- 1. 考虑掷硬币试验。分别使用参数为(a,b)=(1,1)和(a,b)=(10,5)的贝塔分布作为先验,用程序分别画出出现下列正面向上的计数结果时,硬币向上的概率参数的后验分布:
 - (1) 投掷0次, 0次正面向上
 - (2) 投掷1次, 1次正面向上
 - (3) 投掷2次, 2次正面向上
 - (4) 投掷3次, 2次正面向上
 - (5) 投掷8次, 4次正面向上
 - (6) 投掷15次, 6次正面向上
 - (7) 投掷50次, 24次正面向上
 - (8) 投掷500次, 263次正面向上
- 在掷硬币试验中,将参数 θ 的先验分布设定为Beta分布

$$p(\theta) = Beta(\theta|a,b) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}$$
 (1)

• 似然函数为二项分布

$$p(x|\theta) = C_n^x \theta^x (1-\theta)^{n-x} \tag{2}$$

• 通过贝叶斯公式推导得到

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = Beta(\theta|a+x, b+n-x)$$
(3)

- 先验分布参数为(1,1)时,需依次画出 $Beta(\theta|1,1)$ 、 $Beta(\theta|1+1,1+1-1)$ 、 $Beta(\theta|1+2,1+2-2)$ 、 $Beta(\theta|1+2,1+3-2)$ 、 $Beta(\theta|1+4,1+8-4)$ 、 $Beta(\theta|1+6,1+15-6)$ 、 $Beta(\theta|1+24,1+50-24)$ 、 $Beta(\theta|1+263,1+500-263)$
- 先验分布参数为(10,5)时,需依次画出 $Beta(\theta|10,5)$ 、 $Beta(\theta|10+1,5+1-1)$ 、 $Beta(\theta|10+2,5+2-2)$ 、 $Beta(\theta|10+2,5+3-2)$ 、 $Beta(\theta|10+4,5+8-4)$ 、 $Beta(\theta|10+6,5+15-6)$ 、 $Beta(\theta|10+24,5+50-24)$ 、 $Beta(\theta|10+263,5+500-263)$

核心代码如下:

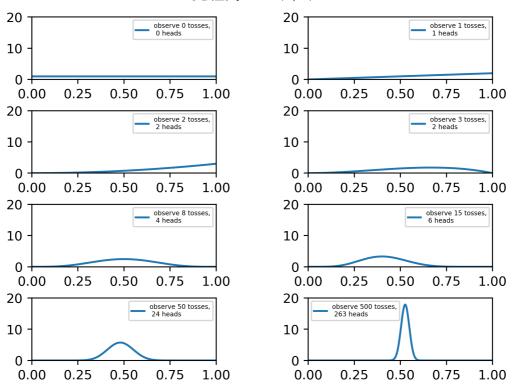
```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import beta

def betaDist(a, b):
    x = np.linspace(0, 1, 1002)[1:-1] # 创建一系列x值
    y = beta(a,b).pdf(x)#使用 scipy.stats 中的 beta 类产生贝塔分布的概率密度函数
    return x,y

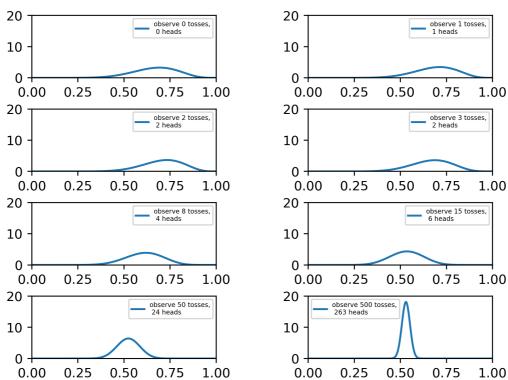
x,y = betaDist(1,1)
plt.plot(x,y)
plt.show()
```

结果:

先验为Beta(1,1)



先验为Beta(10,5)



2. 分别证明:

