(C4.1.) The entropy of a discrete random variable X with pmf $P_X(x)$ is: (1)

$$H(X) = -\sum_{x} p(x)\log p(x) = -E[\log p(x)]$$ → measures the expected uncertainty in X

- A function f is continuous at $x=a$ if and only if:

$$\lim_{x \to a} f(x) = f(a)$$

- We have $0 \leq p \leq 1 \Rightarrow \log p \leq 0 \Rightarrow -p\log p \geq 0$

and thus $H(p) = -\sum_i p_i \log p_i \geq 0$

- Non-negativity:

$$H(p) = 0$$

We can solve the optimization problem:

$$\max H(p) \mid \sum_{i=1}^{N} p_i = 1. \Rightarrow$$

$$\Rightarrow \frac{\partial H}{\partial p_j} - \lambda \frac{\partial \sum_i^N p_i}{\partial p_j} = 0 \Rightarrow -\frac{\partial}{\partial p_j}\left(\sum_i^N p_i \log p_i - \lambda p_i\right) = 0 \Rightarrow$$

$$\Rightarrow -\log p_j - 1 - \lambda = 0 \Rightarrow$$

$$\boxed{p_j = e^{-(1+\lambda)}}$$

Thus, $$\left.\sum_{i=1}^{N} p_i = 1 \xrightarrow{p_i = e^{-(1+\lambda)}} \sum_{i=1}^{N} e^{-(1+\lambda)} = 1 \Rightarrow \lambda = \log N - 1\right\} \Rightarrow p_j = e^{-\log N} = \frac{1}{N}$$

Thus, $\max(H_p) = \sum_{i=1}^{N} \frac{1}{N} \log\frac{1}{N} = \log N$. So, $H(p) \in [0, \log N]$

• $H(p,q) = \sum_x \sum_y p(x)\, q(y) \log[p(x)\, q(y)] \Rightarrow$

$H(p,q) = -\sum_x \sum_y p(x)\, q(y) \log p(x) - \sum_x \sum_y p(x)\, q(y) \log q(y) \Rightarrow$

$H(p,q) = -\sum_x p(x) \log(px) \overset{1}{\cancel{\sum_y q(y)}} - \sum_y q(y) \log q(y) \overset{1}{\cancel{\sum_x p(x)}} \Rightarrow$

$H(p,q) = H(p) + H(q)$

---

from eq. 4.10:

(4.2) $I(x,y) = \sum_x \sum_y p(x,y) \log \frac{p(x,y)}{p(x)\, p(y)} \Rightarrow$

$I(x,y) = \sum_x \sum_y p(x,y) \log p(x,y) - \sum_x \sum_y p(x,y) \log p(x) - \sum_y \sum_x p(x,y) \log p(y) \Rightarrow$

$I(x,y) = \sum_x \sum_y p(x,y) \log p(x,y) - \sum_x p(x) \log p(x) - \sum_y p(y) \log p(y) \Rightarrow$

$I(x,y) = -H(x,y) + H(x) + H(y)$

(4.4) Let X be a random variable with a probability density function p whose support is a set $\mathcal{X}$. The differential entropy is defined as,

$$H(x) = -\int_{-\infty}^{+\infty} p(x) \log p(x)\, dx \Rightarrow$$

$$H(x) = -\langle \log p(x) \rangle \Rightarrow$$

$$H(x) = -\left\langle \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right)\right\rangle \Rightarrow$$

$$H(x) = -\left\langle \log \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{(x-\mu)^2}{2\sigma^2} \right\rangle \Rightarrow$$

$$H(x) = \log\sqrt{2\pi\sigma^2} + \frac{\cancel{\sigma^2}}{2\cancel{\sigma^2}} \Rightarrow$$

$$H(x) = \frac{1}{2}\log(2\pi e\sigma^2)$$

(4.5) (a) We can model this as a Gaussian process:

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

and the capacity of a band-limited Gaussian channel is given by:

$$C = \Delta f \log\left(1 + \frac{S}{N}\right) \Rightarrow$$

$$dB = 20 \log_{10} \frac{X_1}{X_2}$$

$$SNR = 10 \log_{10} \left(\frac{S}{N}\right) \Rightarrow$$

$$20 = 10 \log_{10} \left(\frac{S}{N}\right) \Rightarrow \frac{S}{N} = 10^{20/10}$$

$$C = 3300\, s^{-1} \log\left(1 + 10^{20/10}\right) \Rightarrow$$

$$C = 22 \text{ kbps}$$

(B) $SNR = ?$ | $C = 1\,\text{Gbit/s}$

We can write

$$10^9 = 3300 \log\left(1 + \frac{S}{N}\right) \Rightarrow$$

$$\log\left(1 + \frac{S}{N}\right) = \frac{10^9}{3300} \Rightarrow$$

$$\frac{S}{N} = 2^{10^9/3300} - 1 \approx 2^{3\times10^5}$$

Thus, $SNR = 10 \log_{10} 2^{3\times10^5} \approx 10^6\, dB$

(4.6) The Cramér-Rao inequality states that the mean square error of an unbiased estimator f of a is lower bounded by the reciprocal of the fisher information;

$$\sigma^2(f) \geq \frac{1}{J(a)}$$

mean square error: $\langle f - x_0 \rangle^2 = \left\langle \left[\frac{1}{n}\sum_{i=1}^{n} x_i - x_0\right]\right\rangle^2 =$

$$= \left\langle \left(\frac{1}{n}\sum_{i=1}^{n} x_i - x_0\right)\left(\frac{1}{n}\sum_{k=1}^{n} x_k - x_0\right)\right\rangle =$$

$$= \left\langle \frac{1}{n}\sum_{i=1}^{n}(x_i - x_0)\right\rangle \left\langle \frac{1}{n}\sum_{k=1}^{n}(x_i - x_0)\right\rangle =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\langle (x_i - x_0)\rangle \times \frac{1}{n}\sum_{k=1}^{n}\langle (x_k - x_0)\rangle$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\langle (x_i - x_0)\rangle^2$$

$$= \frac{\sigma^2}{n^2}$$

fisher information:

$$J_n(x_0) = n J(x_0) = n\left\langle \left[\frac{\partial}{\partial x_0}\log(x_0, \sigma^2)\right]^2\right\rangle$$

$$= n\left\langle \partial_{x_0}\left[\log\frac{1}{\sqrt{2\pi\sigma^2}} + \frac{(x-x_0)^2}{2\sigma^2}\right]^2\right\rangle$$

$$= n \left\langle \left[\frac{x - x_0}{\sigma^2} \right]^2 \right\rangle =$$

$$= n \frac{\sigma^2}{\sigma^4} =$$

$$= \frac{n}{\sigma^2}$$, which is the Cramér-Rao lower bound