Lecture Note of Stochastic Hydrology

Introduction to Random Processes

Professor Ke-Sheng Cheng

Department of Bioenvironmental Systems Engineering

National Taiwan University

Introduction to Stochastic Processes

1. Definition of a stochastic process

A random variable is a mapping function which assigns outcomes of a random experiment to real numbers (see Fig. 1). Occurrence of the outcome follows certain probability distribution. Therefore, a random variable is completely characterized by its probability density function (*PDF*).

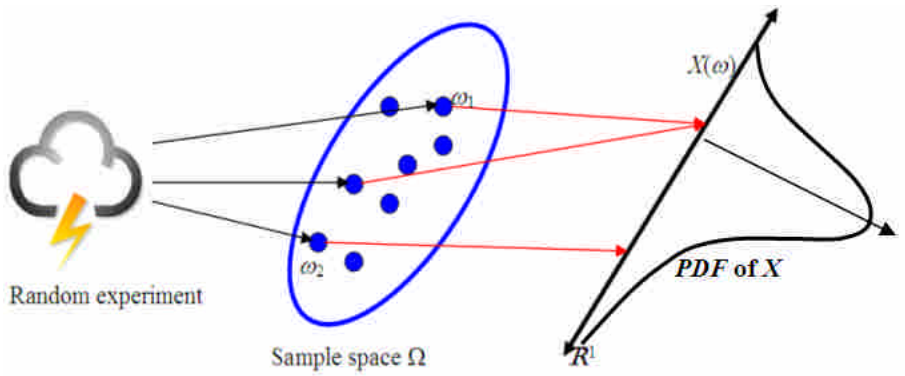


Figure 1. A random variable is a function mapping outcomes of a random experiment to real numbers.

A stochastic (or random) process $\{X(t), t \in T\}$ (or $\{X_t, t \in T\}$) is a family of random variables where the index set *T* may be discrete ($T = \{0,1,2,\cdots\}$) or continuous ($T = [0,\infty)$). The set of possible values which random variables $X(t), t \in T$ may assume is called the *state space* of the process.

A continuous time stochastic process $\{X(t), t \in T\}$ is said to have independent increments if for all choices of $t_0 < t_1 \cdots < t_n$, the *n* random variables $X(t_1) - X(t_0)$, $X(t_2) - X(t_1), \cdots, X(t_n) - X(t_{n-1})$ are independent. The process is said to have *stationary independent increments* if in addition $X(t_i + s) - X(t_{i-1} + s)$ has the same distribution as $X(t_i) - X(t_{i-1})$ for all $t_i, t_{i-1} \in T$ and $s > 0$.

A random variable *X* can be assigned a number $x(\omega)$ based on the outcome ω of a random experiment. Similarly, a stochastic process $\{X(t), t \in T\}$ can assume values $\{x(t,\omega), t \in T\}$ depending on the outcome of a random experiment. Each possible $\{x(t,\omega), t \in T\}$ is called a *realization* of the stochastic process $\{X(t), t \in T\}$. A totality of all realizations is called the *ensemble* of the stochastic process, as illustrated in Figure 2.

The term "stochastic processes" appears mostly in statistical books; however, the term "random processes" are frequently used in books of many engineering applications.

