CHAPTER 3

# Random signals and noise

## 3.1 Introduction

The periodic and transient signals discussed in Chapter 2 are deterministic. This precludes them from conveying information since nothing new can be learned by receiving a signal which is entirely predictable. Unpredictability or randomness is a property which is essential for information bearing signals. (The definition and quantification of randomness is an interesting topic. Here, however, an intuitive and common-sense notion of randomness is all that is required.)

Whilst one type of random signal creates information (i.e. increases knowledge) at a communication receiver another type of random signal destroys it (i.e. decreases knowledge). The latter type of signal is known as noise [Rice]. The distinction between signals and noise is therefore essentially one of their origin (i.e. an information source or elsewhere) and whether reception is intended or not. In this context interference (signals arising from information sources other than the one expected or intended) is a type of noise. From the point of view of describing information signals and noise mathematically, no distinction is necessary. Since signals and noise are random such descriptions must be, at least partly, in terms of probability theory.

## 3.2 Probability theory

Consider an experiment with three possible, random, outcomes $A$, $B$, $C$. If the experiment is repeated $N$ times and the outcome $A$ occurs $L$ times then the probability of outcome $A$ is defined by:

$$P(A) \triangleq \lim_{N \to \infty} \left\{ \frac{L}{N} \right\} \tag{3.1}$$

Note that the error, $\varepsilon$, for $N$ experimental trials does not tend to zero for large $N$ but actually increases (on average) as $\sqrt{N}$, Figure 3.1. The ratio $\varepsilon/N$ does tend to zero, however, for large $N$, Figure 3.2, i.e.:

$$\frac{\varepsilon}{N} \rightarrow 0 \quad \text{as} \quad N \rightarrow \infty$$

(Thus on tossing a coin, for example, the result of achieving close to 50% heads is much more likely to be achieved with a large number of samples, e.g. >50 individual tosses or trials.) Such an experiment could be performed $N$ times with one set of (unchanging) apparatus or $N$ times, simultaneously, with $N$ sets of (identical) apparatus. The former is called a temporal experiment whilst the latter is called an ensemble experiment. If, after $N$ trials, the outcome $A$ occurs $L$ times and the outcome $B$ occurs $M$ times, and if $A$ and $B$ are *mutually exclusive* (i.e. they cannot occur together) then the probability that $A$ or $B$ occurs is:

$$\begin{aligned} P(A \text{ or } B) &= \lim_{N \to \infty} \left\{ \frac{L+M}{N} \right\} \\ &= \lim_{N \to \infty} \left\{ \frac{L}{N} \right\} + \lim_{N \to \infty} \left\{ \frac{M}{N} \right\} \\ &= P(A) + P(B) \end{aligned} \tag{3.2}$$

This is the basic law of additive probabilities which can be used for any number of mutually exclusive events.

Since the outcome of the experimental trials described above is variable and random it is called (unsurprisingly) a random variable. Such random variables can be discrete or continuous. An example of the former would be the score achieved by the throw of a dice. An example of the latter would be the final position of a coin in a game of shove halfpenny. (In the context of digital communications relevant examples might be the voltage of a quantised signal source and the voltage of an, unquantised, noise source.)

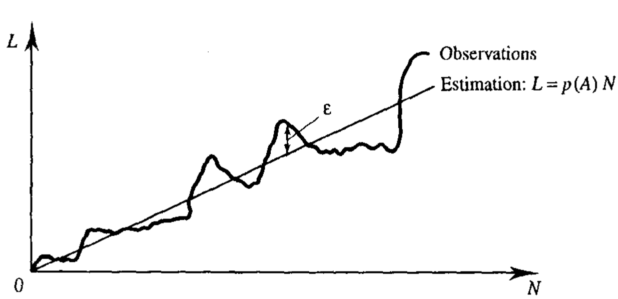



**Figure 3.1** *Observations compared with estimation for L outcomes of A after N random trials.*

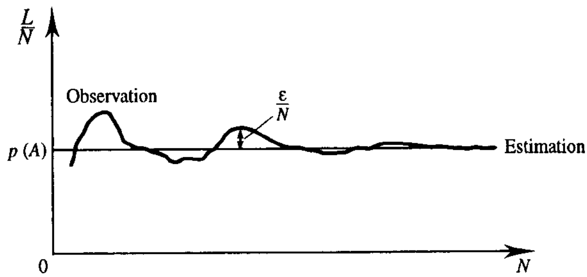



**Figure 3.2** *Observation compared with estimation for the fraction L/N of outcomes A after N random trials.*

---

EXAMPLE 3.1
A dice is thrown once. What is the probability that: (i) the dice shows 3; (ii) the dice shows 6; (iii) the dice shows a number greater than 2; (iv) the dice does not show 5?

(i) Since all numbers between 1 and 6 inclusive are equiprobable, $P(3) = \frac{1}{6}$

(ii) As above, $P(6) = \frac{1}{6}$

(iii) $P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$

$$P(3 \text{ or } 4 \text{ or } 5 \text{ or } 6) = P(3) + P(4) + P(5) + P(6) = \frac{4}{6} = \frac{2}{3}$$

(iv) $P(\text{any number but } 5) = 1 - P(5) = 1 - \frac{1}{6} = \frac{5}{6}$
because total probability, $\Sigma_{i=1}^{6} P(i)$, must sum to 1.

---

### 3.2.1 Conditional probabilities, joint probabilities and Bayes's rule

The probability of event $A$ occurring given that event $B$ is known to have occurred, $P(A|B)$, is called the conditional probability of $A$ on $B$ (or the probability of $A$ conditional on $B$). The probability of $A$ and $B$ occurring together, $P(A, B)$, is called the joint probability of $A$ and $B$. Joint and conditional probabilities are related by:

$$\begin{aligned} P(A, B) &= P(B)P(A|B) \\ &= P(A)P(B|A) \end{aligned} \tag{3.3}$$

Rearranging equation (3.3) gives Bayes's rule:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \tag{3.4}$$

---

EXAMPLE 3.2
Four cards are dealt off the top of a shuffled pack of 52 playing cards. What is the probability that all the cards will be of the same suit?

Given that the first card is a spade the probability that the second card will be a spade is:

$$P(2nd\text{ spade}) = \frac{12}{51} = 0.2353$$

and so on:

$$P(3rd\text{ spade}) = \frac{11}{50} = 0.2200$$

$$P(4th\text{ spade}) = \frac{10}{49} = 0.2041$$

$$\text{Therefore } P(4\text{ spades}) = 0.2353 \times 0.2200 \times 0.2041 = 0.01056$$

This is the probability of all the cards being from the same suit.

Notice that intuitively the suit of the first card has been ignored from a probability point of view. A more formal solution to this problem uses Bayes's rule explicitly:

$$\begin{aligned} P(4\text{ spades}) &= P(\text{spade, 3 spades}) \\ &= P(\text{3 spades})\, P(\text{spade} \mid \text{3 spades}) \\ &= P(\text{2 spades})\, P(\text{spade} \mid \text{2 spades})\, P(\text{spade} \mid \text{3 spades}) \\ &= P(\text{spade})\, P(\text{spade} \mid \text{spade})\, P(\text{spade} \mid \text{2 spades})\, P(\text{spade} \mid \text{3 spades}) \\ &= \frac{13}{52} \times \frac{12}{51} \times \frac{11}{50} \times \frac{10}{49} \\ &= 0.002641 \end{aligned}$$

$$\begin{aligned} P(4\text{ same suit}) &= P(\text{4 spades or 4 clubs or 4 diamonds or 4 hearts}) \\ &= P(\text{4 spades}) + P(\text{4 clubs}) + P(\text{4 diamonds}) + P(\text{4 hearts}) \\ &= 0.002641 \times 4 = 0.01056 \end{aligned}$$

---

### 3.2.2 Statistical independence

Events $A$ and $B$ are statistically independent if the occurrence of one does not affect the probability of the other occurring, i.e.:

$$P(A|B) = P(A) \tag{3.5(a)}$$

and:

$$P(B|A) = P(B) \tag{3.5(b)}$$

It follows that for statistically independent events:

$$P(A, B) = P(A)P(B) \tag{3.6}$$

---

EXAMPLE 3.3

Two cards are dealt one at a time, face up, from a shuffled pack of cards. Show that these two events are not statistically independent.

The unconditional probabilities for both events are:

$$P(A) = P(B) = \frac{1}{52}$$

The probability of event $A$ is:

$$P(A) = \frac{1}{52}$$

The probability of event $B$ is:

$$P(B|A) = \frac{1}{51}$$

The joint probability of the two events is therefore:

$$\begin{aligned} P(A, B) &= P(A)P(B|A) \\ &= \frac{1}{52} \times \frac{1}{51} \\ &\neq P(A)\,P(B) \end{aligned}$$

i.e. the events are not statistically independent.

---

### 3.2.3 Discrete probability of errors in a data block

When we consider the problem of performance prediction in a digital coding system we often ask what is the probability of having more than a given number of errors in a fixed length codeword? This is a discrete probability problem. Assume that the probability of single bit (binary digit) error is $P_e$, that the number of errors is $R'$ and $n$ is the block length, i.e. we require to determine the probability of having more than $R'$ errors in a block of $n$ digits. Now:

$$P(> R' \text{ errors}) = 1 - P(\leq R' \text{ errors}) \tag{3.7(a)}$$

because total probability must sum to 1. We also assume that errors are independent. The above equation may thus be expanded as:

$$P(> R' \text{ errors}) = 1 - [P(0 \text{ error}) + P(1 \text{ error}) + P(2 \text{ errors}) + \cdots + P(R' \text{ errors})] \quad (3.7(b))$$

These probabilities will be calculated individually starting with the probability of no errors. A block representing the codeword is divided into bins labelled 1 to $n$. Each bin corresponds to one digit in the $n$ digit codeword and it is labelled with the probability of the event in question.

Now considering the general case of $j$ errors in $n$ digits with a probability of error per digit of $P_e$ we can generalise the above equations i.e.:

$$P(j \text{ errors}) = (P_e)^j \, (1 - P_e)^{n-j} \, {}^nC_j \tag{3.8}$$

where the binomial coefficient ${}^nC_j$ is given by:

$${}^nC_j = \frac{n!}{j!(n-j)!} = \binom{n}{j} \tag{3.9(a)}$$

This is the probability of $j$ errors in an $n$-digit codeword, but what we are interested in is the probability of having more than $R'$ errors. We can write this using equation (3.7(b)) as:

$$P(> R' \text{ errors}) = 1 - \sum_{j=0}^{R'} P(j) \tag{3.9(b)}$$

For large $n$ we have statistical stability in the sense that the number of errors in a given block will tend to the product $P_e\, n$. Furthermore, the fraction of blocks containing a number of errors that deviates significantly from this value will tend to zero, Figure 3.2. The above statistical stability gained from long blocks is very important in the design of effective error correction codewords and is discussed further in Chapter 10.

---

EXAMPLE 3.4

If the probability of single digit error is 0.01 and hence the probability of correct digit reception is 0.99 calculate the probability of 0 to 2 errors occurring in a ten digit codeword.

If all events (receptions) are independent then the probability of having no errors in the block is $(0.99)^{10} = 0.904382$, Figure 3.3(a).

Now consider the probability of a single error. Initially assume that the error is in the first position. Its probability is 0.01. All other digits are received correctly, so their probabilities are 0.99 and there are 9 of them, Figure 3.3(b). However there are 10 positions where the single error can occur in the 10 digit codeword and therefore the overall probability of a single error occurring is:

$$P(1 \text{ error}) = (0.01)^1 \, (0.99)^9 \, {}^{10}C_1 = 0.091352$$

where ${}^{10}C_1 = 10$, represents the number of combinations of 1 object from 10 objects. We can similarly calculate the probability of 2 errors as:

$$P(2 \text{ errors}) = (0.01)^2 \, (0.99)^8 \, {}^{10}C_2 = 0.00415$$

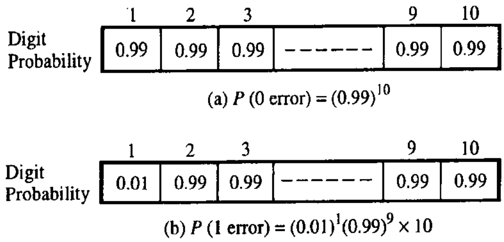



**Figure 3.3** *Discrete probability of codeword errors with a per digit error probability of 1%.*

Thus the probability of three or more errors, equation (3.7), is $1 - 0.904382 - 0.091352 - 0.00415 = 0.000116$. Notice how the probability of $j$ errors in the data block falls off rapidly with $j$.

