Physics 786 Spring 2023
Homework 5 Part II

1. Consider a set of random variables $X_1, \cdots, X_N$, such that the joint probability distribution factorizes:

$$p(x_1, \cdots, x_N) = p(x_N|x_{N-1})p(x_{N-1}|x_{N-2}) \cdots p(x_2|x_1)p(x_1) = p(x_1) \prod_{i=2}^{N} p(x_i|x_{i-1}). \tag{1}$$

   Show that the entropy of the full joint distribution decomposes as

$$S_{X_1,\cdots,X_N} = S_{X_1} + S_{X_2|X_1} + \cdots + S_{X_N|X_{N-1}}, \tag{2}$$

   where $S_{X|Y}$ is the conditional entropy between random variables X and Y.

2. (a) Show that the Gaussian distribution is the maximum entropy distribution for a fixed mean and variance.
   (b) Compute the entropy of the Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$

3. (a) Suppose two random variables X and Y have a Gaussian distribution with mean μ_X and μ_Y, respectively, and covariance matrix Σ. Compute the mutual information $I(X;Y)$.
   (b) Suppose that X has probability distribution given by the Gaussian $\mathcal{N}(0, \sigma^2)$. Y is given in terms of X by the equation $Y = X + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. Compute the mutual information $I(X;Y)$. The interpretation here is that X has gone through a "noisy channel" and converted into Y, and we are measuring the average amount of information from X received at the end of the channel.

4. Evaluate the relative entropy (KL-divergence) between two Gaussian distributions $\mathcal{N}(\mu_1, \sigma_1^2)$ and $\mathcal{N}(\mu_2, \sigma_2^2)$.

5. Suppose the conditional entropy $S_{Y|X}$ between two discrete random variables is 0. Show that for all values of x such that $p(x) > 0$, y must be a function of x. That is, there is only one value of y for which $p(y|x) \neq 0$.

6. Let X, Y be binary random variables taking values in $\{0, 1\}$. Suppose that the joint distribution is $p(0,0) = p(0,1) = p(1,1) = 1/3$, $p(1,0) = 0$. Compute S_X, S_Y, $S_{Y|X}$, $S_{X|Y}$, $I(X;Y)$.

7. (a) Consider the Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$. Viewing the two parameters of the distribution as μ and σ^2, compute the Fisher information matrix.
   (b) Using the Fisher information, compute the distance from the distribution $\mathcal{N}(\mu_1, \sigma_1^2)$ to $\mathcal{N}(\mu_2, \sigma_2^2)$ along the path $\mathcal{N}(\lambda\mu_1 + (1-\lambda)\mu_2, \lambda\sigma_1^2 + (1-\lambda)\sigma_2^2)$.
   (c) Compute the distance between $\mathcal{N}(\mu_1, \sigma_1^2)$ to $\mathcal{N}(\mu_2, \sigma_2^2)$ using the Jensen-Shannon divergence.

8. *Data processing inequality.* Consider three random variables X, Y, and Z, and suppose that Z is conditionally independent of X. This means the joint distribution factorizes as $p(x,y,z) = p(z|y)p(y|x)p(x)$. Show that $I(Y;X) \geq I(Z;X)$. This is the data processing inequality, and implies that we cannot increase our information about X by post-processing on Y.

1. $S_{X|Y} = \sum_y p(y) S_{X|Y=y} = -\sum_{x,y} p(x|y) p(y) \log p(x|y)$

$N=2$: $p(x_1, x_2) = p(x_1) \, p(x_2|x_1)$

$$S_{X_1, X_2} = -\sum_{x_1, x_2} p(x_1) p(x_2|x_1) \log p(x_1) p(x_2|x_1)$$

$$= -\sum_{x_1, x_2} \left[p(x_1) \log p(x_1)\right] \times p(x_2|x_1)$$

$$- \sum_{x_1, x_2} p(x_1) p(x_2|x_1) \times \log p(x_2|x_1)$$

$$= -\sum_{x_1} p(x_1) \log p(x_1) - \sum_{x_1, x_2} p(x_2|x_1) p(x_1) \log p(x_2|x_1)$$

$$= S_{X_1} + S_{Y_2|X_1}$$

N: $S_{X_1, \cdots, X_N} = -\sum_{x_1, \cdots x_N} p(x_1, \cdots, x_N) \log p(x_1, \cdots, x_N)$

$$= -\sum_{x_1, \cdots, x_N} p(x_1, \cdots, x_{N-1}) p(x_N|x_{N-1}) \log p(x_1, \cdots, x_{N-1}) p(x_N|x_{N-1})$$

