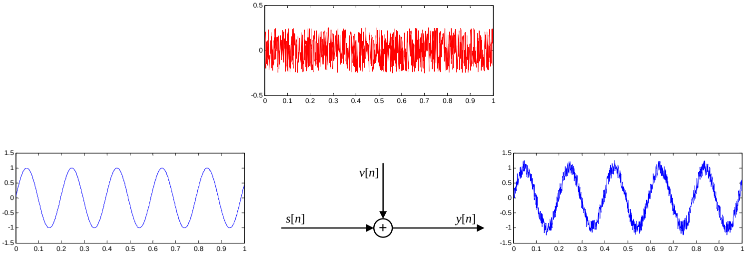
ELEC 3224 — Random Processes

Professor Eric Rogers

Department of Electronics and Computer Science
University of Southampton
etar@ecs.soton.ac.uk
Office: Building 1, Room 2037

Introduction

- Many real world signals yield small amplitudes (e.g. EGG order of $10^{-6}v$), and amplification is required prior to further processing;
- Generally such signals are corrupted by noise. To proceed with analysis in such cases the starting point is often to assume an additive noise model:



Disturbance Rejection

Consider the case when we want to estimate a constant c from measurements corrupted by an additive sequence $d[n]$, i.e.

$$x[n] = c + d[n]$$

Then a common problem is to compare the performance of two (or more) filters in such an application. Suppose that we know the frequency of the disturbance (strong assumption). Then a basic starting point is to evaluate the gains of the filters at this frequency and select the one with the lowest value.
In general, we need a different setting in which to deal with noise.

Introduction

A signal is termed **non-deterministic or random** if its value at some future time **cannot be predicted exactly.**
Often we need to deduce features of the signal (or process) $x(t)$ **based on a single realisation of duration** T seconds.

Introduction

A random variable is a function defined on a **sample space** (Ω) which we write as $X(\omega)$. The **range space** is the elements in the outcome. As a simple example, toss two coins and let X be the number of tails that occur. The sample space here is (H, H), (H, T), (T, H), and (T, T). The range space here is 0, 1, (twice) and 2.

This example shows that the elements of Ω and the values taken by X need not be the same.

Often $X(\omega)$ is abbreviated as X. The values that X can take are denoted by x.

Introduction

Probability is the 'likelihood of occurrence' of an 'event' in an 'experiment of chance.'

The sample space is the set of all possible outcomes — containing n_Ω elements. Event E is a subset of Ω. The probability of occurrence of event E is $P(E)$:

$$P(E) = \frac{n_E}{n_\Omega}$$

where n_E is the number of elements in E. Always

$$0 \leq P(E) \leq 1$$

Introduction

As defined, probability is a theoretical concept that can be computed without experimentation.

Empirical probability is deduced from observed data, i.e. conduct an 'experiment' N times and count the number of times an event E occurs — say n_1. Then the **relative frequency of occurrence** is given by

$$f_E = \frac{n_1}{N}$$

Often this is taken as an estimate of $P(E)$.

Introduction

Types of random variable:

- discrete — the possible outcomes x are countable (e.g. roll a dice).
- continuous — the possible outcomes x are uncountable (e.g. a fluctuating voltage).
- mixed — (e.g. a binary signal in noise).

Probability Distributions

These can be discrete $P[X = x_i]$, e.g. a six faced dice with associated distribution function $F(x) = P[X \leq x]$

Continuous Random Variables The distribution function $F(x) = Prob[X \leq x]$ satisfies

$$F(-\infty) = 0, \quad F(\infty) = 1$$

Hence

$$P[a \leq X \leq b] = F(b) - F(a)$$

Since $P[X = x] = 0$ for continuous processes, we introduce the **probability density function**, i.e. let $a = x, \; b = x + \delta x$.

Probability Density Functions

Then

$$p(x) = \lim_{\delta x \to 0} \frac{F(x + \delta x) - F(x)}{\delta x} = \frac{dF(x)}{dx}$$

Also

$$\int_{-\infty}^{\infty} p(x) = 1$$

Expectations

Rather than using probability distributions **often averages are used.** This leads to the concept of the **expectation of a process,** e.g. the **sample mean** $\overline{x}$. If x_i occurs n_i times in N trials then

$$\overline{x} = \frac{1}{N}\sum_i x_i n_i$$

Here $\frac{n_i}{N} = f_i$ is the empirical probability of occurrence and for a continuous process $\rightarrow p(x_i)\delta x_i$.

Expectations

Hence

$$\overline{x} \to E[X]$$

where

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

The expectation operation also generalises to functions of a random variable, e.g. if $Y = g(X)$ then

$$E[Y] = E[g(X)] = \int_{-\infty}^{\infty} g(x)p(x)\ dx$$

and also to **bivariate processes**

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

Moments

Rather than use probability functions it is often easier to use summarising quantities, e.g.

- location — the mean $\mu_x = E[X]$
- spread — the variance $\sigma_X^2 = E[(X - \mu_x)^2]$
- standard deviation — σ_X
- skewness — $E[(x - \mu_x)^3]$
- kurtosis (a measure of flattening) — $\frac{E[(x-\mu_x)^4]}{\sigma_X^4}$

Gaussian Distribution

Probability density function

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\left(\frac{(x-\mu_X)^2}{2\sigma_X^2}\right)}$$

Here the mean $= \mu_x$, the standard deviation is σ_X, the skewness $= 0$ and the kurtosis $= 3$.