---

## 3.2.4 Cumulative distributions and probability density functions

A cumulative distribution (also called a probability distribution) is a curve showing the probability, $P_X(x)$, that the value of the random variable, $X$, will be less than or equal to some specific value, $x$, Figure 3.4(a), i.e.:

$$P_X(x) = P(X \leq x) \tag{3.10}$$

(Conventionally, upper case letters are used for the name of a random variable and lower case letters for particular values of the random variable. Note, however, that the upper case subscript in the left hand side of equation (3.10) is often omitted.) Some properties of a cumulative distribution (CD) apparent from Figure 3.4(a) are:

1. $0 \leq P_X(x) \leq 1 \quad \text{for} \quad -\infty \leq x \leq \infty$
2. $P_X(-\infty) = 0$
3. $P_X(\infty) = 1$
4. $P_X(x_2) - P_X(x_1) = P(x_1 < X \leq x_2)$
5. $\dfrac{dP_X(x)}{dx} \geq 0$

Exceedance curves are also sometimes used to describe the probability behaviour of random variables. These curves are complementary to CDs in that they give the probability that a random variable exceeds a particular value, i.e.:

$$P(X > x) = 1 - P_X(x) \tag{3.11}$$

Probability density functions (pdfs) give the probability that the value of a random variable, $X$, lies between $x$ and $x + dx$. This probability can be written in terms of a CD as:

$$P_X(x + dx) - P_X(x) = \frac{dP_X(x)}{dx}\, dx \tag{3.12}$$

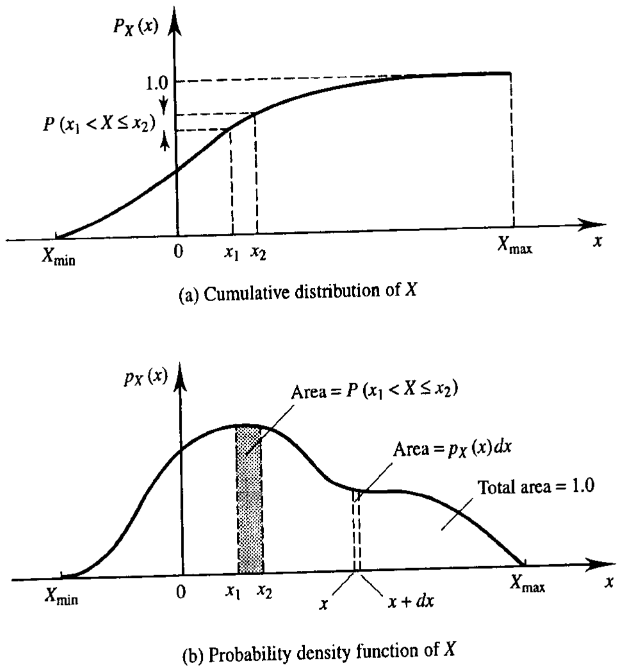



**Figure 3.4** *Descriptions of a continuous random variable (smooth CD and continuous pdf).*

It is the factor $[dP_X(x)]/dx$, normally denoted by $p_X(x)$, which is defined as the pdf of $X$. Figure 3.4(b) shows an example. Pdfs and CDs are therefore related by:

$$p_X(x) = \frac{dP_X(x)}{dx} \tag{3.13(a)}$$

$$P_X(x) = \int_{-\infty}^{x} p_X(x')\, dx' \tag{3.13(b)}$$

The important properties of pdfs are:

$$\int_{-\infty}^{\infty} p_X(x)\, dx = 1 \tag{3.14(a)}$$

and

$$\int_{x_1}^{x_2} p_X(x)\, dx = P(x_1 < X \leq x_2) \tag{3.14(b)}$$

Both CDs and pdfs can represent continuous, discrete or mixed random variables. Discrete random variables, as typified in Example 3.1, have stepped CDs and purely impulsive pdfs, Figure 3.5. Mixed random variables have CDs which contain discontinuities and pdfs which contain impulses, Figure 3.6.

Pdfs can be measured, in principle, by taking an ensemble of identical random variable generators, sampling them all at one instant and plotting a relative frequency histogram of samples. (The pdf is the limit of this histogram as the size interval shrinks

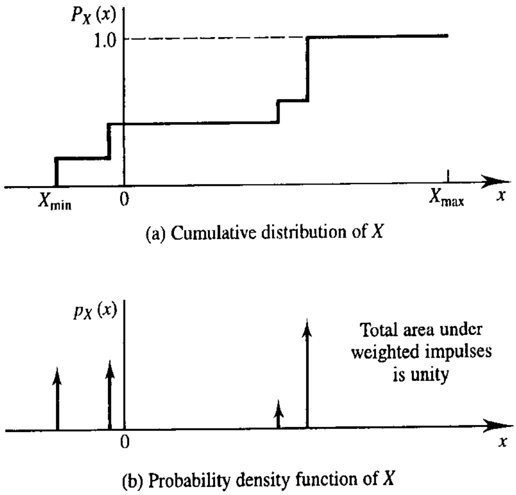



**Figure 3.5** *Probability descriptions of a discrete random variable (stepped CD and discrete pdf).*

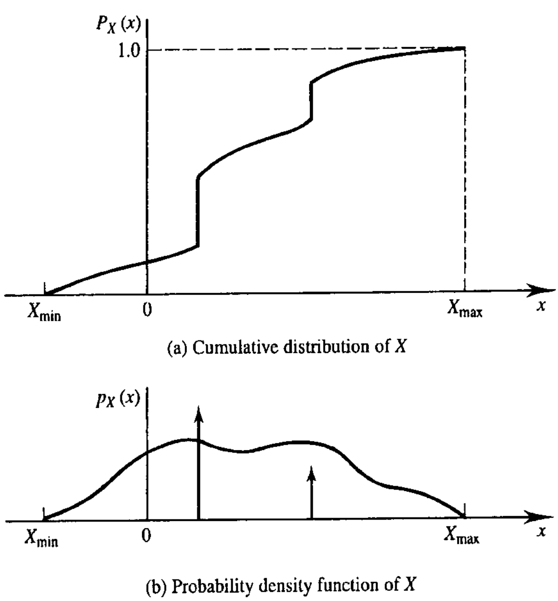



**Figure 3.6** *Probability descriptions of a mixed random variable.*

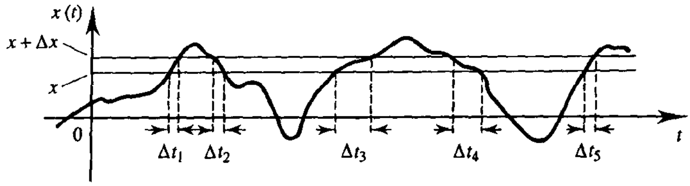



**Figure 3.7** *Determination of pdf for an (ergodic) random process.*

to zero and the number of samples tends to infinity.) Alternatively, at least for ergodic processes (see section 3.3.1), the proportion of time the random variable spends in different value intervals of size $\Delta x$ could be determined, (see Figure 3.7). In this case:

$$p_X(x)\,\Delta x = \sum_{i=1}^{N} \frac{\Delta t_i}{\text{total observation time}} \tag{3.15}$$

and $p_X(x)$ can be found as the limit of equation (3.15) as $\Delta x \rightarrow 0$ (and the observation time and $N \rightarrow \infty$). This allows one to calculate the time that the signal $x(t)$ lies between the voltages $x$ and $\Delta x$.

---

EXAMPLE 3.5

A random voltage has a pdf given by:

$$p(V) = k\, u(V+4)\, e^{-3(V+4)} + 0.25\, \delta(V-2)$$

where $u(\,)$ is the Heaviside step function. (i) Sketch the pdf; (ii) find the probability that $V = 2$ V; (iii) find the value of $k$; and (iv) find and sketch the CD of $V$.

(i) Figure 3.8(a) shows the pdf of $V$.
(ii) Area under impulse is 0.25 therefore $P(V = 2) = 0.25$.
(iii) If the area under impulse part of pdf is 0.25 then the area under exponential part of the pdf must be 0.75, i.e.:

$$\int_{-4}^{\infty} k\, e^{-3(V+4)}\, dV = 0.75$$

Use change of variable:

$$V + 4 = x\,, \quad dx/dV = 1$$

when $V = -4$, $x = 0$, and when $V = \infty$, $x = \infty$

Therefore:

$$\int_{0}^{\infty} k\, e^{-3x}\, dx = 0.75$$

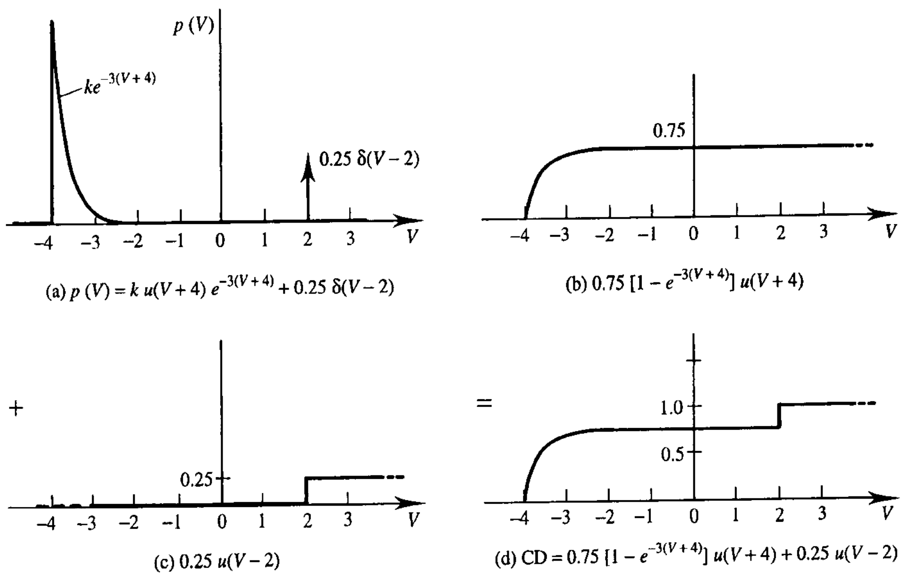



**Figure 3.8** *Pdf of V in Example 3.5, (a) and components (b) and (c) of resulting CD, (d). (The step function is implied in the solution by the lower limit of integration.)*

i.e.:

$$k = \frac{0.75}{[e^{-3x}/-3]_0^\infty}$$

$$= \frac{-3 \times 0.75}{[0-1]} = 2.25$$

(iv)

$$\text{CD} = \int_{-\infty}^{V} p(V')\, dV'$$

$$= \int_{-4}^{V} [2.25e^{-3(V'+4)} + 0.25\delta(V'-2)]\, dV'$$

Using the same substitution as in part (iii):

$$\text{CD} = 2.25\left[\frac{e^{-3x}}{-3}\right]_0^{V+4} + 0.25u(V-2)$$

$$= -2.25/3 \times [e^{-3(V+4)} - 1] + 0.25u(V-2)$$

Figure 3.8 (b) and (c) shows these individual waveforms and (d) shows the combined result.

### 3.2.5 Moments, percentiles and modes

The first moment of a random variable $X$ is defined by:

$$\bar{X} = \int_{-\infty}^{\infty} x \, p(x) \, dx \tag{3.16}$$

where $\bar{X}$ denotes an ensemble *mean*. (Sometimes this quantity is called the expected value of $X$ and is written E[$X$].) The second moment is defined as:

$$\overline{X^2} = \int_{-\infty}^{\infty} x^2 \, p(x) \, dx \tag{3.17(a)}$$

and represents the mean square of the random variable. (The square root of the second moment is thus the RMS value of $X$.)

Higher order moments are defined by the general formula:

$$\overline{X^n} = \int_{-\infty}^{\infty} x^n \, p(x) \, dx \tag{3.17(b)}$$

(The zeroth moment ($n = 0$) is always equal to 1.0 and is therefore not a useful quantity.)

Central moments are the net moments of a random variable taken about its mean. The second central moment is therefore given by:

$$\overline{(X - \bar{X})^2} = \int_{-\infty}^{\infty} (x - \bar{X})^2 \, p(x) \, dx \tag{3.18(a)}$$

and is usually called the *variance* of the random variable. The square root of the second central moment is the *standard deviation* of the random variable (general symbol $s$, although for a Gaussian random variable the symbol $\sigma$ is commonly used). The variance or standard deviation provides a measure of random variable spread, or pdf width. Higher order central moments are defined by the general formula:

$$\overline{(X - \bar{X})^n} = \int_{-\infty}^{\infty} (x - \bar{X})^n \, p(x) \, dx \tag{3.18(b)}$$

(The zeroth central moment is always 1.0 and the first central moment is always zero. Neither are of any practical use, therefore.) The 3rd and 4th central moments divided by $s^3$ and $s^4$ respectively are called *skew* and *kurtosis* [1]. These are a measure of pdf asymmetry and peakiness, the latter being in comparison to a Gaussian function, Figure 3.9. These higher order moments are of current interest for the analysis of non-stationary signals, such as speech, and they are also appropriate for the analysis of non-Gaussian signals.

[1] If kurtosis is defined as $[\overline{(X - \bar{X})^4}/s^4] - 3$ then a Gaussian curve will have zero kurtosis.

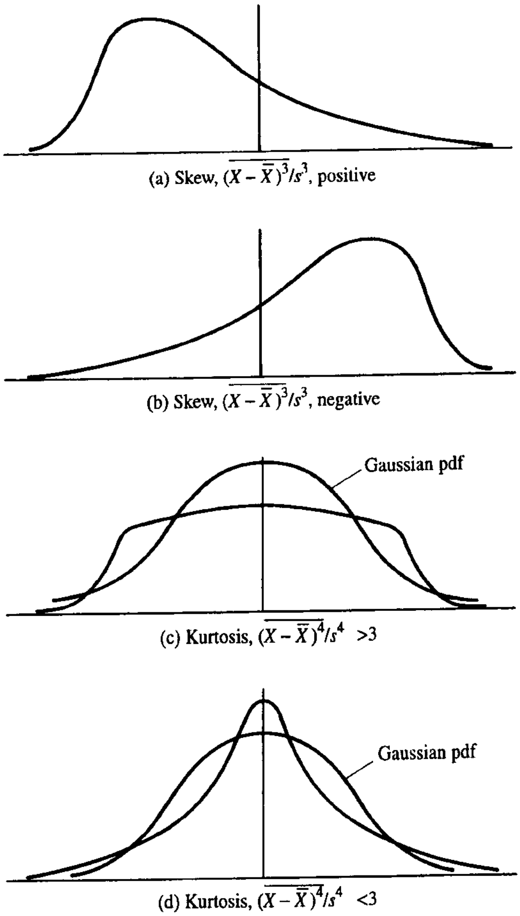



**Figure 3.9** *Illustration of skew and kurtosis as descriptors of pdf shape.*

The $p$th *percentile* is the value of $X$ below which $p$ percent of the total area under the pdf lies, Figure 3.10(a), i.e.:

$$\int_{-\infty}^{x} p(x')\,dx' = \frac{p}{100} \tag{3.19}$$

where $x$ is the $p$th percentile. In the special case of $p = 50$ (i.e. the 50th percentile) the corresponding value of $x$ is called the *median* and the pdf is divided into two equal areas, Figure 3.10(b).