(1) 多项分布的共轭先验是狄利克雷分布,参数为 $heta_1,\ldots, heta_k$,观测值为 x_1,\ldots,x_k 。

多项分布是二项分布的推广扩展,在n次独立试验中只输出k种结果中的一个,且每种结果都有一个确定的概率 θ 。多项分布给出了在多种输出状态的情况下,关于成功次数的各种组合的概率。举个例子,投掷n次骰子,这个骰子共有6种结果输出(k=6),且1点出现的概率为 θ_1 ,2点出现的概率为 θ_2 ,...,在n次试验中,骰子1点出现 x_1 次,2点出现 x_2 次...。这个结果组合出现的概率为 $C_n^{x_1}C_{n-x_1}^{x_2}\dots C_{n-x_1\dots x_5}^{x_6}\theta_1^{x_1}\theta_2^{x_2}\dots \theta_6^{x_6}=\frac{n!}{x_1!\dots x_6!}\theta_1^{x_1}\dots \theta_6^{x_6}$

狄利克雷分布是Beta分布在多项情况下的推广,概率密度函数如下:

$$p\left(heta_{1},\ldots, heta_{k}|lpha_{1},\ldots,lpha_{k}
ight) = rac{\Gamma\left(\sum_{i=1}^{k}lpha_{i}
ight)}{\prod_{i=1}^{k}\Gamma\left(lpha_{i}
ight)}\prod_{i=1}^{k} heta_{i}^{lpha_{i}-1}$$

Beta分布的概率密度函数 $p(\theta)=rac{\Gamma(lpha+eta)}{\Gamma(lpha)\Gamma(eta)} heta^{lpha-1}(1- heta)^{eta-1}$

根据贝叶斯公式,有

$$p\left(heta_{1},\ldots, heta_{k}|x_{1},\ldots,x_{k}
ight) = rac{p\left(x_{1},\ldots,x_{k}| heta_{1},\ldots, heta_{k}
ight)p\left(heta_{1},\ldots, heta_{k}
ight)}{\int p\left(x_{1},\ldots,x_{k}| heta_{1},\ldots, heta_{k}
ight)p\left(heta_{1},\ldots, heta_{k}
ight)d heta_{1},\ldots, heta_{k}}$$

观测值 x_1, \ldots, x_k 关于参数 $\theta_1, \ldots, \theta_k$ 的似然函数:

$$egin{aligned} p\left(x_1,\ldots,x_k| heta_1,\ldots heta_k
ight) &= rac{n!}{x_1!\ldots x_k!} heta_1^{x_1}\ldots heta_k^{x_k} \ &= rac{\Gamma\left(n+1
ight)}{\prod_{i=1}^k\Gamma(x_i+1)}\prod_{i=1}^k heta_i^{x_i} \end{aligned}$$

参数 $\theta_1, \ldots, \theta_k$ 的先验:

$$p\left(heta_{1},\ldots, heta_{k}
ight)=rac{\Gamma\left(\sum_{i=1}^{k}lpha_{i}
ight)}{\prod_{i=1}^{k}\Gamma\left(lpha_{i}
ight)}\prod_{i=1}^{k} heta_{i}^{lpha_{i}-1}$$

计算归一化因子 $p(x_1,\ldots,x_k)$:

$$\begin{split} p(x_1,\ldots,x_k) &= \int p(x_1,\ldots,x_k|\theta_1,\ldots,\theta_k) p(\theta_1,\ldots,\theta_k) d\theta \\ &= \frac{\Gamma(n+1)\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(x_i+1)\Gamma(\alpha_i)} \int \prod_{i=1}^k \theta_i^{\alpha_i+x_i-1} d\theta \\ &= \frac{\Gamma(n+1)\Gamma(\sum_{i=1}^k \alpha_i) \prod_{i=1}^k \Gamma(\alpha_i+x_i)}{\Gamma(\sum_{i=1}^k (\alpha_i+x_i)) \prod_{i=1}^k \Gamma(x_i+1)\Gamma(\alpha_i)} \int \frac{\Gamma\left(\sum_{i=1}^k (\alpha_i+x_i)\right)}{\prod_{i=1}^k \Gamma\left(\alpha_i+x_i\right)} \prod_{i=1}^k \theta_i^{\alpha_i+x_i-1} d\theta \\ &= \frac{\Gamma(n+1)\Gamma(\sum_{i=1}^k \alpha_i) \prod_{i=1}^k \Gamma(\alpha_i+x_i)}{\Gamma(\sum_{i=1}^k (\alpha_i+x_i)) \prod_{i=1}^k \Gamma(\alpha_i+x_i)} \end{split}$$

