Random Processes (aka Stochastic Processes)

A **Random Process** (RP) is a model for an experiment or phenomena whose outcome consists of infinite number of outcomes. Three possibilities.

1) An infinite sequence of real-valued outcomes x(1), x(2), ...

examples: seq's of temp's in one location at different times, or different locations stock market values samples from speech waveform pixels from a video



such experiments are called time-discrete, discrete-in-time, discrete-time

or 2) A waveform of real-valued outcomes $x(t), -\infty < t < \infty$



examples:

voltage waveform from microphone,

number of customers in a queue as a function of time

such experiments are called time-continuous, cont-in-time, cont-time

or more generally 3) an infinite collection of outcomes indexed by some index set.

that is, there is an "index set" T and outcome x(t) for each $t \in T$.

that is, the "outcome" of the experiment is x is a function -- $x: T \to (-\infty, \infty)$

- The discrete-time random processes we consider will have $T = \{1, 2, 3, ...\}$ or $\{..., -2, -1, 0, 1, 2, ...\}$
- The continuous-time random processes we consider will have T = (a,b), an interval or [a,b], (a,b], [a,b)
- We won't consider any other index sets, but there are other interesting ones, such as $T = R^2$, such as when x(s,t) represents an image.

Terminology

The sequence/waveform/function x is called a "sample function" or "realization"

(a random process is a model for such)

Notation:

- A sample function x is also denoted $\{x(t): t \in T\}$ or $\{x(t)\}$ or x(t), though the latter is ambiguous. The notations with curly brackets emphasize that we're talking about the whole function rather than just the value at on particular t.
- In the discrete-time case

x is also denoted $\{x(n): n \in T\}$ or $\{x(n)\}$ or $\{x[n]: n \in T\}$ or $\{x[n]\}$ or x(n), though the latter is ambiguous

or any of the above with x(n) replaced by x[n] or x_n

Notes:

- Usually we model experiments/phenomena that we believe are "random", but "random" is a subjective thing -- what's random to one might not be to another.
- Though random processes with nonnumerical outcomes are possible, we focus exclusively on real-valued ones. Example of random process with nonnumerical values: sequence of letters of English text.
- Some people use the word "parameter" rather than "index", as in: T is the parameter set; the outcomes are parameterized by t; a discrete parameter experiment
- Discrete-time random processes are discussed in Chapter 7 of S&W. **Read** Section 7.1.
- Continuous-time random processes are discussed in Chapters 8, 9 and 10. **Read** Section 8.1, 8.2 and 8.4.

A Random Process is each of the following three things:

(each is a model of, or definition of, a random process) :

1. Random Sample Function

A random experiment in which the "outcome" is an entire sample function.

Probability Model (Ω_X, E_X, P_X, X) where

 $\Omega_X = \{ \text{functions } x : T \to (-\infty, \infty) \}$ (called sample functions)

 $E_{\rm X}$ = event space = collection of subsets of $\Omega_{\rm X}$

$$= \sigma(B)$$

where B = all sets of form {functions x: $x(t) \in A$ }, $A \in$ Borel σ -algebra of $(-\infty, \infty)$

(E_X must be a σ -algebra and must not be "too large", i.e. it must be like the Borel σ -algebra.)

Examples of sets of interest in E_X

$$\begin{array}{l} \{x: \ x(3) = 4\}, \ \{x: \ 2 < x(3) \le \pi\}, \ \{x: \ x(3) \le 4, \ x(5) > 6\}_T \\ \{x: \ x(t) \le 4, \ 0 \le t \le 1\}, \ \{x: \ \int_0^0 x^2(t) \ dt \le 2\}, \ \{x: \lim_{T \to \infty} \frac{1}{T} \int_0^T x(t) \ dt = 0\} \end{array}$$

 P_X is probability measure on E_X .

It characterizes the probability distribution of the random process.

There is no concept of "density function" that applies to random samples functions. (There was for random variables and random vectors.)

2. Infinite Collection of Random Variables

 $\{\mathbf{X}(\mathbf{t}): \mathbf{t} \in T\}$

for each $t \in T$, X(t) is a random variable

probability distribution of the random process

there is no such thing as the joint PDF, pdf or pmf of an infinite number of random variables

so ... probability distribution of the random process is considered to be characterized by specifying:

the joint distribution of each finite subcollection of the random variables;

i.e. by specifying the probability distribution of $X(t_1),...,X(t_N)$ for all N and all $t_1,...,t_N$.

e.g $F_{X(t_1),...,X(t_N)}(x_1,...,x_N)$ for all N, all $t_1,...,t_N$ and all $x_1,...,x_N$

Notes:

Why is this enough? Well in most cases this allows us to calculate what we need. For example, $Pr(X(3.5) \le 6 \text{ and } X(4) \ge 7)$.