The *mode* of a pdf is the value of $x$ for which $p(x)$ is a maximum, Figure 3.11(a). For a pdf with a single mode this can be interpreted as the most likely value of $X$. In general pdfs can be multimodal, Figure 3.11(b). Moments, percentiles and modes are all examples of *statistics*, i.e. numbers which in some way summarise the behaviour of a

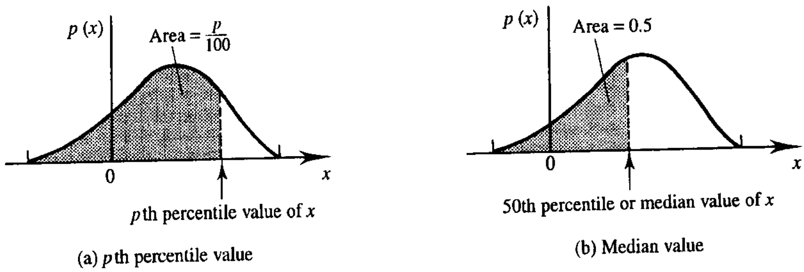



**Figure 3.10** *Illustration of percentiles and median.*

random variable. The difference between probabilistic and statistical models is that statistical models usually give an incomplete description of random variables.

There are some useful electrical interpretations of moments and central moments in the context of random voltages (and currents). These interpretations are summarised below:

| *Moment* | *Familiar name* | *Interpretation* |
|---|---|---|
| 1st | Mean value | DC voltage (or current) |
| 2nd | Mean square value | Total power[1] |
| 2nd central | Variance | AC power[2] |

*Notes:*
(1) This is the total normalised power, i.e. the power dissipated in a 1 Ω load.
(2) This is the power dissipated in a 1 Ω load by the fluctuating (i.e. AC) component of voltage.

The interpretations of 1st and 2nd moments are obvious. The interpretation of 2nd central moment, $s^2$, becomes clear when the left hand side of equation (3.18(a)) is expanded as shown below, i.e.:

$$\begin{aligned} s^2 &= \langle [x(t) - \langle x(t)\rangle]^2 \rangle \\ &= \langle x^2(t) - 2x(t)\,\langle x(t)\rangle + \langle x(t)\rangle^2 \rangle \\ &= \langle x^2(t)\rangle - \langle x(t)\rangle^2 \end{aligned} \tag{3.20}$$

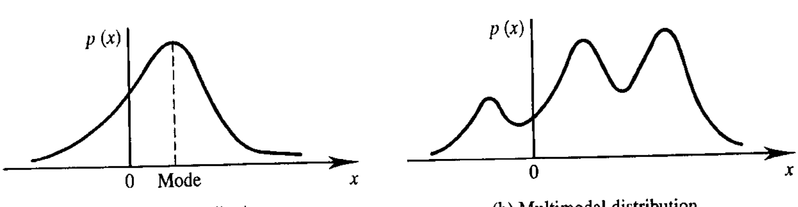



**Figure 3.11** *Illustration of modal values.*

(The angular brackets, ⟨ ⟩, here indicate time averages, equation (2.16), which for most random variables of engineering interest can be equated to ensemble averages, see section 3.3.1.) Equation (3.20) is an expression of the familiar statistical statement that variance is the mean square minus the square mean. This is clearly also equivalent to total power minus DC power which must be the AC (or fluctuating) power. Figure 3.12 illustrates these electrical interpretations.

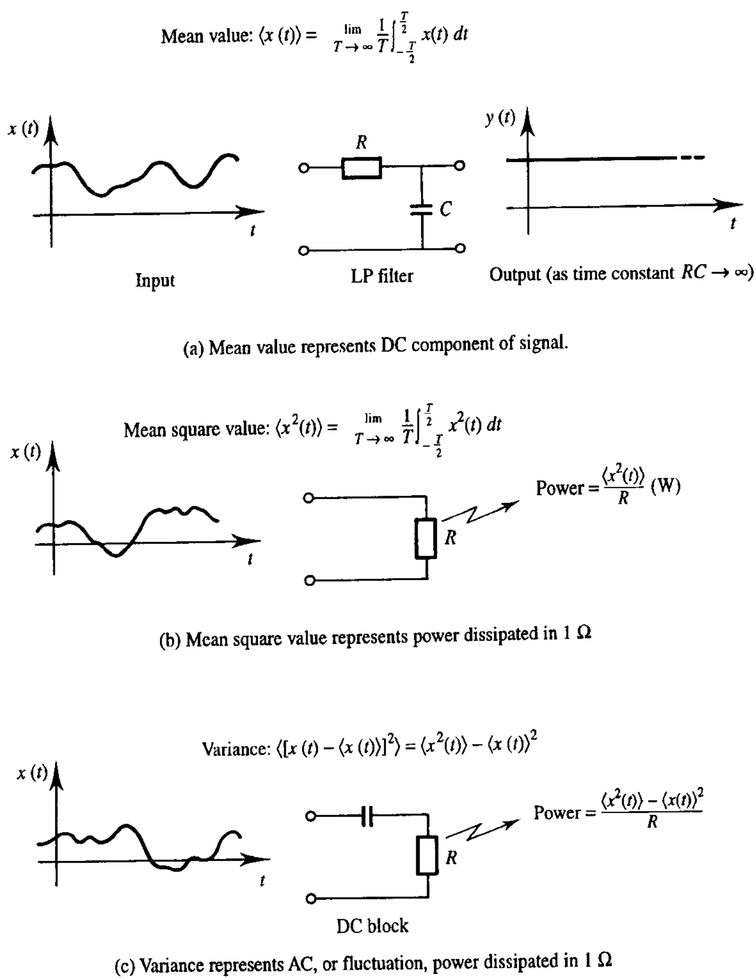



**Figure 3.12** *Engineering interpretations of: (a) mean; (b) mean square; (c) variance.*

---

EXAMPLE 3.6

A random voltage has a pdf given by:

$$p(V) = u(V)\,3\,e^{-3V}$$

Find the DC voltage, the power dissipated in a 1 Ω load and the median value of voltage. What power would be dissipated at the output of an AC coupling capacitor?

Equation (3.16) defines the DC value:

$$\bar{V} = \int_{-\infty}^{\infty} V\,p(V)\,dV = 3\int_{0}^{\infty} V\,e^{-3V}\,dV$$

Using the standard integral [Dwight, equation 567.9]:

$$\int x^n e^{ax}\,dx = e^{ax}\left[\frac{x^n}{a} - \frac{nx^{n-1}}{a^2} + \frac{n(n-1)x^{n-2}}{a^3} \cdots (-1)^{n-1}\frac{n!x}{a^n} + (-1)^n\frac{n!}{a^{n+1}}\right],\quad n \geq 0$$

$$\bar{V} = 3\left[e^{-3V}\left(\frac{V}{-3} - \frac{1\times V^0}{(-3)^2}\right)\right]_0^{\infty}$$

$$= -3\,[0 - 1/9] = 1/3\ \text{V}$$

The power is given by equation (3.17):

$$\overline{V^2} = \int_{-\infty}^{\infty} V^2\,p(V)\,dV = \int_{0}^{\infty} V^2\,3\,e^{-3V}\,dV$$

Again using [Dwight, equation 567.9]:

$$\overline{V^2} = 3\left[e^{-3V}\left(\frac{V^2}{-3} - \frac{2V^1}{(-3)^2} + \frac{2(1)\,V^0}{(-3)^3}\right)\right]_0^{\infty}$$

$$= 3\left[0 - \left(-\frac{2}{27}\right)\right] = \frac{6}{27}\ (\text{or } 0.2222)\ \text{V}^2$$

Median value = 50th percentile, i.e.:

$$\int_{-\infty}^{X_{median}} p(V)\,dV = 0.5$$

Therefore:

$$\int_{0}^{X_{median}} 3\,e^{-3V}\,dV = 0.5$$

i.e.:

$$3\left[\frac{e^{-3V}}{-3}\right]_0^{X_{median}} = 0.5$$

$$1 - e^{-3X_{median}} = 0.5$$

$$X_{median} = \frac{\ln(1 - 0.5)}{-3} = 0.2310 \text{ V}$$

The AC coupling capacitor acts as a DC block and the fluctuating or AC power is given by the variance of the random signal as defined in equation (3.20), i.e.:

$$\begin{aligned} P_{AC} &= s^2 \\ &= \langle v^2(t)\rangle - \langle v(t)\rangle^2 \\ &= \overline{V^2} - \bar{V}^2 \quad \text{(signal assumed ergodic, see section 3.3.1)} \\ &= \frac{6}{27} - \left(\frac{1}{3}\right)^2 = \frac{3}{27} \quad \text{(or 0.1111) V}^2 \end{aligned}$$

---

### 3.2.6 Joint and marginal pdfs, correlation and covariance

If $X$ and $Y$ are two random variables a joint probability density function, $p_{X,Y}(x, y)$, can be defined such that $p_{X,Y}(x, y)\, dx\, dy$ is the probability that $X$ lies in the range $x$ to $x + dx$ and $Y$ lies in the range $y$ to $y + dy$. The joint (or bivariate) pdf can be represented by a surface as shown in Figure 3.13(a). For quantitative work, however, it is often more convenient to display $p_{X,Y}(x, y)$ as a contour plot, Figure 3.13(b), and, when investigating bivariate random variables experimentally, sample values of $(x, y)$ can be plotted as a scattergram, Figure 3.13(c). (Contours of constant point density in Figure 3.13(c) correspond, of course, to contours of constant probability density in Figure 3.13(b).)

Just as the total area under the pdf of a single random variable is 1.0, the volume under the surface representing a bivariate random variable is 1.0, i.e.:

$$\int_X \int_Y p_{X,Y}(x, y)\, dx\, dy = 1.0 \tag{3.21}$$

The probability of finding the bivariate random variable in any particular region, $F$, of the $X,Y$ plane, Figure 3.13(b), is:

$$P([X, Y] \text{ lies within } F) = \iint_F p_{X,Y}(x, y)\, dx\, dy \tag{3.22}$$

If the joint pdf, $p_{X,Y}(x, y)$, of a bivariate variable is known then the probability that $X$ lies in the range $x_1$ to $x_2$ (irrespective of the value of $Y$) is called a marginal probability of $X$ and is found by integrating over all $Y$, Figure 3.14. The marginal pdf of $X$ is therefore given by:

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x, y)\, dy \tag{3.23(a)}$$

Similarly, the marginal pdf of $Y$ is given by:

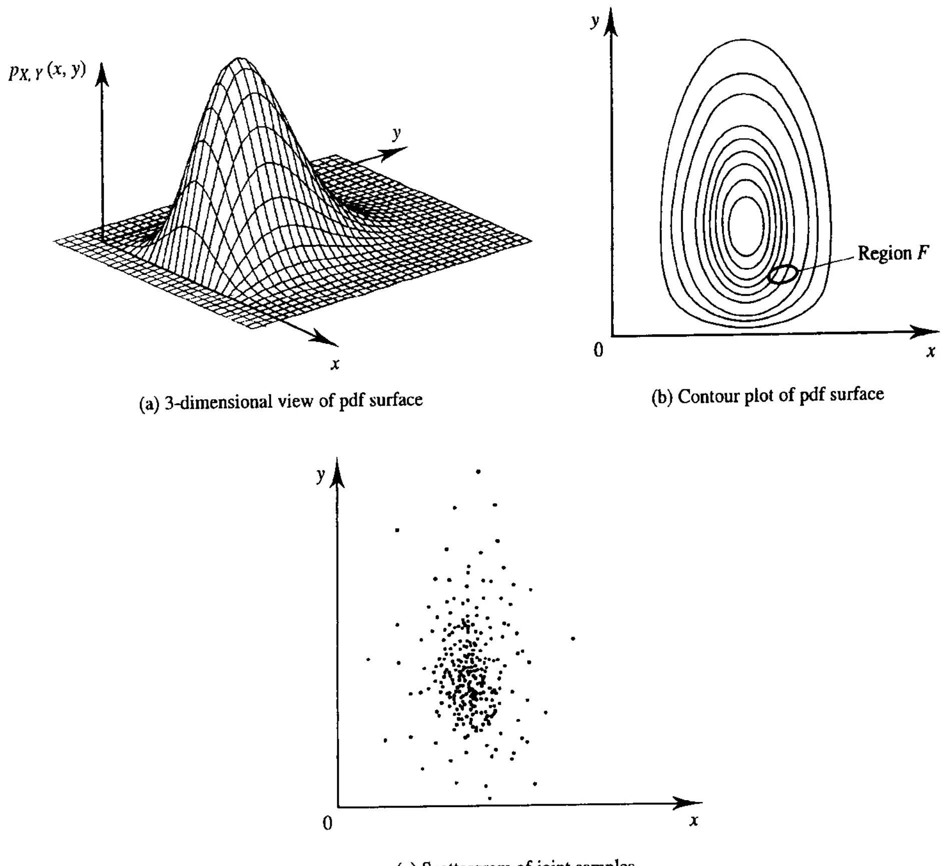



**Figure 3.13** *Representations of pdf for a bivariate joint random variable.*

$$p_Y(y) = \int_{-\infty}^{\infty} p_{X,Y}(x, y)\, dx \qquad (3.23(b))$$

---

EXAMPLE 3.7
Two quantised signals have the following (discrete) joint pdfs:

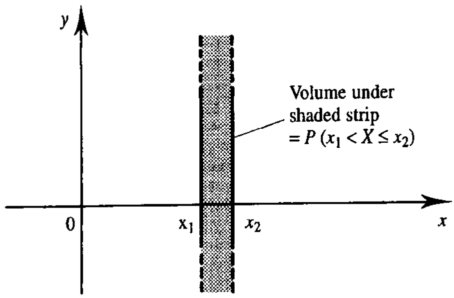



**Figure 3.14** *Relationship between a marginal probability and a joint pdf.*

| | | X | | | |
|---|---|---|---|---|---|
| | | 1.0 | 1.5 | 2.0 | 2.5 |
| Y | −1.0 | 0.15 | 0.08 | 0.06 | 0.05 |
| | −0.5 | 0.10 | 0.13 | 0.06 | 0.05 |
| | 0.0 | 0.04 | 0.07 | 0.05 | 0.05 |
| | 0.5 | 0.01 | 0.02 | 0.03 | 0.05 |

Find and sketch the marginal pdfs of $X$ and $Y$.

