DISCRETE-TIME RANDOM PROCESS (1)

DIGITAL SIGNALS: DETERMINISTIC AND RANDOM

Discrete-time (or digital) signal:

Deterministic and random signals:

RANDOM VARIABLES

Single Random Variables

- Definition of random variables:
- Probability distribution function and probability density function:
- Mean (or mean value, or expected value):
- Variance:
- Moments:

Multiple Random Variables

- Joint distribution function and joint density function:
- Joint moments:
- Correlation and covariance of multiple random variables:
- Correlation coefficient:
- Relationship between random variables (Independence, Uncorrelatedness, and Orthogonality):

Prediction and Estimation

- Parameter estimation, bias and consistency:
  Sample mean and sample variance
- Linear Mean-Square Estimation:
  Orthogonality principle

Two Types of Random Variables: Uniform and Gaussian Random Variables

- Uniform random variable:
- Gaussian random variable:

Some useful MATLAB functions:

DISCRETE-TIME RANDOM PROCESS

DISCRETE-TIME (DIGITAL) SIGNALS: DETERMINISTIC AND RANDOM

Discrete-time (digital) signal:

$$x(n) = x_a(nT_s)\,, \tag{1}$$

that is a function of an *integer-valued* variable *n* and may result from sampling a *continuous-time (analog) signal* $x_a(t)$ with an A/D (analog-to-digital) converter having a sampling interval T_s (or sampling rate $f_s = 1/T_s$) so that continuous time *t* becomes discrete-time nT_s, i.e., $t = nT_s$. For a digital signal *x*(*n*), therefore, time instant *n* means nT_s. The signals that we study here are always digital.

Deterministic and random signals:

Example 1. Measurements of signals using a digital oscilloscope

Digital oscilloscopes are a commonly-used instrument for various signal measurements, and they are usually equipped with an averaging function that is used to reduce noise when measuring a noisy deterministic signal that is a desired deterministic signal interfered with unwanted noise.

(a) *Measurements of an analog sinusoidal signal* $x_a(t) = \sin(2\pi t/T)$ *where T is the period.*

When $x_a(t)$ is recorded with the sampling interval $T_s = T/20$, it has a digital form of $x(n) = \sin(\pi n/10)$. Repeating the measurement *N* times, we will get *N* sinusoidal waveforms $x_k(n)$ (where *k* = 1, 2, …, *N*) that all are exactly the same, i.e., $x_k(n) = \sin(\pi n/10)$ for all *k*, and the mean-value waveform $\frac{1}{N}\sum_{k=1}^{N} x_k(n) = \sin(\pi n/10)$ is the same as all the *N* waveforms. At any time instant *n*, we can determine the signal value *x*(*n*) in a definite way, e.g., for *n*=1 and $n = 5$ we have $x(n) = \sin(\pi/10)$ and $x(n) = 1$, respectively. These show that a deterministic signal *x*(*n*) is repeatable and predictable.

(b) *Measurements of thermal noise* $x_a(t)$*, a random signal.*

The thermal noise voltage generated in a resistor is a commonly-encountered random signal. When the noise signal $x_a(t)$ is recorded by a digital oscilloscope, it has a digital form *x*(*n*). Repeating the record *N* times, the *N* recorded waveforms $x_k(n)$ (where *k* = 1, 2, …, *N*) will be different and the mean-value waveform $\frac{1}{N}\sum_{k=1}^{N} x_k(n)$ approaches to zero when $N \to \infty$, which demonstrates that *x*(*n*) is not repeatable. At time instant *n* the value of *x*(*n*) it may take one of the *N* possible waveforms $x_k(n)$ for *k* = 1, 2, …, *N*. This shows that *x*(*n*) is not predictable. Thus, a random signal *x*(*n*) is unrepeatable and unpredictable.

The characteristics of deterministic and random signals can, in general, be summarized as follows:

- A deterministic signal can be described by a mathematical expression and reproduced exactly with repeated measurements; it is related to one signal.
- A random signal that is associated with a set of (digital) signals is not repeatable and predictable; thus it may only be described probabilistically or in terms of average behaves of the signal set.