代入贝叶斯公式计算:

$$egin{aligned} p\left(heta_1,\ldots, heta_k|x_1,\ldots,x_k
ight) &= rac{p(x_1,\ldots,x_k| heta_1,\ldots, heta_k)p(heta_1,\ldots, heta_k)}{p(x_1,\ldots,x_k)} \ &= rac{\Gamma\left(\sum_{i=1}^k\left(lpha_i+x_i
ight)
ight)}{\prod_{i=1}^k\Gamma\left(lpha_i+x_i
ight)} \prod_{i=1}^k heta_i^{lpha_i+x_i-1} \ &= Dirichlet\left(heta_1,\ldots, heta_k|lpha_1+x_1,\ldots,lpha_k+x_k
ight) \end{aligned}$$

(2) 泊松分布的共轭先验是Gamma分布,参数为 λ ,观测值为x。

泊松分布的随机变量表示某事件在单位时间内随机独立出现的次数

$$P(X = k) = \frac{\lambda^k}{k!} \exp(-\lambda), \ \lambda > 0, \ k = 0, 1, 2, \dots$$

指数分布的随机变量表示独立随机事件发生的时间间隔,即要等到一个随机事件发生,需要经历多久时间

$$p(x) = \begin{cases} \lambda \exp(-\lambda x), & x \geqslant 0 \\ 0, & x < 0 \end{cases}$$

Gamma分布的随机变量表示要等到 α 个随机事件都发生,需要经历多久时间

对Gamma函数做个变形,可以得到如下式子:

$$\Gamma(\alpha) = \int_0^{+\infty} \exp\left(-t\right) t^{\alpha-1} dt \quad \boxed{\rightarrow} \quad \int_0^{+\infty} \frac{t^{\alpha-1} \exp(-t)}{\Gamma(\alpha)} dt = 1$$
做一个变换 $t = \beta x$, $\int_0^{+\infty} \frac{\beta^{\alpha} x^{\alpha-1} \exp(-\beta x)}{\Gamma(\alpha)} d(x) = 1$

取等式左边积分中的函数作为概率密度函数,得到Gamma分布的一般形式:

$$Gamma(x, \alpha, \beta) = \frac{\beta^{\alpha} x^{\alpha - 1} \exp(-\beta x)}{\Gamma(\alpha)}$$

且 α =1时,上式就变成了指数分布,指数分布是Gamma分布的特殊形式。

根据贝叶斯公式,有:

$$p\left(\lambda|x
ight) = rac{p\left(x|\lambda
ight)p\left(\lambda
ight)}{\int p\left(x|\lambda
ight)p\left(\lambda
ight)d\lambda}$$

观测值x关于参数 λ 的似然函数:

$$p\left(x|\lambda
ight)=rac{\lambda^{x}}{x!}\mathrm{exp}\left(-\lambda
ight)$$

参数 λ 的先验:

$$p\left(\lambda\right) = Gamma\left(\lambda|\alpha, eta
ight) = rac{eta^{lpha}\lambda^{lpha-1}\exp\left(-eta\lambda
ight)}{\Gamma\left(lpha
ight)}$$

计算 $p(x|\lambda)p(\lambda)$:

$$p\left(x|\lambda
ight)p\left(\lambda
ight)=rac{eta^{lpha}}{\Gamma\left(x+1
ight)\Gamma\left(lpha
ight)}\lambda^{lpha+x-1}\exp\left(-\left(eta+1
ight)\lambda
ight)$$

计算归一化因子p(x):

$$egin{aligned} p(x) &= \int_0^{+\infty} p(x|\lambda) p(\lambda) d\lambda \ &= rac{eta^{lpha}}{\Gamma\left(x+1
ight)\Gamma\left(lpha
ight)} \int_0^{+\infty} \lambda^{lpha+x-1} \exp\left(-\left(eta+1
ight)\lambda
ight) d\lambda \ &= rac{eta^{lpha} \Gamma\left(lpha+x
ight)}{\Gamma\left(x+1
ight)\Gamma\left(lpha
ight)(eta+1
ight)^{lpha+x}} \int_0^{+\infty} rac{\left(eta+1
ight)^{lpha+x} \lambda^{lpha+x-1} \exp\left(-\left(eta+1
ight)\lambda
ight)}{\Gamma\left(lpha+x
ight)} d\lambda \ &= rac{eta^{lpha} \Gamma\left(lpha+x
ight)}{\Gamma\left(x+1
ight)\Gamma\left(lpha
ight)(eta+1
ight)^{lpha+x}} \end{aligned}$$

