Lecture 24

(11 November 2024)

Random Process

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

Discrete-time random process:

$$(X_t \,;\, t \in \mathbb{N})$$

Continuous-time random process:

$$(X_t \,:\, t \in \mathbb{R})$$

For each t, X_t is a random variable.

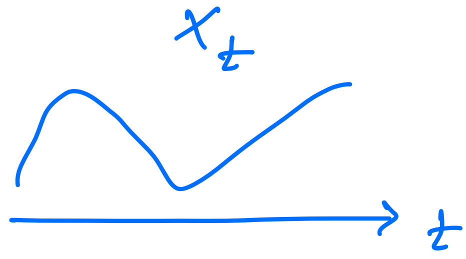
$$\omega \longmapsto X_1(\omega), X_2(\omega), \ldots$$

Discrete-time

$$\omega \longmapsto X_t(\omega) \quad t \in \mathbb{R}$$

Continuous-time

For a fixed $\omega \in \Omega$, $(X_t(\omega)\; t \in T)$ is called the sample path at ω.



E.g. Bernoulli process

E.g. Stock value

Example. Let $X_t = A + Bt$, for all $t \in [0, \infty)$, where A and B are independent Gaussian $N(1,1)$ random variables.

(i) Find the PDF of $Y = X_1$.

(ii) If $Z = X_2$, find $E[YZ]$.

$X_1 = A + B \implies X_1$ is $N(2, 2)$.

$$f_{X_1}(x) = \frac{1}{\sqrt{2\pi\, 2}} e^{-(x-2)^2/4}.$$

$$E[YZ] = E[X_1 X_2]$$

$$= E[(A+B)(A+2B)]$$

$$= E[A^2 + 2B^2 + 3AB]$$

$$= 2 + 4 + 3 = 9.$$

First-order Distribution: $F_{X_t}(x) = P(X_t \le x)$.

Second-order Distribution:

$$F_{X_{t_1}, X_{t_2}}(x_1, x_2) = P(X_{t_1} \le x_1, X_{t_2} \le x_2).$$

n^{th}-order Distribution:

$$F_{X_{t_1}, X_{t_2}, \dots, X_{t_n}}(x_1, x_2, \dots, x_n) = P(X_{t_1} \le x_1, \dots, X_{t_n} \le x_n).$$

<u>Mean Function</u> of a <u>Random Process</u>

For a random process $(X_t, t \in T)$, the mean function is defined as

$$\mu_X(t) = E[X_t].$$

<u>Example</u>. $X_t = A + Bt$, $A, B \sim N(1, 1)$ and A, B are independent.

$$\mu_X(t) = E[X_t] = E[A] + t\,E[B] = t + 1,$$

$$t \in [0, \infty).$$

The mean function $\mu_X(t)$ gives us the expected value of X_t at time t but it does not give us any information about how X_{t_1} and X_{t_2} are related. To get some insight on the relation between X_{t_1} and X_{t_2}, we define correlation and covariance functions.

Correlation function

$$R_X(t_1, t_2) = E[X_{t_1} X_{t_2}]$$

Covariance function

$$C_X(t_1, t_2) = \operatorname{Cov}(X_{t_1}, X_{t_2})$$

$$= E[X_{t_1} X_{t_2}] - E[X_{t_1}]\, E[X_{t_2}]$$

$$= R_X(t_1, t_2) - \mu_X(t_1)\mu_X(t_2).$$

Exercise. $X_t = A + Bt$, A, B are independent and $N(1, 1)$. Show that

(i) $R_X(t_1, t_2) = 2 + t_1 + t_2 + t_1 t_2$

(ii) $C_X(t_1, t_2) = 1 + t_1 t_2$.

Bernoulli Process

$X_1, X_2, \ldots$ are i.i.d. Bernoulli RVs with

$$P(X_i = 1) = p = 1 - P(X_i = 0) \text{ for } i \in \mathbb{N}.$$

This can be visualized as a sequence of independent coin tosses, where the probability of heads in each toss is a fixed number p in the range $0 < p < 1$.

Of course, coin tossing is just a paradigm for a broad range of contexts involving a sequence of independent binary outcomes. For example, a Bernoulli process is often used to model systems involving arrivals of customers or jobs at service centers. Here, time is discretized into periods, and a "success" at the k^{th} trial is associated with the arrival of at least one customer at the service center during the k^{th} period.

