Lecture 2: Review of basic concepts in probability

1) Probability distribution function (cumulative distribution function) - CDF

Let $X$ be a continuous random variable. Then a CDF of $X$ is defined as:

$$\text{cdf}(x) = \Pr\{X \leq x\}$$

**Typical shape of a CDF**

Properties of a CDF:

1. $0 \leq \text{cdf}(x) \leq 1$, for all $x$
2. $\text{cdf}(\pm \infty) = 0$
3. $\text{cdf}(x) = \Pr\{X \leq x\}$
4. $\Pr\{a \leq X \leq b\} = \text{cdf}(b) - \text{cdf}(a)$
5. $\Pr\{X > b\} = 1 - \text{cdf}(b)$

Example: Consider a r.v. with $\text{cdf}(x) = \begin{cases} 1 - \exp(-2x) & x \geq 0 \\ 0 & x < 0 \end{cases}$

Determine $\Pr\{X \leq 3\}$, $\Pr\{X \in (1, 3)\}$

$$\Pr\{X \leq 3\} = \text{cdf}(3) = 1 - \exp(-2 \times 3) = 0.9975.$$  

$$\Pr\{X \in (1, 3)\} = \text{cdf}(3) - \text{cdf}(1) = [1 - \exp(-6)] - [1 - \exp(-2)] = 0.9975 - 0.8647 = 0.1329.$$
Probability density function (PDF)

\[ \text{pdf}(x) = \frac{d}{dx} [\text{cdf}(x)] \]

\[ P_a \leq x \leq b = \int_a^b \text{pdf}(x) \, dx \]

Some properties of PDF

1) \( \text{pdf}(x) \geq 0 \), \( \forall x \)

2) \( \int_{-\infty}^{\infty} \text{pdf}(x) \, dx = 1 \)

3) \( \text{cdf}(x) = \int_{-\infty}^{x} \text{pdf}(x) \, dx \)

4) \( P_a \leq x \leq b = \int_a^b \text{pdf}(x) \, dx \)

Example 2: Find \( \text{pdf}(x) \) for \( \text{cdf}(x) \), in Example 1.

\[ \text{pdf}(x) = \frac{d}{dx} \text{cdf}(x) = \frac{d}{dx} [1 - \exp(-2x)] = 2 \exp(-2x), \quad x > 0 \]

Expectation, mean and moments

Consider a function \( f(x) \), where \( x \) is a random variable with given pdf. The output (i.e. \( f(x) \)) is a random variable as well. Expected value of \( f(x) \) is given by
\[ E_x f(x) = \int_{-\infty}^{\infty} f(x) \cdot p_d(x) \, dx \]

**Some cases of special interest**

1. \( f(x) = x \)  \( E_x x = \int_{-\infty}^{\infty} x \cdot p_d(x) \, dx \) - mean value of \( x \) \( (\mu) \)

2. \( f(x) = x^2 \)  \( E_x x^2 = \int_{-\infty}^{\infty} x^2 \cdot p_d(x) \, dx \) - second moment of \( x \)

3. \( f(x) = x^n \)  \( E_x x^n = \int_{-\infty}^{\infty} x^n \cdot p_d(x) \, dx \) - \( n \) th moment

4. \( f(x) = (x-\mu)^n \)  \( E_x (x-\mu)^n = \int_{-\infty}^{\infty} (x-\mu)^n \cdot p_d(x) \, dx \) - \( n \) th central moment. For \( n = 2 \),  \( E_x (x-\mu)^2 = \sigma^2 \) - variance  \( \sigma \) - standard deviation

**Example 3.** Find mean and standard deviation of a r.v. in Example 1.8.2

\[ p_d(x) = 2e^{x} \]

\[ \mu = E_x x = \int_{-\infty}^{\infty} x \cdot p_d(x) \, dx = \int_{0}^{\infty} x \cdot 2e^{-2x} \, dx = \]

\[ = \frac{1}{2} \int_{0}^{\infty} (2x) \cdot e^{-2x} \, dx = \frac{1}{2} \]

\[ \sigma^2 = E_x (x-\mu)^2 = \int_{0}^{\infty} (x-\frac{1}{2})^2 \cdot 2e^{-2x} \, dx = \int_{0}^{\infty} (x^2 - x + \frac{1}{4}) \cdot 2e^{-2x} \, dx = \]