代入贝叶斯公式:

$$egin{aligned} p\left(\lambda|x
ight) &= rac{p(x|\lambda)p(\lambda)}{p(x)} \ &= rac{\left(eta+1
ight)^{lpha+x}\lambda^{lpha+x-1}\exp\left(-\left(eta+1
ight)\lambda
ight)}{\Gamma\left(lpha+x
ight)} \ &= Gamma\left(\lambda|lpha+x,eta+1
ight) \end{aligned}$$

(3) 指数分布的共轭先验是Gamma分布,参数为 θ ,观测值为x。

根据贝叶斯公式,有:

$$p\left(heta|x
ight) = rac{p\left(x| heta
ight)p\left(heta
ight)}{\int p\left(x| heta)p\left(heta
ight)d heta}$$

观测值x关于参数 θ 的似然函数:

$$p(x|\theta) = \theta \exp(-\theta x), x > 0$$

参数 θ 的先验:

$$p\left(heta
ight) =Gamma\left(heta | lpha, eta
ight) =rac{eta^{lpha} heta^{lpha-1}\exp\left(-eta heta
ight)}{\Gamma\left(lpha
ight)}$$

计算 $p(x|\theta)p(\theta)$:

$$p\left(x| heta
ight)p\left(heta
ight)=rac{eta^{lpha}}{\Gamma\left(lpha
ight)} heta^{lpha}\exp\left(-\left(eta+x
ight) heta
ight)$$

计算归一化因子p(x):

$$egin{aligned} p(x) &= \int p\left(x| heta
ight)p\left(heta
ight)d heta \ &= rac{eta^{lpha}}{\Gamma\left(lpha
ight)}\int_{0}^{+\infty} heta^{lpha}\exp\left(-\left(eta+x
ight) heta
ight)d heta \ &= rac{eta^{lpha}\Gamma\left(lpha+1
ight)}{\Gamma\left(lpha
ight)(eta+x)^{lpha+1}}\int_{0}^{+\infty}rac{\left(eta+x
ight)^{lpha+1} heta^{lpha}\exp\left(-\left(eta+x
ight) heta
ight)}{\Gamma\left(lpha+1
ight)}d heta \ &= rac{eta^{lpha}\Gamma\left(lpha+1
ight)}{\Gamma\left(lpha
ight)(eta+x
ight)^{lpha+1}} \end{aligned}$$

代入贝叶斯公式:

$$\begin{split} p\left(\theta|x\right) &= \frac{p(x|\theta)p(\theta)}{p(x)} \\ &= \frac{\left(\beta + x\right)^{\alpha + 1}\theta^{\alpha}\exp\left(-\left(\beta + x\right)\theta\right)}{\Gamma\left(\alpha + 1\right)} \\ &= Gamma\left(\theta|\alpha + 1, \beta + x\right) \end{split}$$

(4) 方差已知的正态分布的共轭先验是正态分布,参数为均值 μ ,观测值为x。

根据贝叶斯公式,有:

$$p\left(\mu|x
ight) = rac{p\left(x|\mu
ight)p\left(\mu
ight)}{\int p\left(x|\mu
ight)p\left(\mu
ight)d\mu}$$

观测值x关于参数 μ 的似然函数:

$$p(x|\mu) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

参数 $\mu \sim N(u,v^2)$, 先验为:

$$p\left(\mu\right) = rac{1}{v\sqrt{2\pi}} \exp\left(-rac{\left(\mu - u\right)^2}{2v^2}\right)$$

计算 $p(x|\mu)p(\mu)$:

$$p\left(x|\mu
ight)p\left(\mu
ight)=rac{1}{2\pi v\sigma}\mathrm{exp}\left(-rac{\left(x-\mu
ight)^{2}}{2\sigma^{2}}-rac{\left(\mu-u
ight)^{2}}{2v^{2}}
ight)$$

计算归一化因子p(x):

代入贝叶斯公式:

$$egin{aligned} p\left(\mu|x
ight) &= rac{p(x|\mu)p(\mu)}{p(x)} \ &= rac{1}{\sqrt{rac{\sigma^2v^2}{v^2+\sigma^2}}\sqrt{2\pi}} \exp\left(-rac{\left(\mu - rac{xv^2 + u\sigma^2}{\sigma^2 + v^2}
ight)^2}{rac{2\sigma^2v^2}{v^2+\sigma^2}}
ight) \end{aligned}$$