Stochastic Processes

A **stochastic process is a paramterised random variable**, i.e. $X_t(\omega)$ or $X(t)$. The set of **all possible realisations** (an infinite number) for $t \in (-\infty, \infty)$ define the **ensemble** $\{X(\omega)\}$.
The probability density function **for a stochastic process is time dependent.** If N is a finite set of records and n_1 records fall in bandwidth Δx then

$$p(x,t) = \lim_{\Delta x \to 0} \frac{P[x < X(t) \leq x + \Delta x]}{\Delta x} \approx \frac{n_1}{N} \frac{1}{\Delta x}$$

Stochastic Processes

This extends naturally to the joint density function of a bivariate process. The moments here **are time varying as are the underlying probability distributions. Such processes are termed non-stationary stochastic processes.**
We restrict attention here to **stationary processes** but **the problem of deciding from measured data whether a process is stationary or not is often 'difficult'.**

Cross Correlation

This is defined for a bivariate process as

$$R_{xy}(\tau) = E[x(t)y(t+\tau)]$$

and is a measure of cross association.
In general **ensemble averaging is not feasible — usually we only have a single realisation.** Instead we must consider **time averages.**

Critical question: Do time averages along one record give the same results as ensemble averages?

The answer is — sometimes and such averages are termed **ergodic.**

Time Averages

Mean Value

$$\mu_x = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, dt$$

Cross-Correlation

$$R_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)y(t+\tau)\, dt$$

Covariance (Correlation) Functions

It is useful to also look at **joint moments,** e.g. **the autocorrelation function**

$$R_{xx}(t_1, t_2) = E[\{X(t_1) - \mu_x(t_1)\}\{X(t_2) - \mu_x(t_2)\}]$$

This is a measure of the degree of association of the signal $x(t)$ **at time** t_1 **with itself at time** t_2. For **a stationary process this is a function of** $\tau = t_2 - t_1$ **only.** Hence in this case

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

Also

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

Spectra

These are **frequency domain interpretations of correlation functions.**

For an finite length realisation of duration T we have by the Fourier transform that the average power of $x(t)$ is

$$\begin{aligned} \overline{x}^2 &= \frac{1}{T}\int_{-T}^{T} x^2(t)\, dt \\ &= \frac{1}{T}\int_{-\infty}^{\infty} |X_T(f)|^2 df \end{aligned}$$

(by Parseval's theorem).

Spectra

Hence the average power is

$$\int_{-\infty}^{\infty} \frac{1}{T} |X_T(f)|^2 df$$

$$\hat{S}_{ff}^2 = \frac{1}{T} |X_T(f)|^2$$

is known as the **raw power spectral density** because $\hat{S}_{ff}$ **does not converge as** $T \rightarrow \infty$ **and hence is not ergodic.**
If, however, **we average this**, i.e. form $E[\overline{x}^2]$ and let $T \rightarrow \infty$ then

$$Var[x] = E[\overline{x}^2] = \int_{-\infty}^{\infty} S_{xx}(f)\, df$$

Spectra

where

$$S_{xx}(f) = \lim_{T\to\infty} E \frac{|X_T(f)|^2}{T}$$

is the power spectral density.

$S_{xx}(f)$ is a two-sided spectral density (If x is in Volts, this quantity has units of $volts^2/Hz$.)

A function defined for positive frequencies, say $G_{xx}(f)$, can be defined as

$$G_{xx}(f) = \begin{cases} 2S_{xx}(f),\ f > 0 \\ 0,\ f < 0 \\ S_{xx}(0),\ f = 0 \end{cases}$$

Also the power of the process in the band $f_1 \to f_2$ is

$$\int_{f_1}^{f_2} G_{xx}(f)\, df$$

Cross-Spectra

Cross-spectral densities are defined in a similar manner, i.e.

$$S_{xy}(f) = \lim_{T\to\infty} \frac{E\{X_T^*(f)Y_T(f)\}}{T} = |S_{xy}|e^{j\psi_{xy}}$$

Note that this quantity is complex and $|S_{xy}(f)|$ is **a measure of the association of amplitudes in x and y at frequency f.** $\psi_{xy}(f)$ shows the lag/lead between x and y at frequency f. Also

$$S_{xy}(f) = S_{yx}^*(f), \;\; |S_{xy}(f)|^2 \le S_{xx}(f)S_{yy}(f)$$

Spectra and Correlation Functions

Correlation functions and spectra carry equivalent information and are related through the Fourier transform

$$\begin{aligned} S_{xy}(f) &= \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j2\pi f\tau}\, d\tau \\ R_{xy}(\tau) &= \int_{-\infty}^{\infty} S_{xy}(f) e^{j2\pi f\tau}\, df \end{aligned}$$

These are sometimes termed the **Wiener-Khinchin relationships.**

Spectra and Correlation Functions

Examples

$$R_{xx}(\tau) = \sigma^2 e^{-\alpha|\tau|}, \quad S_{xx}(f) = \frac{2\sigma^2\alpha}{(2\pi f)^2 + \alpha^2}$$

$$R_{xx}(\tau) = \delta(\tau), \quad S_{xx}(f) = 1, \text{ for all } f$$