ECE 3800 Probabilistic Methods of Signal and System Analysis, Spring 2015

Course Topics:

1. Probability
2. Random variables
3. Multiple random variables
4. Elements of Statistics
5. Random processes
6. Correlation Functions
7. Spectral Density
8. Responses of Linear Systems

Course Objectives:

This course seeks to develop a mathematical understanding of basic statistical tools and processes as applied to inferential statistics and statistical signal processing (with ABET objectives are listed per course objective).

1. The student will learn how to convert a problem description into a precise mathematical probabilistic statement (a)
2. The student will use the general properties of random variables to solve a probabilistic problem (a, e)
3. The student will be able to use a set of standard probability distribution functions suitable for engineering applications (a)
4. The student will be able to calculate standard statistics from probability mass, distribution and density functions (a)
5. The student will learn how to calculate confidence intervals for a population mean (a)
6. The student will be able to recognize and interpret a variety of deterministic and nondeterministic random processes that occur in engineering (a, b, e)
7. The student will learn how to calculate the autocorrelation and spectral density of an arbitrary random process (a)
8. The student will understand stochastic phenomena such as white, pink and black noise (a)
9. The student will learn how to relate the correlation of and between inputs and outputs based on the autocorrelation and spectral density (a)
10. The student will understand the mathematical characteristics of standard frequency isolation filters (a, e)
11. The student will be exposed to the signal-to-noise optimization principle as applied to filter design (a, e, k)
12. The student will be exposed to Weiner and matched noise filters (a, c, e)
Exam  **Tuesday 10:15-12:15 am.**
Those with permission to take it early Monday, 12:30- , Room B-211.

Expect approximately 6-8 multipart problems

- Problems will be similar or even identical to those from previous exams and homework.
- I will try to insure that an element from every chapter is tested.

The exam is open paper notes. No electronic devices except for a calculator (and I intentionally make calculator use minimal, possibly even useless, or even providing incorrect or incomplete results).

Be familiar with or bring the appropriate math tables as needed (Appendix A p. 419-424 is a good set).

- Appendix D: Normal Probability Distribution Function
- Appendix E: The Q-Function
- Appendix F: Student’s t Distribution Function

Minimal calculator work is required … using rational fractions may be easier.

- I intentionally make calculator use minimal, possibly even useless, or even providing incorrect or incomplete results.

**Additional Materials for studying**

All homework solution sets have been posted and returned.

Skills are available on the web site.
Table of Contents

1. Introduction to Probability

1.1. Engineering Applications of Probability

1.2. Random Experiments and Events

1.3. Definitions of Probability
   - Experiment
   - Possible Outcomes
   - Trials
   - Event
   - Equally Likely Events/Outcomes
   - Objects
   - Attribute
   - Sample Space
   - With Replacement and Without Replacement

1.4. The Relative-Frequency Approach

\[ r(A) = \frac{N_A}{N} \]
\[ \text{Pr}(A) = \lim_{N \to \infty} r(A) \]

Where \( \text{Pr}(A) \) is defined as the probability of event A.

1.  \( 0 \leq \text{Pr}(A) \leq 1 \)
2.  \( \text{Pr}(A) + \text{Pr}(B) + \text{Pr}(C) + \cdots = 1 \), for mutually exclusive events
3.  An impossible event, A, can be represented as \( \text{Pr}(A) = 0 \).
4.  A certain event, A, can be represented as \( \text{Pr}(A) = 1 \).

1.5. Elementary Set Theory
   - Set
   - Subset
   - Space
   - Null Set or Empty Set
   - Venn Diagram
   - Equality
   - Sum or Union
   - Products or Intersection
   - Mutually Exclusive or Disjoint Sets
   - Complement
   - Differences
   - Proofs of Set Algebra

1.6. The Axiomatic Approach

1.7. Conditional Probability

\[ \Pr(A \cap B) = \Pr(A \mid B) \cdot \Pr(B), \text{ for } \Pr(B) > 0 \]

\[ \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}, \text{ for } \Pr(B) > 0 \]

Joint Probability

\[ \Pr(A, B) \neq \Pr(A) \text{ when } A \text{ follows } B \]

\[ \Pr(A, B) = \Pr(B, A) = \Pr(A \mid B) \cdot \Pr(B) = \Pr(B \mid A) \cdot \Pr(A) \]

Marginal Probabilities:

Total Probability

\[ \Pr(B) = \Pr(B \mid A_1) \cdot \Pr(A_1) + \Pr(B \mid A_2) \cdot \Pr(A_2) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n) \]

Bayes Theorem

\[ \Pr(A_i \mid B) = \frac{\Pr(B \mid A_i) \cdot \Pr(A_i)}{\Pr(B \mid A_1) \cdot \Pr(A_1) + \Pr(B \mid A_2) \cdot \Pr(A_2) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n)} \]

1.8. Independence

\[ \Pr(A, B) = \Pr(B, A) = \Pr(A) \cdot \Pr(B) \]

1.9. Combined Experiments

1.10. Bernoulli Trials

\[ \Pr(A \text{ occurring } k \text{ times in } n \text{ trials}) = p_n(k) = \binom{n}{k} p^k q^{n-k} \]

1.11. Applications of Bernoulli Trials
2. Random Variables

2.1. Concept of a Random Variable

2.2. Distribution Functions

Probability Distribution Function (PDF)
- $0 \leq F_X(x) \leq 1$, for $-\infty < x < \infty$
- $F_X(-\infty) = 0$ and $F_X(\infty) = 1$
- $F_X$ is non-decreasing as $x$ increases
- $\Pr(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$

For discrete events
For continuous events

2.3. Density Functions

Probability Density Function (pdf)
- $f_X(x) \geq 0$, for $-\infty < x < \infty$
- $\int_{-\infty}^{\infty} f_X(x) \cdot dx = 1$
- $F_X = \int_{-\infty}^{x} f_X(u) \cdot du$
- $\Pr(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f_X(x) \cdot dx$

Probability Mass Function (pmf)
- $f_X(x) \geq 0$, for $-\infty < x < \infty$
- $\sum_{x=\infty}^{x_{\infty}} f_X(x) = 1$
- $F_X = \sum_{x=-\infty}^{x} f_X(x)$
- $\Pr(x_1 \leq X \leq x_2) = \sum_{x=x_1}^{x_2} f_X(x)$

Functions of random variables

$Y = fn(X)$ and $X = fn^{-1}(Y)$

$f_Y(y) = f_X(fn^{-1}(y)) \cdot \left| \frac{dx}{dy} \right|$
2.4. Mean Values and Moments

1st, general, nth Moments

\[ \overline{X} = E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x) \cdot dx \quad \text{or} \quad \overline{X} = E[X] = \sum_{x=-\infty}^{\infty} x \cdot \Pr(X = x) \]

\[ E[g(X)] = \int_{-\infty}^{\infty} g(X) \cdot f_X(x) \cdot dx \quad \text{or} \quad E[g(X)] = \sum_{x=-\infty}^{\infty} g(X) \cdot \Pr(X = x) \]

\[ \overline{X^n} = E[X^n] = \int_{-\infty}^{\infty} x^n \cdot f_X(x) \cdot dx \quad \text{or} \quad \overline{X^n} = E[X^n] = \sum_{x=-\infty}^{\infty} x^n \cdot \Pr(X = x) \]

Central Moments

\[ \left(\overline{X - \overline{X}}^n\right) = E\left[(X - \overline{X})^n\right] = \int_{-\infty}^{\infty} (x - \overline{X})^n \cdot f_X(x) \cdot dx \]

\[ \left(\overline{X - \overline{X}}^n\right) = E\left[(X - \overline{X})^n\right] = \sum_{x=-\infty}^{\infty} (x - \overline{X})^n \cdot \Pr(X = x) \]

Variance and Standard Deviation

\[ \sigma^2 = (\overline{X - \overline{X}}^2) = E\left[(X - \overline{X})^2\right] = \int_{-\infty}^{\infty} (x - \overline{X})^2 \cdot f_X(x) \cdot dx \]

\[ \sigma^2 = (\overline{X - \overline{X}}^2) = E\left[(X - \overline{X})^2\right] = \sum_{x=-\infty}^{\infty} (x - \overline{X})^2 \cdot \Pr(X = x) \]

2.5. The Gaussian Random Variable

\[ f_X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left(\frac{-(x - \overline{X})^2}{2 \cdot \sigma^2}\right), \text{ for } -\infty < x < \infty \]

where \( \overline{X} \) is the mean and \( \sigma \) is the variance

\[ F_X(x) = \int_{v=-\infty}^{x} \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left(\frac{-(v - \overline{X})^2}{2 \cdot \sigma^2}\right) \cdot dv \]

