INFORMATION AND COMMUNICATION THEORY

Unit-1: Probability Theory
Unit-2: Stochastic Processes
Unit-3: Estimation & Hypothesis Testing
Unit-4: Information Theory
Unit-5: Statistical Modeling of Noise

Module-3-Unit-2

- Stochastic Processes
- Statistical Modeling of Noise

Lecture-1 (Unit-2, Module-3)

Stochastic Process

Stochastic (Random) Process

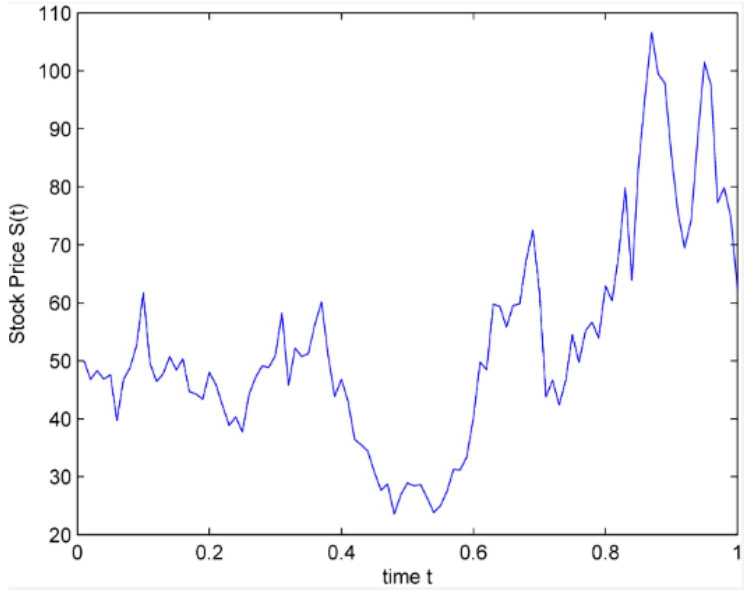
Random variable is a rule to assign a number $x(\xi)$ to every outcome ξ of an experiment S.

A stochastic process is a rule to assign a function $X(t, \xi)$ to every outcome ξ of an experiment S.

Stochastic process is represented by $X(t)$. It has following interpretations –

- It is a ***family (or ensemble)*** of functions $X(t,\xi)$. Here t and ξ are variables.
- It is a ***single time-function*** (or a ***sample*** of the given process). In this case t is a variable but ξ is fixed.
- It is the ***random variable*** equal to the ***state of given process at time t***. In this case ξ is a variable but t is fixed.
- It is a ***number*** if both t and ξ are fixed.

At time
$t_1 \in [0,\infty)$
$t_2 \in [0,\infty)$



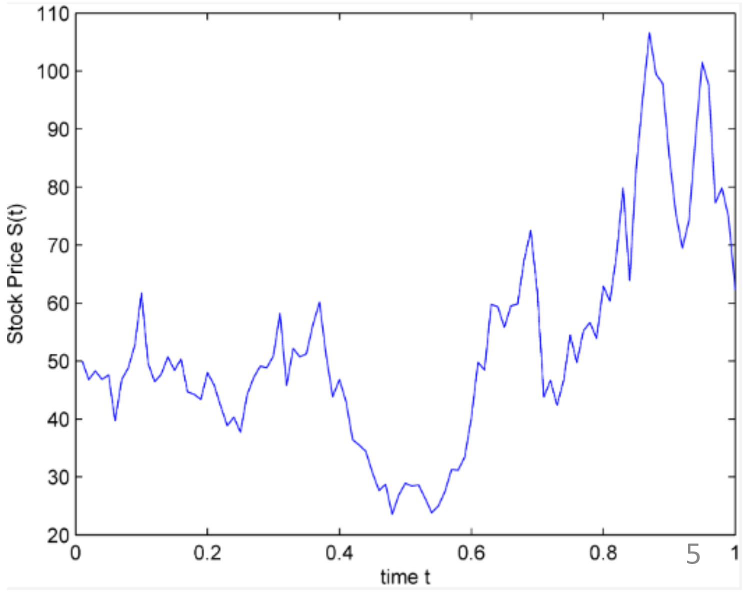
Random Process

In real-life applications, we are often interested in multiple observations of random values over a period of time.