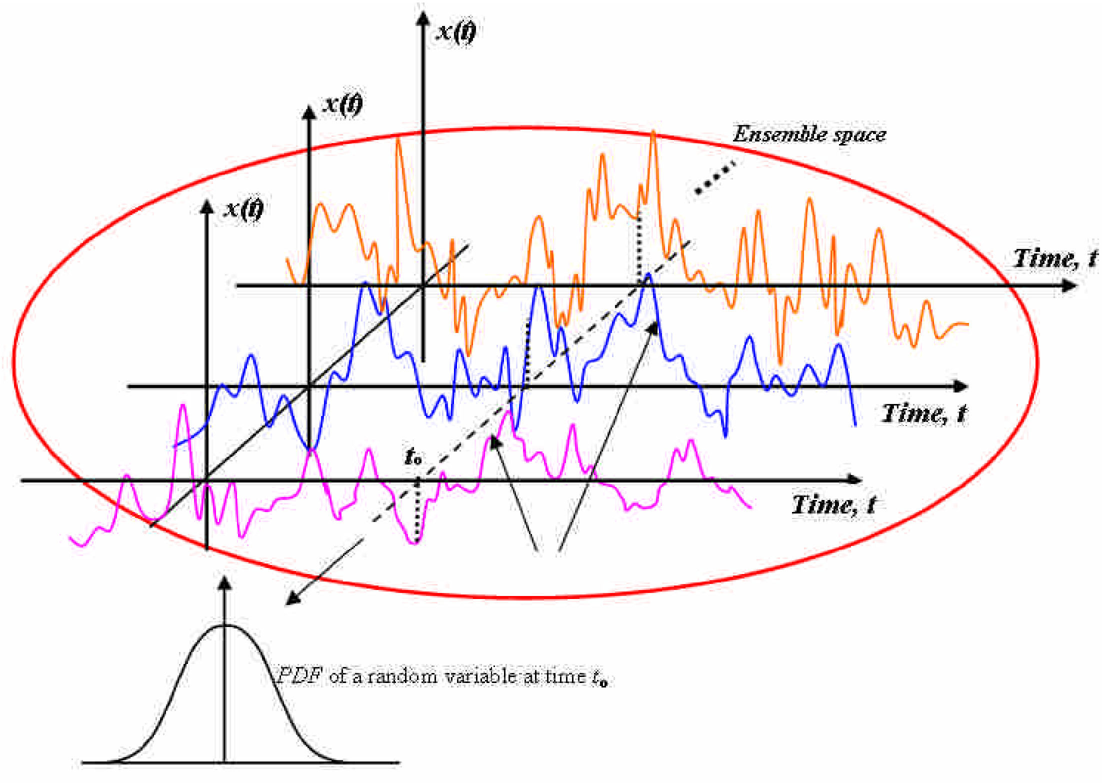


Figure 2. The ensemble space and realizations of a random process.

2. Characterizations of a stochastic processes

A stochastic process is composed of a series of random variables. As a result, its characteristics depend on the properties of not only the probability densities of individual random variables but also the correlation functions of a set of random variables.

2.1 First order densities of a random process

A stochastic process is defined to be ***completely or totally characterized*** if the joint densities for the random variables $X(t_1), X(t_2), \cdots X(t_n)$ are known for all times $t_1, t_2, \cdots, t_n$ and all *n*. In general, a complete characterization is practically impossible, except in rare cases. As a result, it is desirable to define and work with various partial characterizations. Depending on the objectives of applications, a partial characterization often suffices to ensure the desired outputs.

For a specific *t*, *X*(*t*) is a random variable with distribution $F(x,t) = p[X(t) \le x]$. The function $F(x,t)$ is defined as the *first-order distribution* of the random variable *X*(*t*). Its derivative with respect to *x*

$$f(x,t) = \frac{\partial F(x,t)}{\partial x} \tag{1}$$

is the *first-order density* of *X*(*t*). If the first-order densities defined for all time *t*, i.e. *f*(*x*,*t*), are all the same, then *f*(*x*,*t*) does not depend on *t* and we call the resulting density the first-order density of the random process $\{X(t)\}$; otherwise, we have a family of first-order densities. The first-order densities (or distributions) are only a partial characterization of the random process as they do not contain information that specifies the joint densities of the random variables defined at two or more different times.

2.2 Mean and variance of a random process

The first-order density of a random process, $f(x,t)$, gives the probability density of the random variables $X(t)$ defined for all time t. The mean of a random process, $m_X(t)$, is thus a function of time specified by

$$m_X(t) = E[X(t)] = E[X_t] = \int_{-\infty}^{+\infty} x_t f(x_t, t) dx_t \tag{2}$$

For the case where the mean of $X(t)$ does not depend on t, we have

$$m_X(t) = E[X(t)] = m_X \quad \text{(a constant).} \tag{3}$$

The variance of a random process, also a function of time, is defined by

$$\sigma_X^2(t) = E\{[X(t) - m_X(t)]^2\} = E[X_t^2] - [m_X(t)]^2 \tag{4}$$

2.3 Second and *N*th-order densities of a random process

For any pair of two random variables $X(t_1)$ and $X(t_2)$, we define the second-order densities of a random process as $f(x_1, x_2; t_1, t_2)$ or $f(x_1, x_2)$. Similarly, the nth order density functions for $\{X(t)\}$ at times $t_1, t_2, \cdots, t_n$ are given by $f(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n)$ or $f(x_1, x_2, \cdots, x_n)$.

2.4 Autocorrelation and autocovariance functions of random processes

Given two random variables $X(t_1)$ and $X(t_2)$, a measure of linear relationship between them is specified by $E[X(t_1)X(t_2)]$. For a random process, t_1 and t_2 go through all possible values, and therefore, $E[X(t_1)X(t_2)]$ can change and is a function of t_1 and t_2. The *autocorrelation function* of a random process $\{X(t)\}$ is thus defined by

$$R(t_1, t_2) = E[X(t_1)X(t_2)] = R(t_2, t_1) \tag{5}$$

The *autocovariance function* of a random process $\{X(t)\}$ is defined by

$$\begin{aligned} C(t_1, t_2) &= E[(X(t_1) - m_X(t_1)) \cdot (X(t_2) - m_X(t_2))] \\ &= R(t_1, t_2) - m_X(t_1) m_X(t_2) \end{aligned} \tag{6}$$

The normalized autocovariance function is defined by

$$\rho(t_1, t_2) = \frac{C(t_1, t_2)}{\sqrt{C(t_1, t_1) C(t_2, t_2)}} \tag{7}$$

2.5 Stationarity of random processes

A random process $\{X(t)\}$ is called ***strictly stationary*** (or strict-sense stationary,

SSS) if the sets of random variables $X(t_1), X(t_2), \cdots X(t_n)$ and $X(t_1+\Delta), X(t_2+\Delta),$ $\cdots X(t_n+\Delta)$ have the same probability density functions for all t_i, all n and all Δ, i.e.,

$$f(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n) = f(x_1, x_2, \cdots, x_n; t_1+\Delta, t_2+\Delta, \cdots, t_n+\Delta). \qquad (8)$$

Strict-sense stationarity seldom holds for random processes, except for some Gaussian processes. Therefore, weaker forms of stationarity are needed.