Note that random signals can be either unwanted noise signals as in the above example or desired signals like speech signals, ultrasonic signals, light signals, *etc*.

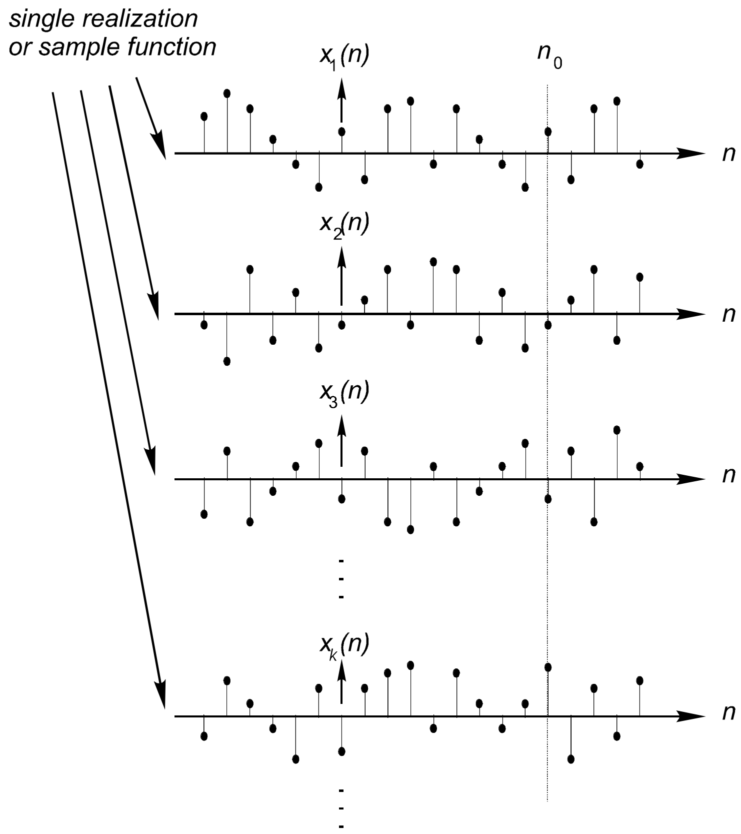

Fig. 1. A random process $x(n)$ is an ensemble of single realizations (or sample functions).

A random signal is represented mathematically by the concept of a random process. For a random signal like the above-mentioned thermal noise signal, the set of all possible waveforms is called an *ensemble* of sample functions (i.e., discrete-time signals) or, equivalently, *a random process*, denoted by $x(n)$ (see Fig. 1). The set (ensemble) can be either finite or infinite. Each recorded waveform (a digital signal) $x_k(n)$ is a *single realization* or a sample function of the random process $x(n)$. Thus, when investigating a random process $x(n)$, we should always consider an ensemble of realizations, $x_k(n)$ for all possible k, and look into their statistical characteristics and average behaves.

Since a random process $x(n)$ is an ensemble of realizations, the process at a fixed time instant $n = n_0$ becomes a random variable $x(n_0)$ that only depends on the realizations, $x_k(n_0)$, and may take one of the possible values $x_k(n_0)$ according to a certain probabilistic law (Fig. 1). A discrete-time random process is, therefore, just an indexed sequence of random variables, and studying random variables may serve as a fundamental step to deal with random processes.

RANDOM VARIABLES

Single Random Variables

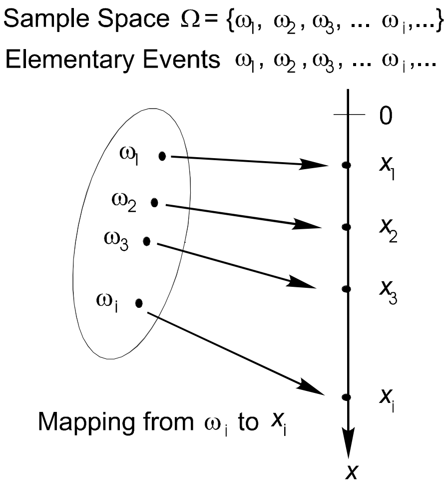

Fig. 2. Definition of a discrete random variable

Definition of random variables:

A random variable x is defined when each elementary event ω_i in the sample space Ω, i.e., $\omega_i \in \Omega$, is mapped to a value of *x*, and is assigned with a certain probability (Fig. 2).

