1. 《Machine Learning A Probabilistic Perspective》这本书的练习题: Exercise 2.16(15分)
2. Exercise 3.6(10分)
3. Exercise 3.7(20分)
4. Exercise 3.15(20分)
5. Exercise 4.3(10分)
6. Exercise 4.4(10分)
7. Exercise 4.5(10分)
8. 求解问题(5分):

$$\max_x \sum p_k log(x_k)$$

$$s.t. \sum x_k = 1$$

1 Ex2.16 解：

$$\begin{aligned}E(\theta) &= \int_{-\infty}^{\infty} \theta \cdot \beta(\theta|a,b) d\theta = \int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\cdot\Gamma(b)} \cdot \theta^{a-1} \cdot (1-\theta)^{b-1} d\theta \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\cdot\Gamma(b)} \cdot \int_0^1 \theta^{(a+1)-1} \cdot (1-\theta)^{b-1} d\theta = \frac{\Gamma(a+b)}{\Gamma(a)\cdot\Gamma(b)} \cdot \frac{\Gamma(a+1)\cdot\Gamma(b)}{\Gamma(a+1+b)} \\ &= \frac{(a+b-1)!\cdot a!}{(a-1)!\cdot(a+b)!} = \frac{a}{a+b}\end{aligned}$$

$$\begin{aligned}E(\theta^2) &= \int_{-\infty}^{\infty} \theta^2 \cdot \frac{\Gamma(a+b)}{\Gamma(a)\cdot\Gamma(b)} \cdot \theta^{a-1} \cdot (1-\theta)^{b-1} d\theta \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\cdot\Gamma(b)} \cdot \int_0^1 \theta^{(a+2)-1} \cdot (1-\theta)^{b-1} d\theta = \frac{\Gamma(a+b)}{\Gamma(a)\cdot\Gamma(b)} \cdot \frac{\Gamma(a+2)\cdot\Gamma(b)}{\Gamma(a+2+b)} \\ &= \frac{(a+b-1)!\cdot(a+1)!}{(a-1)!\cdot(a+b+1)!} = \frac{a(a+1)}{(a+b)(a+b+1)}\end{aligned}$$

$$\begin{aligned}D(\theta) &= E[\theta - E(\theta)]^2 = E[\theta^2 + E^2(\theta) - 2\theta \cdot E(\theta)] = E(\theta^2) - E^2(\theta) \\ &= \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2} = \frac{ab}{(a+b)^2\cdot(a+b+1)}\end{aligned}$$

$\beta(\theta|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\cdot\Gamma(b)}\theta^{a-1}\cdot(1-\theta)^{b-1}$，对 θ 求导，得到 $\beta^{'}(\theta|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\cdot\Gamma(b)}(a-1)\theta^{a-2}\cdot(1-b)(1-\theta)^{b-2}$，令 $\beta^{'}(\theta|a,b) = 0$，则有 $mode(\beta) = \frac{a-1}{a+b-2}$

2 Ex3.6 解：

$$\begin{aligned}L(\lambda) &= log\prod_{i=1}^{n} f(x_i|\lambda) = \sum_{i=1}^{n} log(\frac{e^{-\lambda}\lambda^x}{x!}) \\ &= -n\lambda + (\sum_{i=1}^{n} x_i) log(\lambda) - log(\sum_{i=1}^{n} x_i!)\end{aligned}$$

对 λ 求导并令其等于 0：

$$\frac{d}{d\lambda}L(\lambda) = 0 \Leftrightarrow -n + (\sum_{i=1}^{n} x_i)\frac{1}{\lambda} = 0$$

解得

$$\lambda_{MLE} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

3 Ex3.7 解：

1) 由贝叶斯公式，

$$p(\theta|x) = \frac{p(x|\theta)f(\theta|a,b)}{p(x)}$$

其中 $p(x) = \int g(\theta)p(x|\theta)d\theta$ 与 θ 无关. 那么

$$p(x|\lambda)g(\lambda|a,b) = \frac{\lambda^x e^{-\lambda}}{x!}\frac{b^a\lambda^{a-1}e^{-b\lambda}}{\Gamma(a)} \propto \lambda^{a+x-1}e^{-(\lambda(b+1))} \propto g(\lambda|a+x,b+1)$$

则有

$$p(\lambda|D) = g(\lambda|a+D,b+1) = \Gamma(\lambda|a+D,b+1)$$

2) 当 $a \to 0, b \to 0$ 时，$p(\lambda|D) = \Gamma(\lambda|D,1)$，由于 Gamma 分布 $\Gamma(a,b)$ 的均值为 $\frac{a}{b}$，那么其均值为 D.