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x, y)\, dy$$

For a discrete joint random variable this becomes column summation:

$$P_X(x) = \sum_y P_{X,Y}(x, y)$$

$$P_X(1.0) = 0.15 + 0.10 + 0.04 + 0.01 = 0.30$$

$$P_X(1.5) = 0.08 + 0.13 + 0.07 + 0.02 = 0.30$$

$$P_X(2.0) = 0.06 + 0.06 + 0.05 + 0.03 = 0.20$$

$$P_X(2.5) = 0.05 + 0.05 + 0.05 + 0.05 = 0.20$$

See Figure 3.15(a) for the marginal pdf of $X$. Similarly:

$$P_Y(y) = \sum_x P_{X,Y}(x, y)$$

and by row summation:

$$P_Y(-1.0) = 0.15 + 0.08 + 0.06 + 0.05 = 0.34$$

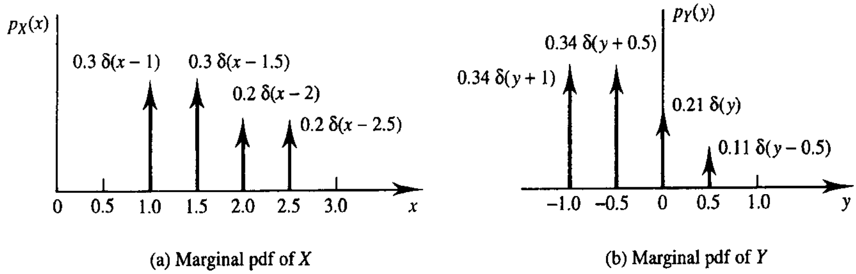



**Figure 3.15** *Marginal pdfs in Example 3.7.*

$$P_Y(-0.5) = 0.10 + 0.13 + 0.06 + 0.05 = 0.34$$

$$P_Y(0.0) = 0.04 + 0.07 + 0.05 + 0.05 = 0.21$$

$$P_Y(0.5) = 0.01 + 0.02 + 0.03 + 0.05 = 0.11$$

Figure 3.15(b) shows the marginal pdf of $Y$.

---

### 3.2.7 Joint moments, correlation and covariance

The joint moments of $p_{X,Y}(x,y)$ are defined by:

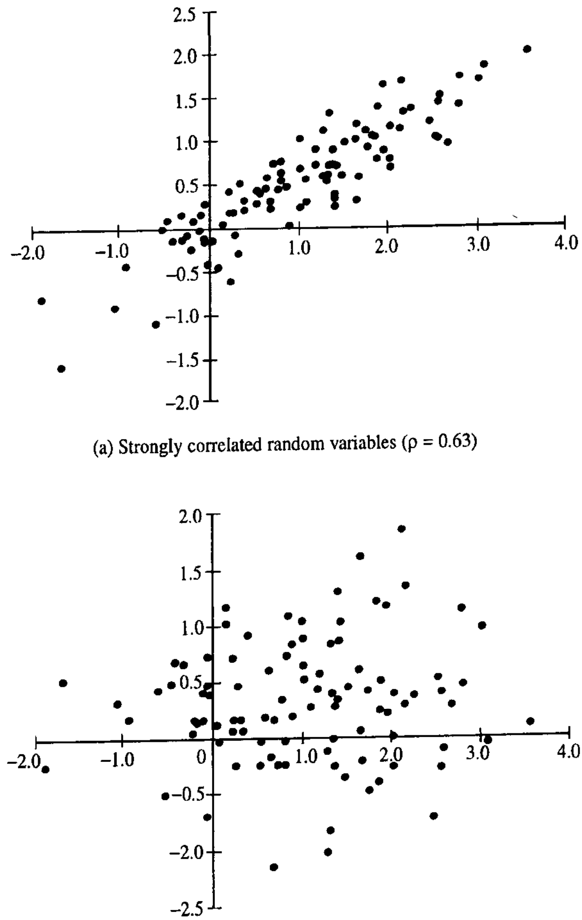
$$\overline{X^n\, Y^m} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^n\, y^m\, p_{X,Y}(x,y)\, dx\, dy \tag{3.24}$$

In the special case of $n = m = 1$ the joint moment is called the *correlation* of $X$ and $Y$ (see previous correlation definition in section 2.6 for transient and periodic signals). A large positive value of correlation means that when $x$ is high then $y$, *on average,* will also be high. A large negative value of correlation means that when $x$ is high $y$, on average, will be low. A small value of correlation means that $x$ gives little information about the magnitude or sign of $y$. The effect of correlation between two random variables, on a scattergram, is shown in Figure 3.16.

The joint central moments of $p_{X,Y}(x, y)$ are defined by:

$$\overline{(X - \bar{X})^n\, (Y - \bar{Y})^m} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \bar{X})^n\, (y - \bar{Y})^m\, p_{X,Y}(x,y)\, dx\, dy \tag{3.25}$$

In this case, when $n = m = 1$, the joint central moment is called the *covariance* of $X$ and $Y$. This is because the mean values of $X$ and $Y$ have been subtracted before *correlating* the resulting zero mean variables. Covariance therefore refers to the correlation of the *varying* parts of $X$ and $Y$. If $X$ and $Y$ are already zero mean variables then the correlation and covariance are identical.




(a) Strongly correlated random variables (ρ = 0.63)

(a) Weakly correlated random variables (ρ = 0.10)

**Figure 3.16** *Influence of correlation on a scattergram.*

The definitions of joint, and joint central, moments can be extended to multivariate (i.e. more than two) random variables in a straightforward way (e.g. $\overline{X^n Y^m Z^l}$ etc.).

The random variables $X$ and $Y$ are said to be uncorrelated if:

$$\overline{XY} = \bar{X}\,\bar{Y} \tag{3.26}$$

Notice that this implies that the *covariance* (not the correlation) is zero, i.e.:

$$\overline{(X-\bar{X})(Y-\bar{Y})} = 0 \tag{3.27}$$

It also follows that the correlation of $X$ and $Y$ can be zero, only if either $\bar{X}$ or $\bar{Y}$ is zero.

It is intuitively obvious that statistically independent random variables (i.e. random variables arising from physically separate processes) must be uncorrelated. It is not generally the case, however, that uncorrelated random variables must be statistically independent. Indeed the concept of correlation can be applied to deterministic signals such as $\cos \omega t$ and $\sin \omega t$. If concurrent samples are taken from these functions and the correlation is subsequently calculated the result will be zero (due to their orthogonality). It is clearly untrue to say that these are independent processes, however, since one can be derived from the other using a simple delay line. It follows that independence is a stronger statistical condition than uncorrelatedness, i.e.:

Independence => uncorrelatedness
Uncorrelatedness ≠> independence

(An exception to the latter rule is when the random variables are Gaussian. In this special case uncorrelatedness does imply statistical independence: see section 3.2.8.)

The normalised correlation coefficient, $\rho$, between two random variables (as used in section 2.6, for transient or periodic signals) is the correlation between the corresponding standardised variables, standardisation in this context implying zero mean and unit standard deviation, i.e.:

$$\rho = \overline{\left(\frac{X - \bar{X}}{s_X}\right)\left(\frac{Y - \bar{Y}}{s_Y}\right)} \tag{3.28}$$

where $s_X$ and $s_Y$ are the standard deviations of $X$ and $Y$ respectively. (Since $\rho$ can also be interpreted as the covariance of the random variables with normalised standard deviation then $\rho = 0$ can be viewed as the defining property for uncorrelatedness.)

---

EXAMPLE 3.8
Find the correlation and covariance of the discrete joint pdf described in Example 3.7.

Correlation is defined in equation (3.24) as:

$$\overline{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \; p_{X,Y}(x, y)\, dx\, dy$$

For a discrete joint pdf this becomes:

$$\overline{XY} = \sum_x \sum_y xy \; P_{X,Y}(x, y)$$

$$\begin{aligned} &= (1.0)\,(-1.0)\,0.15 + (1.0)\,(-0.5)\,0.10 + (1.0)\,(0.0)\,0.04 + (1.0)\,(0.5)\,0.01 \\ &\quad + (1.5)\,(-1.0)\,0.08 + (1.5)\,(-0.5)\,0.13 + (1.5)\,(0.0)\,0.07 + (1.5)\,(0.5)\,0.02 \\ &\quad + (2.0)\,(-1.0)\,0.06 + (2.0)\,(-0.5)\,0.06 + (2.0)\,(0.0)\,0.05 + (2.0)\,(0.5)\,0.03 \\ &\quad + (2.5)\,(-1.0)\,0.05 + (2.5)\,(-0.5)\,0.05 + (2.5)\,(0.0)\,0.05 + (2.5)\,(0.5)\,0.05 \\ &= -0.6725 \end{aligned}$$

Similarly covariance is obtained from equation (3.25) for the discrete pdf as:

$$\overline{(X-\bar{X})(Y-\bar{Y})} = \sum_x \sum_y (x-\bar{X})(y-\bar{Y})\, P_{X,Y}(x,y)$$

Using the marginal pdfs found in Example 3.7:

$$\begin{aligned}\bar{X} &= \frac{1}{N}\sum_x x\, P_X(x)\\ &= \frac{1}{4}\,[(1.0)(0.3)+(1.5)(0.3)+(2.0)(0.2)+(2.5)(0.20)]\\ &= 0.4125\\ \bar{Y} &= \frac{1}{N}\sum_y y\, P_Y(y)\\ &= \frac{1}{4}\,[(-1.0)(0.34)+(-0.5)(0.34)+(0.0)(0.21)+(0.5)(0.11)]\\ &= -0.4550\end{aligned}$$

and by using each of the 16 discrete probabilities in Example 3.7:

$$\begin{aligned}\overline{(X-\bar{X})(Y-\bar{Y})} &= (1.0-0.4125)(-1.0+0.4550)\,0.15\\ &+ (1.0-0.4125)(-0.5+0.4550)\,0.10\\ &+ (1.0-0.4125)(0.0+0.4550)\,0.04\\ &+ (1.0-0.4125)(0.5+0.4550)\,0.01\\ &+ (1.5-0.4125)(-1.0+0.4550)\,0.08\\ &+ (1.5-0.4125)(-0.5+0.4550)\,0.13\\ &+ (1.5-0.4125)(0.0+0.4550)\,0.07\\ &+ (1.5-0.4125)(0.5+0.4550)\,0.02\\ &+ (2.0-0.4125)(-0.1+0.4550)\,0.06\\ &\quad\cdots \qquad \cdots\\ &+ (2.5-0.4125)(0.5+0.4550)\,0.05\\ &= 0.07825\end{aligned}$$

---

### 3.2.8 Joint Gaussian random variables

A bivariate random variable is Gaussian if it can be reduced, by a suitable translation and rotation of axes, to the form:

$$p_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y}\, e^{-\left[\frac{x^2}{2\sigma_X^2}+\frac{y^2}{2\sigma_Y^2}\right]} \tag{3.29}$$