Also, in most nonpathological cases (referred to as separable) $Pr(X(t) \le a, b \le t \le c) = \lim_{N \to \infty} Pr(X(i\frac{c-b}{N}) \le a, i=1,...,N)$

Although this is in some sense a "reduction" in that we only have to specify the joint distribution of finite collections of random variables rather than infinite collections of random variables,

it is still an awfully lot to specify,

and sometimes it is way too much.

3. Infinite collection, indexed by $t \in T$, of functions of some underlying random experiment

(this combines 1 and 2)

Let (Ω_U, E_U, P_U, U) is an underlying probability model.

Then, a random process is a function

 $X(t,\omega), t \in T, \omega \in \Omega_U, i.e. T \times \Omega_U \to (-\infty,\infty)$

For any fixed $t \in T$, $X(t,\omega)$ is a function of ω that is a model for the random variable X(t); i.e for the outcome at time t.

The randomness in the underlying experiment generates the randomness in X(t).

For any fixed $\omega \in \Omega_U$, $X(t,\omega)$ is a function of t, i.e. it is a sample-function

The probability distribution is determined by P_{U} .

Notes:

- a. From any of these three models for a random process one can derive a model of each other type.
- b. There are certain pathological cases where the joint distributions of finite numbers of random variables does not tell the whole story, but we ignore such pathological cases. The book (S&W) gives an example on p. 374.
- c. Model 1 is useful when there is a finite or countably infinite set of sample functions or for example or when there is a simple parameterization of the sample functions, e.g. $X(t) = A \cos (2\pi t + B)$ where A, B are random variables with known joint density. Otherwise events i.e. sets of functions are hard to work with.

Model 2 is the most generally useful when making calculations and doing theory.

Model 3 is, essentially, Models 1 and 2 combined. It is most useful when doing theory

d. Discrete vs. Continuous Valued

In addition to random processes being discrete- or continuous-time, hey are also discrete or continuous valued.

So it is useful to think of each of four different kinds of random variables:

Examples of each:

	Discrete-valued	Continuous-valued
Discrete-time	seq of stock prices	seq of temp's in time or distance
	seq of letters, e.g. from a book	samples of speech, pixels of iamge
Continuous-time	number of customers in line at time	waveform from microphone

Random process theory has a somewhat different flavor in each case. We can't study all. We concentrate more on continuous-valued random processes.

Partial Characterizations of the Distribution of a Random Process

(because a complete characterization is more than we will know in many cases)

- 1st order distribution distribution of X(t) for all t ∈ T e.g. F_{X(t)}(x) for all t ∈ T and all x says nothing about dependence among variables.
- 2. 2nd order distribution joint distribution of X(t),X(s) for all $t,s \in T$
- 3. nth-order distribution this gets to be overwhelming quickly as n increases
- 4. first moment

mean function: $\mu_X(t) = E[X(t)] \quad t \in T$ indicates center of the distribution at each time

5. second moment

autocorrelation function: $R_X(t,s) = E[X(t)X(s)], t,s \in T$ equivalently: covariance function $K_X(t,s) = cov(X(t),X(s)) = R_X(t,s)-\mu_X(t)\mu_X(s)$ this includes the second moments of the variables $R_X(t,t) = E[X^2(t)]$ as well as correlation between different variables.

6. first and second moment this is very useful and very common.

Examples of Random Processes

Discrete-time

- 1. Bernoulli
- 2. IID
- 3. Moving Average (MA)
- 4. Autoregressive random processe
- 5. Gaussian random process

Continuous-time Examples

- 1. R.P. with a finite number of sample functions
- 2. Sinusoidal: $X(t) = A \cos(2\pi ft + \theta)$
- 3. Continuous-time process obtained from discrete-time process produced by digital-to-analog conversion

4. Continuous-time process obtained from discrete-time process by "pulse" modulation.

- 5. Random telegraph
- 6. Poisson counting process
- 7. Gaussian random process (viewpoint 2)

Stationarity (see Sections 7.1 and 8.4 of S&W)

Definition: A random process $\{X(t): t \in T\}$ is (strict-sense) **stationary** if for every n and $t_1, \ldots, t_n \in T$, and τ such that $t_1+\tau, \ldots, t_n+\tau \in T$, $X(t_1), \ldots, X(t_n)$ and $X(t_1+\tau), \ldots, X(t_n+\tau)$ are identical random vectors.

The basic idea is that for a stationary r.p. the probability distributions of random variables (and vectors) do not change with time shifts. The probability of something happening at time is the same as the probability of it happening at any other time.

