Chapter II. Random Signals

1. Definitions

Random variables

- A **random variable** is a variable that holds a value produced by a (partially) random phenomenon (experiment)
- Typically denoted as X, Y etc..
- Examples:
  - The value of a dice
  - The value of the voltage in a circuit
- We get a single value, but
- The opposite = a **constant value**

Sample space and realizations

- **A realization** = a single outcome of the random experiment
- **Sample space** Ω = the set of all values that can be taken by a random variable X
  - i.e. the set of all possible realizations
- Example: rolling a dice
  - we might get a realization $X = 3$
  - but we could have got any value from the sample space

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

- If Ω is a discrete set –> **discrete** random variable
  - Example: value of a dice
- If Ω is a continuous set –> **continuous** random variable
  - Example: a voltage value

Discrete random variable

- For a discrete random variable, the probability that X has x_i is given by the **probability mass function (PMF)** $w(x_i)$

$$w_X(x_i) = P\left\{X = x_i\right\}$$

- Example: what is the PMF of a dice?
- For simplicity we will call it simply the *"distribution"* of X
- The **cumulative distribution function (CDF)** gives the probability that the value of X is smaller or equal than the argument x_i

$$F_X(x_i) = P\left\{X \leq x_i\right\}$$

- In Romanian: *"functie de repartitie"*
- Example: draw the CDF of a dice (= a *staircase* function)

Continuous random variable

- The CDF of a continuous r.v. is in the same way:

$$F_X(x_i) = P\{X \leq x_i\}$$

- The derivative of the CDF is the **probability density function (PDF)**

$$w_X(x_i) = \frac{dF_X(x_i}{dx_i}$$

- The PDF gives the **probability that the value of X is in a small vicinity of around** x_i
- Important: the probability that a continuous r.v. X is **exactly** equal to a value x_i is **zero**
  - because there are an infinity of possibilities (continuous)
  - That's why we can't define a probability mass function like for discrete

Probability and distribution

- Compute probability from PDF (continuous r.v.):

$$P\left\{A \leq X \leq B\right\} = \int_A^B w_X(x)dx$$

- Compute probability from PMF (discrete r.v.):

$$P\left\{A \leq X \leq B\right\} = \sum_{x=A}^{B} w_X(x)$$

- Probability that a r.v. X is between A and B is **the area below the PDF**

Properties of PDF/PMF/CDF

- The CDF is monotonously increasing (non-decreasing)
- The PDF/PMF are always ≥ 0
- The CDF starts from 0 and goes up to 1
- Integral/sum over all of the PDF/PMF $= 1$
- Some others, mention when needed

Examples

- Gaussian PDF
- Uniform PDF
- ...

Multiple random variables

- Consider a system with two random variables X and Y
- Joint cumulative distribution function:

$$F_{XY}(x_i, y_j) = P\left\{X \leq x_i \cap Y \leq y_i\right\}$$

- Joint probability density function:

$$w_{XY}(x_i, y_j) = \frac{\partial^2 P_{XY}(x_i, y_j)}{\partial x \partial y}$$

- The joint PDF gives the probability that the values of the two r.v. X and Y are in a **vicinity** of x_i and y_i simultaneously
- Similar definitions extend to the case of discrete random variables

Random process

- A **random process** = a sequence of random variables indexed in time
- **Discrete-time** random process $f[n]$ = a sequence of random variables at discrete moments of time
  - e.g.: a sequence 50 of throws of a dice, the daily price on the stock market
- **Continuous-time** random process $f(t)$ = a continuous sequence of random variables at every moment
  - e.g.: a noise voltage signal, a speech signal
- Every sample from a random process is a (different) random variable!
  - e.g. $f(t_0)$ = value at time t_0 is a r.v.

Realizations of random processes

- A **realization** of the random process = a particular sequence of values
  - e.g. we see a given noise signal on the oscilloscope, but *we could have seen any other realization just as well*
- When we consider a random process = we consider the set of all possible realizations
- Example: draw on whiteboard

Distributions of order 1 of random processes

- Every sample $f(t_1)$from a random process is a random variable
  - with CDF $F_1(x_i; t_1)$
  - with PDF $w_1(x_i; t_1) = \frac{dF_1(x_i;t_1)}{dx_i}$
- The sample at time t_2 is a different random variable with **possibly different** functions
  - with CDF $F_1(x_i; t_2)$
  - with PDF $w_1(x_i; t_2) = \frac{dF_1(x_i;t_2)}{dx_i}$
- These functions specify how the value of one sample is distributed
- The index w_1 indicates we consider a single random variable from the process $->$ distributions of order 1

