EE 231A HW5 Shengze Ye 205418959

1. Fano's inequality without conditioning

let $Pr(X=i) = p_i$ $i=1,2\dots m$

$p_1 \geq p_2 \geq \dots \geq p_m$. The minimal probability of error predictor of X is $\hat{X}=1$ $P_e = 1-p_1$

(a) choose $p_2, \dots p_m$ so as to maximize $H(X)$ subject to the constraint $1-p_1 = P_e$ to find an upper bound on $H(X)$ that is a function of the constrain P_e.

$P_e = 1-p_1$ is the error probability of predictor $\hat{X}=1$. so, if we fix the constraint P_e, we fix p_1

$H(X) = H(p_1, p_2, p_3, \dots p_m)$

According to grouping theory

$H(X) = H(p_1, p_2, p_3 \dots p_m)$

$$= H(p_1) + (p_2+p_3+\dots+p_m) H\left(\frac{p_2}{p_2+\dots+p_m}, \dots, \frac{p_m}{p_2+\dots+p_m}\right)$$

$\because 1-p_1 = P_e$ $\therefore p_2+p_3+\dots+p_m = 1-p_1 = P_e$

$$H(X) = H(P_e) + P_e H\left(\frac{p_2}{P_e}, \frac{p_3}{P_e}, \frac{p_4}{P_e}, \dots \frac{p_m}{P_e}\right)$$

pe is a fixed number, if we want to maximize $H(X)$, we need to make $H(p_2/pe, \frac{p_3}{pe}, \cdots p_m/pe)$ as large as possible.

$\Rightarrow p_2, \cdots, p_m$ should follow the uniform distribution

$$p_2 = p_3 = \cdots = p_m = \frac{P_e}{m-1}$$

$$H\left(\frac{p_2}{pe}, \frac{p_3}{pe}, \frac{p_4}{pe} \cdots \frac{p_m}{pe}\right) = H\left(\frac{1}{m-1}, \frac{1}{m-1}, \cdots \frac{1}{m-1}\right)$$
$$= \log(m-1)$$

$$\therefore H(X) \leq H(pe) + pe \log(m-1)$$

$\therefore$ The upper bound for $H(X)$ is $H(pe) + pe \log(m-1)$ which could be achieved when

$$p_2 = p_3 = \cdots p_m = \frac{P_e}{m-1} \quad \text{(uniform distribution)}$$

(b) Now upper bound $H(P_e)$ to provide a lower bound on P_e that corresponds to (2.132) but without the conditioning on Y

$$H(P_e) = H(P_e, 1-P_e) \leq 1$$

$$\therefore H(X) \leq 1 + P_e \log(m-1)$$

$$P_e \geq \frac{H(X) - 1}{\log(m-1)}$$

A lower bound on P_e is $\frac{H(X)-1}{\log(m-1)}$

2. Differential entropy

(a) The exponential density, $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$

$$h(f) = -\int_0^\infty \lambda e^{-\lambda x} \ln(\lambda e^{-\lambda x})\,dx$$

$$= -\int_0^\infty \lambda e^{-\lambda x} [\ln\lambda - \lambda x]\,dx$$

$$= \underset{①}{-\int_0^\infty \lambda \ln\lambda\, e^{-\lambda x}\,dx} + \underset{②}{\int_0^\infty \lambda^2 x e^{-\lambda x}\,dx}$$

$$① = -\lambda\ln\lambda \int_0^\infty e^{-\lambda x}\,dx = -\lambda\ln\lambda\left(-\frac{1}{\lambda}\right)e^{-\lambda x}\Big|_0^\infty$$

$$= \ln\lambda\,(0-1) = -\ln\lambda$$

$$② = \lambda^2\int_0^\infty x e^{-\lambda x}\,dx = \lambda^2\left[-\frac{x}{\lambda}e^{-\lambda x}\Big|_0^\infty - \int_0^\infty -\frac{1}{\lambda}e^{-\lambda x}\,dx\right]$$

$$= \lambda^2\left[0 - 0 - \int_0^\infty -\frac{1}{\lambda}e^{-\lambda x}\,dx\right]$$

$$= \lambda^2\left[\frac{1}{\lambda}\left(-\frac{1}{\lambda}\right)e^{-\lambda x}\Big|_0^\infty\right] = -\lambda^2\left[\frac{1}{\lambda^2}(0-1)\right] = 1$$

$$\therefore h(x) = ① + ② = -\ln\lambda + 1 \text{ nats}$$

$$= -\log\lambda + 1 \text{ bits}$$

$$\therefore h(f) = \log e/\lambda \text{ bits}$$

(b) The Laplace density, $f(x) = \frac{1}{2}\lambda e^{-\lambda|x|}$

$$h(f) = -\int_{-\infty}^{+\infty} \frac{1}{2}\lambda e^{-\lambda|x|} \ln\left(\frac{1}{2}\lambda e^{-\lambda|x|}\right) dx$$

$$= -\left[\underbrace{\int_{-\infty}^{0} \frac{1}{2}\lambda e^{\lambda x} \ln\left(\frac{1}{2}\lambda e^{\lambda x}\right) dx}_{①} + \underbrace{\int_{0}^{\infty} \frac{1}{2}\lambda e^{-\lambda x} \ln\left(\frac{1}{2}\lambda e^{-\lambda x}\right) dx}_{②}\right]$$

$$① = \int_{-\infty}^{0} \frac{1}{2}\lambda e^{\lambda x}\left[\ln\frac{1}{2} + \ln\lambda + \lambda x\right] dx$$

