SUYASH GOYAL
2K22/MC/161

Assignment - 1
Stochastic Processes (MC303)

(A1) A Stochastic Process is defined as a collection of random variables $X = \{X_t : t \in T\}$ defined on a common probability space, taking values in a common set S (the state space) and indexed by a set T, often either N or $[0, \infty)$ and thought of as time (discrete or continuous respectively). Stochastic Processes can be divided into 4 types based on nature of state space and index set (S and T):-

(i) Both T and S are discrete.

Eg. X_n represents outcome of n^{th} toss of fair dice.

$S = \{1, 2, 3, 4, 5, 6\}$ and $T = \{1, 2, 3, \ldots\}$.

(ii) T is discrete and S is continuous.

Eg. X_n represents temperature at the end of n^{th} hour.

$S = (0°C, 60°C)$ and $T = \{1, 2, 3, \ldots\}$.

(iii) T is continuous and S is discrete.

Eg. X_n represents no. of phone calls received in interval $(0, t)$.

$S = \{1, 2, 3, \ldots\}$ and $T = (0, t)$.

(iv) Both T and S are continuous.

Eg. X_n represents maximum temperature in interval $(0, t)$.

$S = (10°C, 60°C)$ and $T = (0, t)$.

(A2) A Bernoulli Process is a finite or infinite sequence of binary random variables, so it is a discrete-time stochastic process that takes only two values, canonically 0 and 1. The component Bernoulli Variables X_i are identically distributed and independent.

(i) 7 success in 10 trials (Assuming $P(\text{success}) = p$ & $1 - p = q$)

$P(S = K) = {}^nC_K \, p^K (1-p)^{n-K}$

$= {}^{10}C_7 \, p^7 (1-p)^{10-7} = 120 p^7 (1-p)^3$

(ii) 7^{th} success in 10 trials (Assuming $P(\text{success}) = p$ & $1-p = q$)

$P(T_1 = K) = (1-p)^{K-1} p = p(1-p)^6$

(A3) A merged process of two independent Bernoulli Processes is still a Bernoulli Process with a computed parameter. The independence property is crucial for the merged process, ensuring that the R.V. in different time slots are independent. Collisions between arrivals in the original processes are counted as one arrival in the merged process.

	0	1
1	$(1-p)q$	pq
0	$(1-p)(1-q)$	$p(1-q)$

We are given that there is an arrival in either X_t or Y_t, hence,

$\therefore$ Probability of arrival in merged process $= (1-p)q + p(1-q) = p+q-2pq$

(A4) A counting process $(N(t))_{t\geq 0}$ is said to be a Poisson process with rate (or intensity) λ, $\lambda > 0$ if:

(i) $N(0) = 0$.

(ii) The process has independent increments.

(iii) The number of events in any time interval of length t is Poisson distributed with mean λt. That is, $N((s,t]) \overset{d}{=} Poi(\lambda t)$ $\forall\, s,t \geq 0$:

$$P(N((s,t]) = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n \in \mathbb{N}_0.$$

$$f_{S_n}(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}$$

$$f_{X_1 S_2}(x_1, s_2) = f_{X_1}(x_1)\, f_{X_2}(s_2 - x_1)$$

$$f_{X_1 S_2}(x_1, s_2) = \lambda^2 e^{-\lambda x_1} e^{-\lambda(s_2 - x_1)} = \lambda^2 e^{-\lambda s_2}$$

$$f_{X_1 \dots X_{n-1} S_n}(x_1, \dots, x_n) = \lambda^n \exp(-\lambda x_1 - \lambda x_2 - \dots - \lambda x_n)$$