\[ = \frac{1}{4} \int_{0}^{\infty} x^2 \cdot e^{-2x} \, dx - \frac{1}{4} \int_{0}^{\infty} 2x \cdot e^{-2x} \, dx + \frac{1}{8} \int_{0}^{\infty} e^{-2x} \, dx \]

\[ = \frac{1}{4} \left[ -\frac{e^{-2t}}{2} \right]_{0}^{\infty} - \frac{1}{4} \left[ -\frac{e^{-2t}}{2} \right]_{0}^{\infty} + \frac{1}{8} \left[ -\frac{e^{-2t}}{2} \right]_{0}^{\infty} \]

\[ = \frac{1}{4} - \frac{1}{4} + \frac{1}{4} = \frac{1}{4} \]

\[ \sigma^2 = \frac{1}{4} \Rightarrow \sigma = \frac{1}{2} \]
Normal PDF (Gaussian PDF)

\[ x \sim \mathcal{N}(\mu, \sigma), \text{ read } "x \text{ is a random variable following normal distribution with mean value of } \mu \text{ and standard deviation of } \sigma." \]

\[ \mathcal{N}(\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{(x-\mu)^2}{2\sigma^2} \right) \]

\[ x \sim \mathcal{N}(\mu, \sigma) \Rightarrow E(x \mid x) = \mu \]
\[ E((x-\mu)^2) = \sigma^2 \]

* Normal distribution used to represent many statistical phenomena due to CLT

* CLT: If there is a winch of random variables of arbitrary distributions but at approximately the same power, the sum tends to be distributed in accordance to Gaussian distribution

Random process

* Most signals that we encounter in communications are random signals, i.e.,
random processes

* The actual time domain shape of a random signal is not known in advance
  \( \Rightarrow \) we are designing communication system for unknown signals!
To describe communication systems, we use mathematical models. There are two types of models:

1) Deterministic - no uncertainty, all events depend on behavior.
2) Probabilistic - physical phenomena involve many factors (known & unknown).

We construct probabilistic models describing the processes in communication systems. In most simple cases, we speak about probabilities of certain events.

Random processes have two properties:

1) They are functions of time.
2) They are random in the sense that before the outcome of the process is observed, the actual shape of the process cannot be predicted.

While describing random processes, we usually think in terms of sample space.

Sample space contains all possible realizations of a random process. At any fixed time, the value of the process is a random variable.
Stationary process:

- At any point in time, each sample process takes a value.
- The value taken by a sample process is a random variable described by its pdf(x).
- In general, pdf(x) may depend on the observation time.
- In special cases when pdf(x) does not depend on time, it's a stationary process.

\[ X(t) \]

\[ X(t) \]

\[ Ppdf(x,t) \quad Ppdf(x,t+\Delta t) = Ppdf(x,t) \]

\[ P\omega \quad y.t,A \quad \text{strict sense} \]

\[ \text{Stationary process} \]

\[ E\{x(t)\} = \int \limits_{-\infty}^{\infty} x \cdot pdf(x,t) \, dx = E\{x(t+\Delta t)\} = \int \limits_{-\infty}^{\infty} x \cdot pdf(x,t+\Delta t) \, dx \]

\[ E\{x^2(t)\} = \int \limits_{-\infty}^{\infty} x^2 \cdot pdf(x,t) \, dx = E\{x^2(t+\Delta t)\} = \int \limits_{-\infty}^{\infty} x^2 \cdot pdf(x,t+\Delta t) \, dx \]

- If just the two samples are independent on time, it's W.S.S. process.
Ergodic process: Consider a stochastic stationary process \( x(t) \) and an arbitrary function \( f(x) \). We can define average of \( f(x) \) in two ways:

1) Ensemble average:
\[
\bar{f}(x(t)) = \lim_{T \to \infty} \frac{1}{T} \int_{-T}^{T} f(x) \cdot \text{pdf}(x, t_0) \, dx
\]

Since the process is stationary, the above expression says independent of \( t_0 \), i.e., it is the same for every \( t_0 \).

2) Time average:
\[
< f(x) > = \lim_{T \to \infty} \frac{1}{T} \int_{-T}^{T} f(x) \, dx
\]
- The average is conducted on time
- Only one realization of the process is needed.

The process is called Ergodic if \( E(f(x)) = < f(x) > \)
* Time ensemble average and time domain average are interchangeable

* Ergodic \( \Rightarrow \) Ensemble behavior is the same as time domain behavior.
* For most random signals that we will discuss in this class, we will assume Ergodicity.
* Most natural signals exhibit this property.