$$= \int_{-\infty}^{0} \frac{1}{2}\lambda \ln\frac{1}{2} e^{\lambda x} dx + \int_{-\infty}^{0} \frac{1}{2}\lambda \cdot \ln\lambda\, e^{\lambda x} dx$$

$$+ \int_{-\infty}^{0} \frac{1}{2}\lambda^2 x e^{\lambda x} dx$$

$$= \frac{1}{2}\lambda \ln\frac{1}{2} \left.\frac{1}{\lambda} e^{\lambda x}\right|_{-\infty}^{0} + \frac{1}{2}\lambda \ln\lambda \left.\frac{1}{\lambda} e^{\lambda x}\right|_{-\infty}^{0}$$

$$+ \frac{1}{2}\lambda^2\left[\left.\frac{x}{\lambda} e^{\lambda x}\right|_{-\infty}^{0} - \int_{-\infty}^{0} \frac{1}{\lambda} e^{\lambda x} dx\right]$$

$$= \frac{1}{2}\ln\frac{1}{2}(1-0) + \frac{1}{2}\ln\lambda(1-0) + \frac{1}{2}\lambda^2\left[-\left.\frac{1}{\lambda^2} e^{\lambda x}\right|_{-\infty}^{0}\right]$$

$$= \frac{1}{2}\ln\frac{1}{2} + \frac{1}{2}\ln\lambda + \frac{1}{2}\lambda^2\left(-\frac{1}{\lambda^2}\right)$$

$$= \frac{1}{2}\ln\frac{1}{2} + \frac{1}{2}\ln\lambda - \frac{1}{2}$$

$$② = \int_{0}^{\infty} \frac{1}{2}\lambda e^{-\lambda x}\left[\ln\frac{1}{2} + \ln\lambda - \lambda x\right] dx$$

$$= \int_{0}^{\infty} \frac{1}{2}\lambda \ln\frac{1}{2} e^{-\lambda x} dx + \int_{0}^{\infty} \frac{1}{2}\lambda \ln\lambda\, e^{-\lambda x} dx + \int_{0}^{\infty} \frac{-\lambda^2}{2} x e^{-\lambda x} dx$$

$$= \frac{1}{2}\lambda \ln\frac{1}{2}\left(-\frac{1}{\lambda}\right)e^{-\lambda x}\Big|_0^\infty + \frac{1}{2}\lambda\ln\lambda\left(-\frac{1}{\lambda}\right)e^{-\lambda x}\Big|_0^\infty$$

$$-\frac{\lambda^2}{2}\left[-\frac{x}{\lambda}e^{-\lambda x}\Big|_0^\infty - \int_0^\infty -\frac{1}{\lambda}e^{-\lambda x}dx\right]$$

$$= -\frac{1}{2}\ln\frac{1}{2}(0-1) + \left(-\frac{1}{2}\right)\ln\lambda(0-1) - \frac{\lambda^2}{2}\left[0+\int_0^\infty \frac{1}{\lambda}e^{-\lambda x}dx\right]$$

$$= \frac{1}{2}\ln\frac{1}{2} + \frac{1}{2}\ln\lambda - \frac{\lambda^2}{2}\left[-\frac{1}{\lambda^2}e^{-\lambda x}\Big|_0^\infty\right]$$

$$= \frac{1}{2}\ln\frac{1}{2} + \frac{1}{2}\ln\lambda - \frac{1}{2}$$

$$\therefore h(f) = -(① + ②) = -\left[\frac{1}{2}\ln\frac{1}{2} + \frac{1}{2}\ln\lambda - \frac{1}{2}\right]\times 2$$

$$= 1 - \ln\frac{1}{2} - \ln\lambda \text{ nats}$$

$$h(f) = 1 - \log\frac{1}{2} - \log\lambda \text{ bits}$$

$$\therefore h(f) = \log 2e/\lambda \text{ bits}$$

(c) The sum of X_1 and X_2 where X_1 and X_2 are independent normal random variables with means μ_i and variance σ_i^2 $i=1,2$

$$X_1 \sim N(\mu_1, \sigma_1) \qquad X_2 \sim N(\mu_2, \sigma_2)$$

For Gaussian distribution

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

$$h(X) = -\int f(x)\log f(x)\,dX$$

$$= \int f(x)\left(\tfrac{1}{2}\log 2\pi\sigma^2\right)dX + \int f(x)\log(e)\left(\frac{x^2}{2\sigma^2}\right)dX$$

$$= \tfrac{1}{2}\log 2\pi\sigma^2 + \tfrac{1}{2}\log e \quad = \tfrac{1}{2}\log 2\pi e\sigma^2$$

X_1 and X_2 are independent

$\therefore \quad X_1 + X_2 \sim N(u_1+u_2,\ \sigma_1^2+\sigma_2^2)$

$h(f) = \frac{1}{2}\log 2\pi e(\sigma_1^2+\sigma_2^2)$ bits

3. Exponential Channel

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

$$f_{Y|X}(y|x) = \begin{cases} \frac{3}{2}e^{-y} \text{ for } 0 \le y \le \ln 3 & 0 \text{ otherwise} \quad X=0 \\ \frac{3}{2}e^{-y} \text{ for } \ln\frac{3}{2} \le y \le \infty & 0 \text{ otherwise} \quad X=1 \end{cases}$$

(a) Give the expression for $f_Y(y)$ and show that your expression integrates to 1