In a Bernoulli process we already know the following.

- The number S of successes in n independent trials is Binomial(n, p).

$$p_S(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, ---, n$$

$$E[S] = np,$$

$$Var(S) = np(1-p).$$

- The number T of trials upto the first success is Geometric(p).

$$p_T(t) = (1-p)^{t-1} p, \quad t = 1, 2, ---,$$

$$E[T] = 1/p, \quad Var(T) = \frac{1-p}{p^2}.$$

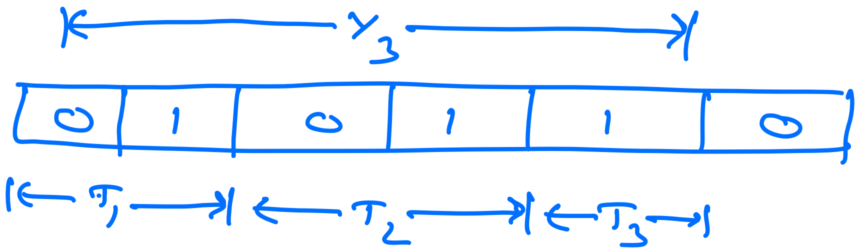
Fresh start Property:

0	1	0	0	1	1	0	- - - -
X_1	X_2	X_3	X_4	X_5	--	X_n	X_{n+1}

For any n, $Y_i = X_{n+i}$ is also a Bernoulli process which is independent of the past RVs.

Arrival time: Let Y_k denote the time of the k^{th} success (or arrival), i.e., k^{th} arrival time.

Interarrival time: Let T_k represent the number of trials following the $(k-1)$ successes until the next success.



$T_1, T_2, \ldots$ are independent and have the same geometric distribution.

$$Y_k = T_1 + T_2 + \ldots + T_k$$

$$E[Y_k] = k/p.$$

$$Var(Y_k) = \frac{k(1-p)}{p^2}$$

We find the PMF of Y_k.

$$Y_k \in \{k, k+1, k+2, \ldots\}.$$

$$p_{Y_k}(t) = P(Y_k = t)$$

$= P($ k^{th} arrival occurs at time $t)$

$= P($ $(k-1)$ arrivals occur in first $t-1$ trials and arrival occurs in t^{th} trial $)$

$= P($ $(k-1)$ arrivals in first $t-1$ trials $)$ $\times$ $P($ arrival occurs in t^{th} trial $)$

(by independence property of Bernoulli process)

$$= \binom{t-1}{k-1} p^{k-1} (1-p)^{t-k} \cdot p$$

$$= \binom{t-1}{k-1} p^{k} (1-p)^{t-k}.$$

$$p_{Y_k}(t) = \binom{t-1}{k-1} p^{k} (1-p)^{t-k} \quad t = k, k+1, \ldots$$

This is known as Pascal pmf of order k.

Poisson Process

Poisson process is a counting process, a random process that counts the number of arrivals from time 0 upto and including time t.

Recall the Poisson random variable with λ.

$$P_X(k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!}, \quad k = 0, 1, 2, 3, \ldots$$

$(N_t, t \in [0, \infty))$ is a Poisson process with rate λ if

(1) $N(0) = 0$;

(2) N_t has independent increments, i.e., for all $0 \le t_1 < t_2 < t_3 < \ldots < t_n$, the RVs N_{t_1}, $N_{t_2} - N_{t_1}$, $N_{t_3} - N_{t_2}$, ---- are mutually independent.

Note that $N_{t_i} - N_{t_{i-1}}$ represents the no. of arrivals in the interval $[t_{i-1}, t_i]$.