Example:

observing the stock price of a company over the next few months. In particular, let $S(t)$ be the stock price at time $t \in [0, \infty), t = 0$ refers to current time.

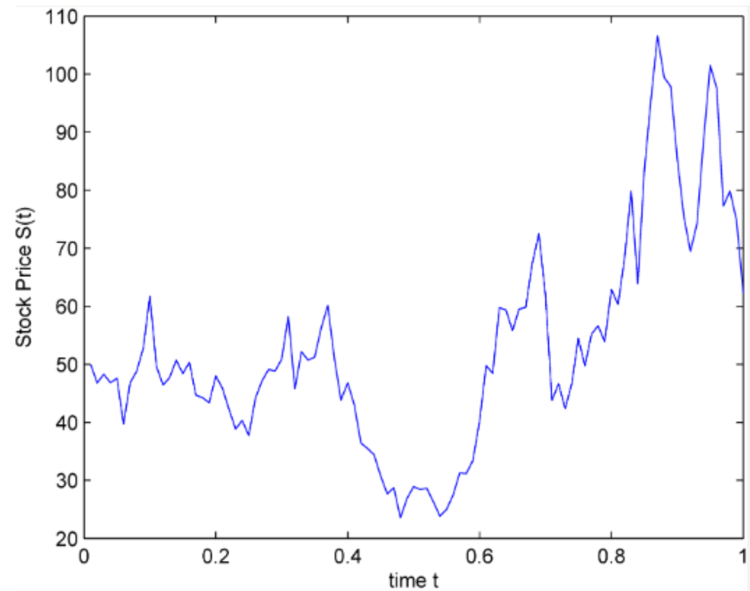
Figure shows a possible outcome of this random experiment from time $t = 0$ to time $t = 1$.



A random process is a collection of random variables usually indexed by time.

At $t_1 \in [0, \infty), S(t_1)$ is a random variable. At time $t_2 \in [0, \infty)$, $S(t_2)$ is another random variable could have a different PDF. The collective values of $S(t)$ for $t \in [0, \infty)$ is a **random process** or a **stochastic process**.

$$\{S(t), t \in [0, \infty)\}$$



The function shown in this figure is just one of the many possible outcomes of this random experiment.

We call each of these possible functions of $S(t)$ a **sample function** or **sample path**.

Continuous-time random processes

Let $N(t)$ be the number of customers who have *visited a bank* from $t = 9$ (when the bank opens at $9:00\ AM$) until time t, on a given day, for $t \in [9,16]$.

Here, we measure t in hours, but t can take any real value between 9 and 16.

We assume that $N(9) = 0$, and $N(t) \in \{0,1,2,\dots\}$ for all $t \in [9,16]$.

Note that for any time t_1, the random variable $N(t_1)$ is a discrete random variable.

Thus, $N(t)$ is a ***discrete-valued random process***.

However, since t can take any real value between 9 and 16, $N(t)$ is a ***continuous- time random process***.

There are uncountable number of random variables. For example, for any given $t_1 \in [9,16]$, $N(t_1)$ is a random variable.

Thus, the ***random process $N(t)$ consists of an uncountable number of random variables***.

Discrete-time random processes

A discrete-time random process is a process

$\{X(t), t \in J\}$, where J is a countable set i.e. $J = \{t_1, t_2, \cdots\}$.

We usually define $X(t_n) = X(n)$ or $X(t_n) = X(n), for\ n = 1,2, \cdots$, (the index values n could be from any countable set such as natural number or integers i.e. $\mathbb{N}\ or\ \mathbb{Z}$).

Therefore, ***a discrete-time random process is just a sequence of random variables***.