According to the conditional distribution of y

$$f_Y(y) = \sum_x f_{Y|X}(y|x)\,P_X(x)$$

$0 \le y \le \ln\frac{3}{2} \rightarrow f_Y(y) = \frac{3}{2}e^{-y}\cdot P_X(0)$

$\ln\frac{3}{2} \le y \le \ln 3 \rightarrow f_Y(y) = \frac{3}{2}e^{-y}\cdot P_X(0) + \frac{3}{2}e^{-y}P_X(1)$

$\ln 3 \le y \le \infty \rightarrow f_Y(y) = \frac{3}{2}e^{-y}P_X(1)$

if X follows the uniform distribution

$$f_Y(y) = \begin{cases} \frac{3}{4}e^{-y} & 0 \le y \le \ln\frac{3}{2} \\ \frac{3}{2}e^{-y} & \ln\frac{3}{2} \le y \le \ln 3 \\ \frac{3}{4}e^{-y} & \ln 3 \le y \le \infty \end{cases}$$

$$\int_0^\infty f_Y(y) = \int_0^{\ln\frac{3}{2}} \frac{3}{4}e^{-y}dy + \int_{\ln\frac{3}{2}}^{\ln 3} \frac{3}{2}e^{-y}dy + \int_{\ln 3}^{\infty} \frac{3}{4}e^{-y}dy$$

$$= \frac{3}{4}(-e^{-y})\Big|_0^{\ln\frac{3}{2}} + \frac{3}{2}(-e^{-y})\Big|_{\ln\frac{3}{2}}^{\ln 3} + \frac{3}{4}(-e^{-y})\Big|_{\ln 3}^{\infty}$$

$$= \frac{3}{4} \times (1 - \frac{2}{3}) + \frac{3}{2}(\frac{2}{3} - \frac{1}{3}) + \frac{3}{4}(\frac{1}{3} - 0)$$

$= \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$ it integrates to 1

(b) Compute $H(X)$ for X equally likely to be 0 or 1

$p(X=0) = \frac{1}{2} = p(X=1)$

$H(X) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1$ bit

$H(X) = 1$ bit

(c) Compute $H(X|Y=y)$ for the three cases

$0 \le y \le \ln\frac{3}{2}$, $\ln\frac{3}{2} \le y \le \ln 3$, $\ln 3 \le y \le \infty$

① $H(X|Y=y)$ $\quad 0 \le y \le \ln \frac{3}{2}$

we are sure that $X=0$ in this situation

$\therefore H(X|Y=y) = 0$

② $H(X|Y=y)$ $\quad \ln \frac{3}{2} \le y < \ln 3$

we are not sure $X=0$ or $X=1$, it depends on the prob that $P(X=0)$ and $P(X=1)$

$\therefore H(X|Y=y) = H(X) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1$

③ $H(X|Y=y)$ $\quad \ln 3 \le y \le \infty$

we are sure that $X=1$ in this case

$\therefore H(X|Y=y) = 0$

(d) Compute $H(X|Y)$

$$H(X|Y) = \int_0^\infty H(X|Y=y) f_Y(y)\,dy$$

$$= H(X|Y=y_1)\Pr(0 \le y_1 \le \ln\tfrac{3}{2}) + H(X|Y=y_2)\Pr(\ln\tfrac{3}{2} \le y_2 \le \ln 3)$$

$$+ H(X|Y=y_3)\Pr(\ln 3 \le y_3 \le \infty)$$

$$= 0 \cdot \frac{1}{4} + H(X)\cdot\frac{1}{2} + 0\cdot\frac{1}{4} = \frac{1}{2}H(X)$$

$H(X|Y) = \frac{1}{2}H(X)$

if X is uniform distribution

$\rightarrow H(X|Y) = \frac{1}{2}$ bit

(e) For X equally likely to be 0 or 1, comput $I(X;Y)$

$I(X;Y) = H(X) - H(X|Y) = H(X) - \frac{1}{2}H(X) = \frac{1}{2}H(X)$

X is uniform distribution $H(X) = 1$

$\therefore I(X;Y) = \frac{1}{2} \cdot 1 = \frac{1}{2}$ bit

(f) Compute $h(Y) = -\int_0^{\infty} f_Y(y) \ln(f_Y(y))\,dy$

$h(Y) = -\Big[\underbrace{\int_0^{\ln\frac{3}{2}} \frac{3}{4}e^{-y} \ln\left(\frac{3}{4}e^{-y}\right)dy}_{①} + \underbrace{\int_{\ln\frac{3}{2}}^{\ln 3} \frac{3}{2}e^{-y} \ln\left(\frac{3}{2}e^{-y}\right)dy}_{②}$

$+ \underbrace{\int_{\ln 3}^{\infty} \frac{3}{4}e^{-y} \ln\left(\frac{3}{4}e^{-y}\right)dy}_{③} \Big]$

$① = \int_0^{\ln\frac{3}{2}} \frac{3}{4}e^{-y}\left[\ln\frac{3}{4} - y\right]dy = \int_0^{\ln\frac{3}{2}} \frac{3}{4}\ln\frac{3}{4}\, e^{-y}dy +$

$\int_0^{\ln\frac{3}{2}} \frac{3}{4} -ye^{-y}dy$

$= \frac{3}{4}\ln\frac{3}{4}(-e^{-y})\Big|_0^{\ln\frac{3}{2}} + \frac{3}{4}\left[ye^{-y}\Big|_0^{\ln\frac{3}{2}} - \int_0^{\ln 3} e^{-y}dy\right]$