A random process is called ***N*th order stationary** (or **stationary of order *N***) if the condition of Eq. (8) holds for all $n \le N$ for N a fixed integer.

A much weaker form of stationarity, even weaker than stationarity of order two, is weak-sense stationarity (or **wide-sense stationarity, *WSS***). A random process $\{X(t)\}$ is wide-sense stationary if

$$E[X(t)] = m \text{ (constant) for all } t. \qquad (9)$$

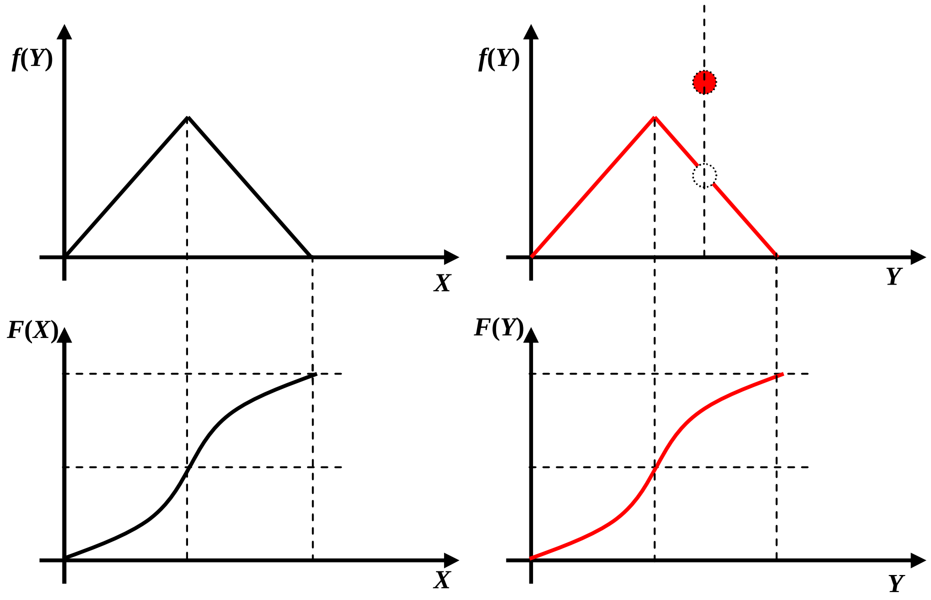
$$R(t_1, t_2) = R(|t_2 - t_1|) = R(|t_2 - t_1|), \text{ for all } t_1 \text{ and } t_2. \qquad (10)$$

2.6 Equality and continuity of random processes

Equality

Two random processes $\{X(t)\}$ and $\{Y(t)\}$ are equal (everywhere) if their respective samples $x(t, \omega_i) = y(t, \omega_i)$ for every t and ω_i.

[Note that "$x(t, \omega_i) = y(t, \omega_i)$ for every ω_i" is not the same as "$x(t, \omega_i) = y(t, \omega_i)$ with probability 1".]



Mean square equality

Two random processes $\{X(t)\}$ and $\{Y(t)\}$ are equal in the mean-square sense if

$$E\left[|X(t)-Y(t)|^2\right]=0 \qquad (11)$$

for every t.

Equality in the MS sense leads to the following conclusions: We denote by $\mathcal{A}_t$ the set of outcomes ω_i such that $x(t, \omega_i) = y(t, \omega_i)$ for a specific t, and by $\mathcal{A}_\infty$ the set of outcomes ω_i such that $x(t, \omega_i) = y(t, \omega_i)$ for every t. From Eq. (11) it follows that $x(t, \omega_i) - y(t, \omega_i) = 0$ with probability 1, hence $P(\mathcal{A}_t) = P(\Omega) = 1$ where Ω represents the sample space. It does not follow, however, that $P(\mathcal{A}_\infty) = 1$. In fact, $P(\mathcal{A}_\infty)$ is the intersection of all sets $\mathcal{A}_t$ as t ranges over the entire axis and $P(\mathcal{A}_\infty)$ might even equal 0.

