Lista de Exercícios referente ao Cap. 2 do livro do Bishop

Exercícios do livro texto:

~~01) 2.1~~
~~02) 2.2~~
~~03) 2.8~~ (fazer apenas a eq. (2.270), que é um resultado bem útil conhecido como Regra da Torre)
~~04) 2.12~~
~~05) 2.13~~
~~06) 2.15~~
~~07) 2.20~~

Dica para o exercício 2.13: Após montar a expressão geral da divergência KL para p(x) e q(x), verifique o que acontece nos seguintes casos particulares:
(a) ambas as pdfs têm mesma média e matriz de covariância (isto é, p(x) = q(x); caso em que já sabemos quanto a divergência KL deve resultar);
(b) ambas as pds têm a mesma média, isto é, **m = \mu.**

Exercícios Extra:

E1) (*Inferência Bayesiana sequencial*) Motivado pela Figura 2.3 do livro, reproduza o experimento da jogada de moeda considerando que foram realizadas 5 jogadas e que a probabilidade de se obter cara é dada por '\mu = 0,7'. Plote a distribuição a priori e todas as 5 distribuições a posteriori geradas ao longo do processo iterativo. Considere que a distribuição a priori é uma Beta com parâmetros 'a' e 'b' escolhidos da seguinte forma:
1º caso: a = b = 1
2º caso: a = b = 2
Compare os resultados obtidos nos 2 casos.

OBS: Para uma comparação justa entre os 2 casos, primeiro gere os 5 dados (saídas do experimento da moeda, amostrados da Bernoulli definida no enunciado) e depois aplique o aprendizado sequencial para os 2 casos (i.e., para ambas as prioris) usando exatamente os mesmos dados gerados.

E2) (*Verificação experimental do Teorema Central do Limite*) Considere a **média** de N variáveis aleatórias iid. Plote o histograma dessa média considerando que as N variáveis aleatórias têm a seguinte pdf:
1º caso: Uniforme(0,1) – uniforme no intervalo 0 a 1;
2º caso: Bernoulli – escolha o valor do parâmetro como quiser;
Note que, para N suficientemente grande, a distribuição da média converge para uma Gaussiana.

OBS: Usei **média** ao invés de **soma** para facilitar a geração do histograma (o eixo horizontal vai ficar fixo, facilitando a comparação para diferentes valores de N, igual na Figura 2.6 do livro).

Lista 2

2.1

$$\text{Bern}(x|\mu) = \mu^{x}(1-\mu)^{1-x}$$

$$x \in \{0,1\}$$

$$\begin{cases} p(x=1|\mu) = \mu \\ p(x=0|\mu) = 1-\mu \end{cases}$$

Definição

$$\sum_{x} p(x) f(x) = E[f] \quad (1.33)$$

$$E[f^2] - E[f]^2 = var[f] \quad (1.39)$$

$$E[(x-\mu)^2] = \sigma^2 \quad \text{ex. 1.8 Cap 1}$$

(i) $\displaystyle\sum_{x=0}^{1} p(x|\mu) = p(x=0) + p(x=1) = 1-\mu+\mu = 1$

(ii) $\displaystyle\sum_{x=0}^{1} x\, p(x|\mu) = 0 \cdot p(x=0) + 1\, p(x=1) = 0(1-\mu) + 1 \cdot \mu = \mu$ //

(iii) $\displaystyle\sum_{x=0}^{1} (x-\mu)^2 p(x|\mu) = \sum_{x=0}^{1} (x^2 - 2x\mu + \mu^2)\, p(x|\mu) =$

$$= \mu^2 p(0) + (1-2\mu+\mu^2)\, p(1)$$

$$= \mu^2(1-\mu) + (1-2\mu+\mu^2)\mu =$$

$$= \mu^2 - \cancel{\mu^3} + \mu - 2\mu^2 + \cancel{\mu^3} = \mu - \mu^2 = \mu(1-\mu) \text{ //}$$

(iv) Definição Entropia

$$H[x] = -\sum_{x} p(x) \ln p(x) \quad (1.98)$$

$$= -\sum_{x=0}^{1} \mu^{x}(1-\mu)^{1-x} \underbrace{\ln\{\mu^{x}(1-\mu)^{1-x}\}}_{\ln\{\mu^{x}\} + \ln\{(1-\mu)^{1-x}\}}$$

$$= \sum_{x=0}^{1} \mu^{x}(1-\mu)^{1-x} \{x \ln\mu + (1-x)\ln(1-\mu)\} = -(1-\mu)\ln(1-\mu) - \mu\ln\mu \text{ //}$$

2.2

$$\mathrm{Bern}(x|\mu) = \mu^{x}(1-\mu)^{1-x} \qquad (2.2)$$
$$x \in \{0,1\}$$

$$p(x|\mu) = \left(\frac{1-\mu}{2}\right)^{\frac{1-x}{2}} \left(\frac{1+\mu}{2}\right)^{\frac{1+x}{2}} \qquad x \in [-1,1]$$
$$\mu \in [-1,1]$$