This idea is illustrated in Figure 3.17. The contours in the $x$, $y$ plane, given by:

$$\frac{x''^2}{2\sigma_{X''}^2} + \frac{y''^2}{2\sigma_{Y''}^2} = \text{constant} \tag{3.30}$$

are ellipses and the double primed and unprimed coordinate systems are related by:

$$x'' = (x - x_0)\cos\theta + (y - y_0)\sin\theta \tag{3.31(a)}$$

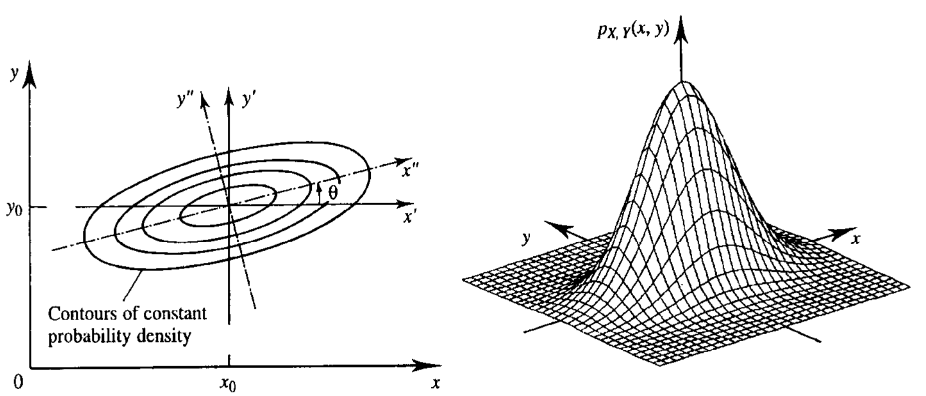
$$y'' = -(x - x_0)\sin\theta + (y - y_0)\cos\theta \tag{3.31(b)}$$

where $x_0$, $y_0$ are the necessary translations and $\theta$ is the necessary rotation. If $\sigma_{X''} = \sigma_{Y''}$ then the ellipses become circles. The translation of axes removes any DC component in the random variables and the rotation has the effect of reducing the correlation between the random variables to zero. This can be seen by writing the probability density function in the original coordinate system, i.e.:

$$p_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{(1-\rho^2)}} \times e^{-\left[\frac{(x-\bar{X})^2}{2\sigma_X^2(1-\rho^2)} - 2\rho\frac{(x-\bar{X})(y-\bar{Y})}{2\sigma_X\sigma_Y(1-\rho^2)} + \frac{(y-\bar{Y})^2}{2\sigma_Y^2(1-\rho^2)}\right]} \tag{3.32}$$

(for $\bar{X} = \bar{Y} = \rho = 0$ this reduces to equation (3.29)) which can then be written as a product of separate functions in $x$ and $y$, i.e.:

$$p_{X,Y}(x, y) = \frac{1}{\sqrt{(2\pi)}\sigma_X} e^{-\frac{x^2}{2\sigma_X^2}} \frac{1}{\sqrt{(2\pi)}\sigma_Y} e^{-\frac{y^2}{2\sigma_Y^2}} = p_X(x)\, p_Y(y) \tag{3.33}$$




(a) Contour plot of pdf

(b) 3-dimensional $p_{X,Y}(x, y)$ surface after translation to remove DC components and rotation to remove correlation

**Figure 3.17** *Joint Gaussian bivariate random variable.*

Equation (3.33) is the necessary and sufficient condition for statistical independence of $X$ and $Y$. For a multivariate Gaussian function, then, uncorrelatedness does imply independence.

### 3.2.9 Addition of random variables and the central limit theorem

If two random variables $X$ and $Y$ are added, Figure 3.18, and their joint pdf, $p_{X,Y}(x,y)$ is known, what is the pdf, $p_Z(z)$, of their sum, $Z$? To answer this question we note the following:

$$Z = X + Y \tag{3.34}$$

Therefore when $Z = z$ (i.e. $Z$ takes on a particular value $z$) then:

$$Y = z - X \tag{3.35(a)}$$

and when $Z = z + dz$ then:

$$Y = z + dz - X \tag{3.35(b)}$$

Equations (3.35), which both represent straight lines in the $X,Y$ plane, are sketched in Figure 3.19. The probability that $Z$ lies in the range $z$ to $z + dz$ is given by the volume contained under $p_{X,Y}(x,y)$ in the strip between these two lines, i.e.:

$$P(z < Z \leq z + dz) = \int_{strip} p_{X,Y}(x, y)\, ds \tag{3.36(a)}$$

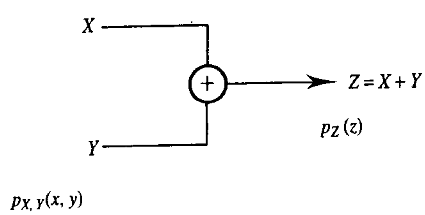



**Figure 3.18** *Addition of random variables.*

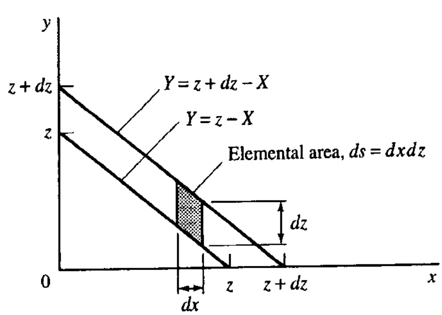



**Figure 3.19** *Strip of integration in XY plane to find $P(z < Z \leq z + dz)$.*

or:

$$p_Z(z)\,dz = \int_{strip} p_{X,Y}(x, z-x)\,dx\,dz \qquad (3.36(b))$$

Therefore:

$$p_Z(z) = \int_{-\infty}^{\infty} p_{X,Y}(x, z-x)\,dx \qquad (3.37)$$

If $X$ and $Y$ are statistically independent then:

$$p_{X,Y}(x, z-x) = p_X(x)\;p_Y(z-x) \qquad (3.38)$$

and equation (3.37) becomes:

$$p_Z(z) = \int_{-\infty}^{\infty} p_X(x) p_Y(z-x)\,dx \qquad (3.39)$$

Equation (3.39) can be recognised as the convolution integral. *The pdf of the sum of independent random variables is therefore the convolution of their individual pdfs.*

The multiple convolution of pdfs which arises when many independent random variables are added has a surprising and important consequence. Since convolution is essentially an integral operation it almost always results in a function which is in some

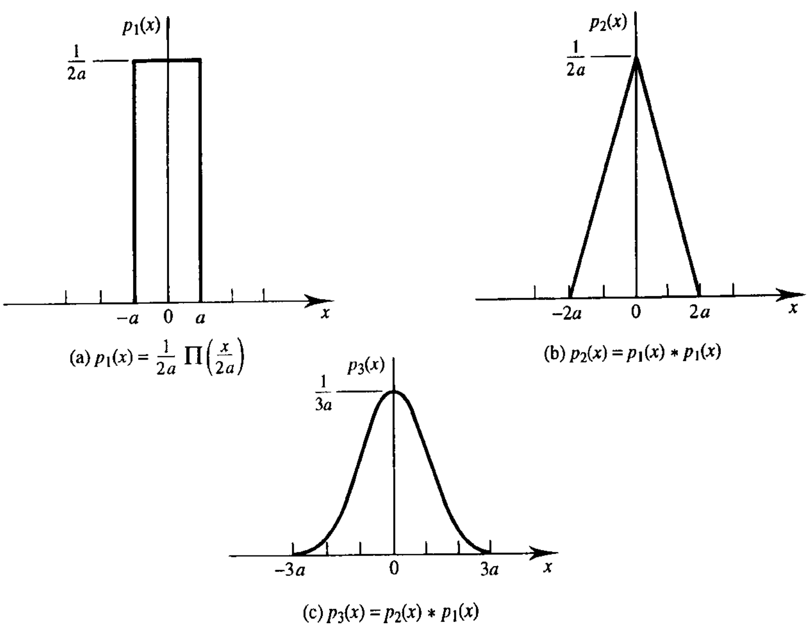



**Figure 3.20** *Multiple self convolution of a rectangular pulse.*

sense smoother (i.e. more gradually varying) than either of the functions being convolved. (This is true provided that the original functions are reasonably smooth which excludes, for instance, the case of impulse functions.) After surprisingly few convolutions this repeated smoothing results in a distribution which approximates a Gaussian function. The approximation gets better as the number of convolutions increases. The tendency for multiple convolutions to give rise to Gaussian functions is called the *central limit theorem* and accounts for the ubiquitous nature of Gaussian noise. It is illustrated for multiple self convolution of a rectangular pulse in Figure 3.20. In the context of statistics the central limit theorem can be stated as follows:

> *If N statistically independent random variables are added, the sum will have a probability density function which tends to a Gaussian function as N tends to infinity, irrespective of the original random variable pdfs.*

A second consequence of the central limit theorem is that the pdf of the *product* of $N$ independent random variables will tend to a log-normal distribution as $N$ tends to infinity, since multiplication of functions corresponds to addition of their logarithms.

If two *Gaussian* random variables are added their sum will also be a Gaussian random variable. In this case the result is exact and holds even if the random variables are correlated. The mean and variance of the sum are given by:

$$\bar{Z} = \bar{X} + \bar{Y} \tag{3.40}$$

and:

$$\sigma_{X\pm Y}^2 = \sigma_X^2 \pm 2\rho\sigma_X\sigma_Y + \sigma_Y^2 \tag{3.41}$$

For uncorrelated (and therefore independent) Gaussian random variables the variances, like the means, are simply added. This is an especially easy case to prove since for independent variables the pdf of the sum is the convolution of two Gaussian functions. This is equivalent to multiplying the Fourier transforms of the original pdfs and then inverse Fourier transforming the result. (The Fourier transform of a pdf is called the *characteristic function* of the random variable.) When a Gaussian pdf is Fourier transformed the result is a Gaussian characteristic function. When Gaussian characteristic functions are multiplied the result remains Gaussian ($e^{-x^2}e^{-x^2} = e^{-2x^2}$). Finally when the Gaussian product is inverse Fourier transformed the result is a Gaussian pdf.

---

EXAMPLE 3.9

$x(t)$ and $y(t)$ are zero mean Gaussian random currents. When applied individually to 1 Ω resistive loads they dissipate 4.0 W and 1.0 W of power respectively. When both are applied to the load simultaneously the power dissipated is 3.0 W. What is the correlation between $X$ and $Y$?

Since $X$ and $Y$ have zero mean their variance is equal to their normalised power, i.e.:

$$\sigma_X^2 = 4.0 \quad \text{and} \quad \sigma_Y^2 = 1.0$$

Their standard deviations are therefore:

$$\sigma_X = 2.0 \quad \text{and} \quad \sigma_Y = 1.0$$

Using equation (3.41):

$$\rho = \frac{\sigma_{X\pm Y}^2 - \sigma_X^2 - \sigma_Y^2}{\pm 2\, \sigma_X\, \sigma_Y}$$

We take the positive sign in the denominator since it is the sum (not difference) which dissipates 3.0 W:

$$\rho = \frac{3.0 - 4.0 - 1.0}{2\,(2.0)\,(1.0)} = -0.5$$

---

EXAMPLE 3.10
Two independent random voltages have a uniform pdf given by:

$$p(V) = \begin{cases} 0.5, & |V| \leq 1.0 \\ 0, & |V| > 1.0 \end{cases}$$

Find the pdf of their sum.

$$p(V_1 + V_2) = p_1(V) * p_2(V)$$

Using $\chi$ to represent the characteristic function, i.e.-

$$\chi(W) = \text{FT}\,\{p(V)\}$$

then from the convolution theorem (Table 2.5) and the Fourier transform of a rectangular function (Table 2.4):

$$\begin{aligned} \chi(W) &= \chi_1(W)\, \chi_2(W) \\ &= 0.5\,[2\,\text{sinc}(2W)]\; 0.5\,[2\,\text{sinc}(2W)] \\ &= \text{sinc}^2(2W) \end{aligned}$$

The pdf of the sum is then found by inverse Fourier transforming $\chi(W)$, i.e.:

$$\begin{aligned} p(V_1 + V_2) &= 0.5\ \Lambda\left(\frac{V}{2}\right) \\ &= \begin{cases} 0.5\left(1 - \dfrac{|V|}{2}\right), & |V| \leq 2.0 \\ 0, & |V| > 2.0 \end{cases} \end{aligned}$$

---

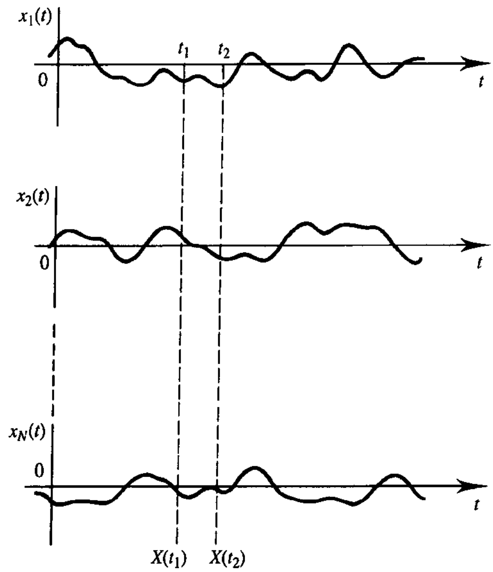



**Figure 3.21** *Random process, $X(t)$, as ensemble of sample functions, $x_i(t)$.*

## 3.3 Random processes

The term random process usually refers to a random variable which is a function of time (or occasionally a function of position) and is strictly defined in terms of an ensemble (i.e. collection) of time functions, Figure 3.21. Such an ensemble of functions may, in principle, be generated using many sets (perhaps an infinite number) of identical sources. The following notation for random processes is adopted here:

1. The random process (i.e. the entire ensemble of functions) is denoted by $X(t)$.
2. $X(t_1)$ or $X_1$ denotes an ensemble of samples taken at time $t_1$ and constitutes a random variable.
3. $x_i(t)$ is the $i$th sample function of the ensemble.

It is often the case, in practice, that only one sample function can be observed, the other sample functions representing what might have occurred (given the statistical properties of the process) but didn't. It is also the case that each sample function $x_i(t)$ is *usually* a random function of time although this does not have to be so. (For example $X(t)$ may be a set of sinusoids each sample function having random phase.)

