Prof. Dr. Rudolf Mathar, Dr.-Ing. Gholamreza Alirezaei, Emilio Balda, Vimal Radhakrishnan

Tutorial 6
- Proposed Solution -

Monday, January 14, 2019

Solution of Problem 1

Differential entropy

Evaluate the differential entropy $h(X)$ for the following:

a) Guassian distributions with density, $f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp(\frac{-(x-\mu)^2}{2\sigma^2})$.

$$\begin{aligned}
h(X) &= -\int f(x)\ln f(x)dx = -\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp(\frac{-(x-\mu)^2}{2\sigma^2})\ln\Big(\frac{1}{\sqrt{2\pi}\sigma}\exp(\frac{-(x-\mu)^2}{2\sigma^2})\Big)dx \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\exp(\frac{-(x-\mu)^2}{2\sigma^2})\ln(\sqrt{2\pi}\sigma\exp(\frac{(x-\mu)^2}{2\sigma^2})dx \\
&= \frac{\sqrt{2}\sigma}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\exp(-t^2)\ln(\sqrt{2\pi}\sigma\exp(t^2)dt \quad (\text{ Substituting } t = \frac{(x-\mu)}{\sqrt{2}\sigma}\ , dx = \sqrt{2}\sigma dt) \\
&= \frac{\ln(\sqrt{2\pi}\sigma)}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp(-t^2)dt + \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}t^2\exp(-t^2)dt \\
&= \frac{\ln(\sqrt{2\pi}\sigma)}{\sqrt{\pi}}\sqrt{\pi} + \frac{1}{\sqrt{\pi}}\frac{\sqrt{\pi}}{2} \\
&\Big(\int_{-\infty}^{\infty}\exp(-at^2)dt = \sqrt{\frac{\pi}{a}} \text{ and } \int_{-\infty}^{\infty}t^2\exp(-at^2)dt = \frac{1}{2}\sqrt{\frac{\pi}{a^3}}\ , a > 0\ \Big) \\
&= \ln(\sqrt{2\pi}\sigma) + \frac{1}{2} \\
&= \frac{1}{2}\ln(2\pi e\sigma^2) \quad \text{nats}
\end{aligned} \tag{1}$$

b) The exponential density, $f(x) = \lambda\exp(-\lambda x), x \geq 0$.

$$\begin{aligned}
h(X) &= -\int f(x)\ln f(x)dx = -\int_0^{\infty}\lambda\exp(-\lambda x)\ln\Big(\lambda\exp(-\lambda x)\Big)dx \\
&= -\int_0^{\infty}\lambda\exp(-\lambda x)\ln\lambda dx + \int_0^{\infty}\lambda\exp(-\lambda x)(\lambda x)dx \\
&= -\lambda\ln\lambda\int_0^{\infty}\exp(-\lambda x)dx + \lambda^2\int_0^{\infty}x\exp(-\lambda x)dx \\
&= -\lambda\ln\lambda\frac{1}{-\lambda}[\exp(-\lambda x)]_0^{\infty} + \lambda^2\frac{1}{\lambda^2}[\exp(-\lambda x)(-\lambda x - 1)]_0^{\infty} \\
&= -\ln\lambda + 1 \text{ nats} \\
&= \log\frac{e}{\lambda} \text{ bits.}
\end{aligned} \tag{2}$$

c) The Laplace density, $f(x) = \frac{1}{2}\lambda\text{exp}(-\lambda|x|)$. Note that the Laplace density is a two sided exponential density, so each side has a differential entropy of the exponential.

$$
\begin{aligned}
h(X) &= -\int f(x)\ln f(x)dx = -\int_{-\infty}^{\infty} \frac{1}{2}\lambda\text{exp}(-\lambda|x|)\ln\Big(\frac{1}{2}\lambda\text{exp}(-\lambda|x|)\Big)dx \\
&= -\int_{-\infty}^{\infty} \frac{1}{2}\lambda\text{exp}(-\lambda|x|)\Big(\ln\frac{1}{2} + \ln\lambda + (-\lambda|x|)\Big)dx \\
&= -2\int_{0}^{\infty} \frac{1}{2}\lambda\text{exp}(-\lambda x)\Big(\ln\frac{1}{2} + \ln\lambda + (-\lambda x)\Big)dx \\
&= -\lambda(\ln\frac{1}{2} + \ln\lambda)\int_{0}^{\infty} \text{exp}(-\lambda x)dx + +\lambda^2\int_{0}^{\infty} x\text{exp}(-\lambda x)dx \\
&= -\ln\frac{1}{2} - \ln\lambda + 1 \\
&= \ln\frac{2e}{\lambda} \text{ nats}
\end{aligned} \tag{3}
$$