The *sample space* is defined by the set of all possible experimental outcomes (elementary events). *An elementary event* is the event consisting of a single element.

A random variable *x* may be of *discrete* type or *continuous* type.

Example 2. Discrete random variables

In the experiment of rolling a fair die, we have six elementary events ω_i (i = 1, 2, … 6), and we assign to the six events ω_i the random variable x=$x(\omega_i)$=i. Thus, the random variable *x* may take one of the six numbers

$x(\omega_1)$=1, $x(\omega_2)$=2, …, $x(\omega_6)$=6

that are mapped from the six elementary events ω_i. The random variable is of discrete type. The sample space contains six elementary events, i.e., Ω={ ω_1, ω_2, …, ω_6 }. The probability that each elementary event may happen is 1/6, and is mathematically expressed as

$\Pr\{x = i\}$ = 1/6, where i=1, 2, …, 6

Example 3. Continuous random variables

In the case of a uniform distribution over interval (b, c), a random variable *x* is of continuous type and assumes a continuum of values in the interval (b, c). The sample space is $\Omega = \{\omega : b < \omega < c\}$. There are an infinite number of elementary events, each taking a value in the interval (b, c). The probability of each elementary event is

$\Pr\{\alpha \le x \le \alpha + \Delta\alpha\} \cong \Delta\alpha/(c-b)$, where $b < \alpha < c$ and $\Delta\alpha$ >0 is a small change, and $\lim_{\Delta\alpha \to 0} \Pr\{\cdot\} = 0$.

Probability distribution function and probability density function:

A random variable *x* is characterized statistically or described probabilistically by means of the *probability distribution function*, $F_x(\alpha)$, given by

$$\boxed{F_x(\alpha) = \Pr\{x \le \alpha\}}, \tag{2}$$

which is a probability that *x* assumes a value that is less than α.

A random variable can also be characterized statistically by the *probability density function*, $f_x(\alpha)$, given by

$$\boxed{f_x(\alpha) = dF_x(\alpha)/d\alpha} \tag{3}$$

We shall say that the *statistics* of a random variable x are known if we can determine the probability distribution function or the probability density function.

Example 4. Probability distribution function and probability density function

(i) For the experiment of rolling a die in which the probability that the random variable assumes one of the six numbers k=1, 2, …, 6 is $\Pr\{x=k\}=1/6$, the distribution function and the density function of the discrete random variable are, respectively,

$$F_x(\alpha)=\begin{cases} 0, & \alpha<1 \\ 1/6, & 1\le\alpha<2 \\ 2/6, & 2\le\alpha<3 \\ \ldots, & \ldots \\ 1, & \alpha\ge 6 \end{cases}, \qquad f_x(\alpha)=\frac{d}{d\alpha}F_x(\alpha)=\frac{1}{6}\sum_{k=1}^{6}\delta(\alpha-k) \tag{4}$$

(ii) For the uniform distribution, the distribution function and density function of the continuous random variable are, respectively,

$$F_x(\alpha)=\begin{cases} 0, & b<\alpha \\ \dfrac{1}{c-b}(\alpha-b), & b<\alpha<c \\ 1, & \alpha\ge c \end{cases}, \qquad f_x(\alpha)=\frac{d}{d\alpha}F_x(\alpha)=\begin{cases} \dfrac{1}{c-b}, & b<\alpha<c \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

Ensemble Averages of random variables

A random variable x can also be characterized by ensemble averages, e.g., mean and variance.