Unit Normal (Appendix D)

\[ \Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{u=-\infty}^{x} \exp\left(\frac{-u^2}{2}\right) \cdot du \]

\[ \Phi(-x) = 1 - \Phi(x) \]

\[ F_X(x) = \Phi\left(\frac{x - \overline{X}}{\sigma}\right) \quad \text{or} \quad F_X(-x) = 1 - \Phi\left(\frac{x - \overline{X}}{\sigma}\right) \]

The Q-function is the complement of the normal function, \( \Phi \): (Appendix E)

\[ Q(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{u=x}^{\infty} \exp\left(\frac{-u^2}{2}\right) \cdot du = 1 - \Phi(x) \]
2.6. Density Functions Related to Gaussian
    Rayleigh
    Maxwell

2.7. Other Probability Density Functions
    Exponential Distribution
    \[
    f_T(t) = \begin{cases} 
    \frac{1}{M} \cdot \exp\left(-\frac{t}{M}\right), & \text{for } 0 \leq t \\ 
    0, & \text{for } t < 0 
    \end{cases}
    \]
    \[
    F_T(t) = \begin{cases} 
    1 - \exp\left(-\frac{t}{M}\right), & \text{for } 0 \leq t \\ 
    0, & \text{for } t < 0 
    \end{cases}
    \]
    \[
    \bar{T} = \mathbb{E}[T] = M \\
    T^2 = \mathbb{E}[T^2] = 2 \cdot M^2 \\
    E\left[(\tau - \bar{T})^2\right] = \sigma_T^2 = T^2 - \mathbb{E}[T]^2 = 2 \cdot M^2 - (M)^2 = M^2 
    \]

    Binomial Distribution
    \[
    f_B(x) = \sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \cdot \delta(x - k) 
    \]
    \[
    F_B(x) = \sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \cdot u(x - k) 
    \]

2.8. Conditional Probability Distribution and Density Functions
    \[
    \Pr(A \cap B) = \Pr(A \mid B) \cdot \Pr(B), \text{ for } \Pr(B) > 0 
    \]
    \[
    \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}, \text{ for } \Pr(B) > 0 
    \]
    \[
    \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A, B)}{\Pr(B)}, \text{ for } \Pr(B) > 0 
    \]

    It can be shown that \(F(x \mid M)\) is a valid probability distribution function with all the expected characteristics:
    - \(0 \leq F(x \mid M) \leq 1, \text{ for } -\infty < x < \infty\)
    - \(F(-\infty \mid M) = 0 \) and \(F(\infty \mid M) = 1\)
    - \(F(x \mid M)\) is non-decreasing as \(x\) increases
    - \(\Pr(x_1 < X \leq x_2 \mid M) = F(x_2 \mid M) - F(x_1 \mid M)\)

2.9. Examples and Applications
3. Several Random Variables

3.1. Two Random Variables

Joint Probability Distribution Function (PDF)
\[ F(x, y) = \Pr(X \leq x, Y \leq y) \]

- \(0 \leq F(x, y) \leq 1\), for \(-\infty < x < \infty\) and \(-\infty < y < \infty\)
- \(F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0\)
- \(F(\infty, \infty) = 1\)
- \(F(x, y)\) is non-decreasing as either \(x\) or \(y\) increases
- \(F(x, \infty) = F_X(x)\) and \(F(\infty, y) = F_Y(y)\)

Joint Probability Density Function (pdf)
\[ f(x, y) = \frac{\partial^2 F_X(x)}{\partial x \partial y} \]

- \(f(x, y) \geq 0\), for \(-\infty < x < \infty\) and \(-\infty < y < \infty\)
- \(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \cdot dx \cdot dy = 1\)
- \(F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) \cdot du \cdot dv\)
- \(f_X(x) = \int_{-\infty}^{\infty} f(x, y) \cdot dy\) and \(f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \cdot dx\)
- \(\Pr(x_1 < X \leq x_2, y_1 < Y \leq y_2) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} f(x, y) \cdot dx \cdot dy\)

Expected Values
\[ E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \cdot f(x, y) \cdot dx \cdot dy \]

Correlation
\[ E[X \cdot Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y \cdot f(x, y) \cdot dx \cdot dy \]
3.2. Conditional Probability--Revisited

\[
F_X(x \mid Y \leq y) = \frac{\Pr(X \leq x \mid M)}{\Pr(M)} = \frac{F(x, y)}{F_Y(y)}
\]

\[
F_X(x \mid Y_1 \leq Y \leq y_2) = \frac{F(x, y_2) - F(x, y_1)}{F_Y(y_2) - F_Y(y_1)}
\]

\[
F_X(x \mid Y = y) = \frac{f(x, y)}{f_Y(y)}
\]

\[
F_Y(y \mid X = x) = \frac{f(x, y)}{f_X(x)}
\]

\[
f(x \mid y) = \frac{f(y \mid x) \cdot f_X(x)}{f_Y(y)}
\]

\[
f(x, y) = f(x \mid Y = y) \cdot f_Y(y) = f(y \mid X = x) \cdot f_X(x)
\]

\[
f(x \mid Y = y) = \frac{f(y \mid X = x) \cdot f_X(x)}{f_Y(y)}
\]

\[
f(y \mid X = x) = \frac{f(x \mid Y = y) \cdot f_Y(y)}{f_X(x)}
\]

3.3. Statistical Independence

\[
f(x, y) = f_X(x) \cdot f_Y(y)
\]

\[
E[X \cdot Y] = E[X] \cdot E[Y] = \overline{X} \cdot \overline{Y}
\]

3.4. Correlation between Random Variables

\[
E[X \cdot Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y \cdot f(x, y) \cdot dx \cdot dy
\]

Covariance

\[
E[(X - E[X]) \cdot (Y - E[Y])] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X) \cdot (y - \mu_Y) \cdot f(x, y) \cdot dx \cdot dy
\]

Correlation coefficient or normalized covariance,

\[
\rho = \frac{E[(X - \mu_X) \cdot (Y - \mu_Y)]}{\sigma_X \cdot \sigma_Y} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X) \cdot (y - \mu_Y) \cdot f(x, y) \cdot dx \cdot dy}{\sigma_X \cdot \sigma_Y}
\]

\[
\rho = \frac{E[x \cdot y] - \mu_X \cdot \mu_Y}{\sigma_X \cdot \sigma_Y}
\]
3.5. Density Function of the Sum of Two Random Variables

\[ Z = X + Y \]

\[ F(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v) \cdot du \cdot dv \]

\[ F_Z(z) = \int_{-\infty}^{\infty} f_Y(y) \cdot \int_{-\infty}^{\infty} f_X(x) \cdot dx \cdot dy \]

\[ f_Z(z) = \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z-x) \cdot dx = \int_{-\infty}^{\infty} f_Y(y) \cdot f_X(z-y) \cdot dy \]

3.6. Probability Density Function of a Function of Two Random Variables

Define the function as

\[ Z = \phi_1(X, Y) \quad \text{and} \quad W = \phi_2(X, Y) \]

and the inverse as

\[ X = \psi_1(Z, W) \quad \text{and} \quad Y = \psi_2(Z, W) \]

The original pdf is \( f(x, y) \) with the derived pdf in the transform space of \( g(z, w) \).

Then it can be proven that:

\[ \Pr(z_1 < Z \leq z_2, w_1 < W \leq w_2) = \Pr(x_1 < X \leq x_2, y_1 < Y \leq y_2) \]

or equivalently

\[ \int_{w_1}^{w_2} \int_{z_1}^{z_2} g(z, w) \cdot dz \cdot dw = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x, y) \cdot dx \cdot dy \]

Using an advanced calculus theorem to perform a transformation of coordinates.

\[ g(z, w) = f(\psi_1(z, w), \psi_2(z, w)) \cdot \left| \begin{array}{cc} \frac{\partial x}{\partial x} & \frac{\partial x}{\partial x} \\ \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \end{array} \right| = f(\psi_1(z, w), \psi_2(z, w)) \cdot |J| \]

\[ \int_{w_1}^{w_2} \int_{z_1}^{z_2} g(z, w) \cdot dz \cdot dw = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(\psi_1(z, w), \psi_2(z, w)) \cdot |J| \cdot dz \cdot dw \]

If only one output variable desired (letting \( W=X \)), integrate for all \( W \) to find \( Z \), do not integrate for \( z \) . . .

\[ g(z) = \int_{-\infty}^{\infty} g(z, w) \cdot dw = \int_{-\infty}^{\infty} f(\psi_1(z, w), \psi_2(z, w)) \cdot |J| \cdot dw \]
3.7. The Characteristic Function (Not covered in class or homework)

\[ \phi(u) = E[\exp(j \cdot u \cdot X)] \]

\[ \phi(u) = \int_{-\infty}^{\infty} f(x) \cdot \exp(j \cdot u \cdot x) \cdot dx \]

The inverse of the characteristic function is then defined as:

\[ f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(u) \cdot \exp(- j \cdot u \cdot x) \cdot du \]

Computing other moments is performed similarly, where:

\[ \frac{d^n[\phi(u)]}{du^n} = \int_{-\infty}^{\infty} f(x) \cdot (j \cdot x)^n \cdot \exp(j \cdot u \cdot x) \cdot dx \]

\[ \left. \frac{d^n[\phi(u)]}{du^n} \right|_{u=0} = \int_{-\infty}^{\infty} (j \cdot x)^n \cdot f(x) \cdot dx = j^n \cdot \int_{-\infty}^{\infty} x^n \cdot f(x) \cdot dx = j^n \cdot E[X^n] \]
4. Elements of Statistics

4.1. Introduction

4.2. Sampling Theory--The Sample Mean

Sample Mean

\[ \hat{X} = \frac{1}{n} \sum_{i=1}^{n} X_i , \]
where \( X_i \) are random variables with a pdf.