Letting $S_n = X_1 + \dots + X_n$ and substituting $S_n - X_1 - \dots - X_{n-1}$ for X_n,

$$f_{X_1 \dots X_{n-1} S_n}(x_1, \dots, x_{n-1}, S_n) = \lambda^n e^{-\lambda S_n}$$ Hence Shown

$$P(N((0,2]) = 4) = \frac{e^{-2\lambda}(2\lambda)^4}{4!} = \frac{2}{3}\lambda^4 e^{-2\lambda}$$

$$P(N((4,7]) = 3) = \frac{e^{-7\lambda}(7\lambda)^3}{3!} = \frac{343}{6}\lambda^3 e^{-7\lambda}$$

(A5) A renewal process is a point process in which the interevent intervals are independent and drawn from the same probability density. More specifically, let T_i be independent, identically distributed interevent times from the probability density $p(t)$. Let $S_r = \sum_{i=1}^{r} T_i$.

The process S_r is a renewal process.

For renewal process $N(t)$, we define the renewal function $H(t)$ by,

$$H(t) = E[N(t)] \quad t \geq 0$$

and the renewal density $h(t)$ by,

$$h(t) = H'(t) \quad t \geq 0.$$

$H(t)$ is a finite non-decreasing function of t with $H(0) = 0$.

When the lifetime distribution is an exponential distribution with parameter λ, so that $f(t) = \lambda e^{-\lambda t}$, $t \geq 0$, then $\{N(t) : t \geq 0\}$ is a Poisson process with rate λ. In this case, S_n has a gamma distribution with parameters n and λ, and $N(t)$ has a Poisson distribution with parameter λt. The exponential distributions are the only ones with the "memoryless" property.

(A6) Let the input space be X and $f : X \rightarrow R$ be a function from the input space to the reals. We then say that f is a Gaussian process if for any vector of inputs $x = [x_1, x_2, \dots, x_n]^T$ s.t. $x_i \in X \; \forall i$, the vector of outputs $f(x) = [f(x_1), f(x_2), \dots, f(x_n)]^T$ is Gaussian distributed.

To prove wide and strict sense stationary are equivalent :-

$X(t)$ is a WSS process so,

$$\mu_x(t_i) = \mu_x(t_j) = \mu_x \; \forall i, j \quad \text{and} \quad C_x(t_i+\Delta, t_j+\Delta) = C_x(t_i, t_j) = C_x(t_i - t_j) \; \forall i, j$$

From above we conclude that the mean vector and covariance matrix of

$$X(t_1), X(t_2), \dots, X(t_r)$$

is the same as the mean vector and covariance matrix of

$$X(t_1+\Delta), X(t_2+\Delta), \dots, X(t_r+\Delta).$$

Hence Proved.

(A7) A stochastic process $B = \{B(t) : t \geq 0\}$ possessing continuous sample paths is called standard Brownian motion if :-

(i) $B(0) = 0$.

(ii) B has both stationary and independent increments.

(iii) $B(t) - B(s)$ has a normal distribution with mean 0 and variance $t-s$, $0 \leq s < t$.

For Brownian motion with variance σ^2 and drift μ, $X(t) = \sigma B(t) + \mu t$, the definition is same except (iii) must be modified :-

(iii) $X(t) - X(s)$ has a normal distribution with mean $\mu(t-s)$ and variance $\sigma^2(t-s)$.

We observe that the vector $(B(t_1), \ldots, B(t_n))$ has a multivariate normal distribution because the event $\{B(t_1) = x_1, \ldots, B(t_n) = x_n\}$ can be re-written in terms of independent increment events, $\{B(t_1) = x_1, B(t_2) - B(t_1) = x_2 - x_1, \ldots, B(t_n) - B(t_{n-1}) = x_n - x_{n-1}\}$, yielding the joint density of $(B(t_1), \ldots, B(t_n))$ as,

$$f(x_1, \ldots, x_n) = f_{t_1}(x_1) f_{t_2 - t_1}(x_2 - x_1) \cdots f_{t_n - t_{n-1}}(x_n - x_{n-1})$$

where $f_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}$ is the density for the $N(0,t)$ distribution.

For a Gaussian Process, $m(t) \stackrel{def}{=} E(X(t))$, $a(s,t) \stackrel{def}{=} cov(X(s), X(t))$, $0 \leq s \leq t$.