Random processes, like other types of signal, can be classified in a number of different ways. For example, they may be:

1. Continuous or discrete.
2. Analogue or digital (or mixed).
3. Deterministic or non-deterministic.

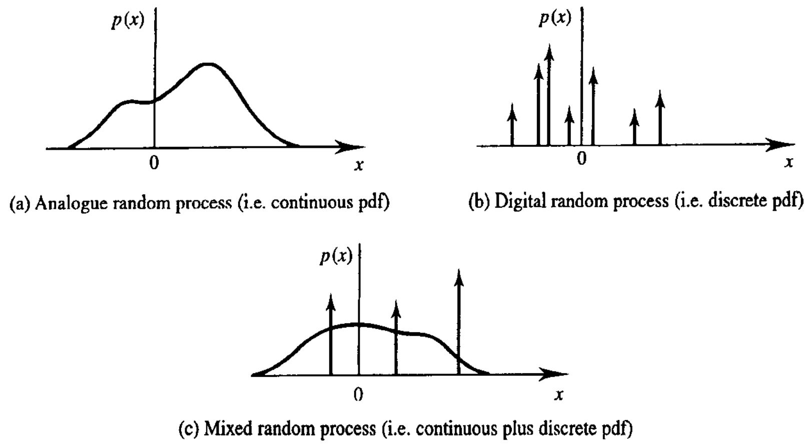



**Figure 3.22** *Pdfs of analogue, digital and mixed random processes.*

4. Stationary or non-stationary.
5. Ergodic or non-ergodic.

The first category refers to continuity or discreteness in time or position. (Discrete time signals are also sometimes called a time series.) The second category could be (and sometimes is) referred to as continuous, discrete or mixed which in this context describes the pdf of the process, Figure 3.22. A deterministic random process seems, superficially, to be a contradiction in terms. It describes a process, however, in which each sample function is deterministic. An example of such a process, $x_i(t) = \sin(\omega t + \theta_i)$ where $\theta_i$ is a random variable with specified pdf, has already been given. Stationarity and ergodicity are concepts which are central to random processes and they are therefore discussed in some detail below.

### 3.3.1 Stationarity and ergodicity

Stationarity relates to the time independence of a random process's statistics. There are two definitions:

(a) A random process is said to be stationary in the *strict* (sometimes called narrow) sense if all its pdfs (joint, conditional and marginal) are the same for any value of $t$, i.e. if none of its statistics change with time.

(b) A random process is said to be stationary in the *loose* (sometimes called wide) sense if its mean value, $\overline{X(t)}$, is independent of time, $t$, and the correlation, $\overline{X(t_1)X(t_2)}$, depends only on time difference $\tau = t_2 - t_1$.

Ergodicity relates to the equivalence of ensemble and time averages. It implies that each sample function, $x_i(t)$, of the ensemble has the same statistical behaviour as any set of ensemble values, $X(t_j)$, Figure 3.23. Thus for an ergodic process:

$$\langle x_i^n(t) \rangle = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x_i^n(t)\, dt$$

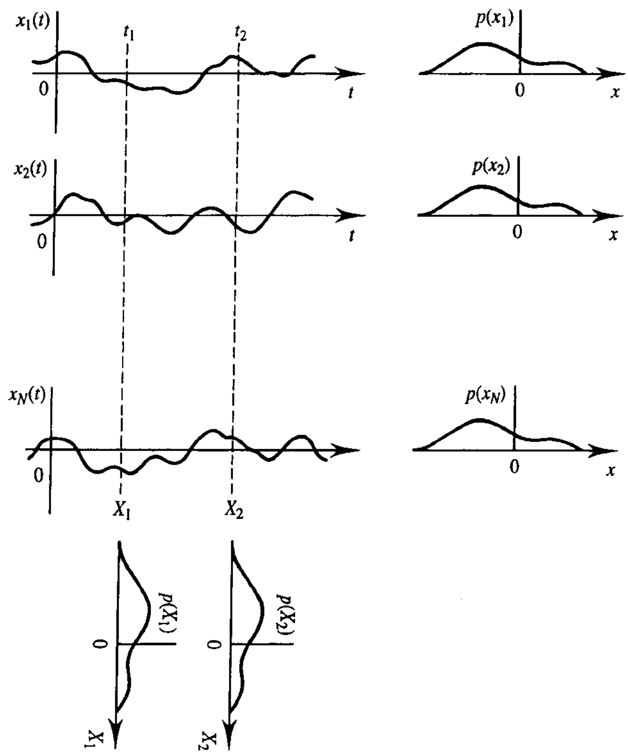



**Figure 3.23** *Identity of sample function pdfs, $p(x_i)$, and ensemble random variable pdfs $p(X_i)$, for an ergodic process.*

$$= \int_{-\infty}^{\infty} X_j^n \, p(X_j) \, dx$$

$$= \overline{X^n(t_j)}, \quad \text{for } any \ i \text{ and } j \tag{3.42}$$

It is obvious that an ergodic process must be statistically stationary. The converse is not true, however, i.e. stationary processes need not be ergodic. Ergodicity is therefore a stronger (more restrictive) condition on a random process than stationarity, i.e.:

Ergodicity => stationarity
Stationarity ≠> ergodicity

### 3.3.2 Strict and loose sense Gaussian processes

A sample function, $x_i(t)$, is said to belong to a Gaussian random process, $X(t)$, in the *strict* sense if the random variables $X_1 = X(t_1)$, $X_2 = X(t_2)$, $\cdots$, $X_N = X(t_N)$ have an $N$-dimensional joint Gaussian pdf, Figure 3.24. For an ergodic process the strict sense

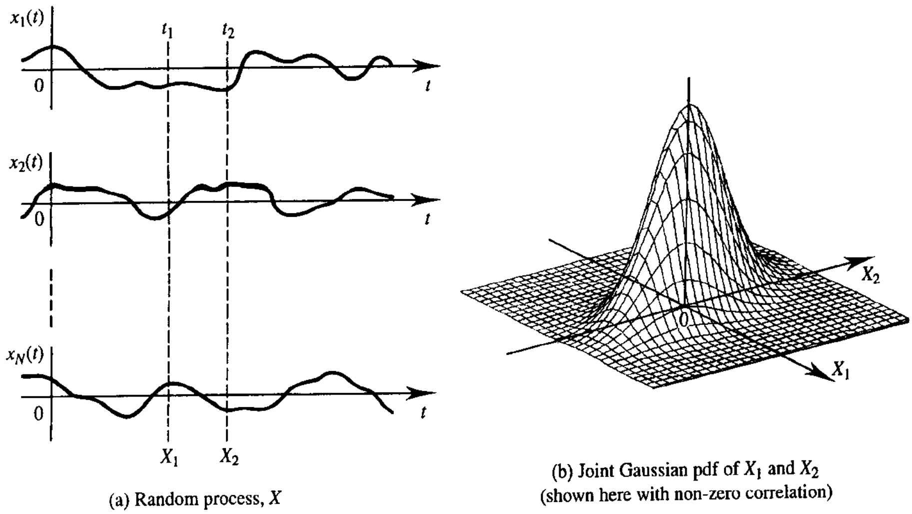
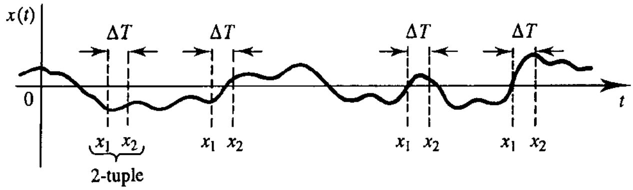



**Figure 3.24** *Example (drawn for $N = 2$) of the joint Gaussian pdf of random variables ($X_1$ and $X_2$) taken from a strict sense Gaussian process.*

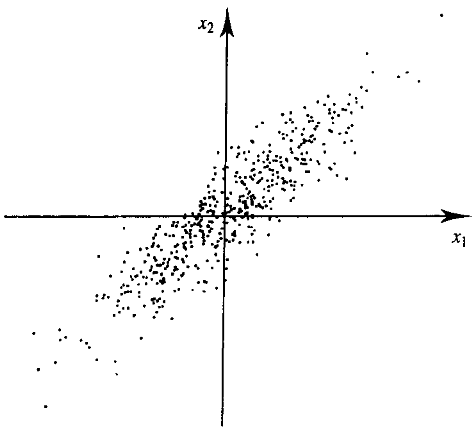
Gaussian condition can be defined in terms of a single sample function. In this case if the joint pdf of multiple sets of $N$-tuple samples, taken with fixed time intervals between the samples of each $N$-tuple, is $N$-variate Gaussian then the process is Gaussian in the strict sense. This definition is illustrated in Figures 3.25(a) and (b) for multiple sets of sample pairs (i.e. $N = 2$).

A sample function, $x_i(t)$, is said to belong to a Gaussian random process in the *loose* sense if isolated samples taken from $x_i(t)$ come from a Gaussian pdf, Figure 3.26. The following points can be made about strict and loose sense Gaussian processes:

1. Being a strict sense Gaussian process is a very strong statistical condition, much stronger than being a loose sense Gaussian process. All strict sense Gaussian processes are, therefore, also loose sense Gaussian processes.
2. Examples do exist of processes which are Gaussian in the loose sense but not the strict sense. They are rare in practice, however.
3. A strict sense Gaussian process is the most structureless, random, or unpredictable statistical process possible. It is also one of the most important processes in the context of communications since it describes thermal noise which is present to some degree in all practical systems.
4. A strict sense $N$-dimensional Gaussian pdf is specified completely by its first and second order moments, i.e. its means, variances and covariances, as all higher moments (Figure 3.9) of the Gaussian pdf are zero.




(a) $N$-tuple ($N = 2$) samples with constant sample separation ($\Delta T$) taken at random times from a sample function, $x(t)$, of the random process $X(t)$




(b) Joint, $N$-variate ($N = 2$), Gaussian scattergram for $p(x_1, x_2)$

**Figure 3.25** *Single sample function definition of ergodic, strict sense Gaussian, random process.*

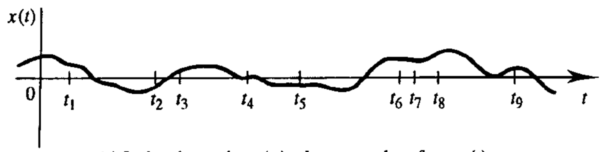
### 3.3.3 Autocorrelation and power spectral density

A simple pdf is obviously insufficient to fully describe a random signal (i.e. a sample function from a random process) because it contains no information about the signal's rate of change.

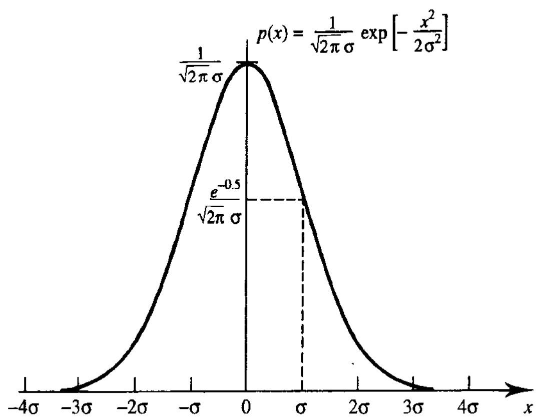
Such information would be available, however, in the joint pdfs, $p(X_1, X_2)$, of random variables, $X(t_1)$ and $X(t_2)$, separated by $\tau = t_2 - t_1$. These joint pdfs are not usually known in full but partial information about them is often available in the form of the correlation, $\overline{X(t_2)X(t_2 - \tau)}$. For ergodic signals the ensemble average taken at any time is equal to the temporal average of any sample function, i.e.:

$$\begin{aligned}\overline{X(t)\,X(t-\tau)} &= \langle x(t)\,x(t-\tau)\rangle \\ &= \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} x(t)\,x(t-\tau)\,dt \\ &= R_x(\tau)\ (\mathrm{V}^2) \qquad (3.43)\end{aligned}$$

The autocorrelation function, $R_x(\tau)$, of a sample function, $x(t)$, taken from a real,




(a) Isolated samples $x(t_n)$ taken at random from $x(t)$




(b) Gaussian distribution of samples from (a)

**Figure 3.26** *Definition of a loose sense Gaussian process.*

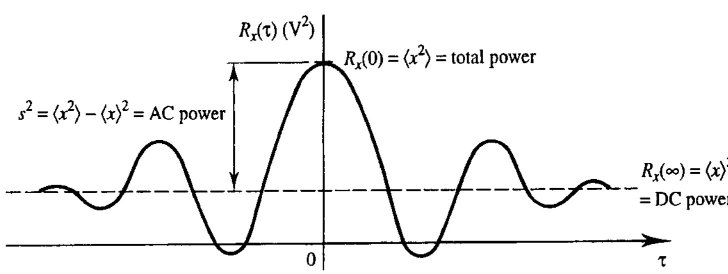



**Figure 3.27** *General behaviour of $R_x(\tau)$ for a random process.*

ergodic, random process has the following properties:

1. $R_x(\tau)$ is real.
2. $R_x(\tau)$ has even symmetry (see Figure 3.27), i.e.:

$$R_x(-\tau) = R_x(\tau) \tag{3.44}$$

3. $R_x(\tau)$ has a maximum (positive) magnitude at $\tau = 0$ which corresponds to the mean square value of (or normalised power in) $x(t)$, i.e.:

$$\langle x^2(t)\rangle \;=\; R_x(0) > |R(\tau)|, \qquad \text{for } all\ \tau \neq 0 \tag{3.45}$$