(i) $$\sum_{x=-1}^{1} p(x|\mu) = p(x=-1) + p(x=1) =$$

$$= \left(\frac{1-\mu}{2}\right)^{1} \cancel{\left(\frac{1+\mu}{2}\right)}^{0} + \cancel{\left(\frac{1-\mu}{2}\right)}^{0} \left(\frac{1+\mu}{2}\right)^{1} = \frac{1-\mu}{2} + \frac{1+\mu}{2} = \frac{2}{2} = 1$$

(ii) $$\sum_{x=-1}^{1} x\, p(x|\mu) = (-1)\left(\frac{1-\mu}{2}\right) + 1\left(\frac{1+\mu}{2}\right) = \frac{-1+\mu}{2} + \frac{1+\mu}{2} = \frac{2\mu}{2} = \mu$$

(iii) $$\sum_{x=-1}^{1} (x-\mu)^2 p(x|\mu) = (-1-\mu)^2\left(\frac{1-\mu}{2}\right) + (1-\mu)^2\left(\frac{1+\mu}{2}\right)$$

$$= (1+2\mu+\mu^2)\left(\frac{1-\mu}{2}\right) + (1-2\mu+\mu^2)\left(\frac{1+\mu}{2}\right) = \frac{1+2\mu+\mu^2-\mu-2\mu^2-\mu^3}{2}$$

$$+ \frac{1-2\mu+\mu^2+\mu-2\mu^2+\mu^3}{2}$$

$$= \frac{2-2\mu^2}{2} = 1-\mu^2$$

(10) Entropia

$$H[x] = -\sum_{n=-1}^{1} p(x) \ln p(x)$$

$$= -\sum_{n=-1}^{1} \left(\frac{1-\mu}{2}\right)^{\frac{1-x}{2}} \left(\frac{1+\mu}{2}\right)^{\frac{1+x}{2}} \ln \left\{ \left(\frac{1-\mu}{2}\right)^{\frac{1-x}{2}} \left(\frac{1+\mu}{2}\right)^{\frac{1+x}{2}} \right\} =$$

$$= \sum_{n=-1}^{1} \left(\frac{1-\mu}{2}\right)^{\frac{1-x}{2}} \left(\frac{1+\mu}{2}\right)^{\frac{1+x}{2}} \left\{ \left(\frac{1-x}{2}\right) \ln\left(\frac{1-\mu}{2}\right) + \left(\frac{1+x}{2}\right) \ln\left(\frac{1+\mu}{2}\right) \right\} =$$

$$= \left(\frac{1-\mu}{2}\right) \ln\left(\frac{1-\mu}{2}\right) + \left(\frac{1+\mu}{2}\right) \ln\left(\frac{1+\mu}{2}\right) //$$

(28)

$$\left.\begin{aligned} E[f(x)] &= \int p(x) f(x)\, dx \\ E[x] &= \int p(x)\, x\, dx \end{aligned}\right\} \text{ Definição (1.34)}$$

$$E_x[x|y] = \int p(x|y)\, x\, dx \quad \therefore \quad E_y[f] = \int f(y)\, p(y)\, dy$$

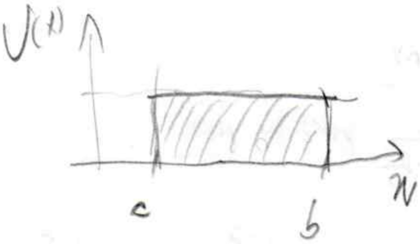
$$E_y[E_x[x|y]] = \int \Big[\underbrace{\int p(x|y)\, x\, dx}_{f(y)} \Big] p(y)\, dy \qquad p(x,y) = p(y|x)\, p(x)$$

$$= \iint x \cdot \underbrace{p(x|y)\, p(y)}_{p(x,y)}\, dy\, dx = \iint x\, p(x,y)\, dx\, dy$$

$$= E[x] //$$

2.12

$U(x|a,b) = \frac{1}{b-a}$



(i) $\int_a^b U(x|a,b)\,dx = \frac{1}{b-a}\int_a^b 1\,dx = \left[\frac{x}{b-a}\right]_a^b = \frac{b}{b-a} - \frac{a}{b-a} = 1$

(ii) $\int_a^b x\,U(x,a,b)\,dx = \frac{1}{b-a}\int_a^b x\,dx = \frac{1}{2(b-a)}\left[x^2\right]_a^b = \frac{b^2-a^2}{2(b-a)}$ $(b-a)(b+a)$

$= \frac{b+a}{2}$ //

$\int (x+c)^2\,dx$

$u = x+c$

$\frac{du}{dx} = 1$

$du = dx$

$\int u^2\,du = \frac{u^3}{3}$

$\frac{(x-c)^3}{3}$

(iii) $\int_a^b \left(x - \frac{b+a}{2}\right)^2 \frac{1}{(b-a)}\,dx = \frac{1}{b-a}\,\frac{1}{3}\left[\left(x - \frac{b+a}{2}\right)^3\right]_a^b$