Stochastic continuity

(1) Continuous in probability

A random process $\{X(t)\}$ is called continuous in probability at t if for any $\varepsilon > 0$,

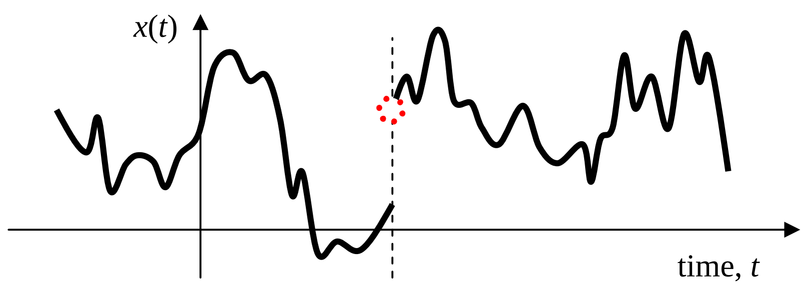
$$P(|X(t+h)-X(t)|>\varepsilon)\underset{h\to 0}{\longrightarrow} 0 \qquad (12)$$

(2) Continuous in mean-square sense

A random process $\{X(t)\}$ is called mean-square (MS) continuous at t if

$$E\{[X(t+\varepsilon)-X(t)]^2\}\underset{\varepsilon\to 0}{\longrightarrow} 0\,. \qquad (13)$$

Is it possible to have a *realization* which is discontinuous at time t whereas the random process is mean-square continuous?



Remarks

(i) A random process $\{X(t)\}$ is MS continuous if its autocorrelation function is continuous.

[Proof]

$$E\{[X(t+\varepsilon)-X(t)]^2\}=R(t+\varepsilon,t+\varepsilon)-2R(t+\varepsilon,t)+R(t,t) \qquad (14)$$

If $R(t_1,t_2)$ is continuous, then the RHS of Eq. (14) approaches zero as $\varepsilon \to 0$.
[Note that $R(t_1,t_2)$ is used in the above equation to simply the expression.]

(ii) Suppose that Eq. (13) holds for every t in an interval I. It follows that almost all samples (or realizations) of $\{X(t)\}$ will be continuous for a particular point of I. It does not follow, however, that these samples (or realizations) of $\{X(t)\}$ will be continuous for every point in I.

(iii) If $\{X(t)\}$ is MS continuous, then its mean is continuous, i.e.,

$$m_X(t+\varepsilon) \underset{\varepsilon\to 0}{\to} m_X(t) \tag{15}$$

[Proof]

$$E\left\{[X(t+\varepsilon)-X(t)]^2\right\} \ge \left\{E[X(t+\varepsilon)-X(t)]\right\}^2 \tag{16}$$

[Note: $Var(X)=E(X^2)-[E(X)]^2 \ge 0$]

Therefore,

$$E[X(t+\varepsilon)-X(t)] \underset{\varepsilon\to 0}{\to} 0$$

(3) Almost-surely continuous

A random process $\{X(t)\}$ is called almost-surely continuous at t if

$$P\left(\omega : \lim_{h\to 0} |X(t+h;\omega)-X(t;\omega)| = 0\right) = 1 \tag{17}$$

If $\{X(t)\}$ is continuous in probability (mean-square continuous, almost-surely continuous) at every t, then it is said to be continuous in probability (mean-square continuous, almost-surely continuous).

Example 1 Suppose that $X(t)$ is a random process with $m_X(t)=5$ and $R(t_1,t_2)=25+3e^{-0.6|t_1-t_2|}$. Determine the mean, the variance and the covariance of the random variables $U=X(6)$ and $V=X(9)$.

[Solution] $E(U)=E[X(6)]=m_X(6)=5$, $E(V)=E[X(9)]=m_X(9)=5$

$$Var(U)=E\{[X(6)]^2\}-\{E[X(6)]\}^2 = R(6,6)-25=28-25=3$$

$$Var(V)=E\{[X(9)]^2\}-\{E[X(9)]\}^2 = R(9,9)-25=28-25=3$$

$$Cov(U,V)=C(6,9)=R(6,9)-E[X(6)]E[X(9)]=3e^{-1.8}=0.496$$

Example 2 The integral of a stochastic process $X(t)$ is a random variable. Let $S=\int_a^b X(t)dt$ it yields

$$m_S = E(S) = E\left[\int_a^b X(t)dt\right] = \int_a^b E[X(t)]dt = \int_a^b m_X(t)dt$$

$$S^2 = \int_a^b \int_a^b X(t_1)X(t_2)dt_1dt_2$$

$$E(S^2) = E\left\{\int_a^b \int_a^b X(t_1)X(t_2)dt_1dt_2\right\} = \int_a^b \int_a^b E\{X(t_1)X(t_2)\}dt_1dt_2$$

$$= \int_a^b \int_a^b R(t_1, t_2)dt_1dt_2$$

2.7 Stochastic Convergence

A *random sequence* or a discrete-time random process is a sequence of random variables $[X_1(\omega), X_2(\omega), \ldots, X_n(\omega)] \equiv \{X_n(\omega)\}$, $\omega \in \Omega$. For a specific ω, $\{X_n(\omega)\}$ is a sequence of numbers that might or might not converge. The notion of convergence of a random sequence can be given several interpretations:

Sure convergence (convergence everywhere)

The sequence of random variables $\{X_n(\omega)\}$ converges surely to the random variable $X(\omega)$ if the sequence of functions $X_n(\omega)$ converges to $X(\omega)$ as $n \to \infty$ for all $\omega \in \Omega$, i.e.,

$$X_n(\omega) \to X(\omega) \quad \text{as } n \to \infty \qquad \text{for all } \omega \in \Omega. \qquad (18)$$

Remarks

(i) The concept of sure convergence of a sequence of random variables can be used to describe the convergence of a sequence of numbers, as illustrated in figure 3. A sequence of real numbers x_n converges to the real number x if, given any $\varepsilon > 0$, we can always specify an integer N such that for all values of n beyond N we can guarantee that $|x_n - x| < \varepsilon$.

