Q1a)Cetral limit thereom

Let $\{X_1, X_2, \ldots, X_n\}$ be a random sample of size n — that is, a sequence of independent and identically distributed random variables with expected values μ and variances σ^2. Suppose we are interested in the behavior of the sample average of these random variables: $S_n = (X_1 + \ldots + X_n)$. Then the central limit theorem asserts that for large n's, the distribution of S_n is approximately normal with mean μ and variance σ^2. The true strength of the theorem is that S_n approaches normality regardless of the shapes of the distributions of individual X_i's. Formally, the theorem can be stated as follows:

> **Lindeberg–Lévy CLT:** suppose $\{X_i\}$ is a sequence of iid random variables with $E[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2$. Then as n approaches infinity, the random variable $\sqrt{\ }(S_n - \mu)$ converges in distribution to a normal $N(0, \sigma^2)$:
>
> $$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right) \xrightarrow{d} \mathcal{N}(0,\ \sigma^2).$$

Convergence in distribution means that the cumulative distribution function of $\sqrt{}(S_n - \mu)$ converges pointwise to the cdf of the $N(0, \sigma^2)$ distribution: for any real number z,

$$\lim_{n\to\infty} \Pr[\sqrt{n}(S_n - \mu) \le z] = \Phi(z/\sigma),$$

where $\Phi(x)$ is the standard normal cdf.

SOLUTION

b) i) Axiomatic definition of probability:

Probability is a set function $P[\cdot]$ that assigns to every event $E \in \mathcal{F}$ a number $P[E]$ called the probability of E such that

1) $P[E] \geqslant 0$
2) $P[\Omega] = 1$
3) $P[E \cup F] = P[E] + P[F]$ if $EF = \phi$

ii) $E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x y \, f_{XY}(x, y) \, dx \, dy$

$\because$ X & Y are independent $\Rightarrow f_{XY}(x, y) = f_X(x) f_Y(y)$

$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x y \, f_X(x) f_Y(y) \, dx$

$= \int_{-\infty}^{\infty} x f_X(x) \, dx \int_{-\infty}^{\infty} y f_Y(y) \, dy$

$= E[X] \, E[Y]$

$= \mu_x \mu_y$

C) a stationary process (or strict stationary process or strong(ly) stationary process) is a stochastic process whose joint probability distribution does not change when shifted in time or space. As a result, parameters such as the mean and variance, if they exist, also do not change over time or position.

Stationarity is used as a tool in time series analysis, where the raw data are often transformed to become stationary; for example, economic data are often seasonal and/or dependent on a non-stationary price level. An important type of non-stationary process that does not include a trend-like behavior is the cyclostationary process.

Note that a "stationary process" is not the same thing as a "process with a stationary distribution". Indeed there are further possibilities for confusion with the use of "stationary" in the context of stochastic processes; for example a "time-homogeneous" Markov chain (which condition is sometimes called by the name "stationary Markov chain") is sometimes said to have "stationary transition probabilities".

Weak or wide-sense stationarity

A weaker form of stationarity commonly employed in signal processing is known as **weak-sense stationarity**, **wide-sense stationarity** (WSS) or **covariance stationarity**. WSS random processes only require that 1st and 2nd moments do not vary with respect to time. Any strictly stationary process which has a mean and a covariance is also WSS.

So, a continuous-time random process $x(t)$ which is WSS has the following restrictions on its mean function

and autocorrelation function

The first property implies that the mean function $m_x(t)$ must be constant. The second property implies that the correlation function depends only on the *difference* between t_1 and t_2 and only needs to be indexed by one variable rather than two variables

d) A **Markov chain**, named for Andrey Markov, is a mathematical system that transits from one state to another (out of a finite or countable number of possible states) in a chainlike manner. It is a random process endowed with the Markov property: that the next state depends only on the current state and not on the past. Markov chains have many applications as statistical models of real-world processes.

Application:Markovian systems appear extensively in thermodynamics and statistical mechanics,

Q. 2 a) Conditions for a function to be a random variable

A random variable X is a process of assigning a number $X(\zeta)$ to every outcome ζ. The resulting function must satisfy the following two conditions

1. The set $\{X \le x\}$ is an event for every x
2. The probabilities of the events $\{X \le \infty\}$ and $\{X = -\infty\}$ equal 0.

The second condition states that, although X is allowed to be $+\infty$ or $-\infty$ for some outcomes, these outcomes form a set with zero probability.

Q2b)Convergence in distribution

A sequence $\{X_1, X_2, \ldots\}$ of random variables is said to **converge in distribution**, or **converge weakly**, or **converge in law** to a random variable X if

$$\lim_{n\to\infty} F_n(x) = F(x),$$

for every number $x \in \mathbf{R}$ at which F is continuous. Here F_n and F are the cumulative distribution functions of random variables X_n and X correspondingly

The basic idea behind this type of convergence is that the probability of an "unusual" outcome becomes smaller and smaller as the sequence progresses
A sequence $\{X_n\}$ of random variables **converges in probability** towards X if for all $\varepsilon > 0$

$$\lim_{n\to\infty} \Pr\left(|X_n - X| \ge \varepsilon\right) = 0.$$

Convergence in mean

We say that the sequence X_n converges **in the *r*-th mean** (or **in the L^r-norm**) towards X, for some $r \ge 1$, if r-th absolute moments of X_n and X exist, and

$$\lim_{n\to\infty} \mathrm{E}\left(|X_n - X|^r\right) = 0,$$

where the operator E denotes the expected value. Convergence in *r*-th mean tells us that the expectation of the *r*-th power of the difference between X_n and X converges to zero.