Or, if the process is defined for all integers, then we may show the process by

$$\{X(n), n \in \mathbb{Z}\} \; or \; \{X_n, n \in \mathbb{Z}\}$$

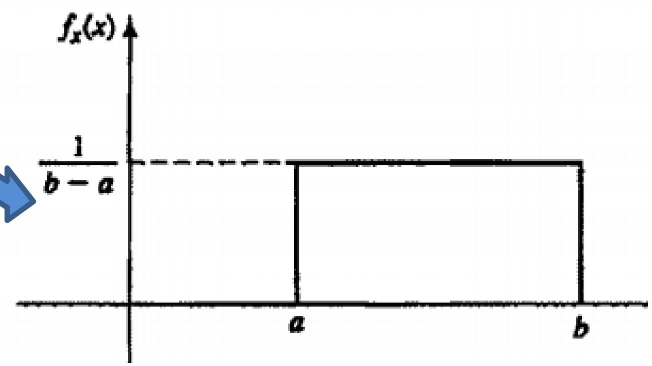
Discrete-time processes are sometimes obtained from continuous-time processes by discretizing time (sampling at specific times).

Example:

You have 1000 dollars to put in an account with interest rate R, compounded annually. That is, if X_n is the value of the account at year n, then

$$X_n = 1000(1+R)^n, \text{ for } n = 0,1,2,\cdots.$$

If $R \sim Uniform(0.04, 0.05)$, determine sample functions for the random process $\{X_n, n = 0,1,2,\ldots\}$ and the expected value of account at year three i.e. $E[X_3]$.



The probability density function of the continuous uniform distribution is:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \le x \le b, \\ 0 & \text{for } x < a \text{ or } x > b \end{cases}$$

Solution:

Here, the randomness in X_n comes from the random variable R.
Here,

$$X_n = 1000(1+R)^n,\ \forall\, n \in \{0,1,2,\cdots\}$$

So sample functions are of the form

$$f(n) = 1000(1+R)^n, \text{ for } n = 0,1,2,\cdots$$

$$\text{where} \quad R \in (0.04, 0.05)$$

For any value of $R \in (0.04, 0.05)$ a sample function of random process X_n can be obtained.

$$X_n = 1000(1+R)^n = 1000Y^n, \text{ thus}$$

$$Y \sim Uniform(1.04, 1.05) \text{ and}$$

$$f_Y(y) = \begin{cases} 100, & 1.04 \le y \le 1.05 \\ 0, & elsewhere \end{cases}$$

Expected value of sample function at n=3

$$E[X_3] = E[1000Y^3]$$

We know that

Continuous random variable

The expected value or mean of a random variable $\boldsymbol{x}$ is given as

$$E\{\boldsymbol{x}\} = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

pdf

$$f_Y(y) = \begin{cases} 100, & 1.04 \le y \le 1.05 \\ 0, & elsewhere \end{cases}$$

$$= 1000 \int_{1.04}^{1.05} 100\, y^3\, dy$$

$$\approx 1141$$

CDF of Random Process

A stochastic process is a non-countable collection of random variables, one for each t.

For a specific t, $X(t)$ is random variable with distribution

$$F(x,t) = P\{X(t) \leq x\}$$

This function depend on t and it equals the probability of the event $\{X(t) \leq x\}$ consisting of all outcomes ξ, such that, at specific time t, the samples of $X(t,\xi)$ of the given process does not exceed x.