Mean (or mean value, or expected value):

For a discrete random variable x that assumes a value of α_k with probability $\Pr\{x=\alpha_k\}$, the mean (expected value) is defined as

$$\boxed{E\{x\}=m_x=\sum_k \alpha_k \Pr\{x=\alpha_k\}} \tag{6}$$

For a continuous random variable x, this expectation may be written in terms of the probability density function $f_x(\alpha)$ as

$$\boxed{E\{x\}=m_x=\int_{-\infty}^{\infty}\alpha f_x(\alpha)d\alpha} \tag{7}$$

The expectation is a linear operation since

$$E\{ax+by\}=aE\{x\}+bE\{y\} \tag{8}$$

For a complex random variable, the mean-square value is

$$E\{zz^*\}=E\{|z|^2\}. \tag{9}$$

Variance:

For a discrete random variable x that assumes a value of α_k with probability $\Pr\{x=\alpha_k\}$, the variance is defined as

$$\boxed{\mathrm{Var}\{x\}=\sigma_x^2=E\{[x-E\{x\}]^2\}=\sum_k[\alpha_k-E\{x\}]^2\Pr\{x=\alpha_k\}} \tag{10a}$$

For a continuous random variable *x* with given probability density function $f_x(\alpha)$, the variance is defined as

$$\boxed{\mathrm{Var}\{x\} = \sigma_x^2 = E\left\{[x - E\{x\}]^2\right\} = \int_{-\infty}^{\infty} [\alpha - E\{x\}]^2 f_x(\alpha) d\alpha} \tag{10b}$$

which is the mean-square value of the random variable $y = x - E\{x\}$. The variance may be expressed

$$\boxed{\mathrm{Var}\{x\} = E\left\{[x - E\{x\}]^2\right\} = E\left\{x^2\right\} - (E\{x\})^2} \tag{11}$$

The square root of the variance, σ_x, is called the *standard deviation*.

For a complex random variable, the variance is

$$\sigma_z^2 = E\{[z - E\{z\}][z^* - E\{z^*\}]\} = E\left\{|z - E\{z\}|^2\right\} \tag{12}$$

The variance represents the average squared deviation of the random variable from the mean.

Example 5. The mean of a random variable

(i) For the random variable *x* defined in the die experiment, the mean value is

$$E(x) = \sum_{k=1}^{6} k \Pr(x = k) = \frac{1}{6}\sum_{k=1}^{6} k = 3.5 \tag{13}$$

(ii) For the continuous random variable *x* having the uniform distribution, the expected value is

$$E(x) = \int_b^c \alpha f_x(\alpha) d\alpha = \frac{1}{c-b}\int_b^c \alpha d\alpha = \frac{b+c}{2} \tag{14}$$

For (*b*, *c*)=(0, 1), *E*(*x*)=0.5; for (*b*, *c*)=(0, 2), *E*(*x*)=1; for (*b*, *c*)=(-1, 2), *E*(*x*)=0.5.

Example 6. The variance of a random variable

(i) For the random variable *x* defined in the die experiment, the variance is

$$Var(x) = \sum_{k=1}^{6} (k - m_x)^2 \Pr(x = k) = \frac{1}{6}\sum_{k=1}^{6} (k - 3.5)^2 = 2.917 \tag{15a}$$

(ii) For the random variable *x* with the uniform distribution, the variance is

$$\mathrm{Var}(x) = \int_b^c (\alpha - m_x)^2 f_x(\alpha) d\alpha = \frac{1}{c-b}\int_b^c (\alpha - m_x)^2 d\alpha = \frac{(c-b)^2}{12} \tag{15b}$$

For (*b*, *c*)=(0, 1), Var(*x*)=0.083; for (*b*, *c*)=(0, 2), Var(*x*)=0.333; for (*b*, *c*)=(-1,2), Var(*x*)=0.75. This shows that the broader the value range taken by *x* the larger the variance.

Example 7. (i) What is the mean value of the mean value of a random variable? That is $E\{m_x\}$=? (ii) What is the mean value of the variance of a random variable? That is $E\{\sigma_x^2\}$=?

Answer: $E\{m_x\} = m_x$, and $E\{\sigma_x^2\} = \sigma_x^2$.

