1. Introduction

Motivation:

* We want to quantify and control amount of uncertainty.
* Sending a message from a cell phone to another cell phone.
* The EM at the received phone is a random signal to environment. We want to study how to best recover the original message.
* We want to model complex phenomena in order to do prediction, optimization, etc.
Probability Manipulation

A. Deterministic signals

\[ x(t) \xrightarrow{\text{#}} y(t) \quad y(t) = h(t) \ast x(t) \]

B. Random Signal

\[ x(t) \xrightarrow{H} Y(t) \]
Probability review:

3.1 Expectation

For continuous random variable (C.R.V)

\[ E[X] = \int_{-\infty}^{\infty} x f_X(x) \, dx \]

For discrete random variable (D.R.V)

\[ E[X] = \sum_{i} x_i \cdot p_X(x_i) \]

Properties of expectation:

1. \( E[cX] = c \cdot E[X] \) \( \cdot c \) is a constant

2. \( E[aX + bY] = aE[X] + bE[Y] \) \( \Rightarrow E \) is a linear operator.

3. \( E[g(X)] \triangleq \int_{-\infty}^{\infty} g(x) f_X(x) \, dx \) (C.R.V)

\[ E[g(X)] \triangleq \sum_{i} g(x_i) p_X(x_i) \] (D.R.V)
(3.2) Moments of Random Variable.

**Def:** The $r^{th}$ moment of R.V. $X$ is formally defined as:

$$m_r = E[X^r] = \int_{-\infty}^{\infty} x^r f_X(x) \, dx \quad (cRV)$$

$$\implies r \geq \sum_{x_i \in X} x_i \cdot f_X(x_i) \quad (dRV)$$

**Example:**

$$m_0 = 1$$

$$m_1 = E[X] = \mu \quad \text{(mean)}$$

$$m_2 = E[X^2] \quad \text{second moment}$$

**Def:** The $r^{th}$ central moment of R.V. $X$ is

$$c_r = E[(X - E[X])^r]$$
Ex:

\[ c_1 = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X - \mathbb{E}[X]] \]

\[ = \mathbb{E}[X] - \mathbb{E}[\mathbb{E}[X]] \]

(Linear property)

\[ = \mathbb{E}[X] - \mathbb{E}[X] = 0 \]

\[ c_2 = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[(X - \mu_X)^2] = \sigma_X^2 \]

(Variance)

\[ c_r = \mathbb{E}[(X - \mathbb{E}[X])^r] = \mathbb{E}\left[ \sum_{i=0}^{r} \binom{r}{i} (-1)^i \mu_X^i X^{r-i} \right] \]

(Binomial)
Ex: Maximum entropy principle.

Suppose we do not know the pdf $f_x(x)$. However, we want to estimate $f_x(x)$ using the known moments. The maximum entropy principle states that a good choice of $f_x(x)$ is the one that maximizes the following quantity:

$$H(X) = - \int_{-\infty}^{\infty} f_x(x) \ln f_x(x) \, dx$$

To apply the maximum entropy principle to a problem at hand, we need to put in more constraints:

1) $f_x(x) \geq 0$, $\forall x$

2) $\int_{-\infty}^{\infty} f_x(x) \, dx = 1$

3) $\int_{-\infty}^{\infty} x f_x(x) \, dx = \mu_X$ (known from (estimated) samples)
4) \[ \int_{-\infty}^{\infty} x^2 f_X(x) \, dx = m_2 \leftarrow \text{known from samples (estimated)} \]

5) \[ \int_{-\infty}^{\infty} x^3 f_X(x) \, dx = m_3 \leftarrow \text{known from samples (estimated)} \]

2.3 Joint Moments

Def: Joint \( i^{th} \) moment of \( X \) and \( Y \) is given by (CK
\[ m_{ij} = E[X^i Y^j] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^i y^j f_{X,Y}(x,y) \, dx \, dy \]

\[ \leq \sum_{k \leq i} x^i y^j P_{X,Y}(x_k y_k) \]
\( \text{Def: central joint } i^{th} \text{ moment} \)

\[
\hat{c}_{ij} = E[(X-M_x)^i (Y-M_y)^j]
\]

\[
= \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} (x_k-M_x)^i (y_l-M_y)^j \cdot s_{xy}(x_k, y_l) \text{d}x_k \text{d}y_l
\]

\[
\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} (x_k-M_x)^i (y_l-M_y)^j \cdot s_{xy}(x_k, y_l) \text{d}x_k \text{d}y_l
\]

\[
m_{11} = E[XY] = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} xy \cdot s_x(x, y) \text{d}x \text{d}y
\]

known as covariance.

\[
c_{11} = E[(X-M_x)(Y-M_y)] = E[XY - M_x \cdot M_y + M_x \cdot M_y]
\]

\[
= E[XY] - E[X] \cdot M_y - M_x \cdot E[Y] + M_x \cdot M_y
\]

\[
= E[XY] - M_x \cdot M_y \leftarrow \text{covariance}
\]

\[
= m_{11} - M_x \cdot M_y
\]

\[
\text{COV}(X, Y)
\]
The correlation coefficient $\rho$ is defined as:

$$\rho = \frac{\text{cov}(X, Y)}{\sqrt{\text{cov}(X, X) \text{cov}(Y, Y)}} = \frac{\text{cov}(X, Y)}{\text{std} X \text{std} Y}$$

$$c_{02} = \text{E}[(Y - \text{E}[Y])^2], \quad c_{20} = \text{E}[(X - \text{E}[X])^2]$$

1. $0 \leq \rho \leq 1$

2. If $\rho = 0 \Rightarrow X$ and $Y$ are called uncorrelated.