$F(x,t)$ is called first-order distribution of the process $X(t)$

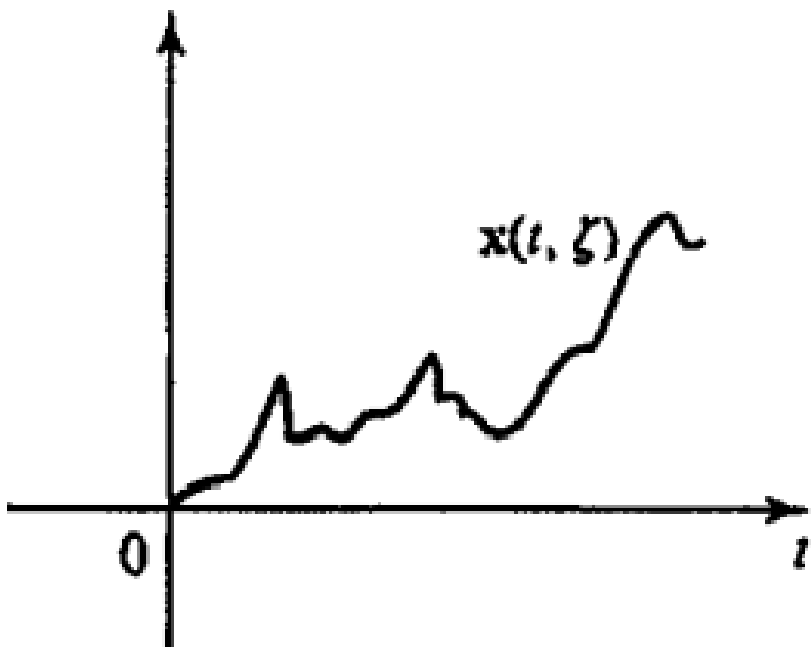
$f(x,t) = \frac{\partial F(x,t)}{\partial x}$ is the first-order density of the process $X(t)$

Continued…

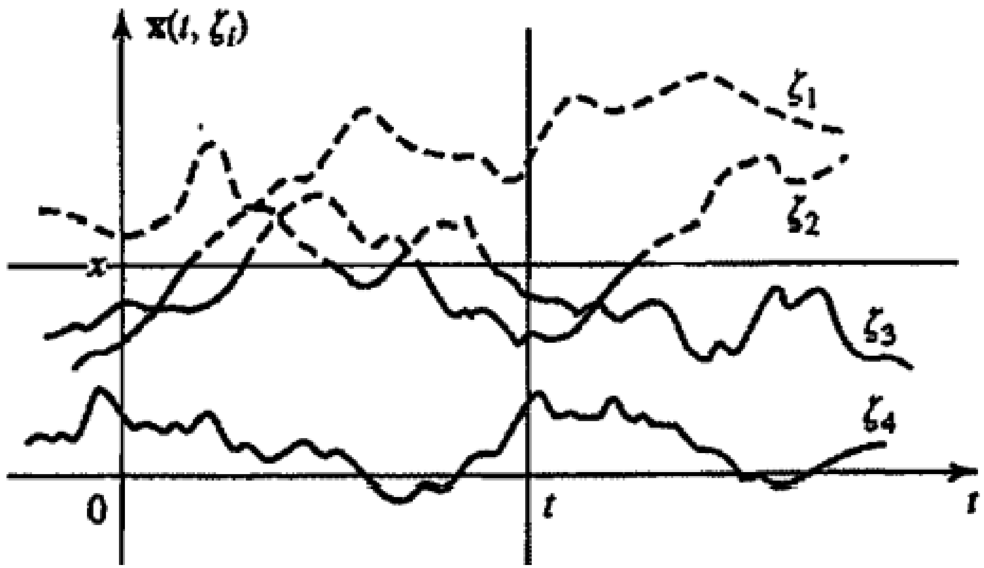
Consider the random process $\{X(t),\ t \in J\}$. For any $t_0 \in J$,

$X(t_0)$ is a random variable, so we can write its CDF as

$$F_{X(t_0)}(x) = P\{X(t_0) \leq x\}.$$



If the experiment is performed $\boldsymbol{n}$ times, then $\boldsymbol{n}$ functions $\boldsymbol{x}(\boldsymbol{t}, \boldsymbol{\varepsilon_i})$ are observed, one for each trial as shown in figure below:



If $t_1, t_2 \in J$, then the joint CDF $X(t_1)$ and $X(t_2)$ is

$$F(x_1, x_2, t_1, t_2) = F_{X(t_1)\,X(t_2)}(x_1, x_2)$$

$$= P\{X(t_1) \leq x_1, X(t_2) \leq x_2\}$$

Example:

Consider the random process $\{X_n, n = 0,1,2,\cdots\}$ in which X_i's are i.i.d. standard normal random variables.