Variance of the sample mean

\[ \text{Var} [\hat{X}] = \frac{1}{n} \left( \frac{n}{n^2} \right) \left( \sum_{i=1}^{n} X_i \right)^2 \]

\[ = \frac{1}{n} \left( n \right) \left( \sum_{i=1}^{n} X_i \right)^2 \]

\[ = \frac{\sigma^2}{n} \left( \frac{N-n}{N-1} \right) \]

4.3. Sampling Theory--The Sample Variance

\[ S^2 = \frac{1}{n} \sum_{i=1}^{n} \left( X_i - \hat{X} \right)^2 \]

\[ = \frac{n}{n} \sum_{i=1}^{n} \left( X_i - \hat{X} \right)^2 \]

\[ = \frac{n-1}{n} \cdot \sigma^2 \]

\[ \text{to make it unbiased} \]

\[ \text{Bias} [S^2] = \frac{n}{n-1} \cdot \text{Bias} [\hat{S}^2] = \frac{n}{n-1} \cdot \frac{1}{n} \sum_{i=1}^{n} \left( X_i - \hat{X} \right)^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left( X_i - \hat{X} \right)^2 \]

When the population is not large, the biased and unbiased estimates become

\[ \text{Bias} [S^2] = \frac{n}{n-1} \cdot \text{Bias} [\hat{S}^2] = \frac{n-1}{n} \cdot \sigma^2 \]

\[ \text{Bias} [\hat{S}^2] = \frac{n}{n-1} \cdot \frac{n-1}{n} \cdot \text{Bias} [S^2] \]
4.4. Sampling Distributions and Confidence Intervals

Gaussian

\[ Z = \frac{\hat{X} - \bar{X}}{\sigma / \sqrt{n}} \]

Student’s t distribution

\[ T = \frac{\hat{X} - \bar{X}}{S / \sqrt{n - 1}} = \frac{\hat{X} - \bar{X}}{\tilde{S} / \sqrt{n}} \]

\[ \bar{X} - \frac{k \cdot \sigma}{\sqrt{n}} \leq \hat{X} \leq \bar{X} + \frac{k \cdot \sigma}{\sqrt{n}} \]

where \( \Gamma(\ ) \) is the gamma function.

\( \Gamma(k + 1) = k \cdot \Gamma(k) \) for any \( k \)

\( = k! \) for \( k \) an integer

and

\[ \Gamma\left( \frac{1}{2} \right) = \sqrt{\pi} \]

Confidence Intervals based on Gaussian and Student’s t distribution

Gaussian

\[ Z = \frac{\hat{X} - \bar{X}}{\sigma / \sqrt{n}} \]

\[ \bar{X} - \frac{k \cdot \sigma}{\sqrt{n}} \leq \hat{X} \leq \bar{X} + \frac{k \cdot \sigma}{\sqrt{n}} \]

Student’s t distribution

\[ T = \frac{\hat{X} - \bar{X}}{\tilde{S} / \sqrt{n - 1}} = \frac{\hat{X} - \bar{X}}{\sqrt{n}} \]

\[ \bar{X} - t \cdot \tilde{S} / \sqrt{n} \leq \hat{X} \leq \bar{X} + t \cdot \tilde{S} / \sqrt{n} \]

4.5. Hypothesis Testing

The null Hypothesis

Accept \( H_0 \): if the computed value “passes” the significance test.
Reject \( H_0 \): if the computed value “fails” the significance test.

One tail or two-tail testing

Using Confidence Interval value computations
4.6. Curve Fitting and Linear Regression

For \( y = a + bx \)

\[
a = \frac{1}{n} \left( \sum_{i=1}^{n} y_i - b \cdot \sum_{i=1}^{n} x_i \right)
\]

and

\[
b = \frac{n \sum_{i=1}^{n} (y_i \cdot x_i) - \left( \sum_{i=1}^{n} x_i \right) \cdot \left( \sum_{i=1}^{n} y_i \right)}{n \sum_{i=1}^{n} x_i^2 - \left( \sum_{i=1}^{n} x_i \right)^2}
\]

Or using 2\(^{nd}\) moment, correlation, and covariance values

\[
a = \frac{R_{YX} \cdot R_{XY} - \left( \frac{\hat{Y}}{\hat{X}} \right) \cdot R_{XY}}{C_{XX}}
\]

\[
b = \frac{R_{XY} - \left( \frac{\hat{Y}}{\hat{X}} \right) \cdot R_{XY}}{C_{XX}}
\]

4.7. Correlation Between Two Sets of Data

\[
\mu_X = E[X] = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i
\]

\[
E[X^2] = R_{XX} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2
\]

\[
\sigma_X^2 = C_{XX} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2 - \left( \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \right)^2 = R_{XX} - \mu_X^2
\]

\[
R_{XY} = E[X \cdot Y] = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \cdot y_i
\]

\[
C_{XY} = E[(X - \bar{X}) \cdot (Y - \bar{Y})] = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i \cdot y_i) - \bar{X} \cdot \bar{Y}
\]

\[
r = \rho_{XY} = \frac{E \left[ \frac{(X - \bar{X})}{\sigma_X} \cdot \frac{(Y - \bar{Y})}{\sigma_Y} \right]}{\sigma_X \cdot \sigma_Y} = \frac{C_{XY}}{\sigma_X \cdot \sigma_Y}
\]
5. Random Processes

5.1. Introduction
A random process is a collection of time functions and an associated probability description.
The entire collection of possible time functions is an ensemble, designated as \( \{x(t)\} \),
where one particular member of the ensemble, designated as \( x(t) \), is a sample function of the ensemble. In general only one sample function of a random process can be observed!

5.2. Continuous and Discrete Random Processes

5.3. Deterministic and Nondeterministic Random Processes
A nondeterministic random process is one where future values of the ensemble cannot be predicted from previously observed values.
A deterministic random process is one where one or more observed samples allow all future values of the sample function to be predicted (or pre-determined).

5.4. Stationary and Nonstationary Random Processes
If all marginal and joint density functions of a process do not depend upon the choice of the time origin, the process is said to be stationary.
Wide-Sense Stationary: the mean value of any random variable is independent of the choice of time, \( t \), and that the correlation of two random variables depends only upon the time difference between them.

5.5. Ergodic and Nonergodic Random Processes
The probability generated means and moments are equivalent to the time averaged means and moments.
A Process for Determining Stationarity and Ergodicity
   a) Find the mean and the 2\textsuperscript{nd} moment based on the probability
   b) Find the time sample mean and time sample 2\textsuperscript{nd} moment based on time averaging.
   c) If the means or 2\textsuperscript{nd} moments are functions of time … non-stationary
   d) If the time average mean and moments are not equal to the probabilistic mean and moments \textbf{or} if it is not stationary, then it is non ergodic.

5.6. Measurement of Process Parameters
The process of taking discrete time measurements of a continuous process in order to compute the desired statistics and probabilities.

5.7. Smoothing Data with a Moving Window Average
An example of using a “FIR filter” to smooth high frequency noise from a slowly time varying signal of interest. 


B.J. Bazuin, Spring 2015 15 of 36 ECE 3800
6. Correlation Functions

6.1 Introduction

The Autocorrelation Function

For a sample function defined by samples in time of a random process, how alike are the different samples?

