_{i=1}^n\theta x_i^{\theta-1}\mathrm{I}_{0<x_i<1}=\theta^n(x_1\cdots x_n)^{\theta-1}\mathrm{I}_{0<x_1,\cdots,x_n<1}=\theta^n\mathrm{e}^{(\theta-1)\ln(x_1\cdots x_n)}\,\mathrm{I}_{0<x_1,\cdots,x_n<1},$$

则样本 $X_1,\cdots,X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{\lambda^\alpha}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}\theta^{n+\alpha-1}\,\mathrm{e}^{-[\lambda-\ln(x_1\cdots x_n)]\theta}\,\mathrm{I}_{0<x_1,\cdots,x_n<1,\,\theta>0},$$

样本 $X_1,\cdots,X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^{+\infty}\frac{\lambda^\alpha}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}\theta^{n+\alpha-1}\,\mathrm{e}^{-[\lambda-\ln(x_1\cdots x_n)]\theta}\,\mathrm{I}_{0<x_1,\cdots,x_n<1}d\theta$$

$$=\frac{\lambda^\alpha}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}\cdot\frac{\Gamma(n+\alpha)}{[\lambda-\ln(x_1\cdots x_n)]^{n+\alpha}}\mathrm{I}_{0<x_1,\cdots,x_n<1},$$

即参数 $\theta$ 的后验分布为

$$\pi(\theta\,|\,x_1,\cdots,x_n)=\frac{h(x_1,\cdots,x_n,\theta)}{m(x_1,\cdots,x_n)}=\frac{[\lambda-\ln(x_1\cdots x_n)]^{n+\alpha}}{\Gamma(n+\alpha)}\theta^{n+\alpha-1}\,\mathrm{e}^{-[\lambda-\ln(x_1\cdots x_n)]\theta}\,\mathrm{I}_{\theta>0},$$

后验分布仍为伽玛分布 $Ga(n+\alpha, \lambda-\ln(x_1\cdots x_n))$，

因 $E(\theta|x_1,\cdots,x_n)=\int_0^1\theta\cdot\pi(\theta|x_1,\cdots,x_n)d\theta=\dfrac{[\lambda-\ln(x_1\cdots x_n)]^{n+\alpha}}{\Gamma(n+\alpha)}\int_0^1\theta^{n+\alpha}\mathrm{e}^{-[\lambda-\ln(x_1\cdots x_n)]\theta}d\theta$

$$=\frac{[\lambda-\ln(x_1\cdots x_n)]^{n+\alpha}}{\Gamma(n+\alpha)}\cdot\frac{\Gamma(n+\alpha+1)}{[\lambda-\ln(x_1\cdots x_n)]^{n+\alpha+1}}=\frac{n+\alpha}{\lambda-\ln(x_1\cdots x_n)},$$

故参数 $\theta$ 的后验期望估计 $\hat{\theta}_B=\dfrac{n+\alpha}{\lambda-\ln(X_1\cdots X_n)}$.

8. 设 $X_1,\cdots,X_n$ 是来自均匀分布 $U(0,\theta)$ 的样本，$\theta$ 的先验分布是帕雷托（Pareto）分布，密度函数为 $\pi(\theta)=\dfrac{\beta\theta_0^\beta}{\theta^{\beta+1}}$，$\theta>\theta_0$，其中 $\beta,\theta_0$ 是两个已知的常数.

（1）验证：帕雷托分布是 $\theta$ 的共轭先验分布；

（2）求 $\theta$ 的贝叶斯估计.

解：（1）参数 $\theta$ 的先验分布是帕雷托分布，密度函数为 $\pi(\theta)=\dfrac{\beta\theta_0^\beta}{\theta^{\beta+1}}\mathrm{I}_{\theta>\theta_0}$，

因样本 $X_1,\cdots,X_n$ 的联合条件分布为

$$p(x_1,\cdots,x_n|\theta)=\prod_{i=1}^n\frac{1}{\theta}\mathrm{I}_{0<x_i<\theta}=\frac{1}{\theta^n}\mathrm{I}_{0<x_1,\cdots,x_n<\theta},$$

则样本 $X_1,\cdots,X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{\beta\theta_0^\beta}{\theta^{n+\beta+1}}\mathrm{I}_{0<x_1,\cdots,x_n<\theta,\ \theta>\theta_0}=\frac{\beta\theta_0^\beta}{\theta^{n+\beta+1}}\mathrm{I}_{x_1,\cdots,x_n>0,\ \theta>\max\{x_1,\cdots,x_n,\theta_0\}},$$