则 $p(\mu|x)$ 服从正态分布 $N\left(rac{xv^2+u\sigma^2}{\sigma^2+v^2},rac{\sigma^2v^2}{v^2+\sigma^2}
ight)$ 。

(5) 均值已知的正态分布的共轭先验是逆Gamma分布,参数为方差 σ^2 ,观测值为x。

逆Gamma分布

$$IG\left(x|lpha,eta
ight)=rac{eta^{lpha}x^{-lpha-1}\exp\left(-rac{eta}{x}
ight)}{\Gamma\left(lpha
ight)}$$

逆Gamma分布与Gamma分布之间的关系:

若随机变量 $X \sim Gamma(\alpha, \beta)$,则 $\frac{1}{X} \sim IG(\alpha, \beta)$

根据贝叶斯公式,有:

$$p\left(\sigma^{2}|x
ight)=rac{p\left(x|\sigma^{2}
ight)p\left(\sigma^{2}
ight)}{\int p\left(x|\sigma^{2}
ight)p\left(\sigma^{2}
ight)d\sigma^{2}}$$

观测值x关于参数 σ^2 的似然函数:

$$p\left(x|\sigma^2
ight) = rac{1}{\sigma\sqrt{2\pi}} \mathrm{exp}\left(-rac{\left(x-\mu
ight)^2}{2\sigma^2}
ight)$$

参数 σ^2 的先验:

$$p\left(\sigma^{2}
ight)=IG\left(\sigma^{2}|lpha,eta
ight)=rac{eta^{lpha}ig(\sigma^{2}ig)^{-lpha-1}\exp\left(-rac{eta}{\sigma^{2}}
ight)}{\Gamma\left(lpha
ight)}$$

计算 $p(x|\sigma^2)p(\sigma^2)$:

$$p\left(x|\sigma^2
ight)p\left(\sigma^2
ight) = rac{eta^lpha}{\sqrt{2\pi} \Gamma\left(lpha
ight)} \left(\sigma^2
ight)^{-lpha-rac{3}{2}} \exp\left(-rac{\left(x-\mu
ight)^2+2eta}{2\sigma^2}
ight)$$

计算归一化因子p(x):

$$\begin{split} p(x) &= \int p(x|\sigma^2)p(\sigma^2)d\sigma^2 \\ &= \frac{\beta^\alpha}{\sqrt{2\pi}\Gamma\left(\alpha\right)} \int \left(\sigma^2\right)^{-\alpha-\frac{3}{2}} \exp\left(-\frac{\left(x-\mu\right)^2+2\beta}{2\sigma^2}\right) d\sigma^2 \\ &= \frac{\beta^\alpha \Gamma(\frac{1}{2}+\alpha)}{\sqrt{2\pi}\Gamma\left(\alpha\right)\left(\frac{\left(x-\mu\right)^2}{2}+\beta\right)^{\frac{1}{2}+\alpha}} \int \frac{\left(\frac{\left(x-\mu\right)^2}{2}+\beta\right)^{\frac{1}{2}+\alpha}\left(\sigma^2\right)^{-\alpha-\frac{3}{2}} \exp\left(-\frac{\left(x-\mu\right)^2+2\beta}{2\sigma^2}\right)}{\Gamma(\frac{1}{2}+\alpha)} d\sigma^2 \\ &= \frac{\beta^\alpha \Gamma(\frac{1}{2}+\alpha)}{\sqrt{2\pi}\Gamma\left(\alpha\right)\left(\frac{\left(x-\mu\right)^2}{2}+\beta\right)^{\frac{1}{2}+\alpha}} \end{split}$$

代入贝叶斯公式:

$$egin{aligned} p\left(\sigma^2|x
ight) &= rac{p(x|\sigma^2)p(\sigma^2)}{p(x)} \ &= rac{\left(rac{(x-\mu)^2}{2} + eta
ight)^{rac{1}{2}+lpha}ig(\sigma^2ig)^{-lpha-rac{3}{2}} \exp\left(-rac{(x-\mu)^2+2eta}{2\sigma^2}
ight)}{\Gamma(rac{1}{2}+lpha)} \end{aligned}$$

则 $p\left(\sigma^2|x
ight)$ 服从逆Gamma分布 $IG\left(\sigma^2|rac{1}{2}+lpha,rac{(x-\mu)^2}{2}+eta
ight)$ 。