Define: \( X_1 = X(t_1) \) and \( X_2 = X(t_2) \)

The autocorrelation is defined as:

\[
R_{XX}(t_1, t_2) = E[X_1 X_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx_1 dx_2 \cdot \{x_1 x_2, f(x_1, x_2)\}
\]

The above function is valid for all processes, stationary and non-stationary.

For WSS processes:

\[
R_{XX}(t_1, t_2) = E[X(t)X(t + \tau)] = R_{XX}(\tau)
\]

If the process is ergodic, the time average is equivalent to the probabilistic expectation, or

\[
\mathbb{R}_{XX}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t) \cdot x(t + \tau) \cdot dt = \langle x(t) \cdot x(t + \tau) \rangle
\]

and

\[
\mathbb{R}_{XX}(\tau) = R_{XX}(\tau)
\]

As a note for things you’ve been computing, the “zeroth lag of the autocorrelation” is

\[
R_{XX}(t_1, t_1) = R_{XX}(0) = E[X_1 X_1] = E[X_1^2] = \int_{-\infty}^{\infty} dx_1 \cdot \{x_1^2, f(x_1)\} = \sigma^2 - \mu^2
\]

\[
\mathbb{R}_{XX}(0) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)^2 \cdot dt = \langle x(t)^2 \rangle
\]

6.2 Example: Autocorrelation Function of a Binary Process
6.3. Properties of Autocorrelation Functions

1) \( R_{XX}(0) = E[X^2] = \overline{X^2} \) or \( R_{XX}(0) = \langle x(t)^2 \rangle \)

The mean squared value of the random process can be obtained by observing the zeroth lag of the autocorrelation function.

2) \( R_{XX}(\tau) = R_{XX}(-\tau) \)

The autocorrelation function is an even function in time. Only positive (or negative) needs to be computed for an ergodic WSS random process.

3) \( |R_{XX}(\tau)| \leq R_{XX}(0) \)

The autocorrelation function is a maximum at 0. For periodic functions, other values may equal the zeroth lag, but never be larger.

4) If \( X \) has a DC component, then \( R_{xx} \) has a constant factor.

\[
X(t) = \overline{X} + N(t) \\
R_{XX}(\tau) = \overline{X}^2 + R_{NN}(\tau)
\]

Note that the mean value can be computed from the autocorrelation function constants!

5) If \( X \) has a periodic component, then \( R_{xx} \) will also have a periodic component of the same period.

Think of:

\[
X(t) = A \cdot \cos(w \cdot t + \theta), \quad 0 \leq \theta \leq 2 \cdot \pi
\]

where \( A \) and \( w \) are known constants and theta is a uniform random variable.

\[
R_{xx}(\tau) = E[X(t)X(t+\tau)] = \frac{A^2}{2} \cdot \cos(w \cdot \tau)
\]

5b) For signals that are the sum of independent random variables, the autocorrelation is the sum of the individual autocorrelation functions.

\[
W(t) = X(t) + Y(t) \\
R_{WW}(\tau) = R_{XX}(\tau) + R_{YY}(\tau) + 2 \cdot \mu_X \cdot \mu_Y
\]

For non-zero mean functions, (let \( w, x, y \) be zero mean and \( W, X, Y \) have a mean)

\[
R_{WW}(\tau) = R_{ww}(\tau) + \mu_w^2 = R_{xx}(\tau) + \mu_X^2 + R_{yy}(\tau) + \mu_Y^2 + 2 \cdot \mu_X \cdot \mu_Y
\]

\[
R_{WW}(\tau) = R_{ww}(\tau) + \mu_w^2 = R_{xx}(\tau) + R_{yy}(\tau) + \mu_X^2 + 2 \cdot \mu_X \cdot \mu_Y + \mu_Y^2
\]

\[
R_{WW}(\tau) = R_{ww}(\tau) + \mu_w^2 = R_{xx}(\tau) + R_{yy}(\tau) + (\mu_X + \mu_Y)^2
\]
Then we have
\[
\mu^2 = (\mu_X + \mu_Y)^2 \\
R_{ww}(\tau) = R_{xx}(\tau) + R_{yy}(\tau)
\]

6) If X is ergodic and zero mean and has no periodic component, then we expect
\[
\lim_{|\tau| \to \infty} R_{XX}(\tau) = 0
\]

7) Autocorrelation functions cannot have an arbitrary shape. One way of specifying shapes permissible is in terms of the Fourier transform of the autocorrelation function. That is, if
\[
\mathcal{F}[R_{XX}(\tau)] = \int_{-\infty}^{\infty} R_{XX}(\tau) \cdot \exp(-j\omega \tau) \cdot d\tau
\]
then the restriction states that
\[
\mathcal{F}[R_{XX}(\tau)] \geq 0 \quad \text{for all } \omega
\]

Additional concept:
\[
X(t) = a \cdot N(t) \\
R_{XX}(\tau) = a^2 \cdot E[N(t) \cdot N(t + \tau)] = a^2 \cdot R_{NN}(\tau)
\]

6.4. Measurement of Autocorrelation Functions

We love to use time average for everything. For wide-sense stationary, ergodic random processes, time average are equivalent to statistical or probability based values.

\[
\mathcal{R}_{XX}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t) \cdot x(t + \tau) \cdot dt = \langle x(t) \cdot x(t + \tau) \rangle
\]

Using this fact, how can we use short-term time averages to generate auto- or cross-correlation functions?

An estimate of the autocorrelation is defined as:
\[
\hat{R}_{XX}(\tau) = \frac{1}{T - \tau} \int_{0}^{T-\tau} x(t) \cdot x(t + \tau) \cdot dt
\]

Note that the time average is performed across as much of the signal that is available after the time shift by tau.

For tau based on the available time step, k, with N equating to the available time interval, we have:
\[
\hat{R}_{XX}(k\Delta t) = \frac{1}{((N + 1)\Delta t) - (k\Delta t)} \sum_{i=0}^{N-k} x(i\Delta t) \cdot x(i\Delta t + k\Delta t) \cdot \Delta t
\]
\[
\hat{R}_{XX}(k\Delta t) = \hat{R}_{XX}(k) = \frac{1}{N+1-k} \sum_{i=0}^{N-k} x(i) \cdot x(i+k)
\]

In computing this autocorrelation, the initial weighting term approaches 1 when \(k=N\). At this point the entire summation consists of one point and is therefore a poor estimate of the autocorrelation. For useful results, \(k<<N\!\)!

As noted, the validity of each of the summed autocorrelation lags can and should be brought into question as \(k\) approaches \(N\). As a result, a biased estimate of the autocorrelation is commonly used. The biased estimate is defined as:

\[
\tilde{R}_{XX}(k) = \frac{1}{N+1} \sum_{i=0}^{N-k} x(i) \cdot x(i+k)
\]

Here, a constant weight instead of one based on the number of elements summed is used. This estimate has the property that the estimated autocorrelation should decrease as \(k\) approaches \(N\).

6.5. Examples of Autocorrelation Functions

Addition

\[W(t) = X(t) + Y(t)\]

\[R_{WW}(\tau) = E[W(t) \cdot W(t+\tau)] = E[(X(t) + Y(t)) \cdot (X(t+\tau) + Y(t+\tau))] = E[X(t) \cdot (X(t+\tau) + Y(t+\tau)) + Y(t) \cdot X(t+\tau) + Y(t) \cdot Y(t+\tau)] = R_{xx}(\tau) + R_{yy}(\tau) + E[x(t) \cdot y(t+\tau)] + E[y(t) \cdot x(t+\tau)]\]

If \(X\) and \(Y\) are independent

\[R_{WW}(\tau) = R_{xx}(\tau) + R_{yy}(\tau) + 2 \cdot \mu_x \cdot \mu_y\]

If either \(X\) or \(Y\) are zero mean (along with being independent)

\[R_{WW}(\tau) = R_{xx}(\tau) + R_{yy}(\tau)\]

Subtraction

\[W(t) = X(t) - Y(t)\]

\[R_{WW}(\tau) = E[W(t) \cdot W(t+\tau)] = E[(X(t) - Y(t)) \cdot (X(t+\tau) - Y(t+\tau))] = E[X(t) \cdot X(t+\tau) - X(t) \cdot Y(t+\tau) - Y(t) \cdot X(t+\tau) + Y(t) \cdot Y(t+\tau)] = R_{xx}(\tau) + R_{yy}(\tau) - E[x(t) \cdot y(t+\tau)] - E[y(t) \cdot x(t+\tau)]\]

If \(X\) and \(Y\) are independent

\[R_{WW}(\tau) = R_{xx}(\tau) + R_{yy}(\tau) - 2 \cdot \mu_x \cdot \mu_y\]

If either \(X\) or \(Y\) are zero mean (along with being independent)

\[R_{WW}(\tau) = R_{xx}(\tau) + R_{yy}(\tau)\]
6.6. Crosscorrelation Functions

For a two sample function defined by samples in time of two random processes, how alike are the different samples?