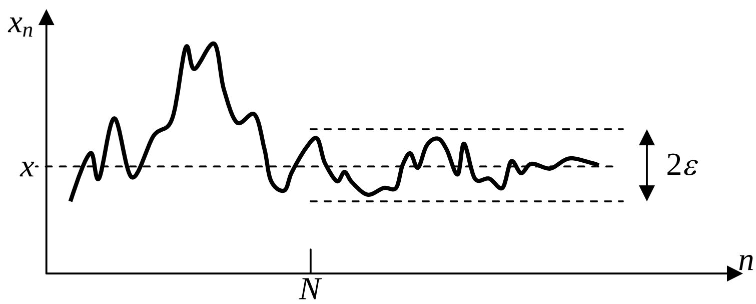


Figure 3. Convergence of a sequence of numbers.

(ii) Sure convergence requires that the sample sequence corresponding to every ω converges. However, it does not require that all the sample sequences converge to the same value; that is the sample sequences for different ω *and* ω' can converge to different values.

Almost-sure convergence (convergence almost everywhere, convergence with probability 1)

The sequence of random variables $\{X_n(\omega)\}$ converges almost surely to a random variable $X(\omega)$ if the sequence of functions $X_n(\omega)$ converges to $X(\omega)$ as $n \to \infty$ for all $\omega \in \Omega$, except possibly on a set of probability zero (see Figure 4); i.e.,

$$P\left[\omega : X_n(\omega) \underset{n\to\infty}{\longrightarrow} X(\omega)\right] = 1. \tag{19}$$

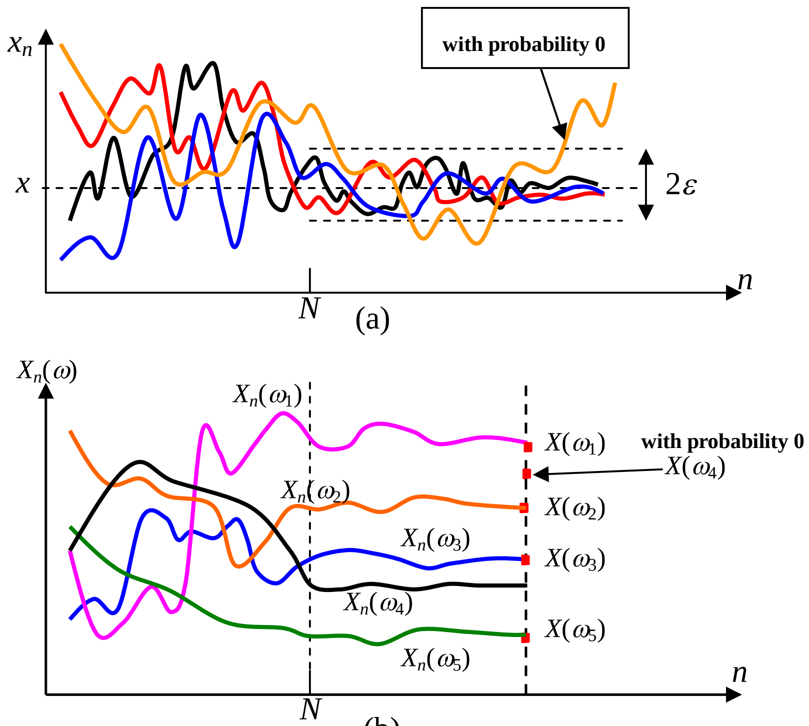


Figure 4. Almost-sure convergence. (a) converging to a constant; (b) converging to a random variable X.

Mean-square convergence

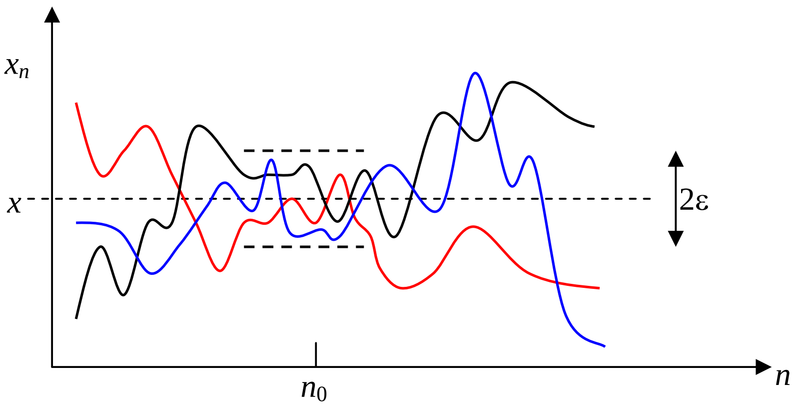
The sequence of random variables $\{X_n(\omega)\}$ converges in the mean square sense to the random variable $X(\omega)$ if

$$E\left[(X_n(\omega) - X(\omega))^2\right] \to 0 \quad \text{as} \quad n \to \infty \tag{20}$$