样本 $X_1,\cdots,X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_{\max\{x_1,\cdots,x_n,\theta_0\}}^{+\infty}\frac{\beta\theta_0^\beta}{\theta^{n+\beta+1}}\mathrm{I}_{x_1,\cdots,x_n>0}d\theta=\beta\theta_0^\beta\cdot\frac{1}{(n+\beta)[\max\{x_1,\cdots,x_n,\theta_0\}]^{n+\beta}}\mathrm{I}_{x_1,\cdots,x_n>0},$$

即参数 $\theta$ 的后验分布为

$$\pi(\theta|x_1,\cdots,x_n)=\frac{(n+\beta)[\max\{x_1,\cdots,x_n,\theta_0\}]^{n+\beta}}{\theta^{n+\beta+1}}\mathrm{I}_{\theta>\max\{x_1,\cdots,x_n,\theta_0\}},$$

后验分布仍为帕雷托分布，其参数为 $n+\beta$ 和 $\max\{x_1,\cdots,x_n,\theta_0\}$，

故帕雷托分布是参数 $\theta$ 的共轭先验分布；

（2）因 $E(\theta|x_1,\cdots,x_n)=\int_{\max\{x_1,\cdots,x_n,\theta_0\}}^{+\infty}\theta\cdot\pi(\theta|x_1,\cdots,x_n)d\theta$

$$=\int_{\max\{x_1,\cdots,x_n,\theta_0\}}^{+\infty}\frac{(n+\beta)[\max\{x_1,\cdots,x_n,\theta_0\}]^{n+\beta}}{\theta^{n+\beta}}d\theta$$

$$=(n+\beta)[\max\{x_1,\cdots,x_n,\theta_0\}]^{n+\beta}\cdot\frac{[\max\{x_1,\cdots,x_n,\theta_0\}]^{-(n+\beta)+1}}{n+\beta-1}=\frac{n+\beta}{n+\beta-1}\max\{x_1,\cdots,x_n,\theta_0\},$$

故 $\theta$ 的贝叶斯估计 $\hat{\theta}_B=\dfrac{n+\beta}{n+\beta-1}\max\{X_1,\cdots,X_n,\theta_0\}$.

9. 设指数分布 $Exp(\theta)$ 中未知参数 $\theta$ 的先验分布为伽玛分布 $Ga(\alpha,\lambda)$，现从先验信息得知：先验均值为 0.0002，先验标准差为 0.01，试确定先验分布.

解：因伽玛分布 $Ga(\alpha,\lambda)$ 密度函数为 $\pi(\theta)=\dfrac{\lambda^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}\mathrm{e}^{-\lambda\theta}\mathrm{I}_{\theta>0}$，

则由 $E(\theta)=\dfrac{\alpha}{\lambda}=0.0002$， $\mathrm{Var}(\theta)=\dfrac{\alpha}{\lambda^2}=(0.01)^2=0.0001$，解得 $\lambda=2$，$\alpha=0.0004$，

故参数 $\theta$ 的先验分布为伽玛分布 $Ga\,(0.0004, 2)$.

10．设 $X_1, \cdots, X_n$ 为来自如下幂级数分布的样本，总体分布密度为

$$p(x_1;c,\theta)=cx_1^{c-1}\theta^{-c}\mathrm{I}_{0\le x_1\le\theta}\quad (c>0,\theta>0),$$

（1）证明：若 $c$ 已知，则 $\theta$ 的共轭先验分布为帕雷托分布；

（2）若 $\theta$ 已知，则 $c$ 的共轭先验分布为伽玛分布.

证：样本 $X_1, \cdots, X_n$ 的联合条件分布为

$$p(x_1,\cdots,x_n\,|\,\theta)=\prod_{i=1}^{n}cx_i^{c-1}\theta^{-c}\mathrm{I}_{0<x_i<\theta}=c^n(x_1\cdots x_n)^{c-1}\theta^{-nc}\mathrm{I}_{0<x_1,\,\cdots,\,x_n<\theta},$$

（1）设参数 $\theta$ 的先验分布是帕雷托分布，密度函数为 $\pi(\theta)=\dfrac{\beta\theta_0^\beta}{\theta^{\beta+1}}\mathrm{I}_{\theta>\theta_0}$，