Define: \( X_1 = X(t_1) \) and \( Y_2 = Y(t_2) \)

The cross-correlation is defined as:

\[
R_{XY}(t_1, t_2) = E[X_1 Y_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 y_2 f(x_1, y_2) dx_1 dy_2
\]

\[
R_{YX}(t_1, t_2) = E[Y_1 X_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 x_2 f(y_1, x_2) dx_2 dy_2
\]

The above function is valid for all processes, jointly stationary and non-stationary.

For jointly WSS processes:

\[
R_{XY}(t_1, t_2) = E[X(t)Y(t + \tau)] = R_{XY}(\tau)
\]

\[
R_{YX}(t_1, t_2) = E[Y(t)X(t + \tau)] = R_{YX}(\tau)
\]

Note: the order of the subscripts is important for cross-correlation!

If the processes are jointly ergodic, the time average is equivalent to the probabilistic expectation, or

\[
\mathfrak{R}_{XY}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t) \cdot y(t+\tau) \cdot dt = \langle x(t) \cdot y(t+\tau) \rangle
\]

\[
\mathfrak{R}_{YX}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} y(t) \cdot x(t+\tau) \cdot dt = \langle y(t) \cdot x(t+\tau) \rangle
\]

and

\[
\mathfrak{R}_{XY}(\tau) = R_{XY}(\tau)
\]

\[
\mathfrak{R}_{YX}(\tau) = R_{YX}(\tau)
\]

6.7. Properties of Crosscorrelation Functions

1) The properties of the zoreth lag have no particular significance and do not represent mean-square values. It is true that the “ordered” crosscorrelations must be equal at 0.

\[
R_{XY}(0) = R_{YX}(0) \quad \text{or} \quad \mathfrak{R}_{XY}(0) = \mathfrak{R}_{YX}(0)
\]

2) Crosscorrelation functions are not generally even functions. However, there is an antisymmetry to the ordered crosscorrelations:

\[
R_{XY}(\tau) = R_{YX}(-\tau)
\]
For
\[ R_{XY}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t) \cdot y(t + \tau) \cdot dt = \langle x(t) \cdot y(t + \tau) \rangle \]

Substitute \( t + \tau = \eta \)
\[ R_{XY}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(\eta - \tau) \cdot y(\eta) \cdot d\eta = \langle x(\eta - \tau) \cdot y(\eta) \rangle \]
\[ R_{XY}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} y(\eta) \cdot x(\eta - \tau) \cdot d\eta = \langle y(\eta) \cdot x(\eta - \tau) \rangle = R_{YX}(\tau) \]

3) \quad The crosscorrelation does not necessarily have its maximum at the zeroth lag. This makes sense if you are correlating a signal with a time delayed version of itself. The crosscorrelation should be a maximum when the lag equals the time delay!

It can be shown however that \[ |R_{XY}(\tau)| \leq \sqrt{R_{XX}(0) \cdot R_{YY}(0)} \]
As a note, the crosscorrelation may not achieve the maximum anywhere …

4) \quad If \( X \) and \( Y \) are statistically independent, then the ordering is not important
\[ R_{XY}(\tau) = E[X(t) \cdot Y(t + \tau)] = E[X(t)] \cdot E[Y(t + \tau)] = \overline{X} \cdot \overline{Y} \]
and
\[ R_{XY}(\tau) = \overline{X} \cdot \overline{Y} = R_{YX}(\tau) \]

5) \quad If \( X \) is a stationary random process and is differentiable with respect to time, the crosscorrelation of the signal and it’s derivative is given by
\[ R_{XX}(\tau) = \frac{dR_{XX}(\tau)}{d\tau} \]
Defining derivation as a limit:
\[ \dot{X}(\tau) = \lim_{e \to 0} \frac{X(t + e) - X(t)}{e} \]
and the crosscorrelation
\[ R_{XX}(\tau) = E[X(t) \cdot \dot{X}(t + \tau)] = E \left[ X(t) \cdot \left( \lim_{e \to 0} \frac{X(t + \tau + e) - X(t + \tau)}{e} \right) \right] \]
\[ R_{XX}(\tau) = \lim_{e \to 0} \frac{E[X(t) \cdot X(t + \tau + e)]}{e} - E[X(t) \cdot X(t + \tau)] \]
\[ R_{XX}(\tau) = \lim_{e \to 0} \frac{E[X(t) \cdot X(t + \tau + e)]}{e} - E[X(t) \cdot X(t + \tau)] \]
\[ R_{XX}(\tau) = \lim_{e \to 0} \frac{R_{XX}(\tau + e) - R_{XX}(\tau)}{e} \]
\[ R_{XX}(\tau) = \frac{dR_{XX}(\tau)}{d\tau} \]
Similarly,

\[ R_{XX}(\tau) = -\frac{d^2 R_{XX}(\tau)}{d\tau^2} \]

6.8. Examples and Applications of Crosscorrelation Functions

RF signal receiver mixing and down conversion, A, B, and theta are R.V with \( \theta \) a uniformly distributed random variable on \([0, 2\pi]\), that is also independent of either A or B.

\[
X(t) = A \cdot \cos(2\pi f_0 t + \theta) + B \cdot \sin(2\pi f_0 t + \theta) \\
Y(t) = \sin(2\pi f_0 t + \theta + \psi)
\]

Forming the cross-correlation

\[
R_{XY}(t, t + \tau) = E \left[ \left( A \cdot \cos(2\pi f_0 t + \theta) + B \cdot \sin(2\pi f_0 t + \theta) \right) \cdot \left( \sin(2\pi f_0 (t + \tau) + \theta + \psi) \right) \right]
\]

\[
R_{XY}(t, t + \tau) = E \left[ \frac{A}{2} \cdot \cos(2\pi f_0 \cdot \tau + \psi) + \frac{B}{2} \cdot \cos(2\pi f_0 \cdot (2t + \tau) + 2\theta + \psi) \right]
\]

The mixing phase can modify the output

\[
R_{XY}(\tau) = E \left[ \frac{A}{2} \cdot \sin(2\pi f_0 \cdot \tau + \psi) + \frac{B}{2} \cdot \cos(2\pi f_0 \cdot \tau + \psi) \right]
\]

6.9. Correlation Matrices For Sampled Functions

For \( \psi = 0 \), \( \psi = n \cdot \pi \) or the cross correlation becomes

\[
R_{XY}(\tau) = E \left[ \frac{A}{2} \cdot \sin(2\pi f_0 \cdot \tau) + \frac{B}{2} \cdot \cos(2\pi f_0 \cdot \tau) \right]
\]

For \( \psi = \pm \frac{\pi}{2} \), \( \psi = \pm \frac{\pi}{2} + n \cdot \pi \) or the cross correlation becomes

\[
R_{XY}(\tau) = E \left[ \frac{A}{2} \cdot \cos(2\pi f_0 \cdot \tau) - \frac{B}{2} \cdot \sin(2\pi f_0 \cdot \tau) \right]
7. Spectral Density

7.1. Introduction

7.2. Relation of Spectral Density to the Fourier Transform

This function is also defined as the spectral density function (or power-spectral density) and is defined for both \( f \) and \( w \) as:

\[
S_{YY}(w) = \lim_{T \to \infty} \frac{E[|Y(w)|^2]}{2T} \quad \text{or} \quad S_{YY}(f) = \lim_{T \to \infty} \frac{E[|Y(f)|^2]}{2T}
\]

The 2\(^{nd}\) moment based on the spectral densities is defined, as:

\[
\bar{Y}^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{YY}(w) \cdot dw \quad \text{and} \quad \bar{Y}^2 = \int_{-\infty}^{\infty} S_{YY}(f) \cdot df
\]

Note: The result is a power spectral density (in Watts/Hz), not a voltage spectrum as (in V/Hz) that you would normally compute for a Fourier transform.