Find $f_{X_n}(x)$ and $f_{X_m,X_n}(x_1,x_2)$

Solution: Since $X_n \sim N(0,1)$,

PDF of normal RV is

$$f_{X_n}(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(x-\eta)^2/2\sigma^2}$$

$$f_{X_n}(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$$

Since it is i.i.d.

$$f_{X_m,X_n}(x_1, x_2) = f_{X_m}(x_1) f_{X_n}(x_2)$$

$$= \left(\frac{1}{\sqrt{2\pi}}\, e^{-{x_1}^2/2}\right)\left(\frac{1}{\sqrt{2\pi}}\, e^{-{x_2}^2/2}\right)$$

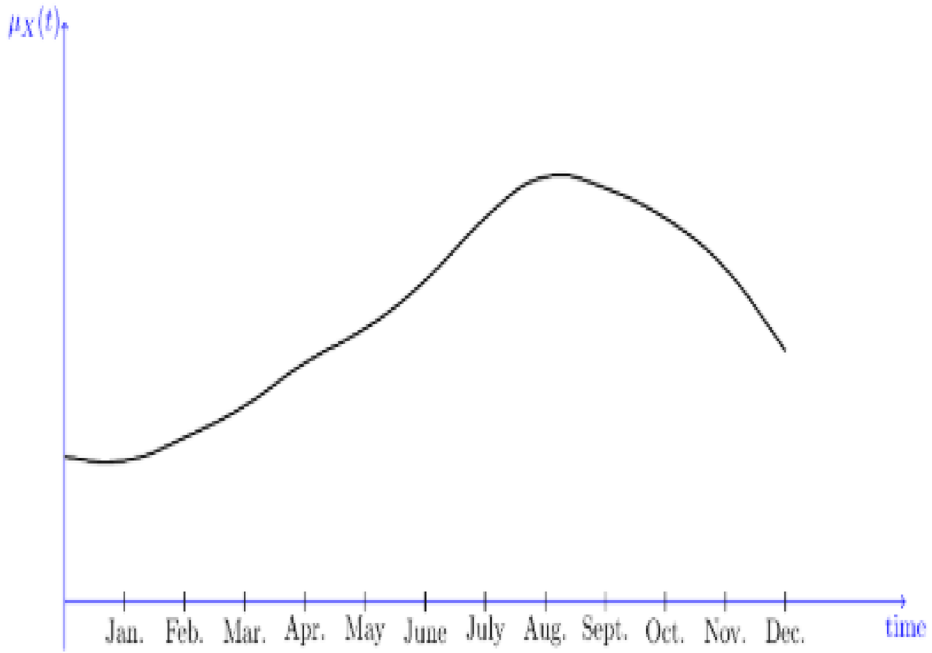
$$= \frac{1}{\sqrt{2\pi}}\, e^{-({x_1}^2+{x_2}^2)/2}$$

Mean Functions

The mean function gives an idea about how the random process behaves on average as time evolves.

$$\mu_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x\, f(x,t)\, dx \qquad \mu_X(n) = E[X_n]$$

For example, if $X(t)$ is the temperature in a certain city, the mean function $\mu_X(t)$ might look like the function shown in Figure below. *(the expected value of $X(t)$ is lowest in the winter and highest In the summer)*



Correlation and Covariance Functions

The mean function $\mu_X(t)$ gives the expected value of $X(t)$ at time t, but it does not give any information about how $X(t_1)$ and $X(t_2)$ are related.

The **autocorrelation function**, $R_X(t_1, t_2)$ is defined for a random process $\{X(t), t \in J\}$

We know that: $$\mu_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x\, f(x,t)\, dx$$

$$R_X(t_1, t_2) = E[X(t_1)\, X(t_2)]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1\, x_2 f(x_1, x_2, t_1, t_2) dx_1\, dx_2 \text{ for } \qquad t_1, t_2 \in J$$

x1 and x2 are two values of X at two different time t1 and t2

The auto covariance function

We know that:

$$\sigma^2 = E\{(\boldsymbol{x}-\eta)^2\} = \int_{-\infty}^{\infty} (x-\eta)^2 f(x)\, dx$$

$$\sigma^2 = E\{\boldsymbol{x}^2\} - E^2\{\boldsymbol{x}\}$$

How?