则样本 $X_1, \cdots, X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{\beta\theta_0^\beta c^n(x_1\cdots x_n)^{c-1}}{\theta^{nc+\beta+1}}\mathrm{I}_{0<x_1,\,\cdots,\,x_n,\,\theta_0<\theta}=\frac{\beta\theta_0^\beta c^n(x_1\cdots x_n)^{c-1}}{\theta^{nc+\beta+1}}\mathrm{I}_{x_1,\,\cdots,\,x_n>0,\,\theta>\max\{x_1,\,\cdots,\,x_n,\,\theta_0\}},$$

样本 $X_1, \cdots, X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_{\max\{x_1,\,\cdots,\,x_n,\theta_0\}}^{+\infty}\frac{\beta\theta_0^\beta c^n(x_1\cdots x_n)^{c-1}}{\theta^{nc+\beta+1}}\mathrm{I}_{x_1,\,\cdots,\,x_n>0}d\theta$$

$$=\beta\theta_0^\beta c^n(x_1\cdots x_n)^{c-1}\frac{[\max\{x_1,\cdots,x_n,\theta_0\}]^{-(nc+\beta)}}{nc+\beta}\mathrm{I}_{x_1,\,\cdots,\,x_n>0},$$

即参数 $\theta$ 的后验分布为

$$\pi(\theta\,|\,x_1,\cdots,x_n)=\frac{(nc+\beta)[\max\{x_1,\cdots,x_n,\theta_0\}]^{nc+\beta}}{\theta^{nc+\beta+1}}\mathrm{I}_{\theta>\max\{x_1,\,\cdots,\,x_n,\,\theta_0\}},$$

后验分布仍为帕雷托分布，其参数为 $nc+\beta$ 和 $\max\{x_1, \cdots, x_n, \theta_0\}$，

故帕雷托分布是参数 $\theta$ 的共轭先验分布；

（2）设参数 $c$ 的先验分布为伽玛分布 $Ga\,(\alpha, \lambda)$，密度函数为 $\pi(c)=\dfrac{\lambda^\alpha}{\Gamma(\alpha)}c^{\alpha-1}\,\mathrm{e}^{-\lambda c}\,\mathrm{I}_{c>0}$，

则样本 $X_1, \cdots, X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1,\cdots,x_n,c)=\frac{\lambda^\alpha}{\Gamma(\alpha)}c^{n+\alpha-1}(x_1\cdots x_n)^{c-1}\,\mathrm{e}^{-\lambda c}\,\theta^{-nc}\mathrm{I}_{0<x_1,\,\cdots,\,x_n<\theta,\,c>0}$$

$$=\frac{\lambda^\alpha}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}c^{n+\alpha-1}\,\mathrm{e}^{-[\lambda+n\ln\theta-\ln(x_1\cdots x_n)]c}\,\mathrm{I}_{0<x_1,\,\cdots,\,x_n<\theta,\,c>0},$$

样本 $X_1, \cdots, X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^{+\infty}\frac{\lambda^\alpha}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}c^{n+\alpha-1}\,\mathrm{e}^{-[\lambda+n\ln\theta-\ln(x_1\cdots x_n)]c}\,\mathrm{I}_{0<x_1,\,\cdots,\,x_n<\theta}d\theta$$

$$=\frac{\lambda^\alpha}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}\cdot\frac{\Gamma(n+\alpha)}{[\lambda+n\ln\theta-\ln(x_1\cdots x_n)]^{n+\alpha}}\mathrm{I}_{0<x_1,\,\cdots,\,x_n<\theta},$$

即参数 $\theta$ 的后验分布为

$$\pi(c\,|\,x_1,\cdots,x_n)=\frac{[\lambda+n\ln\theta-\ln(x_1\cdots x_n)]^{n+\alpha}}{\Gamma(n+\alpha)}c^{n+\alpha-1}\,\mathrm{e}^{-[\lambda+n\ln\theta-\ln(x_1\cdots x_n)]c}\,\mathrm{I}_{c>0}\,,$$

后验分布仍为伽玛分布，其参数为 $n+\alpha$ 和$\lambda+n\ln\theta-\ln(x_1\cdots x_n)$，

故伽玛分布是参数 $c$ 的共轭先验分布.