Wiener-Khinchine relation

For WSS random processes, the autocorrelation function is time based and, for ergodic processes, describes all sample functions in the ensemble! In these cases the Wiener-Khinchine relations is valid that allows us to perform the following.

\[
S_{XX}(w) = \Im[R_{XX}(\tau)] = \int_{-\infty}^{\infty} E[X(t) \cdot X(t+\tau)] \cdot \exp(-i\omega \tau) \cdot d\tau
\]

For an ergodic process, we can use time-based processing to arrive at an equivalent result …

\[
\Re[R_{XX}(\tau)] = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t) \cdot x(t+\tau) \cdot dt = \langle x(t) \cdot x(t+\tau) \rangle
\]

We can define a power spectral density for the ensemble as:

\[
S_{XX}(w) = \Im[R_{XX}(\tau)] = \int_{-\infty}^{\infty} R_{XX}(\tau) \cdot \exp(-i\omega \tau) \cdot d\tau
\]
Based on this definition, we also have

\[
S_{XX}(w) = \mathcal{Z}[R_{XX}(\tau)] \quad R_{XX}(\tau) = \mathcal{Z}^{-1}[S_{XX}(w)]
\]

\[
R_{XX}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(w) \cdot \exp(\jmath \omega t) \cdot dw
\]

7.3. Properties of Spectral Density

The power spectral density as a function is always

- real,
- positive,
- and an even function in \( w \).

As an even function, the PSD may be expected to have a polynomial form as:

\[
S_{XX}(w) = S_0 \frac{w^{2n} + a_{2n-2}w^{2n-2} + a_{2n-4}w^{2n-4} + \cdots + a_2w^2 + a_0}{w^{2m} + b_{2m-2}w^{2m-2} + b_{2m-4}w^{2m-4} + \cdots + b_2w^2 + b_0}
\]

where \( m > n \).

Notice the squared terms, any odd power would define an anti-symmetric element that, by definition and proof, can not exist!

Finite property in frequency. The Power Spectral Density must also approach zero as \( w \) approached infinity …. Therefore,

\[
S_{XX}(w \to \infty) = \lim_{w \to \infty} S_0 \frac{w^{2n} + a_{2n-2}w^{2n-2} + \cdots + a_2w^2 + a_0}{w^{2m} + b_{2m-2}w^{2m-2} + \cdots + b_2w^2 + b_0} \Rightarrow \lim_{w \to \infty} S_0 \frac{w^{2n}}{w^{2m}} = \lim_{w \to \infty} S_0 \frac{1}{w^{2(m-n)}} = 0
\]

For \( m > n \), the condition will be met.

7.4. Spectral Density and the Complex Frequency Plane

7.5. Mean-Square Values From Spectral Density

\[
R_{XX}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(w) \cdot \exp(\jmath \omega t) \cdot dw
\]

\[
R_{XX}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(w) \cdot dw
\]

Works for discrete time as well, but watch the integral bounds.
7.6. Relation of Spectral Density to the Autocorrelation Function

The power spectral density as a function is always

- real,
- positive,
- and an even function in w/f.

You can convert between the domains using:

The Fourier Transform in w

\[ S_{XX}(w) = \int_{-\infty}^{\infty} R_{XX}(\tau) \cdot \exp(-iw\tau) \cdot d\tau \]

\[ R_{XX}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(w) \cdot \exp(iwt) \cdot dw \]

The Fourier Transform in f

\[ S_{XX}(f) = \int_{-\infty}^{\infty} R_{XX}(\tau) \cdot \exp(-i2\pi f\tau) \cdot d\tau \]

\[ R_{XX}(t) = \int_{-\infty}^{\infty} S_{XX}(f) \cdot \exp(i2\pi ft) \cdot df \]

The 2-sided Laplace Transform

\[ S_{XX}(s) = \int_{-\infty}^{\infty} R_{XX}(\tau) \cdot \exp(-s\tau) \cdot d\tau \]

\[ R_{XX}(t) = \frac{1}{j2\pi} \int_{-j\infty}^{j\infty} S_{XX}(s) \cdot \exp(st) \cdot ds \]

**Deriving the Mean-Square Values from the Power Spectral Density**

The mean squared value of a random process is equal to the 0th lag of the autocorrelation

\[ E[X^2] = R_{XX}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(w) \cdot \exp(iw \cdot 0) \cdot dw = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(w) \cdot dw \]

\[ E[X^2] = R_{XX}(0) = \int_{-\infty}^{\infty} S_{XX}(f) \cdot \exp(i2\pi f \cdot 0) \cdot df = \int_{-\infty}^{\infty} S_{XX}(f) \cdot df \]

Therefore, to find the second moment, integrate the PSD over all frequencies.
7.7. White Noise

Noise is inherently defined as a random process. You may be familiar with “thermal” noise, based on the energy of an atom and the mean-free path that it can travel.

- As a random process, whenever “white noise” is measured, the values are uncorrelated with each other, not matter how close together the samples are taken in time.
- Further, we envision “white noise” as containing all spectral content, with no explicit peaks or valleys in the power spectral density.

As a result, we define “White Noise” as

\[
R_{XX}(\tau) = S_0 \cdot \delta(t)
\]

\[
S_{XX}(w) = S_0 = \frac{N_0}{2}
\]

This is an approximation or simplification because the area of the power spectral density is infinite!

For typical applications, we are interested in Band-Limited White Noise where

\[
S_{XX}(w) = \begin{cases} 
S_0 = \frac{N_0}{2} & |f| \leq W \\
0 & W < |f|
\end{cases}
\]

The equivalent noise power is then:

\[
E[X^2] = R_{XX}(0) = \int_{-W}^{W} S_0 \cdot dw = 2 \cdot W \cdot S_0 = 2 \cdot W \cdot \frac{N_0}{2} = N_0 \cdot W
\]

7.8. Cross-Spectral Density

The Fourier Transform in \( w \)

\[
S_{XY}(w) = \int_{-\infty}^{\infty} R_{XY}(\tau) \cdot \exp(-iw\tau) \cdot d\tau \quad \text{and} \quad S_{YX}(w) = \int_{-\infty}^{\infty} R_{YX}(\tau) \cdot \exp(-iw\tau) \cdot d\tau
\]

\[
R_{XY}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XY}(w) \cdot \exp(iwt) \cdot dw \quad \text{and} \quad R_{YX}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{YX}(w) \cdot \exp(iwt) \cdot dw
\]

Properties of the functions

\[ S_{XY}(w) = \text{conj}(S_{YX}(w)) \]

Since the cross-correlation is real,

- the real portion of the spectrum is even
- the imaginary portion of the spectrum is odd
**Generic Example of a Correlation and Power Spectral Density**

\[ X(t) = A + B \cdot \sin(2\pi \cdot f_1 \cdot t + \theta_1) + C \cdot \cos(2\pi \cdot f_2 \cdot t + \theta_2) \]

where the phase angles are uniformly distributed R.V from 0 to 2\( \pi \).

\[
R_{xx}(\tau) = E[X(t)X(t+\tau)]
\]

\[
= E[(X(t) = A + B \cdot \sin(2\pi \cdot f_1 \cdot (t + \tau) + \theta_1) + C \cdot \cos(2\pi \cdot f_2 \cdot (t + \tau) + \theta_2)) \cdot (X(t) = A + B \cdot \sin(2\pi \cdot f_1 \cdot (t + \tau) + \theta_1) + C \cdot \cos(2\pi \cdot f_2 \cdot (t + \tau) + \theta_2))]
\]

\[
R_{xx}(\tau) = E\left[ A^2 + AB \cdot \sin(2\pi \cdot f_1 \cdot t + \theta_1) \cdot \sin(2\pi \cdot f_1 \cdot (t + \tau) + \theta_1) + B^2 \cdot \sin(2\pi \cdot f_1 \cdot t + \theta_1) \cdot \sin(2\pi \cdot f_1 \cdot (t + \tau) + \theta_1) + AC \cdot \cos(2\pi \cdot f_2 \cdot t + \theta_2) \cdot \cos(2\pi \cdot f_2 \cdot (t + \tau) + \theta_2) + C \cdot \cos(2\pi \cdot f_2 \cdot t + \theta_2) \cdot \cos(2\pi \cdot f_2 \cdot (t + \tau) + \theta_2) + BC \cdot \sin(2\pi \cdot f_1 \cdot t + \theta_1) \cdot \sin(2\pi \cdot f_1 \cdot (t + \tau) + \theta_1) + BC \cdot \cos(2\pi \cdot f_2 \cdot t + \theta_2) \cdot \sin(2\pi \cdot f_1 \cdot (t + \tau) + \theta_1) \right]
\]

Assuming all possible R.V.s are independent and zero mean (cross products zero)

\[
R_{xx}(\tau) = E[A^2] + E[B^2] \cdot E\left[ \frac{1}{2} \cdot \cos(2\pi \cdot f_1 \cdot (2t + \tau) + 2\theta_1) \right] + E[C^2] \cdot E\left[ \frac{1}{2} \cdot \cos(2\pi \cdot f_2 \cdot (2t + \tau) + 2\theta_2) \right]
\]

which lead to (assuming A,B, and C take on values)

\[
R_{xx}(\tau) = A^2 + B^2 \cdot \cos(2\pi \cdot f_1 \cdot \tau) + C^2 \cdot \cos(2\pi \cdot f_2 \cdot \tau)
\]