This example shows that the mean and the variance (that is the mean of $x - E\{x\}$) of a random variable are deterministic. Note that the mean of a random variable is the mean of a complete set (ensemble) of all the elementary events in the sample space.

<u>Moments</u>:

The quantities *moments* are of interest in the study of a random variable because they cover more general ensemble descriptions of a random variable. The moments are defined in different types: moments, central moments, absolute moments.

The *n*th *moment* of a random variable is defined by

$$\boxed{E\{x^n\} = \int_{-\infty}^{\infty} \alpha^n f_x(\alpha) d\alpha} \qquad (16)$$

which is the expected value of x^n. The first order moment, obviously, is the mean value m_x.

The *mean-square value* is $E\{x^2\}$, the second moment, that is an important statistical average and often used as a measure for the quality of an estimate.

The *n*th *central moment* of a random variable is defined by

$$\boxed{E\{(x - m_x)^n\} = \int_{-\infty}^{\infty} (\alpha - m_x)^n f_x(\alpha) d\alpha} \qquad (17)$$

which is the expected value of $(x - m_x)^n$. The variance σ_x^2 is the second central moment.

The *n*th *absolute moment* and *absolute central moment* of a random variable are defined, respectively, as

$$E\{|x|^n\} \text{ and } E\{|x - m_x|^n\} \qquad (18)$$

Note that for a complex variable *z*, the second central moment $E\{(z - m_z)^2\}$ is different form the second absolute central moment $E\{|z - m_z|^2\} = E\{(z - m_z)(z - m_z)^*\}$, where * means the conjugate of a complex value.

Multiple Random Variables

When a problem is involved with two or more random variables, it is necessary to study the statistical dependencies (or relations) that may exist between the random variables. The statistical dependencies can be joint distribution and density functions, and correlation and covariance.

Joint distribution function and joint density function:

Given two random variables $x(1)$, and $x(2)$, the *joint distribution function* for the random variables is

$$\boxed{F_{x(1),x(2)}(\alpha_1,.\alpha_2) = \Pr\{x(1) \le \alpha_1, x(2) \le \alpha_2\}} \qquad (19)$$

and the *joint density function* is

$$\boxed{f_{x(1),x(2)}(\alpha_1,\alpha_2) = \frac{\partial^2}{\partial\alpha_1\partial\alpha_2} F_{x(1),x(2)}(\alpha_1,\alpha_2)} \qquad (20)$$

For a *complex random variable* $z = x + jy$ the distribution function for *z* is given by the joint distribution function

$$F_z(\alpha) = \Pr\{x \le a, y \le b\}\text{, where } \alpha = a + jb \qquad (21)$$

Joint moments:

Like in the case of a single random variable, the statistical dependencies can also be described by ensemble averages. The joint moments are just used for this purpose. The correlation and covariance of two random variables are the two joints most often used in this course.

The *joint moment* of the random variables *x* and *y* is defined by

$$\boxed{E\{x^k y*^l\}} \qquad (22)$$

where $y*$ is the conjugate of *y*.

The *joint central moment* of the random variables x and y is defined by

$$E\{(x-m_x)^k (y-m_y)^{*l}\} \tag{23}$$

The orders of the moments are $k+l$.

Correlation and covariance of two random variables:

Correlation: $\boxed{r_{xy} = E\{xy^*\}}$, (24)

that is a second order joint moment

Covariance: $\boxed{c_{xy} = \mathrm{Cov}(x,y) = E\{[x-m_x][y-m_y]^*\} = E\{xy^*\} - m_x m_y^* = r_{xy} - m_x m_y^*}$, (25)

that is the second order joint central moment. The correlation and covariance are used to statistically characterize the relationship between two random variables, and they play an important role in studying signal modeling, spectrum estimation, and Wiener filters.