$$= S_{X_1, \cdots, X_{N-1}} - \sum_{x_1, \cdots, x_{N-1}, x_N} p(x_1, \cdots, x_{N-1}) p(x_N|x_{N-1}) \log p(x_N|x_{N-1})$$

$$= S_{X_1, \cdots, X_{N-1}} - \sum_{X_{N-1}, X_N} p(x_{N-1}) \, p(x_N|x_{N-1}) \log p(x_N|x_{N-1})$$

$$= S_{X_1, \cdots, X_{N-1}} + S_{X_N|X_{N-1}}$$

$$= S_{X_1} + S_{X_2|X_1} + \cdots + S_{X_N|X_{N-1}}$$: proof by induction

2. (a) $S = -\int dx\, p(x) \log p(x) + \lambda_1 \left(\int p(x) dx - 1 \right)$

$$+ \lambda_2 \left[\int dx\, p(x) (x-\mu)^2 - \sigma^2 \right]$$

$$\frac{\delta S}{\delta p(x)} = -1 - \log p(x) + \lambda_1 + \lambda_2 (x-\mu)^2 = 0$$

$$\Rightarrow \quad p(x) = \exp\left(-1 + \lambda_1 + \lambda_2 (x-\mu)^2 \right)$$

$$1 = \int dx\, p(x) = \exp(-1+\lambda_1) \int dx\, \exp(\lambda_2 (x-\mu)^2)$$

$$= \exp(-1+\lambda_1) \sqrt{\frac{\pi}{-\lambda_2}}$$

$$\sigma^2 = \int dx\, p(x)(x-\mu)^2 = \exp(-1+\lambda_1) \int dx\, (x-\mu)^2 \exp(\lambda_2 (x-\mu)^2)$$

$$= \exp(-1+\lambda_1) \int dx\, x^2 \exp(\lambda_2 x^2)$$

$$= \exp(-1+\lambda_1) \frac{\sqrt{\pi}}{2(-\lambda_2)^{3/2}}$$

$$\Rightarrow \quad \sigma^2 = \frac{1}{-2\lambda_2} \qquad 1 = \exp(-1+\lambda_1) \sqrt{2\pi\sigma^2}$$

$$\lambda_1 = 1 + \log \sqrt{\frac{1}{2\pi\sigma^2}}$$

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

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

(b) $S = -\int dx\, \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right) \left(\log \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{(x-\mu)^2}{2\sigma^2} \right)$

$$= -\int dx\, p(x) \left(\log \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{x^2}{2\sigma^2} \right)$$

$$= \frac{1}{2} \left(1 + \log 2\pi\sigma^2 \right)$$

3. (a) $I(X;Y) = S_X + S_Y - S_{X,Y}$

$$z \equiv \begin{bmatrix} x \\ y \end{bmatrix} \qquad \mu \equiv \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}$$

$$= 1 + \frac{1}{2}\left(\log 2\pi \Sigma_{xx} + \log 2\pi \Sigma_{yy}\right)$$

$$+ \int dx\,dy \quad \frac{1}{2\pi\sqrt{\det\Sigma}}\, e^{-\frac{1}{2}(z-\mu)^T \Sigma^{-1}(z-\mu)}$$

$$\times\left(-\frac{1}{2}(z-\mu)^T \Sigma^{-1}(z-\mu) + \log \frac{1}{2\pi}\frac{1}{\sqrt{\det\Sigma}}\right)$$

$$= 1 + \frac{1}{2}\left(\log 2\pi\Sigma_{xx} + \log 2\pi\Sigma_{yy}\right)$$

$$- 1 - \log 2\pi - \frac{1}{2}\log\det\Sigma$$

$$= \frac{1}{2}\log\frac{\Sigma_{xx}\Sigma_{yy}}{\det\Sigma}$$

(b) $Y \sim N(0, \sigma^2 + \sigma_\epsilon^2)$

$$\Sigma = \begin{bmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + \sigma_\epsilon^2 \end{bmatrix}$$

$$\Rightarrow I(X,Y) = \frac{1}{2}\log\frac{\sigma^2(\sigma^2+\sigma_\epsilon^2)}{\sigma^2\sigma_\epsilon^2}$$

$$= \frac{1}{2}\log\frac{\sigma^2+\sigma_\epsilon^2}{\sigma_\epsilon^2}$$

4.

$$S_{KL} = \int dx \; p_1(x) \log \frac{p_1(x)}{p_2(x)}$$

$$= \int dx \; \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right) \log\left[\frac{\sigma_2}{\sigma_1} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2} + \frac{(x-\mu_2)^2}{2\sigma_2^2}\right)\right]$$

$$= \int dx \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right) \times \left(\log\frac{\sigma_2}{\sigma_1} + \frac{(x-\mu_2)^2}{2\sigma_2^2} - \frac{(x-\mu_1)^2}{2\sigma_1^2}\right)$$

$$= \log\frac{\sigma_2}{\sigma_1} - \frac{1}{2} + \frac{1}{2\sigma_2^2}\left((\mu_1-\mu_2)^2 + \sigma_1^2\right)$$