Solution of Problem 2

We can expand the mutual information

$$I(X;Y) = h(Y) - h(Y|X) = h(Y) - h(Z) \tag{4}$$

and $h(Z) = \log 2$, since $Z \sim U(-1,1)$.
The output Y is a sum a of a discrete and a continuous random variable, and if the probabilities of X are $p_{-2}, p_{-1}, ..., p_2$, then the output distribution of Y has a uniform distribution with weight:

$$
\begin{aligned}
\frac{p_{-2}}{2} &\text{ for } -3 \le Y \le -2, \\
\frac{p_{-2}+p_{-1}}{2} &\text{ for } -2 \le Y \le -1, \\
\frac{p_{-1}+p_{0}}{2} &\text{ for } -1 \le Y \le 0, \\
\frac{p_{0}+p_{1}}{2} &\text{ for } 0 \le Y \le 1, \\
\frac{p_{1}+p_{2}}{2} &\text{ for } 1 \le Y \le 2, \\
\frac{p_{2}}{2} &\text{ for } 2 \le Y \le 3,
\end{aligned} \tag{5}
$$

Given that Y ranges from -3 to 3, the maximum entropy that it can have is an uniform over this range.

$$\frac{p_{-2}}{2} = \frac{p_{-2}+p_{-1}}{2} = \frac{p_{-1}+p_{0}}{2} = \frac{p_{0}+p_{1}}{2} = \frac{p_{1}+p_{2}}{2} = \frac{p_{2}}{2}. \tag{6}$$

From the above equation we get

$$p_{-1} = p_1 = 0; p_{-2} = p_0 = p_2. \tag{7}$$

We know

$$p_0 + p_{-1} + p_1 + p_{-2} + p_2 = 1. \tag{8}$$

Then, we get

$$p_{-2} = p_0 = p_2 = \frac{1}{3}. \tag{9}$$

The distribution of X that can achieve the maximum entropy of Y is $(\frac{1}{3}, 0, \frac{1}{3}, 0, \frac{1}{3})$. Then , the maximum entropy of Y is $h(Y) = \log 6$ and the capacity of this channel is $C = \log 6 - \log 2 = \log 3$ bits.

Solution of Problem 3

The differential entropy of an exponentially distributed random variable with mean $\frac{1}{\lambda}$ is $\log \frac{e}{\lambda}$ bits. If the median is 80 years, then

$$\begin{aligned} \int_0^{80} \lambda e^{-\lambda x} dx &= \frac{1}{2} \\ 1 - e^{-80\lambda} &= \frac{1}{2}. \end{aligned} \tag{10}$$

We get

$$\lambda = \frac{\ln 2}{80} = 0.00866, \tag{11}$$

and the differential entropy is $\log \frac{e}{\lambda} = 8.3$ bits.

In general, $h(X) + n$ is the number of bits on the average required to describe X to n-bit accuracy. So, to represent the random variable to 3 digits (≈ 10 bits accuracy) would need $\log \frac{e}{\lambda} + 10$ bits $= 18.3$ bits.

Solution of Problem 4

We are interested in the set $\{(x_1, x_2, ..., x_n) \in \mathcal{R}^n : f(x_1, x_2, ..., x_n) \in (2^{-n(h+\epsilon)}, 2^{-n(h-\epsilon)})\}$. This is:

$$2^{-n(h+\epsilon)} \leq f(x_1, x_2, ..., x_n) \leq 2^{-n(h-\epsilon)}. \tag{12}$$

Since X_i are i.i.d., $f(x_1, x_2, ..., x_n) = c^n e^{-(x_1^4 + ... + x_n^4)}$. Plugging this in for $f(x_1, x_2, ..., x_n)$ in the above inequality, we get

$$2^{-n(h+\epsilon)} \leq c^n e^{-(x_1^4 + ... + x_n^4)} \leq 2^{-n(h-\epsilon)}. \tag{13}$$

which is equivalent to

$$\begin{aligned} -n(h+\epsilon)\ln 2 \leq n \ln c - (x_1^4 + ... + x_n^4) \leq -n(h-\epsilon)\ln 2, \\ n((h+\epsilon)\ln 2 + \ln c) \geq (x_1^4 + ... + x_n^4) \geq n((h-\epsilon)\ln 2 + \ln c). \end{aligned} \tag{14}$$

The typical set can be written as

$$A_\epsilon^n = \{(x_1, x_2, ..., x_n) \in \mathcal{R}^n : n((h-\epsilon)\ln 2 + \ln c) \leq \sum_i x_i^4 \leq n((h+\epsilon)\ln 2 + \ln c)\} \tag{15}$$

So the shape of the typical set is the shell of a 4-norm ball :

$$\{(x_1, x_2, ..., x_n) : \|(x_1, x_2, ..., x_n)\|_4 \in (n((h \pm \epsilon)\ln 2 + \ln c))^{\frac{1}{4}}\}. \tag{16}$$

.