$= \frac{3}{4}\ln\frac{3}{4} \times \frac{1}{3} + \frac{3}{4}\left[\frac{2}{3}\ln\frac{3}{2} + \left(\frac{2}{3} - 1\right)\right]$

$= \frac{1}{4}\ln\frac{3}{4} + \frac{1}{2}\ln\frac{3}{2} - \frac{1}{4}$ nats

$$② = \int_{\ln\frac{3}{2}}^{\ln 3} \frac{3}{2}e^{-y}\ln\left(\frac{3}{2}e^{-y}\right)dy = \int_{\ln\frac{3}{2}}^{\ln 3} \frac{3}{2}e^{-y}\left[\ln\frac{3}{2} - y\right]dy$$

$$= \frac{3}{2}\ln\frac{3}{2}\int_{\ln\frac{3}{2}}^{\ln 3} e^{-y}dy + \frac{3}{2}\int_{\ln\frac{3}{2}}^{\ln 3} -ye^{-y}dy$$

$$= \frac{3}{2}\ln\frac{3}{2}\left(-e^{-y}\right)\Big|_{\ln\frac{3}{2}}^{\ln 3} + \frac{3}{2}\left[ye^{-y}\Big|_{\ln\frac{3}{2}}^{\ln 3} - \int_{\ln\frac{3}{2}}^{\ln 3} e^{-y}dy\right]$$

$$= \frac{3}{2}\ln\frac{3}{2} \times \frac{1}{3} + \frac{3}{2}\left[\frac{1}{3}\ln 3 - \frac{2}{3}\ln\frac{3}{2} + \left(\frac{1}{3} - \frac{2}{3}\right)\right]$$

$$= \frac{1}{2}\ln\frac{3}{2} + \left[\frac{1}{2}\ln 3 - \ln\frac{3}{2} - \frac{1}{2}\right]$$

$$= \frac{1}{2}\ln 3 - \frac{1}{2}\ln\frac{3}{2} - \frac{1}{2} \text{ nats}$$

$$③ = \int_{\ln 3}^{\infty} \frac{3}{4}e^{-y}\ln\left(\frac{3}{4}e^{-y}\right)dy = \int_{\ln 3}^{\infty} \frac{3}{4}e^{-y}\left[\ln\frac{3}{4} - y\right]dy$$

$$= \frac{3}{4}\ln\frac{3}{4}\int_{\ln 3}^{\infty} e^{-y}dy + \frac{3}{4}\int_{\ln 3}^{\infty} -ye^{-y}dy$$

$$= \frac{3}{4}\ln\frac{3}{4}\left(-e^{-y}\right|_{\ln 3}^{\infty} + \frac{3}{4}\left[ye^{-y}\Big|_{\ln 3}^{\infty} - \int_{\ln 3}^{\infty} e^{-y}dy\right]$$

$$= \frac{3}{4}\ln\frac{3}{4} \times \frac{1}{3} + \frac{3}{4}\left[0 - \frac{1}{3}\ln 3 + e^{-y}\Big|_{\ln 3}^{\infty}\right]$$

$$= \frac{1}{4}\ln\frac{3}{4} + \frac{3}{4}\left[-\frac{1}{3}\ln 3 + \left[0 - \frac{1}{3}\right]\right]$$

$$= \frac{1}{4}\ln 3 - \frac{3}{4}\ln 3 - \frac{1}{4} \text{ nats}$$

$h(Y) = -(① + ② + ③)$

$= -\left(\frac{1}{4}\ln\frac{3}{4} + \frac{1}{2}\ln\frac{3}{2} - \frac{1}{4} + \frac{1}{2}\ln 3 - \frac{1}{2}\ln\frac{3}{2} - \frac{1}{2}\right.$

$\left. + \frac{1}{4}\ln\frac{3}{4} - \frac{1}{4}\ln 3 - \frac{1}{4}\right)$ nats

$= 1 - \frac{1}{2}\ln\frac{3}{4} - \frac{1}{4}\ln 3$ nats

1 bit = ln2 nats $\qquad$ 1 nat = $\frac{1}{\ln 2}$ bit

$\therefore h(Y) = \frac{1}{\ln 2} - \frac{1}{2}\log\frac{3}{4} - \frac{1}{4}\log 3$ bits

(g) $h(Y|X)$

$h(Y|X) = \underset{①}{h(Y|X=0)}\, p(X=0) + \underset{②}{h(Y|X=1)}\, p(X=1)$

$① = h(Y|X=0) = -\int_0^{\ln 3} \frac{3}{2}e^{-y} \ln\left(\frac{3}{2}e^{-y}\right)dy$

$= -\int_0^{\ln 3} \frac{3}{2}e^{-y}\left[\ln\frac{3}{2} - y\right]dy = -\left\{\frac{3}{2}\ln\frac{3}{2}\int_0^{\ln 3} e^{-y}dy + \frac{3}{2}\int_0^{\ln 3} -ye^{-y}dy\right\}$

$= -\left\{\frac{3}{2}\ln\frac{3}{2} \times \frac{2}{3} + \frac{3}{2}\left[ye^{-y}\Big|_0^{\ln 3} - \int_0^{\ln 3} e^{-y}dy\right]\right\}$

$= -\left[\ln\frac{3}{2} - \frac{3}{2}\left[\frac{1}{3}\ln 3 + e^{-y}\Big|_0^{\ln 3}\right]\right]$