11．某人每天早上在汽车站等公共汽车的时间（单位：min）服从均匀分布 $U(0,\theta)$，其中$\theta$未知，假设$\theta$的先验分布为

$$\pi(\theta)=\begin{cases}\dfrac{192}{\theta^4}, & \theta\ge 4;\\ 0, & \theta<4.\end{cases}$$

假如此人在三个早上等车的时间分别为 5, 3, 8 分钟，求$\theta$后验分布.

解：参数$\theta$的先验分布为$\pi(\theta)=\dfrac{192}{\theta^4}\mathrm{I}_{\theta>4}$，

因样本 $X_1,\cdots,X_n$ 的联合条件分布为

$$p(x_1,\cdots,x_n\,|\,\theta)=\prod_{i=1}^{n}\frac{1}{\theta}\mathrm{I}_{0<x_i<\theta}=\frac{1}{\theta^n}\mathrm{I}_{0<x_1,\cdots,x_n<\theta}\,,$$

则样本 $X_1,\cdots,X_n$ 和参数$\theta$的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{192}{\theta^{n+4}}\mathrm{I}_{0<x_1,\cdots,x_n<\theta,\,\theta>4}=\frac{192}{\theta^{n+4}}\mathrm{I}_{x_1,\cdots,x_n>0,\,\theta>\max\{x_1,\cdots,x_n,4\}}\,,$$

样本 $X_1,\cdots,X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_{\max\{x_1,\cdots,x_n,4\}}^{+\infty}\frac{192}{\theta^{n+4}}\mathrm{I}_{x_1,\cdots,x_n>0}d\theta=\frac{192}{(n+3)[\max\{x_1,\cdots,x_n,4\}]^{n+3}}\mathrm{I}_{x_1,\cdots,x_n>0}\,,$$

即参数$\theta$的后验分布为

$$\pi(\theta\,|\,x_1,\cdots,x_n)=\frac{(n+3)[\max\{x_1,\cdots,x_n,4\}]^{n+3}}{\theta^{n+4}}\mathrm{I}_{\theta>\max\{x_1,\cdots,x_n,4\}}\,,$$

后验分布仍为帕雷托分布，其参数为 $n+3$ 和 $\max\{x_1,\cdots,x_n,4\}$，

因样本观测值为 5, 3, 8，即 $\max\{x_1,\cdots,x_n,4\}=8$，$n=3$，

故参数$\theta$的后验分布为帕雷托分布，其参数为 6 和 8，密度函数为

$$\pi(\theta\,|\,x_1,x_2,x_3)=\frac{6\times 8^6}{\theta^7}\mathrm{I}_{\theta>8}\,.$$

12．从正态分布 $N(\theta,2^2)$中随机抽取容量为 100 的样本，又设$\theta$的先验分布为正态分布，证明：不管先验分布的标准差为多少，后验分布的标准差一定小于 1/5.

解：设$\theta$的先验分布为正态分布 $N(\mu,\sigma^2)$，根据书上 P336 例 6.5.3 的结论可知，$\theta$的后验分布为

$$N\left(\frac{2^{-2}n\overline{X}+\mu\sigma^{-2}}{2^{-2}n+\sigma^{-2}},\frac{1}{2^{-2}n+\sigma^{-2}}\right)=N\left(\frac{25\overline{X}+\mu\sigma^{-2}}{25+\sigma^{-2}},\frac{1}{25+\sigma^{-2}}\right),$$

故后验分布的标准差为$\sqrt{\dfrac{1}{25+\sigma^{-2}}}<\dfrac{1}{5}$.

13．设随机变量 $X$ 服从负二项分布，其概率分布为

$$f(x\,|\,p)=\binom{x-1}{k-1}p^k(1-p)^{x-k},\quad x=k,k+1,\cdots,$$

证明其成功概率 $p$ 共轭先验分布族为贝塔分布族.

证：设参数 $p$ 的先验分布是贝塔分布 $Be(a, b)$，密度函数为 $\pi(p)=\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{a-1}(1-p)^{b-1}\mathrm{I}_{0<p<1}$，

因样本 $X_1,\cdots,X_n$ 的联合条件分布为

$$p(x_1,\cdots,x_n\mid p)=\prod_{i=1}^{n}\binom{x_i-1}{k-1}p^k(1-p)^{x_i-k}=\prod_{i=1}^{n}\binom{x_i-1}{k-1}\cdot p^{nk}(1-p)^{\sum\limits_{i=1}^{n}x_i-nk},$$