Correlation coefficient:

The *correlation coefficient*, a normalized covariance,

$$\boxed{\rho_{xy} = \frac{c_{xy}}{\sigma_x \sigma_y} = \frac{r_{xy} - m_x m_y^*}{\sigma_x \sigma_y}} \tag{26}$$

An interesting property of the correlation coefficient is that

$$\boxed{|\rho_{xy}| \le 1}, \text{ or } |c_{xy}| \le \sigma_x \sigma_y \tag{27}$$

Relationship between random variables (Independence, Uncorrelatedness, and Orthogonality):

Two random variables x and y are said to be *statistically independent* if

$$f_{x,y}(\alpha,\beta) = f_x(\alpha) f_y(\beta) \tag{28}$$

Two random variables x and y are said to be *statistically uncorrelated* if

$$E\{xy^*\} = E\{x\}E\{y^*\} \quad \text{or} \quad r_{xy} = m_x m_y^* \tag{29}$$

which is a weaker form of the independence. In this case, the covariance is zero, $c_{xy} = r_{xy} - m_x m_y^* = 0$. Thus, two random variables x and y will be s*tatistically uncorrelated* if their covariance is zero, $c_{xy} = 0$.

Note that *statistically independent variables are always uncorrelated, but the converse is not true in general*.

A useful property of *uncorrelated* random variables is the following

$$\mathrm{Var}\{x+y\} = \mathrm{Var}\{x\} + \mathrm{Var}\{y\} \tag{30}$$

since $\mathrm{Var}\{x+y\} = E\{[x+y-m_x-m_y][x+y-m_x-m_y]^*\} = \mathrm{Var}\{x\} + c_{xy} + (c_{xy})^* + \mathrm{Var}\{y\}$.

Two random variables x and y are *orthogonal* if their correlation is zero, $E\{xy^*\} = 0$ or $r_{xy} = 0$.

Prediction and Estimation

Prediction and estimation are two general classes of problems encountered in statistical investigations.

In the prediction case, the probabilistic model (e.g., a certain distribution, or density, function) of the problem is assumed to be known, and predictions are made concerning future observations. For example, in the experiment of flipping a coin, we know the probabilities of showing heads and tails and we wish to predict the number of occurrences of showing heads.

In the estimation case, a sequence of observed values of a random variable are know, and a parameter (of a model) is estimated from the observed values; or the problem concerned is an estimation of a random variable *y* in terms of an observation of another random variable *x* (this problem generally arises when *y* cannot be directly measured or observed so a related random variable is measured and used to estimate *y*). The goal of the estimation is to find the best estimate of *y* in terms of the known.

For example, we have observed *N* values of a random variable, and wish to estimate its mean value. Here we focus ourselves on the estimation issue.

Parameter estimation, bias and consistency:

Example 8. *The Sample Mean* and *the Sample Variance*

Given *N* observations x_n of a random variable *x*, the sample mean is

$$\hat{m}_x = \frac{1}{N}\sum_{n=1}^{N} x_n \tag{31}$$

and the sample variance is

$$\hat{\sigma}_x^2 = \frac{1}{N}\sum_{n=1}^{N} (x_n - \hat{m}_x)^2 \ . \tag{32}$$

The sample mean $\hat{m}_x$ and the sample variance $\hat{\sigma}_x^2$ are the estimates of the mean m_x and the variance σ_x^2 of the random variable *x*, respectively. Note that the estimates (e.g., $\hat{m}_x$ and $\hat{\sigma}_x^2$) themselves are random variables because every estimate is a function of observations.

The estimates themselves are random variables. Therefore, in classifying the effectiveness of a particular estimator, it is important to characterize its statistical properties. The statistical properties of interest include the bias and the variance.

The *bias* is the difference between the expected value of the estimate, $\hat{\theta}_N$, and the actual value, θ, and it is denoted by *B*,

$$B = \theta - E\{\hat{\theta}_N\} \tag{33}$$

where the estimate $\hat{\theta}_N$ is a function of *N* random variables (or *N* observations of a random variable).

The estimate $\hat{\theta}_N$ is said to be *unbiased* if *B*=0 or $E\{\hat{\theta}_N\} = \theta$.