Convergence in probability

The sequence of random variables $\{X_n(\omega)\}$ converges in probability to the random variable $X(\omega)$ if, for any $\varepsilon > 0$,

$$P\left[|X_n(\omega) - X(\omega)| > \varepsilon\right] \to 0 \quad \text{as} \quad n \to \infty \tag{21}$$



Convergence in probability for the case where the limiting random variable is a constant x

Convergence in distribution

The sequence of random variables $\{X_n(\omega)\}$ with cumulative distribution functions $\{F_n(x)\}$ converges in distribution to the random variable $X(\omega)$ with cumulative distribution functions $F(x)$ if

$$F_n(x) \to F(x) \quad \text{as} \quad n \to \infty \tag{22}$$

for all x at which $F(x)$ is continuous.

Remarks

(i) The principal difference in the definitions of convergence in probability and convergence with probability one is that the limit is outside the probability in the former and inside the probability in the latter.

(ii) Convergence with probability one applies to the individual realizations of the random process. Convergence in probability does not.

(iii) The weak law of large numbers is an example of convergence in probability.

(iv) The strong law of large numbers is an example of convergence with probability 1 whereas the central limit theorem is an example of convergence in distribution.

Example 3 Let ω be selected at random from the interval S = [0, 1], where we assume that the probability that ω is in a subinterval of S is equal to the length of the subinterval. For n = 1, 2, … we define the following five sequences of random variables:

$$U_n(\omega) = \omega / n, \quad V_n(\omega) = \omega\left(1 - \frac{1}{n}\right), \quad W_n(\omega) = \omega \cdot e^n,$$

$$Y_n(\omega) = \cos 2n\pi\omega, \quad Z_n(\omega) = e^{-n(n\omega - 1)}$$

Determine the stochastic convergence of these random sequences and identify the limiting random variable.

Weak Law of Large Numbers (WLLN)

Let $f(\cdot)$ be a density with finite mean μ and finite variance. Let $\overline{X}_n$ be the sample mean of a random sample of size n from $f(\cdot)$, then for any $\varepsilon > 0$,

$$P[-\varepsilon < \overline{X}_n - \mu < \varepsilon] \to 1 \quad \text{as} \quad n \to \infty$$

Strong Law of Large Numbers (SLLN)

Let $f(\cdot)$ be a density with finite mean μ and finite variance. Let $\overline{X}_n$ be the sample mean of a random sample of size n from $f(\cdot)$, then for any $\varepsilon > 0$,

$$P\left[\lim_{n\to\infty} \overline{X}_n = \mu\right] = 1$$

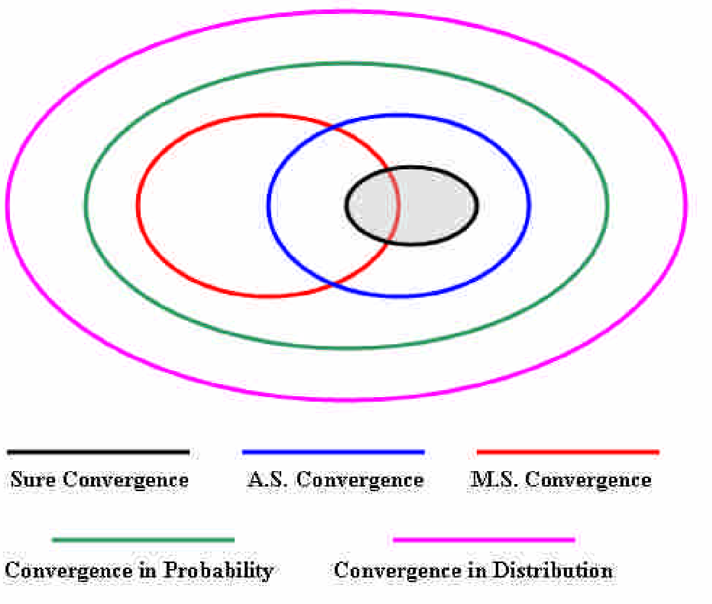
The Central Limit Theorem

Let $f(\cdot)$ be a density with mean μ and finite variance σ^2.

Let $\overline{X}_n$ be the sample mean of a random sample of size n from $f(\cdot)$. Then

$$Z_n = \frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}}$$

approaches the standard normal distribution as n approaches infinity. [Note: It is equivalent to say that $\overline{X}_n$ approaches a normal distribution with expected value μ and variance σ^2/n as n approaches infinity.]



[Solution]

$U_n(\omega) \underset{n\to\infty}{\to} U(\omega) = 0$ for every $\omega \in S$. Therefore, it converges surely to a constant 0.