则样本 $X_1,\cdots,X_n$ 和参数 $p$ 的联合分布为

$$h(x_1,\cdots,x_n,p)=\prod_{i=1}^{n}\binom{x_i-1}{k-1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{nk+a-1}(1-p)^{\sum\limits_{i=1}^{n}x_i-nk+b-1}\mathrm{I}_{0<p<1},$$

样本 $X_1,\cdots,X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^1\prod_{i=1}^{n}\binom{x_i-1}{k-1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{nk+a-1}(1-p)^{\sum\limits_{i=1}^{n}x_i-nk+b-1}dp$$

$$=\prod_{i=1}^{n}\binom{x_i-1}{k-1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\cdot\frac{\Gamma(nk+a)\cdot\Gamma\left(\sum\limits_{i=1}^{n}x_i-nk+b\right)}{\Gamma\left(\sum\limits_{i=1}^{n}x_i+a+b\right)},$$

即参数 $p$ 的后验分布为

$$\pi(p\mid x_1,\cdots,x_n)=\frac{\Gamma\left(\sum\limits_{i=1}^{n}x_i+a+b\right)}{\Gamma(nk+a)\cdot\Gamma\left(\sum\limits_{i=1}^{n}x_i-nk+b\right)}p^{nk+a-1}(1-p)^{\sum\limits_{i=1}^{n}x_i-nk+b-1}\mathrm{I}_{0<p<1},$$

后验分布仍为贝塔分布，其参数为 $nk+a$ 和 $\sum\limits_{i=1}^{n}x_i-nk+b$，

故贝塔分布是参数 $p$ 的共轭先验分布.

14．从一批产品中抽检 100 个，发现 3 个不合格，假定该产品不合格率 $\theta$ 的先验分布为贝塔分布 $Be(2, 200)$，求 $\theta$ 的后验分布.

解：参数 $\theta$ 的先验分布是贝塔分布 $Be(2, 200)$，密度函数为 $\pi(\theta)=\dfrac{\Gamma(202)}{\Gamma(2)\Gamma(200)}\theta(1-\theta)^{199}\mathrm{I}_{0<\theta<1}$，

因样本 $X_1,\cdots,X_n$ 的联合条件分布为

$$p(x_1,\cdots,x_n\mid\theta)=\prod_{i=1}^{n}\theta^{x_i}(1-\theta)^{1-x_i}=\theta^{\sum\limits_{i=1}^{n}x_i}(1-\theta)^{n-\sum\limits_{i=1}^{n}x_i},$$

则样本 $X_1,\cdots,X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{\Gamma(202)}{\Gamma(2)\Gamma(200)}\theta^{1+\sum\limits_{i=1}^{n}x_i}(1-\theta)^{n+199-\sum\limits_{i=1}^{n}x_i}\mathrm{I}_{0<\theta<1},$$

样本 $X_1,\cdots,X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^1\frac{\Gamma(202)}{\Gamma(2)\Gamma(200)}\theta^{1+\sum\limits_{i=1}^{n}x_i}(1-\theta)^{n+199-\sum\limits_{i=1}^{n}x_i}d\theta$$

$$=\frac{\Gamma(202)}{\Gamma(2)\Gamma(200)}\cdot\frac{\Gamma\left(2+\sum_{i=1}^{n}x_i\right)\Gamma\left(n+200-\sum_{i=1}^{n}x_i\right)}{\Gamma(n+202)},$$

即参数$\theta$的后验分布为

$$\pi(\theta\mid x_1,\cdots,x_n)=\frac{\Gamma(n+202)}{\Gamma\left(2+\sum_{i=1}^{n}x_i\right)\Gamma\left(n+200-\sum_{i=1}^{n}x_i\right)}\theta^{1+\sum_{i=1}^{n}x_i}(1-\theta)^{n+199-\sum_{i=1}^{n}x_i}\mathrm{I}_{0<\theta<1},$$

后验分布仍为贝塔分布，其参数为$2+\sum_{i=1}^{n}x_i$和$n+200-\sum_{i=1}^{n}x_i$，

因$n=100$，$\sum_{i=1}^{n}x_i=3$，

故参数$\theta$的后验分布为贝塔分布$Be(5, 297)$，密度函数为

$$\pi(\theta\mid x_1,\cdots,x_n)=\frac{\Gamma(302)}{\Gamma(5)\Gamma(297)}\theta^4(1-\theta)^{296}\mathrm{I}_{0<\theta<1}.$$