$= \frac{1}{3(b-a)}\left[\left(b - \frac{b+a}{2}\right)^3 - \left(a - \frac{b+a}{2}\right)^3\right]$

$= \frac{1}{3(b-a)}\left[\left(\frac{b-a}{2}\right)^3 - \left(\frac{a-b}{2}\right)^3\right] = \frac{2(b-a)^{\cancel{3}\,2}}{\cancel{8}^{4} \times 3\cancel{(b-a)}} = \frac{(b-a)^2}{12}$ //

$c - (-c) = 2c$

2.13

Entropia Relativa

$$KL(p\|q) = -\int p(x)\ln q(x)\,dx - \left(-\int p(x)\ln p(x)\,dx\right) = -\int p(x)\ln\left\{\frac{q(x)}{p(x)}\right\}dx$$

a) $q(x) = p(x)$

$KL(p\|q) = 0$, não há divergência

b) $$KL(p\|q) = -\int p(x)\ln\left\{\frac{\mathcal{N}(x|\mu,\Sigma)}{\mathcal{N}x(m,L)}\right\}dx \quad \mu = m$$

$$\ln p(x) = -\frac{D}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma| - \frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) \quad (2.118)$$

$$\ln q(x) = -\frac{D}{2}\ln(2\pi) - \frac{1}{2}\ln|L| - \frac{1}{2}(x-\mu)^T L^{-1}(x-\mu)$$

$$\ln q(x) - \ln p(x) = -\frac{1}{2}\ln|L| + \frac{1}{2}\ln|\Sigma| - \frac{1}{2}(x-\mu)^T L^{-1}(x-\mu) + \frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)$$

Multiplicando os termos por $p(x)$ e integrando:

$$-\frac{1}{2}\ln|L| + \frac{1}{2}\ln|\Sigma| - \frac{1}{2}\int p(x)(x-\mu)^T L^{-1}(x-\mu)\,dx + \frac{1}{2}\int(x-\mu)^T\Sigma^{-1}(x-\mu)p(x)\,dx$$

Das eq 2.45, 2.50 e 2.56 do Bishop

$$(x-\mu)^T L^{-1}(x-\mu)\,dx = \Delta^2 = \sum_{i=1}^{D}\frac{y_i^2}{\lambda_i}, \quad y_i = u_i^T(x-\mu)$$

$$\int\Delta^2 p(x)\,dx = \int\underbrace{\left(\sum_{i=1}^{D}\frac{y_i^2}{\lambda_i}\right)}_{\Delta^2}\underbrace{\prod_{j=1}^{D}\frac{1}{(2\pi\lambda_j)^{1/2}}e^{\frac{-y_j^2}{2\lambda_j}}}_{p(y)}dy$$

$$\int \Delta^2 p(x)\,dx = \sum_{i=1}^{D} \int \frac{y_i^2}{\lambda_i} \prod_{j=1}^{D} \frac{1}{(2\pi\lambda_j)^{1/2}} e^{\frac{-y_j^2}{2\lambda_j}}\,dy$$

Conforme ex 1.7 da lista 1:

$$\frac{1}{(2\pi\sigma^2)^{1/2}} \int e^{\frac{-x^2}{2\sigma^2}} = \frac{1}{(2\pi\sigma^2)^{1/2}} \int \frac{x^2}{\sigma^2} e^{\frac{-x^2}{2\sigma^2}} = 1$$

$$\int \Delta^2 p(x)\,dx = \sum_{i=1}^{D} 1 = D$$

Portanto

$$KL(p\|q) = -\frac{1}{2}\ln|L| + \frac{1}{2}\ln|\Sigma| - \cancel{\frac{D}{2}} + \cancel{\frac{D}{2}}$$

$$KL(p\|q) = -\frac{1}{2}\ln|L| + \frac{1}{2}\ln|\Sigma|.$$

(2.15)

$$N(x|\mu,\Sigma) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)} \quad (2.43)$$

Entropia

$$H[x] = -\int p(x) \ln p(x)\, dx \quad (1.104)$$

$$H[x] = -\int p(x) \left\{ -\frac{D}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma| - \frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu) \right\} dx$$

$$H[x] = -\frac{D}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma| - \frac{D}{2}$$ //

(2.20)

$$\Sigma u_i = \lambda_i u_i \quad (2.45)$$

Para que Σ seja positivo definido, $\lambda_i \geq 0$

$a^T \Sigma a = a^T \Sigma \alpha_i u_i$ onde $a = \sum_i \alpha_i u_i$

$$a^T \Sigma a = a^T \sum_i \alpha_i \Sigma u_i = a^T \sum_i \alpha_i \lambda_i u_i$$

$$= \sum_j \alpha_j u_j^T \sum_i \alpha_i \lambda_i u_i = \sum_i \alpha_i^2 \lambda_i \quad // \quad \left(u_i^T u_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \right)$$

$$a^T \Sigma a > 0, \quad \lambda_i > 0$$