$V_n(\omega) \underset{n\to\infty}{\to} V(\omega) = \omega$ for every $\omega \in S$. Therefore, it converges surely to a random variable which is uniformly distributed over [0, 1].

$$E\left[(V_n(\omega)-\omega)^2\right] = E\left[\left(\frac{\omega}{n}\right)^2\right] = \int_0^1 \frac{\omega^2}{n^2} d\omega = \frac{1}{3n^2}.$$

$E\left[(V_n(\omega)-\omega)^2\right] \underset{n\to\infty}{\to} 0$. Thus, the sequence $V_n(\omega)$ converges in the mean-square sense.

$W_n(\omega)$ converges to 0 for $\omega = 0$, but diverges to infinite for all other values of ω. Therefore, it does not converge.

$Y_n(\omega)$ converges to 1 for $\omega = 0$ and $\omega = 1$, but oscillates between –1 and 1 for all other values of ω. Therefore, it does not converge.

$Z_n(\omega = 0) = e^n \underset{n\to\infty}{\to} +\infty$, $Z_n(\omega) \underset{n\to\infty}{\to} 0$ for $\omega > (1/n)$.

$P[\omega > 0] = 1$. Thus, $Z_n(\omega)$ converges almost surely to 0.

$$E\left[(Z_n(\omega)-0)^2\right] = E\left[e^{-2n(n\omega-1)}\right] = e^{2n}\int_0^1 e^{-2n^2\omega} d\omega = \frac{e^{2n}}{2n^2}\left(1-e^{-2n^2}\right).$$

As n approaches infinity, the rightmost term in the above equation approaches infinity. Therefore, the sequence $Z_n(\omega)$ does not converge in the mean square sense even though it converges almost surely.

2.8 Ergodic Theorem

A discrete time random process $\{X_n, n = 0, 1, 2, \dots\}$ is said to satisfy an ergodic theorem if there exists a random variable X such that in some sense

$$\sum_{i=0}^{n-1} X_i / n \underset{n\to\infty}{\to} X \qquad (23)$$

The type of convergence determines the type of the ergodic theorem. For example, if the convergence is in mean square sense, the result is called a mean ergodic theorem. If the convergence is with probability one, it is called an almost sure ergodic theorem.

A continuous time random process $\{X(t)\}$ is said to satisfy an ergodic theorem if there exists a random variable X such that

$$\frac{1}{T}\int_0^T X(t)dt \underset{T\to\infty}{\to} X \tag{24}$$

where again the type of convergence determines the type of the ergodic theorem.

Note that we only require the time average to converge, however, it does not need to converge to some constant, for example the common expectation of the random process. In fact, ergodic theorem can hold even for nonstationary random processes where $E[X(t)]$ does depend on time t.

The Mean-Square Ergodic Theorem

Let $\{X_n\}$ be a random process with mean function $E[X_n]$ and covariance function $C_X(k,j)$. (The process need not to be even weakly stationary.) Necessary and sufficient conditions for the existence of a constant m such that

$$E\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i - m\right)^2\right] \to 0 \quad \text{as} \quad n\to\infty \tag{25}$$

are that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} E(X_i) = m \tag{26}$$

and

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{i=1}^{n}\sum_{k=1}^{n} C_X(i,k) = 0 \tag{27}$$

The above theorem shows that one can expect a sample average to converge to a constant in mean square sense if and only if the average of the means converges and if the memory dies out asymptotically, that is, if the covariance decreases as the lag increases in the sense of Eq. (27).

Mean-Ergodic Processes

A random process $\{X(t)\}$ with constant mean μ is said to be mean-ergodic if it satisfies

$$P\left\{\frac{1}{2T}\int_{-T}^{T} x(t)dt \underset{T\to\infty}{\longrightarrow} \mu\right\} = 1. \tag{28}$$

Strong or Individual Ergodic Theorem

Let $\{X_n\}$ be a strictly stationary random process with $E[X_n] < \infty$. Then the sample mean $\sum_{i=1}^{n} X_i / n$ converges to a limit with probability one.

Remarks

(i) Let $\{X(t)\}$ be a wide-sense stationary random process with constant mean μ and covariance function C(t). Then $\{X(t)\}$ is mean-ergidic if and only if

$$\frac{1}{2T}\int_{-2T}^{2T} C(\tau)\left(1-\frac{|\tau|}{2T}\right)d\tau \underset{T\to\infty}{\longrightarrow} 0 \tag{29}$$

or equivalently,

$$\frac{1}{T}\int_{0}^{2T} C(\tau)\left(1-\frac{\tau}{2T}\right)d\tau \underset{T\to\infty}{\longrightarrow} 0 \tag{30}$$

(ii) Let $\{X(t)\}$ be a wide-sense stationary random process with constant mean μ and covariance function C(t) and $\int_{-\infty}^{\infty} |C(\tau)|\, d\tau < \infty$, then $\{X(t)\}$ is mean-ergidic.

(iii) Let $\{X(t)\}$ be a wide-sense stationary random process with constant mean μ and covariance function $C(t)$ and $C(0) < \infty$ and $C(\tau) \underset{|\tau|\to\infty}{\longrightarrow} 0$, then $\{X(t)\}$ is mean-ergidic.