Forming the PSD

And then taking the Fourier transform

\[
S_{xx}(f) = A^2 \cdot \delta(f) + \frac{B^2}{2} \cdot \left( \frac{1}{2} \cdot \delta(f + f_1) + \frac{1}{2} \cdot \delta(f - f_1) \right) + \frac{C^2}{2} \cdot \left( \frac{1}{2} \cdot \delta(f + f_2) + \frac{1}{2} \cdot \delta(f - f_2) \right)
\]

\[
S_{xx}(f) = A^2 \cdot \delta(f) + \frac{B^2}{4} \cdot (\delta(f + f_1) + \delta(f - f_1)) + \frac{C^2}{4} \cdot (\delta(f + f_2) + \delta(f - f_2))
\]
7.9. Autocorrelation Function Estimate of Spectral Density

7.10. Periodogram Estimate of Spectral Density

7.11. Examples and Applications of Spectral Density

Determine the autocorrelation of the binary sequence.

\[ x(t) = \sum_{k=-\infty}^{\infty} A_k \cdot p(t - t_0 - k \cdot T) \]

\[ x(t) = p(t) \ast \sum_{k=-\infty}^{\infty} A_k \cdot \delta(t - t_0 - k \cdot T) \]

Determine the auto correlation of the discrete time sequence (leaving out the pulse for now)

\[ y(t) = \sum_{k=-\infty}^{\infty} A_k \cdot \delta(t - t_0 - k \cdot T) \]

\[ E[y(t) \cdot y(t + \tau)] = E\left[\sum_{k=-\infty}^{\infty} A_k \cdot \delta(t - t_0 - k \cdot T) \cdot \sum_{j=-\infty}^{\infty} A_j \cdot \delta(t + \tau - t_0 - j \cdot T)\right] \]

\[ R_{yy}(\tau) = E[y(t) \cdot y(t + \tau)] = E\left[\sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} A_k \cdot A_j \cdot \delta(t - t_0 - k \cdot T) \cdot \delta(t + \tau - t_0 - j \cdot T)\right] \]

\[ R_{yy}(\tau) = \sum_{k=-\infty}^{\infty} E[A_k^2] \cdot E[\delta(t - t_0 - k \cdot T) \cdot \delta(t + \tau - t_0 - k \cdot T)] \]

\[ R_{yy}(\tau) = \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} E[A_k \cdot A_j] \cdot E[\delta(t - t_0 - k \cdot T) \cdot \delta(t + \tau - t_0 - j \cdot T)] \]

\[ R_{yy}(\tau) = E[A_k^2] \cdot \frac{1}{T} \cdot \delta(\tau) + E[A_k]^2 \cdot \frac{1}{T} \cdot \sum_{m=-\infty}^{\infty} \delta(\tau - m \cdot T) \]

\[ R_{yy}(\tau) = [E[A_k^2] - E[A_k]^2] \cdot \frac{1}{T} \cdot \delta(\tau) + E[A_k]^2 \cdot \frac{1}{T} \cdot \sum_{m=-\infty}^{\infty} \delta(\tau - m \cdot T) \]

\[ R_{yy}(\tau) = \sigma_A^2 \cdot \frac{1}{T} \cdot \delta(\tau) + \mu_A^2 \cdot \frac{1}{T} \cdot \left\{ \sum_{m=-\infty}^{\infty} \delta(\tau - m \cdot T) \right\} \]

\[ S_{yy}(f) = \sigma_A^2 \cdot \frac{1}{T} + \mu_A^2 \cdot \frac{1}{T} \cdot \left\{ \frac{1}{T} \cdot \sum_{m=-\infty}^{\infty} \delta\left(f - \frac{m}{T}\right) \right\} \]

From here, it can be shown that
\[ S_{xx}(f) = |P(f)|^2 \cdot S_{yy}(f) \]

\[ S_{xx}(f) = |P(f)|^2 \cdot \left[ \sigma_A^2 \cdot \frac{1}{T} + \mu_A^2 \cdot \frac{1}{T^2} \cdot \sum_{m=-\infty}^{\infty} \delta \left( f - \frac{m}{T} \right) \right] \]

\[ S_{xx}(f) = |P(f)|^2 \cdot \sigma_A^2 + |P(f)|^2 \cdot \mu_A^2 \cdot \sum_{m=-\infty}^{\infty} \delta \left( f - \frac{m}{T} \right) \]

This is a magnitude scaled version of the power spectral density of the pulse shape and numerous impulse responses with magnitudes shaped by the pulse at regular frequency intervals based on the signal periodicity.

The result was picture in the textbook as …

![Figure 7-2 Random amplitude pulse sequence.](image)

![Figure 7-3 Spectral density for rectangular pulse sequence with random amplitudes.](image)
8. Response of Linear Systems to Random Inputs

8.1. Introduction
Linear transformation of signals: convolution in the time domain
\[ y(t) = h(t) * x(t) \quad Y(s) = H(s) \cdot X(s) \]

The convolution Integrals (applying a causal filter)
\[ y(t) = \int_{0}^{\infty} x(t - \lambda) \cdot h(\lambda) \cdot d\lambda \quad \text{or} \quad y(t) = \int_{-\infty}^{t} h(t - \lambda) \cdot x(\lambda) \cdot d\lambda \]

8.2. Analysis in the Time Domain

8.3. Mean and Mean-Square Value of System Output
The Mean Value at a System Output
\[ E[Y(t)] = E\left[ \int_{0}^{\infty} X(t - \lambda) \cdot h(\lambda) \cdot d\lambda \right] \]
\[ E[Y(t)] = \int_{0}^{\infty} E[X(t - \lambda)] \cdot h(\lambda) \cdot d\lambda \]

For a wide-sense stationary process, this result in
\[ E[Y(t)] = \int_{0}^{\infty} E[X] \cdot h(\lambda) \cdot d\lambda = E[X] \int_{0}^{\infty} h(\lambda) \cdot d\lambda \]

The coherent gain of a filter is defined as:
\[ h_{gain} = \int_{0}^{\infty} h(t) \cdot dt \]

Therefore,
\[ E[Y(t)] = E[X] \cdot h_{gain} \]
The Mean Square Value at a System Output

\[ E[Y(t)^2] = E \left[ \int_0^\infty X(t - \lambda) \cdot h(\lambda) \cdot d\lambda \right]^2 \]

\[ E[Y(t)^2] = \int_0^\infty h(\lambda) \cdot [R_{XX}(\lambda) \ast h(\lambda)] \cdot d\lambda \]

Therefore, take the convolution of the autocorrelation function and then sum the filter-weighted result from 0 to infinity.

8.4. Autocorrelation Function of System Output

\[ R_{YY}(\tau) = E[Y(t) \cdot Y(t + \tau)] = E[(h(t) \ast x(t)) \cdot (h(t + \tau) \ast x(t + \tau))] \]

\[ R_{YY}(\tau) = E \left[ \int_0^\infty x(t - \lambda_1) \cdot h(\lambda_1) \cdot d\lambda_1 \right] \cdot \left[ \int_0^\infty x(t + \tau - \lambda_2) \cdot h(\lambda_2) \cdot d\lambda_2 \right] \]

\[ R_{YY}(\tau) = \int_0^\infty h(\lambda_1) \cdot \left[ \int_0^\infty R_{XX}(\tau + \lambda_1 - \lambda_2) \cdot h(\lambda_2) \cdot d\lambda_2 \right] \cdot d\lambda_1 \]

\[ R_{YY}(\tau) = \int_0^\infty h(\lambda_1) \cdot [R_{XX}(\tau + \lambda_1) \ast h(\tau + \lambda_1)] \cdot d\lambda_1 \]

\[ R_{YY}(\tau) = \int_{-\infty}^0 h(-\lambda_1) \cdot [R_{XX}(\tau - \lambda_1) \ast h(\tau - \lambda_1)] \cdot d\lambda_1 \]

Notice that first, you convolve the input autocorrelation function with the filter function and then you convolve the result with a time inversed version of the filter!

8.5. Crosscorrelation between Input and Output

\[ R_{XY}(\tau) = E[X(t) \cdot Y(t + \tau)] = E[x(t) \cdot (h(t + \tau) \ast x(t + \tau))] \]

\[ R_{XY}(\tau) = \int_0^\infty R_{XX}(\tau - \lambda_1) \cdot h(\lambda_1) \cdot d\lambda_1 \]

This is the convolution of the autocorrelation with the filter.