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

$= -\ln\frac{3}{2} - \frac{3}{2}\ln 3 + 1$ nats

$② = h(Y|X=1) = -\int_{\ln\frac{3}{2}}^{\infty} \frac{3}{2}e^{-y}\ln\left(\frac{3}{2}e^{-y}\right)dy$

$$= -\left\{\frac{3}{2}\ln\frac{3}{2}\int_{\ln\frac{3}{2}}^{\infty} e^{-y}dy + \frac{3}{2}\int_{\ln\frac{3}{2}}^{\infty} -ye^{-y}dy\right\}$$

$$= -\left[\frac{3}{2}\ln\frac{3}{2}\times\frac{2}{3} + \frac{3}{2}\left[ye^{-y}\Big|_{\ln\frac{3}{2}}^{\infty} - \int_{\ln\frac{3}{2}}^{\infty} e^{-y}dy\right]\right]$$

$$= -\left[\ln\frac{3}{2} + \frac{3}{2}\left[0 - \frac{2}{3}\ln\frac{3}{2} + e^{-y}\Big|_{\ln\frac{3}{2}}^{\infty}\right]\right]$$

$$= -\left\{\ln\frac{3}{2} + \frac{3}{2}\left[-\frac{2}{3}\ln\frac{3}{2} + \left(0 - \frac{2}{3}\right)\right]\right\}$$

$$= \ln\frac{3}{2} - \ln\frac{3}{2} + 1 = 1 \text{ nat}$$

$$h(Y|X) = h(Y|X=0)\,p(X=0) + h(Y|X=1)\,p(X=1)$$

$$= P_X(X=0)\cdot\left(1 - \ln\frac{3}{2} - \frac{1}{2}\ln 3\right) + 1\cdot P(X=1)$$

if $p(X=0) = p(X=1) = \frac{1}{2}$ uniform distribution

$$h(Y|X) = 1 - \frac{1}{2}\ln\frac{3}{2} - \frac{1}{4}\ln 3 \text{ nats}$$

$$h(Y|X) = \frac{1}{\ln 2} - \frac{1}{2}\log\frac{3}{2} - \frac{1}{4}\log 3 \text{ bits}$$

(h) For X equally likely to be 0 or 1 compute $I(X;Y)$ using $h(Y)$ and $h(Y|X)$

$$I(X;Y) = h(Y) - h(Y|X)$$

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

$$= \frac{1}{2}\log\frac{3}{2} - \log\frac{3}{4} = \frac{1}{2}\log\frac{3}{2}\Big/\frac{3}{4} = \frac{1}{2}\log 2 = \frac{1}{2}\text{ bit}$$

$\therefore I(X;Y) = \frac{1}{2}$ bit

4. Conditional entropy of product

(a) For a discrete random variable Y, expreess $H(aY)$ in terms of $H(Y)$, assume $a \neq 0$.

$H(aY) = H(Y)$

Since the values of outcomes do not affect the entropy.

$\because a \neq 0$ $\therefore$ $Y \to aY$ is a one-to-one mapping.

The probability distribution remains the same

$\therefore H(aY) = H(Y)$

(b) Find a simplified expression for $H(XY|X)$ involving $H(Y|X)$, $P(X=0)=0$

$$H(XY|X) = \sum_i H(x_i Y | X=x_i)\, p(X=x_i)$$

$(P(X_i=0)=0)$

$$= \sum_i H(Y|X=x_i)\, p(X=x_i) = H(Y|X)$$

$\therefore\ H(XY|X) = H(Y|X)$

(c) Now consider a continuous random variable Y with $f(Y)$. Find a simplified expression for $h(XY|X)$ involving $h(Y|X)$

$$h(XY|X) = \sum_i h(x_i Y | X=x_i)\, p(X=x_i)$$

$$= \sum_i [h(Y|X=x_i) + \log|x_i|]\, p(X=x_i)$$

$$= \sum_i h(Y|X=x_i)\, p(X=x_i) + \sum_i \log|x_i|\, p(X=x_i)$$

$$= h(Y|X) + E[\log|X|]$$

$$\therefore\ h(XY|X) = h(Y|X) + E[\log|X|]$$

5. Data Processing and Entropy

(a) As a brief review, prove that $H(g(X)) \le H(X)$ for any deterministic function $g(\cdot)$

$$H(g(X), X) = H(X) + H(g(X)|X)$$

since $g(X)$ is a deterministic function

$H(g(X)|X) = 0$

$\therefore\ H(g(X), X) = H(X)$ ①

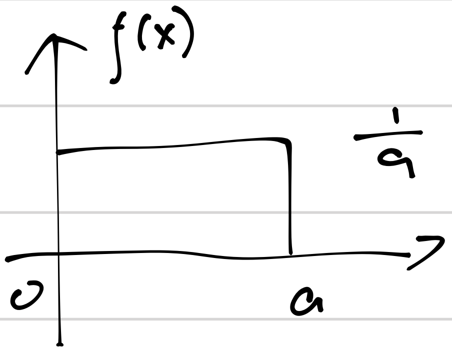
Also $H(g(X), X) = H(g(X)) + H(X \mid g(X))$
$= H(X)$

$\because H(X \mid g(X)) \geq 0$

$\therefore H(g(X)) \leq H(X)$

1b) Show that the inequality of part (a) does not hold for differential entropy by providing a simple example whe $h(g(x)) > h(x)$