The following are some of the consequences of stationarity:

$$\begin{split} f_{X(t)}(x) &= f_{X(s)}(x) \text{ all } t, s, x \\ f_{X(t)X(t+\tau)}(x_1, x_2) &= f_{X(t)X(t+\tau)}(x_1, x_2) \text{ all } t, \tau, x_1, x_2 \\ \mu_X(t) \text{ is the same for all } t \\ R_X(t, t+\tau) \text{ does not depend on } t. \\ E \ g(X(t_1), X(t_2), \dots, X(t_n)) &= E \ g(X(t_1+\tau), X(t_2+\tau), \dots, X(t_n+\tau)) \end{split}$$

Definition: A random process $\{X(t): t \in T\}$ is wide-sense stationary (WSS) if $\mu_X(t)$ and $R_X(t,t+\tau)$ do not depend on t.

Fact: Stationarity \Rightarrow wide-sense stationary. The converse is false.

So wide-sense stationarity is a weak kind of stationarity that is easier to check and work with, since it only depends on the mean and autocorrelation functions.

Properties of the Autocorrelation Function of a Wide-Sense Stationary Random Processes

- 1. Symmetry: $R_X(-\tau) = R_X(\tau)$
- 2. $R_X(0) \ge |R_X(\tau)|$ for all τ
- 3. Positive Definite

For any n and any set of n complex numbers $a_1,...,a_n$ and any $t_1,...,t_n \in T$,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} R_X(t_i, t_j) a_i^* a_j \text{ is real and nonnegative}$$

Actually, it can be shown the $3 \Rightarrow 1$ and 2.

Property 3 implies that the Fourier transform of $R_X(\tau)$ is real, nonnegative and symmetric.

Ergodicity (see Section 9.4 of S&W)

Definition: (not the standard mathematical definition, but the one that we shall adopt)

A discrete-time stationary random process $\{X(n):n = 1,2,...\}$ is (strict-sense) ergodic if

$$\frac{1}{n}\sum_{i=1}^n g(X(i+1),\ldots,X(i+m) \ \rightarrow E \ g(X(1),\ldots,X(m)) \ \text{ almost surely as } n \rightarrow \infty$$

for any m and any function $g(x_1,...,x_m)$ such that $E g(X_1,...,X_m)$ is well-defined.

A continuous-time stationary random process $\{X(t): t \in [0,\infty)\}$ is (strict-sense) ergodic if

$$\frac{1}{T}\int_{0}^{T}g(X(t+\tau_{i}),...X(t+\tau_{m})) \rightarrow E \ g(X(\tau_{1}),...,X(\tau_{m})) \quad \text{almost surely as} \ T \rightarrow \infty$$

for any m, $\tau_1,...,\tau_m$ and any function $g(x_1,...,x_m)$ such that $Eg(X(\tau_1),...,X(\tau_m))$ is well-defined.

For "two-sided" discrete- and continuous-time random processes, the above averages are replaced by

$$\frac{1}{2n+1}\sum_{i=-n}^{n}$$
 and $\frac{1}{2T}\int_{-T}^{1}$, respectively.

The basic idea is that for ergodic processes, time averages converge to expected values.

For example, the following are consequences of ergodicity

$$\begin{split} \frac{1}{n} & \sum_{i=1}^{n} X(i) \to EX, \qquad \frac{1}{T} \int_{0}^{T} X(t) \ dt \to EX \\ \frac{1}{n} & \sum_{i=1}^{n} X^{2}(i) \to EX^{2}, \qquad \frac{1}{T} \int_{0}^{T} X^{2}(t) \ dt \to EX^{2} \\ \frac{1}{n} & \sum_{i=1}^{n} X(i)X(i+1) \to R_{X}(1), \qquad \frac{1}{n} \sum_{i=1}^{n} X(i)X(i+m) \to R_{X}(m), \\ \frac{1}{T} \int_{0}^{T} X(t)X(t+\tau) \ dt \to R_{X}(\tau) \\ \frac{n_{A}}{N} \to P(A) \ \text{where } A \ \text{is any event and } n_{A} \ \text{is the number of times } A \ \text{occurs in} \\ X(1), \dots, X(n) \end{split}$$

For stationary processes that are not ergodic, time averages such as those above converge, but not to the expected value. Instead, all that we can say is

$$E\left(\frac{1}{n} \sum_{i=1}^{n} g(X(i+1),...,X(i+m))\right) \to E g(X(1),...,X(m)) \text{ as } n \to \infty$$
$$E\left(\frac{1}{T} \int_{0}^{T} g(X(t+\tau_{i}),...,X(t+\tau_{m})) \to E g(X(\tau_{1}),...,X(\tau_{m}))\right) \text{ as } T \to \infty$$