5. $S_{Y|X} = S_{X,Y} - S_X = 0$

$$= -\sum_{x,y} p(y|x)\, p(x) \log p(y|x) = 0$$

$p(x) \geq 0$, $p(y|x) \geq 0$ & $-\log p(y|x) \geq 0$

$p(x) > 0 \Rightarrow -\sum_y p(y|x) \log p(y|x)$ should be zero

with $\sum_y p(y|x) = 1$

only possibility: $p(y|x) = \begin{cases} 1 & \text{for some } y \\ 0 & \text{else} \end{cases}$

: only one value of y for which $p(y|x) \neq 0$

actually 1.

$\Rightarrow$ y is a function of x.

6. $S_X = -\sum_x p(x) \log p(x)$ $\quad p(x,y) = p(x|y)\,p(y)$

$P_X(0) = P(0,0) + P(0,1) = \frac{2}{3}$

$P_X(1) = \frac{1}{3}$

$\Rightarrow S_X = -\frac{2}{3}\log\frac{2}{3} - \frac{1}{3}\log\frac{1}{3}$

$P_Y(0) = P(0,0) + P(1,0) = \frac{1}{3}$

$P_Y(1) = \frac{2}{3}$

$\Rightarrow S_Y = -\frac{1}{3}\log\frac{1}{3} - \frac{2}{3}\log\frac{2}{3}$

$$S_{Y|X} = -\sum_{x,y} p(x,y) \log \frac{p(x,y)}{P_X(x)}$$

$$= -\frac{1}{3}\log\frac{\frac{1}{3}}{\frac{2}{3}} - \frac{1}{3}\log\frac{\frac{1}{3}}{\frac{2}{3}} - \frac{1}{3}\log\frac{\frac{1}{3}}{\frac{1}{3}}$$

$$= \frac{2}{3}\log 2$$

$$S_{X|Y} = -\sum_{x,y} p(x,y) \log \frac{p(x,y)}{P_Y(y)} = -\frac{1}{3}\log\frac{\frac{1}{3}}{\frac{1}{3}} - \frac{1}{3}\log\frac{\frac{1}{3}}{\frac{2}{3}}$$

$$- \frac{1}{3}\log\frac{1/3}{2/3}$$

$$= \frac{2}{3}\log 2$$

$$I(X;Y) = S_X - S_{X|Y} = -\frac{2}{3}\log\frac{2}{3} - \frac{1}{3}\log\frac{1}{3} - \frac{2}{3}\log 2$$

$$= -\frac{4}{3}\log 2 - \log\frac{1}{3}$$

$$= \log 3 - \frac{4}{3}\log 2$$

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

$$S_{p\|q} = \sum_i p_i \log \frac{p_i}{q_i}$$

$$= \int dx \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right) \left(\frac{1}{2}\log\frac{\sigma_2^2}{\sigma_1^2} - \frac{(x-\mu_1)^2}{2\sigma_1^2} + \frac{(x-\mu_2)^2}{2\sigma_2^2}\right)$$

$$\frac{\partial^2 S_{p\|q}}{\partial\mu_2\partial\mu_2} = \int dx\, p(x;\mu_1,\sigma_1) \frac{1}{\sigma_2}\Big|_{\substack{\mu=\mu \\ \sigma=\sigma}} = \frac{1}{\sigma^2}$$

$$\frac{\partial^2 S_{p\|q}}{\partial\sigma_2^2\partial\sigma_2^2} = \int dx\, p(x;\mu_1,\sigma_1)\left(-\frac{1}{2\sigma_2^4} + \frac{(x-\mu_2)^2}{\sigma_2^6}\right)\Big|_{\substack{\mu=\mu \\ \sigma=\sigma}}$$

$$= -\frac{1}{2\sigma^4} + \frac{1}{\sigma^4} = \frac{1}{2\sigma^4}$$

$$\frac{\partial^2 S_{p\|q}}{\partial\mu_2\partial\sigma_2^2} = \int dx\, p(x;\mu_1,\sigma_1)\left(-\frac{\mu-x_2}{\sigma_2^4}\right)\Big|_{\substack{\mu=\mu \\ \sigma=\sigma}} = 0$$