Consider the uniform distribution



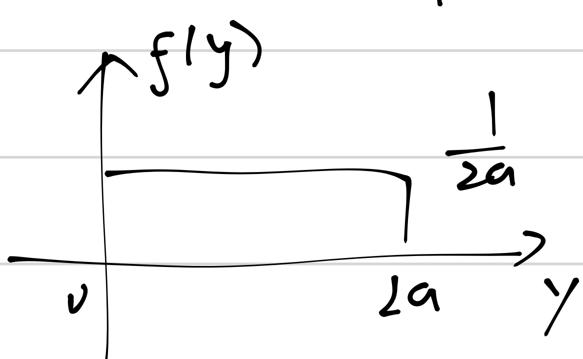
the PDF is

$$f(x) = \begin{cases} \frac{1}{a} & 0 \leq x \leq a \\ 0 & \text{else} \end{cases}$$

$$h(x) = -\int_0^a \frac{1}{a} \log \frac{1}{a} \, dx = \log a$$

suppose $y = g(x) = 2X$

$\therefore$ The PDF of $g(x) = y$ is



$$f(y) = \begin{cases} \frac{1}{2a} & 0 \leq x \leq 2a \\ 0 & \text{else} \end{cases}$$

$$h(Y) = -\int_0^{2a} \frac{1}{2a} \log \frac{1}{2a} \, dy = \log 2a$$

when $a=1$ $\quad h(X)=0$

$$h(Y) = h(g(x)) = \log 2 = 1$$

$$h(Y) = h(g(x)) > h(x)$$

(c) Show that $h(g(x)) \le h(x)$ for the many-to-one mapping $g(\cdot)$ which maps the real line to the interval $(-0.5, 0.5]$

$$g(x) = x + n(x) \qquad g(x) \in (-0.5, 0.5]$$

$g(x) = x + n(x)$ and $g(x) \in (-0.5, 0.5]$

$$\therefore g(x) = \begin{cases} \vdots & \\ x+2 & x \in (-2.5, -1.5] \\ x+1 & x \in (-1.5, -0.5] \\ x+0 & x \in (-0.5, 0.5] \\ x-1 & x \in (0.5, 1.5] \\ x-2 & x \in (1.5, 2.5] \\ \vdots & \end{cases}$$

$$\downarrow$$

$$g(x) = x - i \qquad (i-0.5 < x \le i+0.5)$$

Suppose X have a PDF $f_X(x)$

$\rightarrow Y = g(x)$ $\quad y = g(x)$ so y has a PDF, we denote it as $f_Y(y)$

when $(i-0.5 < x \le i+0.5)$

$$F_Y(y) = P(Y < y) = P(X - i < y) = P(X < y + i)$$

$\therefore f_Y(y) = f_X(y+i)$ when $i-0.5 < x \le i+0.5$

$$\therefore f_{Y|X}(y|x) = \begin{cases} \vdots \\ f_X(y-2)/Pr(-2.5 < X \le -1.5) & x \in (-2.5, -1.5] \\ f_X(y-1)/Pr(-1.5 < X \le -0.5) & x \in (-1.5, -0.5] \\ f_X(y)/Pr(-0.5 < X \le 0.5) & x \in (-0.5, 0.5] \\ f_X(y+1)/Pr(0.5 < X \le 1.5) & x \in (0.5, 1.5] \\ f_X(y+2)/Pr(1.5 < X \le 2.5) & x \in (1.5, 2.5] \\ \vdots \\ f_X(y+i) & x \in (i-0.5, i+0.5] \end{cases}$$

$$f_Y(y) = \sum_x f_{Y|X}(y|x)\, p_X(x)$$

$$= \cdots + f_X(y-1) + f_X(y) + f_X(y+1) + \cdots$$

$$= \sum_{i=-\infty}^{+\infty} f_X(y+i) = f_Y(y)$$

since $g(x) \in (-0.5, 0.5]$

$$h(Y) = h(g(X)) = -\int_{-0.5}^{0.5} f_Y(y) \log f_Y(y)\, dy$$

$$= -\int_{-0.5}^{0.5} \sum_{i=-\infty}^{+\infty} f_X(y+i) \log\left(\sum_{i=-\infty}^{+\infty} f_X(y+i)\right) dy$$

From the hint

$$\sum_{i=-\infty}^{+\infty} f(x+i)\log(f(x+i)) \le \left(\sum_{i=-\infty}^{+\infty} f(x+i)\right)\log\left(\sum_{i=-\infty}^{+\infty} f(x+i)\right)$$

$$\therefore -\sum_{i=-\infty}^{+\infty} f(x+i)\log(f(x+i)) \geq -\left(\sum_{i=-\infty}^{+\infty} f(x+i)\right)\log\left(\sum_{i=-\infty}^{+\infty} f(x+i)\right)$$

$$\therefore h(Y) = -\int_{-0.5}^{0.5} \sum_{i=-\infty}^{+\infty} f_X(y+i) \log\left(\sum_{i=-\infty}^{+\infty} f_X(y+i)\right) dy$$

$$\le -\int_{-0.5}^{0.5} \sum_{i=-\infty}^{+\infty} f_X(y+i) \log f_X(y+i) \, dy$$

$$= \sum_{i=-\infty}^{+\infty} \int_{-0.5}^{0.5} f_X(y+i) \log(f_X(y+i)) \, dy$$

let $t = y+i$ $dy = dt$

$$= \sum_{i=-\infty}^{+\infty} \int_{-0.5+i}^{0.5+i} f_X(t) \log f_X(t) \, dt$$

$$= \int_{-\infty}^{+\infty} f_X(t) \log(f_X(t)) dt = \int_{-\infty}^{+\infty} f_X(x)\log(f_X(x))dx$$

$$= h(X)$$

$$\therefore h(Y) = h(g(X)) \le h(X)$$

6. More Modulo Micheif

(a) For positive a and b show that

$$a\log a + b\log b \le (a+b)\log(a+b)$$

$$a\log a + b\log b + a\log(a+b) + b\log(a+b) - a\log(a+b) - b\log(a+b)$$

$$= (a+b)\log(a+b) + a\log\left(\frac{a}{a+b}\right) + b\log\left(\frac{a}{a+b}\right)$$