Distributions of order 2

- A pair of random variables $f(t_1)$ and $f(t_2)$ sampled from the random process $f(t)$ have
  - joint CDF $F_2(x_i, x_j; t_1, t_2)$
  - joint PDF $w_2(x_i, x_j; t_1, t_2) = \frac{\partial^2 F_2(x_i, x_j; t_1, t_2)}{\partial x_i \partial x_j}$
- These functions specify how the pair of values is distributed (are distributions of order 2)
- Marginal integration

$$w_1(x_i; t_1) = \int_{\infty}^{\infty} w_2(x_i, x_j; t_1, t_2) dx_j$$

- (integrate over one variable –> disappears –> only the other one remains)

Distributions of order n

- Generalize to n samples of the random process
- A set of n random variables $f(t_1), ...f(t_n)$ sampled from the random process $f(t)$ have
  - joint CDF $F_n(x_1, ...x_n; t_1, ...t_n)$
  - joint PDF $w_n(x_1, ...x_n; t_1, ...t_n) = \frac{\partial^2 F_n(x_1, ...x_n; t_1, ...t_n)}{\partial x_1 ... \partial x_n}$
- These functions specify how the whole set of n values is distributed (are distributions of order n)

Statistical averages

We characterize random processes using statistical / temporal averages (*moments*)

1. Average value

$$\overline{f(t_1)} = \mu(t_1) = \int_{-\infty}^{\infty} x \cdot w_1(x; t_1) dx$$

2. Average squared value (*valoarea patratica medie*)

$$\overline{f^2(t_1)} = \int_{-\infty}^{\infty} x^2 \cdot w_1(x; t_1) dx$$

Statistical averages - variance

3. Variance (= *dispersia*)

$$\sigma^2(t_1) = \overline{\{f(t_1) - \mu(t_1)\}^2} = \int_{-\infty}^{\infty} (x - \mu(t_1)^2 \cdot w_1(x; t_1) dx$$

- The variance can be computed as:

$$\sigma^2(t_1) = \overline{\{f(t_1) - \mu(t_1)\}^2} = \overline{f(t_1)^2 - 2f(t_1)\mu(t_1) + \mu(t_1)^2} = \overline{f^2(t_1)} - \mu($$

Statistical averages - autocorrelation

4. The autocorrelation function

$$R_{ff}(t_1, t_2) = \overline{f(t_1)f(t_2)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 w_2(x_1, x_2; t_1, t_2) dx_1 dx_2$$

5. The correlation function (for different random processes $f(t)$ and $g(t)$)

$$R_{fg}(t_1, t_2) = \overline{f(t_1)g(t_2)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 y_2 w_2(x_1, y_2; t_1, t_2) dx_1 dy_2$$

- **Note 1**:
  - all these values are calculated across all realizations, at a single time t_1
  - all these characterize only the r.v. at time t_1
  - at a different time t_2, the r. v. $f(t_2)$ is different so *all average values might be different*

Temporal averages

- What to do when we only have access to a single realization?
- Compute values **for a single realization $f^{(k)(t)}$, across all time moments**

1. Temporal average value

$$\overline{f^{(k)}(t)} = \mu^{(k)} = \lim_{T \to \infty} \frac{1}{T} \int_{T/2}^{T/2} f^{(k)}(t) dt$$

- This value does not depend on time t

2. Temporal average squared value

$$\overline{[f^{(k)}(t)]^2} = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} [f^{(k)}(t)]^2 dt$$

Temporal variance

3. Temporal variance

$$\sigma^2 = \overline{\left\{f^{(k)}(t) - \mu^{(k)}\right\}^2} = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} (f^{(k)}(t) - \mu^{(k)})^2 dt$$

- The variance can be computed as:

$$\sigma^2 = \overline{[f^{(k)}(t)]^2} - [\mu^{(k)}]^2$$

Temporal autocorrelation

4. The temporal autocorrelation function

$$R_{ff}(t_1, t_2) = \overline{f^{(k)}(t_1 + t)f^{(k)}(t_2 + t)}$$

$$R_{ff}(t_1, t_2) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f^{(k)}(t_1 + t)f^{(k)}(t_2 + t)dt$$

5. The temporal correlation function (for different random processes $f(t)$ and $g(t)$)

$$R_{fg}(t_1, t_2) = \overline{f^{(k)}(t_1 + t)g^{(k)}(t_2 + t)}$$

$$R_{fg}(t_1, t_2) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f^{(k)}(t_1 + t)g^{(k)}(t_2 + t)dt$$

Stationary random processes

- All the statistical averages are dependent on the time t_1
  - i.e. they might be different for a sample at t_2
- **Stationary** random process = when statistical averages are identical upon shifting the time origin (e.g. delaying the signal
- The PDF are identical when shifting the time origin:

$$w_n(x_1, ...x_n; t_1, ...t_n) = w_n(x_1, ...x_n; t_1 + \tau, ...t_n + = tau)$$

- Strictly stationary / strongly stationary / strict-sense stationary:
  - relation holds for every n
- Weakly stationary / wide-sense stationary:
  - relation holds only for $n = 1$ and $n = 2$ (the most used)