4. If $x(t)$ has units of V then $R_x(\tau)$ has units of $V^2$ (i.e. normalised power).
5. $R_x(\infty)$ is the square mean value of (or normalised DC power in) $x(t)$, i.e.:

$$R_x(\infty) = \langle x(t)\rangle^2 \tag{3.46}$$

6. $R_x(0) - R_x(\infty)$ is the variance, $s^2$, of $x(t)$, i.e.:

$$R_x(0) - R_x(\infty) = \langle x^2(t)\rangle - \langle x(t)\rangle^2 = s^2 \tag{3.47}$$

7. The autocorrelation function and *two sided* power spectral density of $x(t)$ form a Fourier transform pair, i.e.:

$$R_x(\tau) \overset{\text{FT}}{\Leftrightarrow} G_x(f) \tag{3.48}$$

(This is the Wiener-Kintchine theorem [Papoulis] which, although not proved here, can be readily accepted since a similar theorem has been proved in Chapter 2 for transient signals.) Properties of the corresponding power spectral density (most of which are corollaries of the above) include the following:

1. $G_x(f)$ has even symmetry about $f = 0$, i.e.:

$$G_x(-f) = G_x(f) \tag{3.49}$$

2. $G_x(f)$ is real.
3. The area under $G_x(f)$ is the mean square value of (or normalised power in) $x(t)$, i.e.:

$$\int_{-\infty}^{\infty} G_x(f)\, df \;=\; \langle x^2(t)\rangle \tag{3.50}$$

4. If $x(t)$ has units of V then $G_x(f)$ has units of $V^2$/Hz.
5. The area under any impulse in $G_x(f)$ occurring at $f = 0$ is the square mean value of (or normalised DC power in) $x(t)$, i.e.:

$$\int_{0-}^{0+} G_x(f)\, df \;=\; \langle x(t)\rangle^2 \tag{3.51}$$

6. The area under $G_x(f)$, excluding any impulse function at $f = 0$, is the variance of $x(t)$ or the normalised power in the fluctuating component of $x(t)$, i.e.:

$$\int_{-\infty}^{0-} G_x(f)\, df + \int_{0+}^{\infty} G_x(f)\, df \;=\; \langle x^2(t)\rangle - \langle x(t)\rangle^2 \tag{3.52}$$

7. $G_x(f)$ is positive for all $f$, i.e.:

$$G_x(f) \geq 0, \qquad \text{for } all\ f \tag{3.53}$$

(White noise is a random signal with particularly extreme spectral and autocorrelation properties. It has no self similarity with any time shifted version of itself so its autocorrelation function consists of a single impulse at zero delay, and its power spectral density is flat.)

A normalised autocorrelation function, $\rho_x(\tau)$, can be defined by subtracting any DC value present in $x(t)$, dividing by the resulting RMS value and autocorrelating the result. This is equivalent to:

$$\rho_x(\tau) = \frac{\langle x(t)x(t-\tau) - \langle x(t)\rangle^2\rangle}{\langle x^2(t)\rangle - \langle x(t)\rangle^2} \tag{3.54}$$

The normalised function, Figure 3.28, is clearly an extension of the normalised correlation coefficient (equation (3.28)) and has the properties:

$$\rho_x(0) = 1 \tag{3.55(a)}$$

$$\rho_x(\pm\infty) = 0 \tag{3.55(b)}$$

It can be interpreted as the fraction of $x(t-\tau)$ which is contained in $x(t)$ neglecting DC components. This is easily demonstrated as follows:

Let $f(t)$ be a zero mean stationary random process, i.e.:

$$f(t) = x(t) - \langle x(t)\rangle \tag{3.56}$$

If the new function:

$$g(t) = f(t) - \rho f(t-\tau) \tag{3.57}$$

is formed then the value of $\rho$ which minimises $\langle g^2(t)\rangle$ will be the fraction of $f(t-\tau)$ contained in $f(t)$. Expanding $\langle g^2(t)\rangle$:

$$\begin{aligned}\langle g^2(t)\rangle &= \langle [f(t) - \rho f(t-\tau)]^2\rangle \\ &= \langle f^2(t) - 2\rho f(t)f(t-\tau) + \rho^2 f^2(t-\tau)\rangle \\ &= \langle f^2(t)\rangle - 2\rho\langle f(t)f(t-\tau)\rangle + \rho^2\langle f^2(t-\tau)\rangle\end{aligned} \tag{3.58}$$

The value of $\rho$ which minimises $\langle g^2(t)\rangle$ is found by solving $d\langle g^2(t)\rangle/d\rho = 0$, i.e.:

$$0 - 2\,R_f(\tau) + 2\rho\,\langle f^2(t-\tau)\rangle = 0 \tag{3.59}$$

giving:

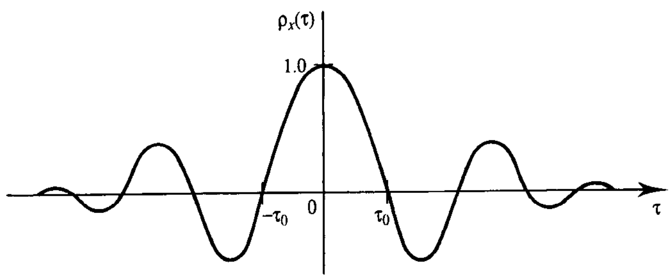



**Figure 3.28** *General behaviour of $\rho_x(\tau)$ for a random process. (Shows first null definition of decorrelation time, $\tau_0$.)*

$$\rho = \frac{R_f(\tau)}{\langle f^2(t)\rangle} = \frac{\langle [x(t) - \langle x(t)\rangle][x(t-\tau) - \langle x(t)\rangle]\rangle}{\langle [x(t) - \langle x(t)\rangle]^2\rangle}$$

$$= \frac{\langle x(t)x(t-\tau)\rangle - \langle x(t)\rangle^2}{\langle x^2(t)\rangle - \langle x(t)\rangle^2} \tag{3.60}$$

---

EXAMPLE 3.11

Find and sketch the autocorrelation function of the stationary random signal whose power spectral density is shown in Figure 3.29(a).

Using the triangular function, $\Lambda$, with $f$ measured in Hz:

$$G_x(f) = 3.0\left[\Lambda\left(\frac{f-5000}{1000}\right) + \Lambda\left(\frac{f+5000}{1000}\right)\right]$$

$$= 3.0\,\Lambda\left(\frac{f}{1000}\right) * [\delta(f-5000) + \delta(f+5000)]$$

$$R_x(\tau) = \mathrm{FT}^{-1}\{G_x(f)\}$$

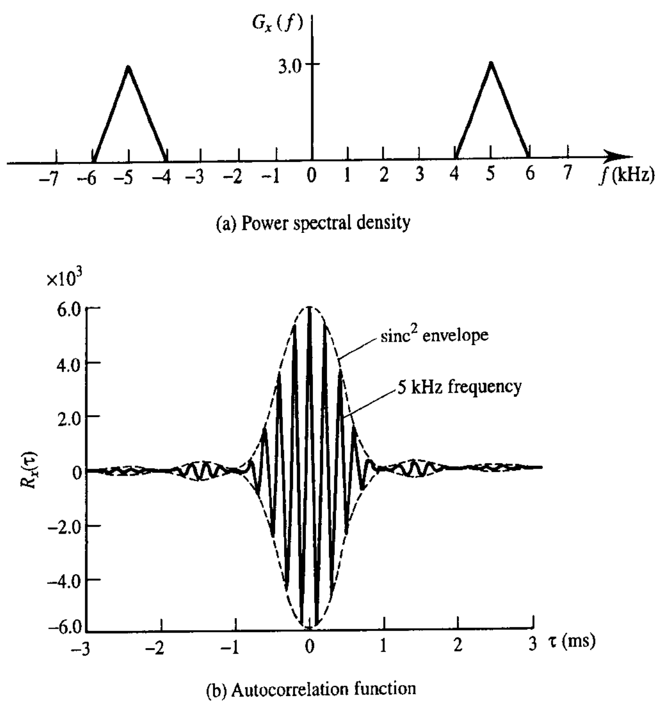



**Figure 3.29** *Spectral and temporal characteristics of stationary random signal, Example 3.12.*

$$= 3.0\,\mathrm{FT}^{-1}\left\{\Lambda\left(\frac{f}{1000}\right)\right\}\mathrm{FT}^{-1}\,\{\delta(f-5000)+\delta(f+5000)\}$$

Using Tables 2.4 and 2.5:

$$R_x(\tau) = 3.0 \times 1000\,\mathrm{sinc}^2(1000\tau)\,2\cos(2\pi\,5000\tau)$$

Figure 3.29(b) shows a sketch for the solution of Example 3.11.

---

### 3.3.4 Signal memory, decorrelation time and white noise

It is physically obvious that practical signals must have a finite memory, i.e. samples taken close enough together must be highly correlated. The decorrelation time, $\tau_0$, of a signal provides a quantitative measure of this memory and is defined as the minimum time shift, $\tau$, required to reduce $\rho_x(\tau)$ to some predetermined, or reference, value, Figure 3.28. The reference value depends on the application and/or preference and can be somewhat arbitrary in the same way as the definition of bandwidth, $B$ (section 2.2.5). Popular choices, however, are $\rho_x(\tau_0) = 1/\sqrt{2}$, 0.5, $1/e$ and 0. Due to the Wiener-Kintchine theorem there is clearly a relationship between $B$ and $\tau_0$, i.e.:

$$B \propto \frac{1}{\tau_0}\ \mathrm{Hz} \tag{3.61}$$

(The constant of proportionality depends on the exact definitions adopted but for reasonably consistent choices is of the order of unity.) Equation (3.61) requires a careful interpretation if the random signal has a passband spectrum (see Example 3.12).

For a random signal or noise with a white power spectral density, equation (3.61) implies that $\tau_0 = 0$, i.e. that the signal has zero memory. In particular the autocorrelation function of white noise will be impulsive, i.e.:

$$R_x(\tau) = C\delta(\tau) \tag{3.62}$$

This means that adjacent samples taken from a white noise process are uncorrelated no matter how closely the samples are spaced. As this is physically impossible it means that white noise, whilst important and useful conceptually, is not practically realisable. (The same conclusion is obvious when considering the total power in a white noise process.)

The common assumption of white, Gaussian, noise processes sometimes gives the impression that Gaussianness and whiteness are connected. This is not true. Noise may be Gaussian or white, or both, or neither. If noise is Gaussian *and* white (and thermal noise, for example, is often modelled in this way) then the fact that adjacent samples from the process are uncorrelated (irrespective of separation) means that they are also independent.

---

EXAMPLE 3.12

Stating the definition you use, find the decorrelation time of the random signal described in Example 3.11.

Referring to Figure 3.29(b) and using $\rho_x(\tau_0) = 0$, the decorrelation time of $x(t)$ in Example 3.11 is 0.25 cycle of the 5 kHz signal, i.e.:

$$\rho_x(\tau_0) = 0.25\,T = 0.25\,\frac{1}{f_c} = \frac{0.25}{5000}$$
$$= 5 \times 10^{-5}\ \text{s} \quad (\text{or } 50\ \mu\text{s})$$

(Note that the decorrelation time of the *envelope* of $R_x(\tau)$ in Figure 3.29(b) is 1.0 ms and it is this quantity which is of the order of the reciprocal of the bandwidth.)

---

### 3.3.5 Cross correlation of random processes

The cross correlation of functions, taken from two *real* ergodic random processes is:

$$R_{xy}(\tau) = \langle x(t)y(t-\tau)\rangle$$
$$= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)y(t-\tau)\,dt \qquad (3.63)$$

Some of the properties of this cross correlation function are:

1. $R_{xy}(\tau)$ is real. (3.64)
2. $R_{xy}(-\tau) = R_{yx}(\tau)$ (3.65)

   (Note that, in general, $R_{xy}(-\tau) \neq R_{xy}(\tau)$.)
3. $[R_x(0)R_y(0)]^{1/2} > |R_{xy}(\tau)|$ for *all* $\tau$. (3.66)

   (Note that the maximum value of $R_{xy}(\tau)$ can occur anywhere.)
4. If $x(t)$ and $y(t)$ have units of V the $R_{xy}(\tau)$ has units of $V^2$ (i.e normalised power) and, for this reason, it is sometimes called a *cross-power*.
5. $[R_x(0) + R_y(0)]/2 > |R_{xy}(\tau)|$, for *all* $\tau$. (3.67)

   (This follows from property 3 since the geometric mean of two real numbers cannot exceed their arithmetic mean.)
6. For *statistically independent* random processes:

   $$R_{xy}(\tau) = R_{yx}(\tau) \qquad (3.68(a))$$

   and if either process has zero mean then:

   $$R_{xy}(\tau) = R_{yx}(\tau) = 0 \quad \text{for } all\ \tau. \qquad (3.68(b))$$

7. The Fourier transform of $R_{xy}(\tau)$ is often called a cross-power spectral density, $G_{xy}(f)$, since its units are $V^2$/Hz:

$$R_{xy}(\tau) \overset{FT}{\Leftrightarrow} G_{xy}(f) \tag{3.69}$$

If the functions $x(t)$ and $y(t)$ are complex then the cross correlation is defined by:

$$R_{xy}(\tau) = \langle x(t)y^*(t-\tau)\rangle = \langle x^*(t)y(t+\tau)\rangle \tag{3.70}$$

and many of the properties listed above do not apply.