$\because a>0,\ b>0 \quad \therefore \quad 0 < \frac{a}{a+b} \le 1$

$\therefore a\log\left(\frac{a}{a+b}\right) \le 0, \quad b\log\left(\frac{a}{a+b}\right) \le 0$

$$\therefore (a+b)\log(a+b) + a\log\left(\frac{a}{a+b}\right) + b\log\left(\frac{a}{a+b}\right)$$

$$= a\log a + b\log b \le (a+b)\log(a+b) \quad \square$$

(b) Use the generalization of part(a) to prove the hint of problem 4 on problem set 5 as follows:

$$\sum_{i=-\infty}^{+\infty} f(x+i)\log(f(x+i)) \le \left(\sum_{i=-\infty}^{+\infty} f(x+i)\right)\log\left(\sum_{i=-\infty}^{+\infty} f(x+i)\right)$$

let $a_i = f(x+i)$

$a_i \ge 0 \quad \therefore a_1\log a_1 + a_2\log a_2 \le (a_1+a_2)\log(a_1+a_2)$

suppose for i

$$a_1\log a_1 + a_2\log a_2 + \cdots + a_i\log a_i \le (a_1+\cdots+a_i)\log(a_1+\cdots+a_i)$$

for $i+1$

$$a_1\log a_1 + \cdots + a_i\log a_i + a_{i+1}\log a_{i+1} \le (a_1+\cdots+a_i)\log(a_1+\cdots+a_i) + a_{i+1}\log(a_{i+1})$$

$$= b_i\log b_i + a_{i+1}\log a_{i+1} \quad (b_i = a_1+\cdots+a_i)$$

$$\le (b_i+a_i)\log(b_i+a_i)$$

$$= (a_1+\cdots+a_i+a_{i+1})\log(a_1+\cdots+a_i+a_{i+1})$$

$\therefore$ The inequality always exists as i grows large

$$\therefore \sum_{i=-\infty}^{+\infty} a_i\log a_i \le \sum_{i=-\infty}^{+\infty} a_i \sum_{i=-\infty}^{+\infty}\log a_i$$

$$\therefore \sum_{i=-\infty}^{+\infty} f(x+i)\log(f(x+i)) \le \left(\sum_{i=-\infty}^{+\infty} f(x+i)\right)\log\left(\sum_{i=-\infty}^{+\infty}(f(x+i))\right)$$ [?]

(c)

$$a\log a + b\log b + \underbrace{a\log(a+b) + b\log(a+b)} - a\log(a+b) - b\log(a+b)$$

$$= (a+b)\log(a+b) + a\log\left(\frac{a}{a+b}\right) + b\log\left(\frac{a}{a+b}\right)$$

(d) For a continuous random variable X with pdf $f(x)$, define $Y=g(x)$ $Z=h(X)$

Show that $h(X) = h(Y) + H(Z|Y)$

From c, we know that

$$a\log a + b\log b = (a+b)\log(a+b) + a\log(a/a+b) + b\log(b/a+b)$$

This can also be generalized to the hint from 5c

$$\therefore \sum_{i=-\infty}^{+\infty} f(x+i)\log[f(x+i)] = \left(\sum_{i=-\infty}^{+\infty} f(x+i)\right)\log\left(\sum_{i=-\infty}^{+\infty} f(x+i)\right)$$

$$+\sum_{i=-\infty}^{+\infty} f(x+i)\log\frac{f(x+i)}{\sum_{i=-\infty}^{+\infty} f(x+i)}$$

From Problem 5c we know that

$$h(Y) = -\int_{-0.5}^{0.5}\sum_{i=-\infty}^{+\infty} f_X(y+i) \log[f_X(y+i)] dy$$

According to the formula above

$$h(Y) = -\int_{-0.5}^{0.5}\sum_{i=-\infty}^{+\infty} f_X(y+i) \log[f_X(y+i)] dy \quad ①$$

$$-\int_{-0.5}^{0.5}\sum_{i=-\infty}^{+\infty} f_X(y+i)\log\frac{f_X(y+i)}{\sum_{i=-\infty}^{+\infty} f_X(y+i)} dy \quad ②$$

$$① = -\int_{-0.5}^{0.5} \sum_{i=-\infty}^{+\infty} f_X(y+i) \log [f_X(y+i)] dy$$

let $x = y+i$ $\quad dx = dy$

$$\therefore ① = \int_{i-0.5}^{i+0.5} \sum_{i=-\infty}^{+\infty} f_X(x) \log f_X(x) dx = \int_{-\infty}^{+\infty} f_X(x) \log f_X(x) dx$$

$$= h(X)$$

for ②

from the hint we know that

$$P(Z=z \mid Y=y) = \frac{f(y-z)}{\sum_{i=-\infty}^{+\infty} f(y-i)}$$

$$H(Z \mid Y=y) = \sum_z P(Z=z \mid Y=y) \log P(Z=z \mid Y=y)$$

$$= \sum_{z=-\infty}^{+\infty} \frac{f(y-z)}{\sum_{i=-\infty}^{+\infty} f(y-i)} \log \frac{f(y-z)}{\sum_{i=-\infty}^{+\infty} f(y-i)}$$