For a standard BM, $m(t) = 0$ and $a(s,t) = var(B(s)) + 0 = s$.

Thus standard BM is the unique Gaussian Process with $m(t) = 0$ and $a(s,t) = \min\{s, t\}$.

(A8) A random walk is a stochastic process that describes a path that consists of a succession of random steps on some mathematical space which starts at 0 and at each step moves +1 or -1 with equal probability. Real work processes which can be realized as random walk are path of a molecule, search path of a foraging animal, price of a fluctuating stock and the financial status of a gambler. Types of random walk include Unrestricted, Restricted with Absorbing barriers, Restricted with Reflecting barriers, 2-D, 3-D, etc. Problem of interest is to determine P(ruin) for given x_0.

(A9) P(absorption occurs at a) = $\begin{cases} p^a \dfrac{p^b - q^b}{p^{a+b} - q^{a+b}} & (p \neq q) \\ \dfrac{b}{a+b} & (p = q) \end{cases}$

- P(absorption occurs at b) = 1 − P(absorption occurs at a)

Here $a = 5$, $b = 4$ ([redacted] Because formula assumes $-b$ [redacted]), $p = 0.3$ and $q = 0.4$.

$\Rightarrow$ P(absorption at a) = $(0.3)^{-5} \dfrac{(0.3)^4 - (0.4)^4}{(0.3)^9 - (0.4)^9} = 0.1754$

$\Rightarrow$ P(absorption at b) = 1 − P(absorption at a) = 0.8246

~~Also, P(absorption occurs) = 1 for any N. So for N = 10, P(absorption occurs) = 1.~~

- $E(N) = \dfrac{ap^a(p^b - q^b) - bq^b(p^a - q^a)}{(p-q)(p^{a+b} - q^{a+b})}$ $(p \neq q)$

$$= \frac{5(0.3)^5(0.3^4 - 0.4^4) - 4(0.4)^4(0.3^5 - 0.4^5)}{(0.3 - 0.4)(0.3^9 - 0.4^9)}$$

$$= 24.21$$

This mean absorption is likely to occur by the time the particle takes 25 steps.

(A10) To find the MGF of X_n, $E(e^{\alpha X_n})$, we'll first calculate the MGF of a single step z_i.

$M_{z_i}(\alpha) = E(e^{\alpha z_i}) = e^{\alpha}p + e^{-\alpha}q$

Since X_n is sum of n IID random variables z_i, the MGF of X_n is: $E($[redacted]$e^{\alpha X_n}) = (E(e^{\alpha z_i}))^n = (e^{\alpha}p + e^{-\alpha}q)^n$

The [redacted] asymptotic behaviour of X_n largely depends on the value of p and q :-

(i) If $p = q$ then X_n behaves like a normal distribution by the CLT for large n with mean 0 and var n. Hence $X_n \approx \sqrt{n}\, Z$.

(ii) If $p \neq q$ then X_n will drift with mean $n(2p-1)$ and its variance remains n, so it tends to a normal distribution with non-zero mean and variance n.

(A11) For $m < n$:-

$X_m = \sum_{i=1}^{m} z_i$ and $X_n = \sum_{i=1}^{n} z_j = X_m + \sum_{i=m+1}^{n} z_i$

Given X_n, the distribution of the sum of the increments from $m+1$ to n is s.t. the increments are evenly distributed. In particular $E(X_m | X_n)$ should be a linear function of X_n.

To compute $E(X_m | X_n)$, note that the increments between times m and n will average out. Specifically, the walk is equally likely to have any value from m to n due to symmetry and linearity.

$\therefore E(X_m | X_n) = \frac{m}{n} X_n$.

For $m > n$:-

In this case X_n is already determined so X_m is just X_n plus the increments from $n+1$ to m which are independent of X_n and hence have no effect on $E(X_m | X_n)$.

$\therefore E(X_m | X_n) = X_n$.