Consequences of stationarity

- For $n = 1$:

$$w_1(x_i; t_1) = w_1(x_i; t_2) = w_1(x_i)$$

- Consequence: the average value, average squared value, variance of a sample are all **identical** for any time t

$$\overline{f(t)} = constant, \forall t$$

$$\overline{f^2(t)} = constant, \forall t$$

$$\sigma^2(t) = constant, \forall t$$

- For $n = 2$:

$$w_2(x_i, x_j; t_1, t_2) = w_2(x_i, x_j; 0, t_2 - t_1) = w_2(x_i, x_2; t_2 - t_1)$$

- Consequence: the autocorrelation / correlation functions depend only on the **time difference** $t_2 - t_1$ between the samples, no matter where they are located

$$R_{ff}(t_1, t_2) = R_{ff}(t_2 - t_1) = R_{ff}(\tau)$$

$$R_{fg}(t_1, t_2) = R_{fg}(t_2 - t_1) = R_{fg}(\tau)$$

Ergodic random processes

- In practice, we have access to a single realization
- **Ergodic** random process = when the temporal averages on any realization are **equal** to the statistical averages
- We can compute all averages from a single realization
  - the realization must be very long (length $\rightarrow \infty$)
  - a realization is characteristic of the whole process
  - realizations are all similar to the others, statistically
- Most random processes we are about are ergodic and stationary
  - e.g. noises
- Example of non-ergodic process:
  - throw a dice, then the next 50 values are identical to the first
  - a single realization is not characteristic

Practical distributions

- Some of the most encountered probability density functions:
- The uniform distribution $U[a, b]$
  - insert expression here
- The normal (gaussian) distribution $N(\mu, \sigma^2)$

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

  - has average value μ
  - has variance (*"dispersia"*) σ^2
  - has the familiar "bell" shape
  - variance controls width
  - narrower = taller, fatter = shorter

Computation of probabilities for normal functions

- We sometimes need to compute $\int_a^b$ of a normal function
- Use *the error function*:

$$erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$$

- The cumulative distribution function of a normal distribution $N(\mu, \sigma^2)$

$$F(X) = \frac{1}{2}(1 + erf(\frac{x - \mu}{\sigma\sqrt{2}}))$$

- The error function can be simply calculated on Google, e.g. search $erf(0.5)$
- Also, we might need:
  - $erf(-\infty) = -1$
  - $erf(\infty) = 1$
- Examples at blackboard

Properties of the auto-correlation function

- For a stationary random process:

$$R_{ff}(\tau) = \overline{f(t)f(t+\tau)}$$

$$R_{ff}(t_1, t_2) = R_{ff}(\tau = t_2 - t_1)$$

- Is the average value of a product of two samples time τ apart
- Depends on a single value τ = time difference of the two samples

The Wiener-Khinchin theorem

- *Rom: teorema Wiener-Hincin*
- **The Fourier transform of the autocorr function = power spectral density of the process**

$$S_{ff}(\omega) = \int_{-\infty}^{\infty} R_{ff}(\tau) e^{-j\omega\tau} d\tau$$

$$R_{ff}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{ff}(\omega) e^{j\omega\tau} d\omega$$

- No proof
- The power spectral density
  - tells the average power of the process at every frequency
- Some random processes have low frequencies (they vary rather slowly)
- Some random processes have high frequencies (they vary rather fast)

White noise

- White noise = random process with autocorr function = a Dirac

$$R_{ff}(\tau) = \delta(\tau)$$

- Any two different samples are not correlated
  - all samples are absolutely independent one of the other
- Power spectral density = a constant
  - has equal power at all frequencies
- In real life, power goes to 0 for very high frequencies
  - e.g. samples which are very close are necessarily correlated
  - = *limited white noise*

Properties of the autocorrelation function

1. Is even

$$R_{ff}(\tau) = R_{ff}(-\tau)$$

- Proof: change variable in definition

2. At infinite it goes to a constant

$$R_{ff}(\infty) = \overline{f(t)}^2 = const$$

- Proof: two samples separated by ∞ are independent

3. Is maximum in 0

$$R_{ff}(0) \geq R_{ff}(\tau)$$

- Proof: start from $\overline{(f(t) - f(t+\tau))^2} \geq 0$
- Interpretation: different samples might vary differently, by a sample is always identical with itself

Properties of the autocorrelation function

4. Value in 0 = the power of the random process

$$R_{ff}(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{ff}(\omega)d\omega$$

- Proof: Put $\tau = 0$ in inverse Fourier transform of Wiener-Khinchin theorem

5. Variance = difference between values at 0 and ∞

$$\sigma^2 = R_{ff}(0) - R_{ff}(\infty)$$

- Proof: $R_{ff}(0) = \overline{f(t)^2}$, $R_{ff}(\infty) = \overline{f(t)}^2$