The estimate $\hat{\theta}_N$ is said to be *biased* if $B \neq 0$.

The estimate $\hat{\theta}_N$ is said to be *asymptotically unbiased* if $\lim_{N\to\infty} E\{\hat{\theta}_N\} = \theta$.

In general, it is desirable that an estimator is either unbiased or asymptotically unbiased. However, the bias is not the only statistical measure of importance.

Consistency means that an estimate is said to be consistent if it converges, in some sense, to the true value of the parameter. Mathematically, an estimate is consistent if the variance of the estimate goes to zero,

$$\lim_{N\to\infty} \mathrm{Var}\{\hat{\theta}_N\} = \lim_{N\to\infty} E\left\{\left|\hat{\theta}_N - E\{\hat{\theta}_N\}\right|^2\right\} = 0 \tag{34}$$

Example 9. *The bias and consistency of the sample mean* (Example 3.2.3 on p. 74)
Let x be a random variable with a mean m_x and variance σ_x^2. Given N uncorrelated observations of x that are denoted by x_n, the sample mean $\hat{m}_x = \frac{1}{N}\sum_{n=1}^{N} x_n$ has the expected value

$$E\{\hat{m}_x\} = \frac{1}{N}\sum_{n=1}^{N} E\{x_n\} = \frac{1}{N}\sum_{n=1}^{N} m_x = m_x \tag{35}$$

which is an unbiased estimator. The variance of the sample mean is

$$\mathrm{Var}\{\hat{m}_x\} = \mathrm{Var}\left\{\sum_{n=1}^{N}\frac{x_n}{N}\right\} = \sum_{n=1}^{N}\mathrm{Var}\left\{\frac{x_n}{N}\right\} = \sum_{n=1}^{N} E\left\{\left(\frac{x_n - m_x}{N}\right)^2\right\} = \frac{1}{N^2}\sum_{n=1}^{N}\mathrm{Var}\{x_n\} = \frac{\sigma_x^2}{N} \tag{36}$$

which goes to zero as $N \to \infty$. Therefore, the sample mean is an unbiased and consistent estimator.

Linear Mean-Square Estimation:

The linear mean-square estimator $\hat{y}$ of a random variable y in terms of a random variable x is of the form,

$$\hat{y} = ax + b \tag{37}$$

and the goal is to find the values for a and b that minimizes the mean-square error

$$\xi = E\left\{(y - \hat{y})^2\right\} = E\left\{(y - ax - b)^2\right\}. \tag{38}$$

To minimize ξ, we differentiate ξ with respect to a and b and set zeros to the derivatives as follows,

$$\frac{\partial \xi}{\partial a} = E\left\{\frac{\partial}{\partial a}(y - ax - b)^2\right\} = E\{2(y - ax - b)(-x)\} = -2\left[E\{xy\} + aE\left\{x^2\right\} + bm_x\right] = 0 \tag{39}$$

$$\frac{\partial \xi}{\partial b} = E\left\{\frac{\partial}{\partial b}(y - ax - b)^2\right\} = E\{2(y - ax - b)(-1)\} = -2\left[m_y + am_x + b\right] = 0 \tag{40}$$

Solving the two equations for a and b, we find

$$a = \rho_{xy}\frac{\sigma_y}{\sigma_x}, \text{ and } b = m_y - am_x \tag{41}$$

and further the optimum linear estimator for y is

$$\hat{y} = \rho_{xy}\frac{\sigma_y}{\sigma_x}(x - m_x) + m_y \tag{42}$$

Inserting $\hat{y}$ into $\xi = E\left\{(y - \hat{y})^2\right\}$, we find the minimum mean-square error of the estimator

$$\xi_{\min} = \sigma_y^2(1 - \rho_{xy}^2) \tag{43}$$

In one extreme case when x and y are uncorrelated, we have $\rho_{xy} = 0$ so that $a = 0$ and $b = E\{y\}$. Thus, the estimator for y is $\hat{y} = E\{y\}$ and $\xi_{\min} = \sigma_y^2$, which shows that the best estimator with the minimum mean-square error $\xi_{\min} = \sigma_y^2$ is the mean of y, and also shows that x is not used in the estimation of y, so that knowing x does not affect the estimate of y. In another extreme case when $|\rho_{xy}| = 1$, $\xi_{\min} = 0$, and it follows that $y = ax + b$. The random variables x and y are linearly related to each other.