(3) The no. of arrivals in any interval of length $T>0$ has Poisson(λT) distribution, i.e., $N_{t+T} - N_t \sim$ Poisson(λT) for all $t \in [0,\infty)$.

$$E[N(t)] = \lambda t, \quad Var(N(t)) = \lambda t.$$

$$R_N(t_1, t_2) = E[N_{t_1} N_{t_2}] \qquad (\text{Let } t_1 < t_2)$$

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

$$= E[N_{t_1}(N_{t_2} - N_{t_1})] + E[N_{t_1}^2]$$

$$= E[N_{t_1}]\, E[N_{t_2} - N_{t_1}] + E[N_{t_1}^2]$$

(because of independent increments)

$$= \lambda t_1 \; \lambda(t_2 - t_1) + \lambda t_1 + \lambda^2 t_1^2$$

$$= -\cancel{\lambda^2 t_1^2} + \lambda^2 t_1 t_2 + \lambda t_1 + \cancel{\lambda^2 t_1^2}$$

$$= \lambda t_1 + \lambda^2 t_1 t_2.$$

$$C_N(t_1, t_2) = \lambda t_1 + \cancel{\lambda^2 t_1 t_2} - \cancel{\lambda^2 t_1 t_2} = \lambda t_1.$$

(assuming $t_1 < t_2$)

Interarrival and Arrival Times

Let X_1 denote the time of the first arrival.

$$P(X_1 > t) = P(\text{no arrival in } [0\ t])$$
$$= P(N(t) = 0)$$
$$= e^{-\lambda t}.$$

$$\Rightarrow F_{X_1}(t) = \begin{cases} 1-e^{-\lambda t}, & t>0 \\ 0, & \text{otherwise.} \end{cases}$$

$\therefore X_1 \sim \text{Exponential}(\lambda)$.

Let X_k be the time lapsed between $(k-1)^{th}$ arrival and k^{th} arrival.

S_k denote the time of k^{th} arrival.

We compute the joint CDF of $S_1 S_2$.

For $t_1 < t_2$,

$$F_{S_1,S_2}(t_1, t_2) = P(S_1 \le t_1, S_2 \le t_2)$$

$$= P(N_{t_1} \ge 1, N_{t_2} \ge 2)$$

$$= P(N_{t_1} = 1, N_{t_2} \ge 2) + P(N_{t_1} \ge 2, N_{t_2} \ge 2)$$

$$= P(N_{t_1}=1,\ N_{t_2}-N_{t_1}\geq 1) + P(N_{t_1}\geq 2)$$

$$(\because N_{t_2}\geq N_{t_1})$$

$$= \lambda t_1 e^{-\lambda t_1}\left(1-e^{-\lambda(t_2-t_1)}\right)$$

$$+\left(1-e^{-\lambda t_1}-\lambda t_1 e^{-\lambda t_1}\right)$$

$$= 1-e^{-\lambda t_1}-\lambda t_1 e^{-\lambda t_2}.$$

$$f_{S_1,S_2}(t_1,t_2)=\begin{cases}\lambda^2 e^{-\lambda t_2}, & 0<t_1<t_2,\\ 0, & \text{otherwise.}\end{cases}$$

$$X_1 = S_1 = g_1(S_1,S_2), \qquad X_2 = S_2 - S_1 = g_2(S_1,S_2)$$

$$S_1 = h_1(X_1,X_2) = X_1, \qquad S_2 = h_2(X_1,X_2) = X_1+X_2$$

$$f_{X_1,X_2}(x_1,x_2) = \frac{f_{S_1,S_2}(x_1,x_1+x_2)}{\left|J(x_1,x_1+x_2)\right|},$$

where

$$J(x_1, x_1+x_2) = \left| \begin{matrix} \frac{\partial}{\partial s_1} s_1 & \frac{\partial}{\partial s_2} s_1 \\ \frac{\partial}{\partial s_1}(s_2 - s_1) & \frac{\partial}{\partial s_2}(s_2 - s_1) \end{matrix} \right|_{s_1 = x_1,\ s_2 = x_1 + x_2}$$

$$= \left| \begin{matrix} 1 & 0 \\ -1 & 1 \end{matrix} \right| = 1.$$

so $f_{X_1 X_2}(x_1, x_2) = f_{S_1 S_2}(x_1, x_1+x_2)$

$$= \begin{cases} \lambda^2 e^{-\lambda(x_1+x_2)}, & x_1, x_2 \in \mathbb{R}_+. \\ 0, & \text{otherwise.} \end{cases}$$

$\Rightarrow$ X_1 and X_2 are independent and exponentially distributed.

(since $f_{XY}(x,y) = g(x)h(y) \Rightarrow$ X & Y are independent)

Similarly, $X_1, X_2, X_3 \ldots$ are i.i.d. and exponentially distributed.