3. Examples of Stochastic Processes

iid random process

A discrete time random process $\{X(t), t = 1, 2, \dots\}$ is said to be independent and identically distributed (*iid*) if any finite number, say k, of random variables $X(t_1)$, $X(t_2), \dots, X(t_k)$ are mutually independent and have a common cumulative distribution function $F_X(\cdot)$. The joint cdf for $X(t_1), X(t_2), \dots, X(t_k)$ is given by

$$\begin{aligned} F_{X_1,X_2,\cdots,X_k}(x_1,x_2,\cdots,x_k) &= P\left(X_1 \le x_1, X_2 \le x_2, \cdots, X_k \le x_k\right) \\ &= F_X(x_1)F_X(x_2)\cdots F_X(x_k) \end{aligned} \tag{31}$$

It also yields

$$p_{X_1,X_2,\cdots,X_k}(x_1,x_2,\cdots,x_k) = p_X(x_1)\,p_X(x_2)\cdots p_X(x_k) \tag{32}$$

where $p(x)$ represents the common probability mass function.

Let $\{X_n, n = 0, 1, 2, \dots\}$ be a sequence of independent Bernoulli random variables with parameter p. It is therefore an iid Bernoulli random process and $E[X_n] = p$ and $Var[X_n] = p(1-p)$.

Random walk process

Let $\xi_1, \xi_2, \dots$ be integer-valued random variables having common probability mass function $f(\cdot)$. Let X_0 be an integer-valued random variable that is independent of the ξ_i's and let X_n be the sum of these random variables, i.e.,

$$X_n = X_0 + \sum_{i=1}^{n} \xi_i \quad . \tag{33}$$

The sequence $\{X_n , n \ge 0\}$ is called a random walk process.

Let π_0 denote the probability mass function of X_0. The joint probability of X_0, X_1, … X_n is

$$\begin{aligned}
&P(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n) \\
&= P\left(X_0 = x_0, \xi_1 = x_1 - x_0, \cdots, \xi_n = x_n - x_{n-1}\right) \\
&= P(X_0 = x_0)P(\xi_1 = x_1 - x_0)\cdots P(\xi_n = x_n - x_{n-1}) \\
&= \pi_0(x_0) f(x_1 - x_0)\cdots f(x_n - x_{n-1}) \\
&= \pi_0(x_0)P(x_1 \mid x_0)\cdots P(x_n \mid x_{n-1})
\end{aligned} \tag{34}$$

Thus,

$$\begin{aligned}
&P(X_{n+1} = x_{n+1} \mid X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n) \\
&= \frac{P(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n, X_{n+1} = x_{n+1})}{P(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n)} \\
&= \frac{\pi_0(x_0)P(x_1 \mid x_0)\cdots P(x_n \mid x_{n-1}) \cdot P(x_{n+1} \mid x_n)}{\pi_0(x_0)P(x_1 \mid x_0)\cdots P(x_n \mid x_{n-1})} \\
&= P(x_{n+1} \mid x_n)
\end{aligned} \tag{35}$$

The property

$$P(X_{n+1} = x_{n+1} \mid X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n) = P(X_n = x_{n+1} \mid X_n = x_n) \tag{36}$$

is known as the Markov property. A special case of random walk is the Brownian motion.

Gaussian process

A random process $\{X(t)\}$ is said to be a Gaussian random process if all finite collections of the random process, $X_1=X(t_1)$, $X_2=X(t_2)$, …, $X_k=X(t_k)$, are jointly Gaussian random variables for all k, and all choices of t_1, t_2, …, t_k.

Joint pdf of jointly Gaussian random variables X_1, X_2, …, X_k:

$$f_{X_1,\cdots,X_k}(x_1,\cdots,x_k) = \frac{1}{\sqrt{(2\pi)^k |C|}} \exp\left[-\frac{1}{2}\left(X - m\right)^T C^{-1}(X - m)\right] \tag{37}$$

where $X^T = (X_1, \cdots, X_k)$, $m^T = (E(X_1), \cdots, E(X_k))$, and

$$C = \begin{bmatrix} C(t_1,t_1) & C(t_1,t_2) & \cdots & C(t_1,t_k) \\ C(t_2,t_1) & C(t_2,t_2) & \cdots & C(t_2,t_k) \\ \vdots & \vdots & \ddots & \vdots \\ C(t_k,t_1) & C(t_k,t_1) & \cdots & C(t_k,t_k) \end{bmatrix} \tag{38}$$

Time series – AR and MA random processes

A wide-sense stationary Autoregressive (AR(k)) model

$$X(t) = \sum_{i=1}^{k} a_i X(t-i) + \varepsilon(t) \tag{39}$$

where $E\left[X(t)\right]=0$, $R\left[X(t),X(s)\right]=R\left(\left|t-s\right|\right)$, and $\varepsilon(t)\sim N(0,\sigma_{\varepsilon}^{2})$.