$$F = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{bmatrix}$$

(b) $ds = \sqrt{g_{\alpha\beta}\, dx^\alpha dx^\beta}$

$$\mu(\lambda) = \lambda\mu_1 + (1-\lambda)\mu_2 \qquad \sigma^2(\lambda) = \lambda\sigma_1^2 + (1-\lambda)\sigma_2^2$$

$$\frac{d\mu}{d\lambda} = \mu_1 - \mu_2 \qquad \frac{d\sigma^2}{d\lambda} = \sigma_1^2 - \sigma_2^2$$

$$ds = \sqrt{g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}}\, d\lambda$$

$$= \sqrt{\frac{1}{\sigma^2}(\mu_1-\mu_2)^2 + \frac{1}{2\sigma^4}(\sigma_1^2-\sigma_2^2)^2} \quad d\lambda$$

$$= \sqrt{\frac{(\mu_1-\mu_2)^2}{\lambda\sigma_1^2+(1-\lambda)\sigma_2^2} + \frac{(\sigma_1^2-\sigma_2^2)^2}{2(\lambda\sigma_1^2+(1-\lambda)\sigma_2^2)^2}} \; d\lambda$$

$$\text{distance} = \int_0^1 d\lambda \sqrt{\frac{(\mu_1-\mu_2)^2}{\lambda\sigma_1^2+(1-\lambda)\sigma_2^2} + \frac{(\sigma_1^2-\sigma_2^2)^2}{2(\lambda\sigma_1^2+(1-\lambda)\sigma_2^2)^2}}$$

(c) $JSD(P,Q) = S_{P \| \frac{P+Q}{2}} + S_{Q \| \frac{P+Q}{2}}$

$$= S_{P, \frac{1}{2}(P+Q)} + S_{Q, \frac{1}{2}(P+Q)} - S_P - S_Q$$

$$= 2S_{\frac{P+Q}{2}, \frac{P+Q}{2}} - (S_P + S_Q)$$

$$= -2\int_0^1 dx \left(\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} \right)$$

$$\times \log\left(\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} \right)$$

8. WTS: $S_{X,Y|Z} = S_{X|Z} + S_{Y|X,Z}$

pf) $S_{X,Y|Z} = -\sum_z p(z) \sum_{x,y} p(x,y|z) \log p(x,y|z)$

$$p(x,y,z) = p(z)\, p(x,y|z)$$

$$= -\sum_{x,y,z} p(x,y,z) \log p(x,y|z)$$

$$p(x,y|z) = p(x|z)\, p(y|x,z)$$

$$\Rightarrow S_{X,Y|Z} = -\sum_{x,y,z} p(x,y,z) \log p(x|z)\, p(y|x,z)$$

$$= -\sum_{x,y,z} p(x,y,z) \log p(x|z)$$

$$- \sum_{x,y,z} p(x,y,z) \log p(y|x,z)$$

$$= S_{X|Z} + S_{Y|X,Z}$$

$I(XY;Z) = S_{XY} - S_{XY|Z} = S_X + S_{Y|X} - S_{X|Z} - S_{Y|X,Z}$

$I(Z;XY)$ $\qquad = I(X;Z) + I(Y;Z|X)$

$I(X;YZ) = I(X;Z) + I(X;Y|Z)$

$I(X;YZ) = I(X,Z) + I(X;Y|Z)$

$= I(X,Y) + \underbrace{I(X;Z|Y)}_{=0}$ by Z is conditionally independent

Since $I(X;Y|Z) \geq 0$, $\quad \underbrace{I(X;Y)}_{=I(Y;X)} \geq \underbrace{I(X;Z)}_{=I(Z;X)}$

Another proof

$I(Z,X) = S_X + S_Z - S_{Z,X}$, $I(Y,X) = S_Y + S_X - S_{X,Y}$

$$I(Y,X) - I(Z,X) = (S_Y - S_{X,Y}) - (S_Z - S_{Z,X})$$

$$= S_{Z|X} - S_{Y|X}$$

$$= - \sum_{x,y,z} p(z|y)p(y|x)p(x) \log p(z|y)p(y|x)p(x)$$

$$+ \sum_{x,y} p(y|x)p(x) \log p(y|x)p(x)$$

$$= - \sum_{x,y,z} \frac{p(y|z)p(z)}{p(y)} p(y|x)p(x) \log p(z|y)p(y|x)p(x)$$

$$+ \sum_{x,y} p(y|x)p(x) \log p(y|x)p(x)$$

$$\geq - \sum_{x,y,z} \frac{p(y|z)p(z)}{p(y)} p(y|x)p(x) \log p(y|x)p(x)$$

$$+ \sum_{x,y} p(y|x)p(x) \log p(y|x)p(x)$$

$$= I(X, Y|Z) \geq 0.$$