4 Ex3.15 解：

$$E(\theta)=m=\frac{\alpha_1}{\alpha_1+\alpha_2},\qquad var(\theta)=v=\frac{\alpha_1\alpha_2}{(\alpha_1+\alpha_2)^2(\alpha_1+\alpha_2+1)}$$

$$\Rightarrow \alpha_1=\frac{m^2(1-m)}{v}-m,\quad \alpha_2=\frac{m(1-m)^2}{v}+m-1$$

当 $m=0.7,\ v=0.2^2$ 时，$\alpha_1=2.975,\ \alpha_2=1.275.$

5 Ex4.3 解：

$$Cov(X,Y)^2=E[(X-\mu_x)(Y-\mu_y)]^2\le E[(X-\mu_x)^2]E[(X-\mu_y)^2]=Var(X)Var(Y)$$

$$Cov(X,Y)^2\le Var(X)Var(Y)\Rightarrow -1\le\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}=\rho(X,Y)\le 1$$

6 Ex4.4 解：

$$\begin{aligned}Cov(X,Y)&=E[XY]-E[X]E[Y]=E[aX^2+bX]-E[X](aE[X]+b)\\&=aE[X^2]+bE[X]-aE[X]^2-bE[X]=a(E[X^2]-E[X]^2)\\&=aVar(X)\end{aligned}$$

$$\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{aVar(X)}{\sqrt{a^2Var(X)^2}}=\frac{aVar(X)}{|a|Var(X)}$$

故容易知道，当 $a>0$ 时，$\rho=1$；当 $a<0$ 时，$\rho=-1.$

7 Ex4.5 解：

$$\int exp(-\frac{1}{2}(x-\mu)^t\Sigma^{-1}(x-\mu))dx=\int exp(-\frac{1}{2}(x-\mu)^tU\Lambda^{-1}U^t(x-\mu))dx$$

令 $y=U^t(x-\mu)$，

$$\int exp(-\frac{1}{2}(x-\mu)^tU\Lambda^{-1}U^t(x-\mu))dx=\int exp(-\frac{1}{2}y^t\Lambda^{-1}y)dx=$$

$$\int exp(-\frac{1}{2}\sum_i\frac{y_i^2}{\lambda_i})dx=\int exp(-\frac{1}{2}\sum_i\frac{y_i^2}{\lambda_i})\frac{\partial(x_1,x_2,\cdots,x_n)}{\partial(y_1,y_2,\cdots,y_n}dy$$

计算雅各比：

$$y=U^t(x-\mu)\Rightarrow x=Uy+\mu$$

$$J_{ij}=\frac{\partial x_i}{\partial y_j}=u_{ij}$$

因此有 $J=U,\ \frac{\partial(x_1,x_2,\cdots,x_n)}{\partial(y_1,y_2,\cdots,y_n}=det(U)=1$

$$\begin{aligned}\int exp(-\frac{1}{2}\sum_i\frac{y_i^2}{\lambda_i})dy&=\prod_i\int exp(-\frac{1}{2}\frac{y_i^2}{\lambda_i})dy_i\\&=\prod_{i=1}^{D}\sqrt{(2\pi\lambda_i)}=\sqrt{(2\pi)}^D\prod_{i=1}^{D}\sqrt{\lambda_i}\\&=(2\pi)^{D/2}|\Sigma|^{1/2}\end{aligned}$$

8 解：令 $F(x_1,\cdots,x_k,\lambda)=\sum_{i=1}^k p_ilog(x_i)+\lambda(\sum_{i=1}^k x_i-1)$，对 $x_1,\cdots,x_k,\lambda$ 求导，得 $\frac{p_1}{x_1}+\lambda=0,\cdots,\frac{p_k}{x_k}+\lambda=0,\ \sum_{i=1}^k x_i=1$。解得 $\lambda=-(x_1+\cdots+x_k)$，则

$$x_i=\frac{p_i}{p_1+\cdots+p_k},i=1,\cdots,k$$