Proof:

$$\begin{aligned}\sigma^2 = E\{(\boldsymbol{x}-\eta)^2\} &= E\{\boldsymbol{x}^2 - 2\boldsymbol{x}\eta + \eta^2\} \\ &= E\{\boldsymbol{x}^2\} - 2\eta E\{\boldsymbol{x}\} + \eta^2 \\ &= E\{\boldsymbol{x}^2\} - 2\eta\,\eta + \eta^2 \\ &= E\{\boldsymbol{x}^2\} - 2\,\eta^2 + \eta^2 \\ &= E\{\boldsymbol{x}^2\} - \eta^2\end{aligned}$$

$$\begin{aligned}C_X(t_1,t_2) &= Cov[X(t_1)\,X(t_2)] \\ &= E[\{X(t_1) - E[X(t_1)]\}\ \{X(t_2) - E[X(t_2)]\}] \\ &= R_X(t_1,t_2) - \mu_X(t_1)\mu_X(t_2) \qquad \text{for} \qquad t_1,t_2 \in J\end{aligned}$$

How?

HELP:

$$C_X(t_1,t_2) = Cov[X(t_1)\,X(t_2)]$$
$$= E[\{X(t_1) - E[X(t_1)]\}\{X(t_2) - E[X(t_2)]\}]$$
$$= R_X(t_1,t_2) - \mu_X(t_1)\mu_X(t_2)$$
for $t_1, t_2 \in J$

$$= cov[X(t_1)\,X(t_2)]$$

$$= E[\{X(t_1) - E[X(t_1)]\}\{X(t_2) - E[X(t_2)]\}]$$

$$= E[\{X(t_1)X(t_2) - X(t_1)E[X(t_2)] - X(t_2)E[X(t_1)] + E[X(t_1)]E[X(t_2)]\}]$$

$$= E[\{X(t_1)X(t_2) - X(t_1)\mu_2 - X(t_2)\mu_1 + \mu_1\mu_2\}]$$

$$= E[X(t_1)X(t_2)] - \mu_2 E[X(t_1)] - \mu_1 E[X(t_2)] + \mu_1\mu_2$$

$$= Rx(t_1,t_2) - \mu_2\mu_1 - \cancel{\mu_1\mu_2} + \cancel{\mu_1\mu_2}$$

$$= Rx(t_1,t_2) - \mu_2\mu_1$$

Continued…

The **autocovariance function**, $C_X(t_1, t_2)$ is defined for a random process $\{X(t), t \in J\}$

Special condition:

If $t_1 = t_2 = t$

$$R_X(t,t) = E[X(t)\,X(t)]$$
$$= E[X(t)^2]$$

$$C_X(t,t) = E[\{X(t) - E[X(t)]\}\ \{X(t) - E[X(t)]\}]$$
$$= Var[X(t)]$$

$$= R_X(t_1, t_1) - \mu_X(t_1)\mu_X(t_1)$$

$$\text{for} \quad t_1 \in J$$

Continued…

- Intuitively, $C_X(t_1, t_2)$ shows how $X(t_1)$ and $X(t_2)$ move relative to each other.

- If large values of $X(t_1)$ tend to imply large values of $X(t_2)$, then $\{X(t_1) - E[X(t_1)]\}\,\{X(t_2) - E[X(t_2)]\}$ is positive on average. In this case, $C_X(t_1, t_2)$ is positive, and we say $X(t_1)$ and $X(t_2)$ are positively correlated.

- On the other hand, if large values of $X(t_1)$ imply small values of $X(t_2)$, then $\{X(t_1) - E[X(t_1)]\}\,\{X(t_2) - E[X(t_2)]\}$ is negative on average, and we say $X(t_1)$ and $X(t_2)$ are negatively correlated.