$$H(Z \mid Y) = \int_{-0.5}^{0.5} H(Z \mid Y=y) f_Y(y) dy$$

$$= \int_{-0.5}^{0.5} \sum_{z=-\infty}^{+\infty} \frac{f(y-z)}{\sum_{i=-\infty}^{+\infty} f(y-i)} \log \frac{f(y-z)}{\sum_{i=-\infty}^{+\infty} f(y-i)} \sum_{i=-\infty}^{+\infty} f(y+i) dy$$

$$= \int_{-0.5}^{0.5} \sum_{z=-\infty}^{+\infty} f(y-z) \log \frac{f(y-z)}{\sum_{i=-\infty}^{+\infty} f(y-i)} dy$$

we note that if we chang i to z

$$② = \int_{-0.5}^{0.5} \sum_{i=-\infty}^{+\infty} f(y+i) \log \frac{f(y+i)}{\sum_{i=-\infty}^{+\infty} f(y+i)} dy$$

$$\Downarrow$$

$$= \int_{-0.5}^{0.5} \sum_{z=-\infty}^{+\infty} f(y-z) \log \frac{f(y-z)}{\sum_{i=-\infty}^{+\infty} f(y-i)} dy = H(Z|Y)$$

(since the summation is from $-\infty$ to $+\infty$)

$$\sum_{i=-\infty}^{+\infty} f(y+i) = \sum_{i=-\infty}^{+\infty} f(y-i)$$

$$\therefore\ h(X) = ① + ② = h(Y) + H(Z|Y)$$

$$h(X) = h(Y) + H(Z|Y) \quad \square$$

7. Mutual information for a mixed distribution

X is binary $X \in [0, 1]$ equally

when $X = 0 \Rightarrow Y$ is also 0

when $X = 1 \rightarrow Y$ is uniformly distributed on the closed interval $[\frac{1}{2}, \frac{3}{2}]$

(a) find $H(X)$

$$H(X) = \sum p(x) \log p(x) = -\frac{1}{2} \log \frac{1}{2} - \frac{1}{2} \log \frac{1}{2} = 1 \text{ bit}$$

$H(X) = 1$ bit

(b) Find $H(X|Y)$

$$H(X|Y) = H(X|Y=0)\,p(Y=0) + H(X|\tfrac{1}{2} \le y \le \tfrac{3}{2})\,p(\tfrac{1}{2} \le y \le \tfrac{3}{2})$$

if $y=0$, x could only be 0

if $\frac{1}{2} \le y \le \frac{3}{2}$, x could only be 1

$\therefore H(X|Y=0) = 0$

$H(X|\frac{1}{2} \le y \le \frac{3}{2}) = 0$

$H(X|Y) = 0 + 0 = 0$

(c) i $h(Y|X=0)$

when $X=0$ $y=0$ with probability 1

we could calculate $h(Y|X=0)$ by taking a limit of a rectangular pdf as width goes to 0

$$\lim_{a \to 0} \int_0^a \frac{1}{a} \log \frac{1}{a}\, dy = \lim_{a \to 0} \log a = -\infty$$

$\therefore h(Y|X=0) = -\infty$

ii $h(Y|X=1)$

when $x=1$, y is uniform distribution

$$f_Y(y) = \begin{cases} \frac{1}{\frac{3}{2}-\frac{1}{2}} & \frac{1}{2} \le y \le \frac{3}{2} \\ 0 & \text{else} \end{cases} = \begin{cases} 1 & \frac{1}{2} \le y \le \frac{3}{2} \\ 0 & \text{else} \end{cases}$$

whe $X=1$

$$h(Y|X=1) = -\int_{\frac{1}{2}}^{\frac{3}{2}} 1 \log 1 \, dy = 0$$

$$\therefore h(Y|X=1) = 0$$

iii) $h(Y|X) = h(Y|X=0)\,p(X=0) + h(Y|X=1)\,p(X=1)$

$$= -\infty \times \frac{1}{2} + 0 \times \frac{1}{2}$$

$$= -\infty$$

$$h(Y|X) = -\infty$$

(d) Find $h(Y)$

$$f_Y(y) = \sum f(y|x)\,p_X(x)$$

$$= \begin{cases} \frac{1}{2} & y=0 \\ \frac{1}{2} & \frac{1}{2} \le y \le \frac{3}{2} \end{cases}$$

$$h(Y) = -\int_0^{\infty} f_Y(y) \log f_Y(y)\, dy$$

$$= -\lim_{a\to 0} \int_0^{a} \frac{1}{2a} \log \frac{1}{2a}\, dy + \int_{\frac{1}{2}}^{\frac{3}{2}} \frac{1}{2} \log \frac{1}{2}\, dy$$

$$= \lim_{a\to 0} \frac{1}{2}\log 2a + \frac{1}{2}\log\frac{1}{2}$$

$$= -\infty + \frac{1}{2} \quad = -\infty$$

$\therefore h(Y) = -\infty$

(e) Find $I(X;Y)$

$$I(X;Y) = H(X) - H(X|Y) = 1 - 0 = 1 \text{ bit}$$

(Since $h(Y)$, $h(Y|X)$ contains minus infinity,

we cannot use $I(X;Y) = h(Y) - h(Y|X)$) □