Two notes from this example:

(i) The correlation coefficient provides a measure of the linear predictability between random variables. The closer $|\rho_{xy}|$ is to 1, the more accurate the estimate $\hat{y}$ is to the random variable y.

(ii) The *orthogonality principle*: $E\{(y-\hat{y})x\} = E\{ex\} = 0$ (where $e = y - \hat{y}$ is called *estimation error* that is a random variable, and $E\{ex\}$ is just the correlation between *e* and *x*, namely, $E\{ex\} = r_{ex}$), which follows from $\partial\xi/\partial a = E\{2(y-ax-b)(-x)\} = -2E\{(y-\hat{y})x\} = 0$. This principle states that for the optimum linear predictor the estimation error will be orthogonal to the data *x* because of $E\{ex\} = r_{ex}$ =0. It is fundamental in mean-square estimation problems.

Two Types of Random Variables: Uniform and Gaussian Random Variables

Uniform and Gaussian random variables are two types of important random variables in probability theory. Gaussian random variables are also called normal random variables. The random processes that are made up of a sequence of such random variables play an important role in statistical signal processing.

Uniform random variable:

The statistical properties of a uniform random variable, i.e., its probability, probability distribution and probability density functions, its mean and variance, have been investigated in Examples 3-6, respectively.

Gaussian random variable:

The density function of a Gaussian random variable *x* is of the form

$$f_x(\alpha) = \frac{1}{\sigma_x\sqrt{2\pi}}\exp\left\{-\frac{(\alpha-m_x)^2}{2\sigma_x^2}\right\}$$

We can, from the definitions, find that the mean $E\{x\}$ equals m_x and the variance $Var(x)$ equals σ_x^2. This reveals that the density function of a Gaussian random variable is completely defined once the mean and the variance are given.

The joint density function of Gaussian random variables *x* and *y* is given by

$$f_{x,y}(\alpha,\beta) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}}\exp\left\{-\left[\frac{(\alpha-m_x)^2}{\sigma_x^2} - 2\rho_{xy}\frac{(\alpha-m_x)(\beta-m_y)}{2\sigma_x^2} + \frac{(\beta-m_y)^2}{\sigma_y^2}\right]\right\}$$

Gaussian random variables have a number of important properties as follows:

Property 1. If *x* and *y* are jointly Gaussian, then for any constants *a* and *b* the random variable

$$z = ax + by$$

is Gaussian with mean

$$m_z = am_x + bm_y,$$

and variance

$$\sigma_z^2 = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_x\sigma_y\rho_{xy}.$$

Property 2. If two jointly Gaussian random variables *x* and *y* are uncorrelated, ρ_{xy} =0, then they are statistically independent, $f_{x,y}(\alpha,\beta) = f_x(\alpha)f_y(\beta)$.

Property 3. If *x* and *y* are jointly Gaussian random variables then the optimum nonlinear mean-square estimator for *y*,

$$\hat{y} = g(x),$$

that minimizes the mean-square error

$$\xi = E\{(y-g(x))^2\}$$

is a linear estimator

$$\hat{y} = ax + b$$

Some useful MATLAB functions

Some MATLAB functions useful for studying random variables are given below. Use help to look at the detailed descriptions of the functions.

```
>> rand() % creates a sequence of uniformly distributed random numbers.
>> randn() % creates a sequence of Gaussian (normal) random numbers.
>> hist(x, M) % produces a histogram of x with M bins.
>> mean(x) % gives the mean value of the elements in x
>> var(x) % returns the variance of x
>> str(x) % returns the standard deviation of x
>> corrcoef(x) % is a matrix of correlation coefficients formed from array x
>> cov(x) % produces a covariance matrix
```