- If $C_X(t_1, t_2) = 0$, then $X(t_1)$ and $X(t_2)$ are uncorrelated.

Example:

Find the mean, correlation and covariance function for random process $X_n = 1000(1+R)^n$,

for $n = 0,1,2,\cdots$. $R \sim Uniform(0.04, 0.05)$.

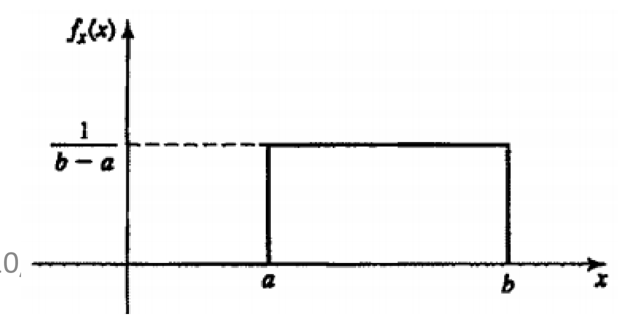
Solution:

For any value of $R \in (0.04, 0.05)$ a sample function of random process X_n can be obtained.

$X_n = 1000(1+R)^n = 1000Y^n$, thus

$Y \sim Uniform(1.04, 1.05)$ and

$$f_Y(y) = \begin{cases} 100, & 1.04 \le y \le 1.05 \\ 0, & elsewhere \end{cases}$$



The probability density function of the continuous uniform distribution is:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \le x \le b, \\ 0 & \text{for } x < a \text{ or } x > b \end{cases}$$

Continued…

We know that:

$$f_Y(y) = \begin{cases} 100, & 1.04 \le y \le 1.05 \\ 0, & elsewhere \end{cases}$$

Mean:

$$\mu_X(n) = E[X_n] = 1000E[Y^n]$$

$$= 1000 \int_{1.04}^{1.05} 100\, y^n\, dy$$

$$= \frac{10^5}{n+1}\left[1.05^{n+1} - 1.04^{n+1}\right]$$

Correlation:

$$R_X(m,n) = E[X_m, X_n]$$
$$= 10^6 E[Y^m\, Y^n]$$
$$= 10^6 \int_{1.04}^{1.05} 100\, y^{(m+n)}\, dy$$
$$= \frac{10^8}{m+n+1}\left[1.05^{m+n+1} - 1.04^{m+n+1}\right]$$

We know that:

$$\mu_X(n) = E[X_n] = 1000E[Y^n]$$

$$f_Y(y) = \begin{cases} 100, & 1.04 \le y \le 1.05 \\ 0, & elsewhere \end{cases}$$

Covariance:

$$C_X(m,n) = R_X(m,n) - \mu_X(m)\mu_X(n)$$

$$= \frac{10^8}{m+n+1}\left[1.05^{m+n+1} - 1.04^{m+n+1}\right]$$
$$- \left\{\frac{10^5}{m+1}\left[1.05^{m+1} - 1.04^{m+1}\right]\right\}\left\{\frac{10^5}{n+1}\left[1.05^{n+1} - 1.04^{n+1}\right]\right\}$$

Multiple Random Processes

For two random processes $\{X(t), t \in J\}$ and $\{Y(t), t \in J\}$

The **cross-correlation function**, $R_{XY}(t_1, t_2)$ is defined as

$$R_{XY}(t_1, t_2) = E[X(t_1)\, Y(t_2)] \qquad \text{for} \qquad t_1, t_2 \in J$$

The **cross-covariance function**, $C_X(t_1, t_2)$ is defined as

$$C_{XY}(t_1, t_2) = Cov[X(t_1)\, Y(t_2)]$$

$$= R_{XY}(t_1, t_2) - \mu_X(t_1)\, \mu_Y(t_2) \qquad \text{for} \qquad t_1, t_2 \in J$$

Reference

- Athanasios Papoulis, “Probability, Random Variables, and Stochastic Processes,” 3rd edition, McGraw Hill Publication.
- https://www.probabilitycourse.com