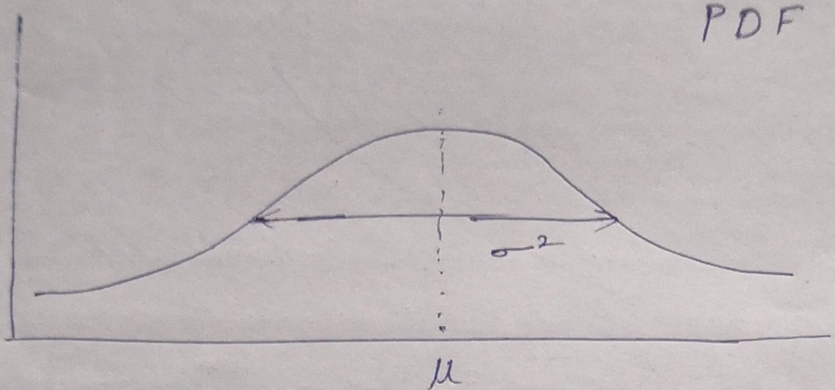
Differential Entropy of gaussian Random Variable

Consider a source which produce a continuous set of symbols which follows Gaussian Probability distribution



This is the most commonly occuring output of an information source and has a fundumental role. Because this gaussian distribution occurs frequently in nature.

PDF of a gaussian distribution is given by

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

Differential Entropy of the source is given by

$$h(X) = \int_{-\infty}^{\infty} f_X(x) \log_2 \frac{1}{f_X(x)} dx \quad \text{———} (1)$$

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

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

$$\log_2 \frac{1}{f_X(x)} = \log_{e} \frac{1}{f_X(x)} \log_2 e$$

$$= \log_2 e \cdot \ln \sqrt{2\pi\sigma^2} \cdot e^{\frac{(x-\mu)^2}{2\sigma^2}}$$

$$= \log_2 e \left\{ \frac{1}{2} \ln 2\pi\sigma^2 + \frac{1}{2\sigma^2} \cdot (x-\mu)^2 \right]$$

Substitute this in eqn (1)

$$h(x) = \int_{-\infty}^{\infty} f_X(x) \cdot \log_2 e \left[\frac{1}{2} \ln 2\pi\sigma^2 + \frac{1}{2\sigma^2} (x-\mu)^2 \right]$$

$$= \log_2 e \int_{-\infty}^{\infty} f_X(x) \cdot \frac{1}{2} \ln 2\pi\sigma^2 \, dx \; +$$

$$\log_2 e \int_{-\infty}^{\infty} f_X(x) \cdot \frac{(x-\mu)^2}{2\sigma^2}$$

$$= \log_2 e \frac{\ln 2\pi\sigma^2}{2} \underbrace{\int_{-\infty}^{\infty} f_X(x)\, dx}_{\text{Area under the PDF curve} = 1} + \log_2 e \cdot \frac{1}{2\sigma^2} \underbrace{\int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\, dx}_{\text{Variance of any random variable} = \sigma^2}$$

$$= \log_2 e \frac{1}{2} \ln 2\pi\sigma^2 \times 1 + \log_2 e \frac{1}{2\cancel{\sigma^2}} \cdot \cancel{\sigma^2}$$

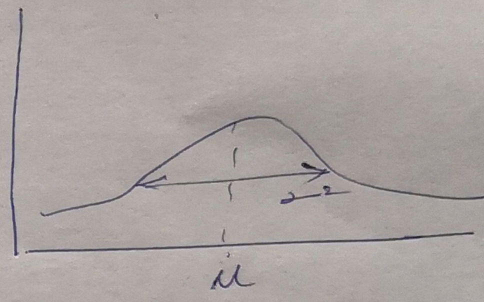
Finally Differential Entropy of a Gaussian random variable is

$$h(x) = \frac{1}{2}\log_2(2\pi\sigma^2 e)$$

From this we can know that differential Entropy of gaussian random variable depends only on variance(σ^2)

It doesn't depends upon μ (mean)

So in the graph as the spread or variance increases randomness of the variable increases.



If the variance decreases the randomness decreases and it concentrate towards μ.

As the randomness increases (variance σ^2) the differential Entropy increases. So the uncertainity will increases.

Conditional Entropy.

$$H(X) = -\sum_{i=1}^{m} P(x_i) \log_2 P(x_i)$$ ⟶ input Entropy

$$H(Y) = -\sum_{j=1}^{n} P(y_j) \log_2 P(y)$$ ⟶ output entropy

then the conditional entropy is given by

$$H(Y/X) = -\sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i, y_j) \log P(y_j/x_i)$$

This is telling about the average information of Y we will get after ~~recieving~~ transmitting X.

Similarly

$$H(X/y) = \sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i, y_j) \log_2 P(x_i/y_j)$$

From this we will come to know the average information about X; after recieving Y.

Joint Entropy

Let $X = (x_1, x_2 \dots x_j)$ bel the input to the channel and Let $Y = \{y_1, y_2 \dots y_k\}$ be the channel output.

Let $P(X = x_j, Y = y_k\}$ be the joint pdf of X and Y

ie: Consider the joint occurance of x and y

The joint Entropy of X and Y represents the uncertainity in the joint event (X, Y) and is defined as

$$H(X,Y) = -\sum_{j=1}^{n}\sum_{k=1}^{m} P(x_j, y_k) \log_2 P(x_j, y_k)$$

$$= \sum_{j=1}^{n}\sum_{k=1}^{m} P(x_j; y_k) \log_2 \frac{1}{P(x_j, y_k)}$$

H(X, Y) represent the amount of uncertainty in the joint event ie; the occurance of input X and output Y

We can derive an equivalent representation of joint Entropy

$$H(X,Y) = -\sum_X\sum_Y P(x,y) \log_2 P(x,y)$$

$$= -\sum_X\sum_Y P(x) P(y/x) \log_2 P(x) P(y/x)$$

$$= -\sum_X\sum_Y P(x) P(y/x) \left[\log_2 P(x) + \log_2 P(y/x)\right]$$

$$= -\sum_X P(x) \sum_Y P(y/x) \log_2 P(x) - \sum_{XY} P(x) P(y/x) \log_2 P(y/x)$$

$$= -\sum_{x} P(x) \sum_{y} P(y/x) \log_2 P(x) - \sum_{xy} P(x,y) \log_2 P(y/x)$$

$\underbrace{\sum_{y} P(y/x)}$ → Marginal density function of y

for summation over all values of $Y = 1$

$$= -\sum_{x} P(x) \log_2 P(x) - \sum_{xy} P(x,y) \log_2 P(Y/x)$$

$$= H(X) + H(Y/X)$$

$$\therefore \boxed{\begin{aligned} H(X,Y) &= H(X) + H(Y/X) \\ H(X,Y) &= H(Y) + H(X/Y) \end{aligned}}$$

ie; the uncertainty in the joint event (X,Y) is the sum of the uncertainty in X plus the remaining uncertainty in y after X is known.

when X and Y are independant

$$H(X,Y) = H(Y) + H(X)$$

The uncertainty in the joint event (X,Y) is the sum of uncertainties in X and Y when they are independant.