## 3.4 Summary

Variables are said to be random if their particular value at specified future times cannot be predicted. Information about their probable future values is often available, however, from a probability model. The (unconditional) probability that any of the events, belonging to a subset of mutually exclusive possible events, occurs as the outcome of a random experiment or trial is the sum of the individual probabilities of the events in the subset. The joint probability of a set of statistically independent events is the product of their individual probabilities. A conditional probability is the probability of an event given that some other, specified, event is known to have occurred. Bayes's rule relates joint, conditional and unconditional probabilities.

Cumulative distributions give the probability that a random variable will be less than, or equal to, any particular value. Pdfs are the derivative of the cumulative distribution. The definite integral of a pdf is the probability that the random variable will lie between the integral's limits. Exceedances are the complement of cumulative distributions.

Moments, central moments and modes are statistics of random variables. In general they give partial information about the shape and location of pdfs. Joint pdfs (on definite integration) give the probability that two or more random variables will concurrently take particular values between the specified limits. A marginal pdf is the pdf of one random variable irrespective of the value of any other random variable. The correlation of two random variables is their mean product. The covariance is the mean product of their fluctuating (zero mean) components only, being zero for uncorrelated signals. Statistically independent random variables are always uncorrelated but the converse is not true. The (normalised) correlation coefficient of two random variables is the correlation of their fluctuating components (i.e. covariance) after the standard deviations of both variables have been normalised to 1.0. The pdf of the sum of independent random variables is the convolution of their individual pdfs and, for the sum of many independent random variables, this results in a Gaussian pdf. This is called the central limit theorem. If the random variables are independent then the mean of their sum is the sum of their means and the variance of their sum is the sum of their variances.

Random processes are random variables which change with time (or spatial position). They are defined strictly by an ensemble of functions. Both ensemble and temporal (or spatial) statistics can therefore be defined. A random process is said to be (statistically) stationary in the strict, or narrow, sense if all its statistics are invariant with time (or space). It is said to be stationary in the loose, or wide, sense if its ensemble mean is invariant with time and the correlation between its random variables at different times depends only on time difference. A random process is said to be ergodic if its ensemble and time averages are equal. Random processes which are ergodic are statistically stationary but the converse is not necessarily true.

Gaussian processes are extremely common and important due to the action of the central limit theorem. A process is said to be Gaussian in the strict sense if any pair of (ensemble) random variables has a joint Gaussian pdf. Any sample function of a random process is said to be Gaussian in the loose sense if samples from it are Gaussianly distributed. Not all loose sense Gaussian sample functions belong to strict sense Gaussian processes. Gaussian processes are completely specified by their first and second order moments.

Signal memory is characterised by the signal's autocorrelation function. This function gives the correlation between the signal and a time shifted version of the signal for all possible time shifts. The decorrelation time of a signal is that time shift for which the autocorrelation function has fallen to some prescribed fraction of its peak value. The Wiener-Kintchine theorem identifies the power and energy spectral densities of power and energy signals with the Fourier transform of these signals' autocorrelation functions. The normalised autocorrelation can be interpreted as the fraction of a signal contained within a time shifted version of itself. Signal memory (i.e. decorrelation time) and signal bandwidth are inversely proportional. White noise, with an impulsive autocorrelation function, is physically unrealisable and is memoryless.

Cross correlation relates to the similarity between a pair of different functions, one offset from the other by a time shift. The Fourier transform of a cross correlation function is a cross energy, or power, spectral density depending on whether the function pair represent energy or power signals.

## 3.5 Problems

3.1. A box contains 30 resistors. 15 of the resistors have nominal values of 1.0 kΩ, 10 have nominal values of 4.7 kΩ and 5 have nominal values of 10 kΩ. 3 resistors are taken at random and connected in series. What is the probability that the 3 resistor combination will have a nominal resistance of: (i) 3 kΩ; (ii) 15.7 kΩ; and (iii) 19.4 kΩ? [0.1121, 0.1847, 0.0554]

3.2. A transceiver manufacturer buys power amplifiers from three different companies (A, B, C). Assembly line workers pick power amplifiers from a rack at random without noticing the supplier. Customer claims, under a one year warranty scheme, show that 8% of all power amplifiers (irrespective of supplier) fail within one year and that 25%, 35% and 40% of all failed power amplifiers were supplied by companies A, B and C respectively. The purchasing department records that power amplifiers have been supplied by companies A, B and C in the proportions 50:40:10 respectively. What is the probability of failure within one year of amplifiers supplied by each company? [0.04, 0.07, 0.32]

3.3. The cumulative distribution function for a continuous random variable, $X$, has the form:

$$P_X(x) = \begin{cases} 0, & -\infty < x \leq -2 \\ a\,(1 + \sin(bx)), & -2 < x \leq 2 \\ c, & x > 2 \end{cases}$$

Find: (a) the values of $a$, $b$ and $c$ that make this a valid CD; (b) the probability that $x$ is negative; and (c) the corresponding probability density function.

3.4. A particular random variable has a cumulative distribution function given by:

$$P_X(x) = \begin{cases} 0, & -\infty < x \leq 0 \\ 1 - e^{-x}, & 0 \leq x < \infty \end{cases}$$

Find: (a) the probability that $x > 0.5$; (b) the probability that $x \leq 0.25$; and (c) the probability that $0.3 < x \leq 0.7$. [0.6065, 0.2212, 0.2442]

3.5. The power reflected from an aircraft of complicated shape that is received by a radar can be described by an exponential random variable, $w$. The pdf of $w$ is:

$$p(w) = \begin{cases} (1/w_o)e^{-w/w_o}, & \text{for } w > 0 \\ 0, & \text{for negative } w \end{cases}$$

where $w_o$ is the average amount of received power. What is the probability that the power received by the radar will be greater than the average received power? [0.368].

3.6. An integrated circuit manufacturer tests the propagation delays of all chips of one particular batch. He discovers that the pdf of the delays is well approximated by a triangular distribution with mean value 8 ns, maximum value 12 ns and minimum value 4 ns. Find: (a) the variance of this distribution; (b) the standard deviation of the distribution; and (c) the percentage of chips which will be rejected if the specification for the device is 10 ns. [2.66, 1.63, 12.5%]

3.7. A bivariate random variable has the joint pdf:

$$p(x, y) = A(x^2 + 2xy)\,\Pi(x)\Pi(y),$$

Find: (a) the value of $A$ which makes this a valid pdf; (b) the correlation of $X$ and $Y$; (c) the marginal pdfs of $X$ and $Y$; (d) the mean values of $X$ and $Y$; and (e) the variances of $X$ and $Y$. [12, 0.05, 1, 0, 0.15, 0.0833]

3.8. (a) For the zero mean Gaussian pdf: $p_X = (1/\sqrt{2\pi}\sigma)e^{-x^2/(2\sigma^2)}$, prove explicitly that the RMS value $\sqrt{\overline{X^2}}$ is $\sigma$.

$$\text{(Hint: } \int_a^b x^2 e^{-x^2}\,dx = -\frac{d}{d\lambda}\left[\int_a^b e^{-\lambda x^2}\,dx\right]_{\lambda = 1}$$

and remember that a complete integral (with limits $\pm\infty$) may be found using the Fourier transform DC value theorem, Table 2.5), and (b) show that, for the Gaussian pdf as defined in part (a), $\overline{X^4} = 3\sigma^4$.

$$\text{(Hint: } \int_0^\infty x^{n-1}e^{-x}\,dx = \Gamma(n), \quad \text{any } n > 0$$

and the gamma function $\Gamma(n)$ has the properties:

$$\Gamma(n+1) = n\Gamma(n) \quad \text{and} \quad \Gamma(½) = \sqrt{\pi}\;)$$

3.9. A random signal with uniform pdf, $p_X(x) = \Pi(x)$ is added to a second, independent, random signal with one sided exponential pdf, $p_Y(y) = 3\,u(y)e^{-3y}$. Find the pdf of the sum.

3.10. $v(t)$, $w(t)$, $x(t)$ and $y(t)$ are independent random signals which have the following pdfs:

$$P_V(v) = \frac{2}{1 + (2\pi v)^2} \qquad P_W(w) = \frac{1}{1 + (\pi w)^2}$$

$$P_X(x) = \frac{2/3}{1 + [(2/3)\pi x]^2} \qquad P_Y(y) = \frac{½}{1 + (½\pi y)^2}$$

Use characteristic functions to find the pdf of their sum.

3.11. For a tossed dice:

(a) Use convolution to deduce the probabilities of the sum of two thrown dice being 2, 3 etc.

(b) 24 is the largest sum possible on throwing 4 dice. What is the probability of this event from the joint probability of independent events? Check your answer by convolution.

(c) What is the most probable sum for 4 dice? Use convolution to find the probability of this event.

(d) A box containing 100 dice is spilled on the floor. Make as many statements as you can about the sum of the uppermost faces by extending the patterns you see developing in the convolution in parts (a), (b) and (c). [(a) 1/36, 2/36, ⋯, 6/36, 5/36, 4/36, ⋯, 1/36]; [(b) $7.7 \times 10^{-4}$]; [(c) 14, $1.13 \times 10^{-1}$]; [(d) $\Sigma_{min} = 100$, $\Sigma_{max} = 600$, 501 possible $\Sigma$'s, most likely $\Sigma = 350$, PDF is truncated Gaussian approximation].

3.12. Two, independent, zero mean, Gaussian noise sources ($X$ and $Y$) each have an RMS output of 1.0 V. A cross-coupling network, is to be used to generate two noise signals ($U$ and $V$), where $U = (1 - \alpha)X + \alpha Y$ and $V = (1 - \alpha)Y + \alpha X$, with a correlation coefficient of 0.2 between $U$ and $V$. What must the (voltage) cross-coupling ratio, $\alpha$, be? [0.8536 or 0.1465]

3.13. A periodic time function, $x(t)$, of period $T$ is defined as a sawtooth waveform with a random 'phase' (i.e. positive gradient zero crossing point), $\tau$, over the period nearest the origin i.e.:

$$x(t) = \frac{2V}{T}(t - \tau), \qquad -T/2 + \tau \le t < T/2 + \tau$$

The pdf of the random variable is:

$$p_T(\tau) = \begin{cases} 1/T, & |\tau| \le T/2 \\ 0, & |\tau| > T/2 \end{cases}$$

Show that the function is ergodic.

3.14. Consider the following time function:

$$X(t) = A\cos(\omega t - \Theta)$$

The phase angle, $\Theta$, is a random variable whose pdf is given as:

$$p_\Theta(\theta) = \frac{1}{2\pi} \qquad \text{for } 0 \le \theta < 2\pi \text{ and zero elsewhere}$$

Find the mean value and variance of $\Theta$ and of $X(t)$.

3.15. Given that the autocorrelation function of a certain stationary process is:

$$R_{xx}(\tau) = 25 + \frac{4}{1 + 6\tau^2}$$

Find: (a) the mean value, and (b) the variance of the process. [±5, 4]

3.16. $X(t)$ is a deterministic random process defined by:

$$X(t) = \cos(2\pi f t + \Theta) + 0.5$$

where $\Theta$ is a uniformly distributed random variable in the range [$-\pi$, $\pi$], but remains fixed for a given sample waveform of the random process. Calculate $R_{xx}(\tau)$, and identify the source of each term in your answer.

3.17. A stationary random process has an autocorrelation function given by:

$$R(\tau) = \begin{cases} 10(1 - |\tau|/0.05), & |\tau| \leq 0.05 \\ 0, & \text{elsewhere} \end{cases}$$

Find: (a) the variance; and (b) the power spectral density of this process. State the relation between bandwidth and decorrelation time of this random process both being defined by the first zero crossing (or touching) point in their respective domains.

3.18. A stationary random process has a power spectral density given by:

$$G(f) = \begin{cases} 5, & 10/2\pi \leq |f| \leq 20/2\pi \\ 0, & \text{elsewhere} \end{cases}$$

Find: (a) the mean square value; and (b) the autocorrelation function of the process. (If you have access to simple plotting software plot the autocorrelation function.)

3.19. A stationary random process has a bilateral (i.e. double sided) power spectral density given by:

$$G_{xx}(\omega) = \frac{32}{\omega^2 + 16}$$

Find: (a) the average power (on a per-ohm basis) of this random process; and (b) the average power (on a per-ohm basis) of this random process in the range –4 rad/s to 4 rad/s. [4, 2]

3.20. A random variable, $Z(t)$, is defined to be:

$$Z(t) = X(t) + X(t + \tau)$$

$X(t)$ is a stationary process whose autocorrelation function is:

$$R_{xx}(\tau) = e^{-\tau^2}$$

Derive an expression for the autocorrelation of the random process $Z(t)$.

3.21. Show that the autocorrelation function of a non-zero mean random process, $X(t)$, may be written as:

$$R_{xx}(\tau) = R_{x'x'}(\tau) + \mathrm{E}[(X(t))^2]$$

where $R_{x'x'}(\tau)$ is the autocorrelation function of a zero-mean random process and E[.] is the expectation operator as defined in equation (3.16).

3.22. The stationary random process $X(t)$ has a power spectral density $G_{xx}(f)$. What is the power spectral density of $Y(t) = X(t - T)$.

3.23. Two jointly stationary random processes are defined by:

$$X(t) = 5\cos(10t + \Theta) \quad \text{and} \quad Y(t) = 20\sin(10t + \Theta)$$

where $\Theta$ is a random variable that is uniformly distributed from 0 to $2\pi$. Find the cross correlation function $R_{xy}(\tau)$ of these two processes.