What about the other Autocorrelation?

\[ R_{YY}(\tau) = E[Y(t) \cdot X(t + \tau)] = E[(h(t) \ast x(t)) \cdot x(t + \tau)] \]

\[ R_{YY}(\tau) = \int_0^\infty R_{XX}(\tau + \lambda_1) \cdot h(\lambda_1) \cdot d\lambda_1 \]

This is the correlation of the autocorrelation with the filter, inherently different than the previous, but equal at \( \tau=0 \).
8.6. Example of Time-Domain System Analysis

8.7. Analysis in the Frequency Domain

8.8. Spectral Density at the System Output

\[ S_{YY}(w) = \mathfrak{Re}[R_{YY}(\tau)] = \int_{-\infty}^{\infty} \int_{0}^{\infty} d\lambda_1 \left[ \int_{0}^{\infty} d\lambda_2 \cdot h(\lambda_1) \cdot h(\lambda_2) \cdot R_{XX}(\tau + \lambda_1 - \lambda_2) \right] \cdot \exp(-i\omega \tau) \cdot d\tau \]

Therefore

\[ S_{YY}(w) = \mathfrak{Re}[R_{YY}(\tau)] = S_{XX}(w) \cdot H(w) \cdot H(-w) \]

8.9. Cross-Spectral Densities between Input and Output

\[ S_{XY}(s) = S_{XX}(s) \cdot H(s) \]

\[ S_{YX}(s) = S_{XX}(s) \cdot H(-s) \]

8.10. Examples of Frequency-Domain Analysis

Noise in a linear feedback system loop.

\[ Y(s) = \frac{A}{s \cdot (s+1)} \cdot [X(s) - Y(s)] + N(s) \]

\[ Y(s) \cdot \left[ 1 + \frac{A}{s \cdot (s+1)} \right] = \frac{A}{s \cdot (s+1)} \cdot X(s) + N(s) \]

\[ Y(s) \cdot \left[ \frac{s^2 + s + A}{s \cdot (s+1)} \right] = \frac{A}{s \cdot (s+1)} \cdot X(s) + N(s) \]

\[ Y(s) = \frac{A}{s^2 + s + A} \cdot X(s) + \frac{s^2 + s}{s^2 + s + A} \cdot N(s) \]

There are effectively two filters, one applied to X and a second apply to N.
\[ H_x(s) = \frac{A}{s^2 + s + A} \quad \text{and} \quad H_n(s) = \frac{s^2 + s}{s^2 + s + A} \]

\[ Y(s) = H_x(s) \cdot X(s) + H_n(s) \cdot N(s) \]

Generic definition of output Power Spectral Density:

\[ S_{yy}(w) = |H_x(w)|^2 \cdot S_{XX}(w) + |H_n(w)|^2 \cdot S_{NN}(w) \]

8.11. Numerical Computation of System Output

9. **Optimum Linear Systems**

9.1. Introduction

9.2. Criteria of Optimality

9.3. Restrictions on the Optimum System

9.4. Optimization by Parameter Adjustment

9.5. Systems That Maximize Signal-to-Noise Ratio

SNR is defined as

\[
\frac{P_{\text{Signal}}}{P_{\text{Noise}}} = \frac{E[s(t)^2]}{N_0 \cdot B_{EQ}}
\]

For a linear system, we have:

\[ s_o(t) + n_o(t) = \int_0^\infty h(\lambda) \cdot [s(t - \lambda) + n(t - \lambda)] \, d\lambda \]

The output SNR can be described as

\[
\text{SNR}_{out} = \frac{P_{\text{Signal}}}{P_{\text{Noise}}} = \frac{E[s_o(t)^2]}{E[n_o(t)^2]} = \frac{E[s_o(t)^2]}{N_o \cdot B_{EQ}}
\]

or substituting

\[
\text{SNR}_{out} = \frac{E \left[ \int_0^\infty h(\lambda) \cdot s(t - \lambda) \, d\lambda \right]^2}{N_o \cdot \frac{1}{2} \int_0^\infty h(t)^2 \, dt}
\]

To achieve the maximum SNR, the equality condition of Schwartz’s Inequality must hold, or
\[
\left( \int_{0}^{\infty} h(\lambda) \cdot s(t - \lambda) \cdot d\lambda \right)^2 = \left( \int_{0}^{\infty} h(\lambda)^2 \cdot d\lambda \right) \cdot \left( \int_{0}^{\infty} s(t - \lambda)^2 \cdot d\lambda \right)
\]

This condition can be met for \( h(\lambda) = K \cdot s(t - \lambda) \cdot u(\lambda) \)

where K is an arbitrary gain constant.

The desired impulse response is simply the time inverse of the signal waveform at time t, a fixed moment chosen for optimality. If this is done, the maximum filter power (with K=1) can be computed as

\[
\int_{0}^{\infty} h(\lambda)^2 \cdot d\lambda = \int_{0}^{\infty} s(t - \lambda)^2 \cdot d\lambda = \int_{-\infty}^{t} s(\tau)^2 \cdot d\tau = \varepsilon(t)
\]

\[
\max[SNR_out] = \frac{2}{N_o} \cdot \varepsilon(t)
\]

This filter concept is called a **matched filter**.

### 9.6. Systems That Minimize Mean-Square Error

The error function \( Err(t) = X(t) - Y(t) \)

where \( Y(t) = \int_{0}^{\infty} h(\lambda) \cdot [X(t - \lambda) + N(t - \lambda)] \cdot d\lambda \)

Computing the error power:

\[
E[Err^2] = \frac{1}{j2\pi} \cdot \int_{-\infty}^{\infty} \left\{ S_{XX}(s) \cdot (1 - H(s)) \cdot (1 - H(-s)) + S_{NN}(s) \cdot H(s) \cdot H(-s) \right\} \cdot ds
\]

\[
E[Err^2] = \frac{1}{j2\pi} \cdot \int_{-\infty}^{\infty} \left\{ [S_{XX}(s) + S_{NN}(s)] \cdot H(s) \cdot H(-s) - S_{XX}(s) \cdot H(s) - S_{XX}(s) \cdot H(-s) + S_{XX}(s) \right\} \cdot ds
\]

Defining \( F_C(s) \cdot F_C(-s) = S_{XX}(s) + S_{NN}(s) \)

Step 1:

So, let’s see about some interpretations. First, let \( H_1(s) = \frac{1}{F_C(s)} \), which should be causal.

Note that for this filter \( [S_{XX}(s) + S_{NN}(s)] \cdot \frac{1}{F_C(s)} \cdot \frac{1}{F_C(-s)} = 1 \)

This is called a whitening filter as it forces the signal plus noise PSD to unity (white noise).
Step 2:
Letting \[ H(s) = H_1(s) \cdot H_2(s) = \frac{H_2(s)}{F_C(s)} \]

\[
E[Err^2] = \frac{1}{j2\pi} \int_{-j\infty}^{j\infty} \left\{ \frac{F_C(s)}{F_C(s)} \cdot \frac{H_2(s)}{F_C(-s)} \cdot \left[ \frac{s_{XX}(s)}{F_C(-s)} \right] \cdot \left[ \frac{H_2(-s)}{F_C(-s)} - \frac{s_{XX}(s)}{F_C(-s)} \right] \right\} \cdot ds
\]

Minimizing the terms containing \( H_2(s) \), we must focus on \( H_2(s) - \frac{s_{XX}(s)}{F_C(-s)} \) and \( H_2(-s) - \frac{s_{XX}(s)}{F_C(-s)} \).

Letting \( H_2 \) be defined for the appropriate Left or Right half-plane poles

Let \( H_2(s) = \left[ \frac{s_{XX}(s)}{F_C(-s)} \right]_{LHP} \) and \( H_2(-s) = \left[ \frac{s_{XX}(s)}{F_C(s)} \right]_{RHP} \)

The composite filter is then

\[ H(s) = H_1(s) \cdot H_2(s) = \frac{1}{F_C(s)} \cdot \left[ \frac{s_{XX}(s)}{F_C(-s)} \right]_{LHP} \]

This solution is often called a **Wiener Filter** and is widely applied when the signal and noise statistics are known a-priori!
Appendices

A. Mathematical Tables

A.1. Trigonometric Identities
A.2. Indefinite Integrals
A.3. Definite Integrals
A.4. Fourier Transform Operations
A.5. Fourier Transforms
A.6. One-Sided Laplace Transforms

B. Frequently Encountered Probability Distributions

B.1. Discrete Probability Functions
B.2. Continuous Distributions

C. Binomial Coefficients

D. Normal Probability Distribution Function

E. The Q-Function

F. Student's t Distribution Function

G. Computer Computations

H. Table of Correlation Function--Spectral Density Pairs

I. Contour Integration