This type of convergence is often denoted by adding the letter L^r over an arrow indicating convergence:

$$X_n \xrightarrow{L^r} X.$$

Read more: http://www.answers.com/topic/convergence-of-random-variables#ixzz1NY2NfgdI

To say that the sequence X_n converges **almost surely** or **almost everywhere** or **with probability 1** or **strongly** towards X means that

$$\Pr\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

This means that the values of X_n approach the value of X, in the sense (see almost surely) that events for which X_n does not converge to X have probability 0. Using the probability space $(\Omega, \mathcal{F}, P)$and the concept of the random variable as a function from Ω to **R**, this is equivalent to the statement

Q3a) Given a probability space , a **stochastic process** (or **random process**) with state space X is a collection of X-valued random variables indexed by a set T ("time"). That is, a stochastic process F is a collection where each F_t is an X-valued random variable.

Given a probability space $(\Omega, \mathcal{F}, P)$, a **stochastic process** (or **random process**) with state space X is a collection of X-valued random variables indexed by a set T ("time"). That is, a stochastic process F is a collection

$$\{F_t : t \in T\}$$

where each F_t is an X-valued random variable.

A **modification** G of the process F is a stochastic process on the same state space, with the same parameter set T such that

$$P(F_t = G_t) = 1 \qquad \forall t \in T.$$

A modification is **indistinguishable** if

$$P(\forall t \in T, F_t = G_t) = 1.$$

Kolmogorov extension

The Kolmogorov extension proceeds along the following lines: assuming that a probability measure on the space of all functions $f : X \to Y$ exists, then it can be used to specify the joint probability distribution of finite-dimensional random variables $f(x_1), \dots, f(x_n)$. Now, from this n-dimensional probability distribution we can deduce an $(n - 1)$-dimensional marginal probability distribution for $f(x_1), \dots, f(x_{n-1})$. Note that the obvious compatibility condition, namely, that this marginal probability distribution be in the same class as the one derived from the full-blown stochastic process, is not a requirement. Such a condition only holds, for example, if the stochastic process is a Wiener process (in which case the marginals are all gaussian distributions of the exponential class) but not in general for all stochastic processes. When this condition is expressed in terms of probability densities, the result is called the Chapman–Kolmogorov equation.

The Kolmogorov extension theorem guarantees the existence of a stochastic process with a given family of finite-dimensional probability distributions satisfying the Chapman-Kolmogorov compatibility condition.

of a function be of little interest, but the really bad news is that virtually all concepts of calculus are of this sort. For example:

1. boundedness
2. continuity
3. differentiability

all require knowledge of uncountably many values of the function.

One solution to this problem is to require that the stochastic process be separable. In other words, that there be some countable set of coordinates $\{f(x_i)\}$ whose values determine the whole random function *f*.

The Kolmogorov continuity theorem guarantees that processes that satisfy certain constraints on the moments of their increments have continuous modifications.

Examples and special cases

Time

A notable special case is where the time is a discrete set, for example the nonnegative integers {0, 1, 2, 3, ...}. Another important special case is $T = \mathbb{R}$.

Stochastic processes may be defined in higher dimensions by attaching a multivariate random variable to each point in the index set, which is equivalent to using a multidimensional index set. Indeed a multivariate random variable can itself be viewed as a stochastic process with index set T = {1, ..., *n*}.

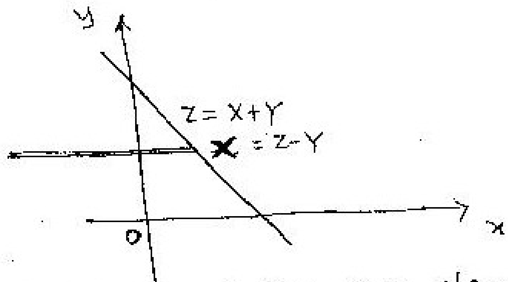
Examples

The paradigm of continuous stochastic process is that of the Wiener process. In its original form the problem was concerned with a particle floating on a liquid surface, receiving "kicks" from the molecules of the liquid. The particle is then viewed as being subject to a random force which, since the molecules are very small and very close together, is treated as being continuous and, since the particle is constrained to the surface of the liquid by surface tension, is at each point in time a vector parallel to the surface. Thus the random force is described by a two component stochastic process; two real-valued random variables are associated to each point in the index set, time, (note that since the liquid is viewed as being homogeneous the force is independent of the spatial coordinates) with the domain of the two random variables being **R**, giving the *x* and *y* components of the force. A treatment of Brownian motion generally also includes the effect of viscosity, resulting in an equation of motion known as the Langevin equation.

If the index set of the process is **N** (the natural numbers), and the range is **R** (the real numbers), there are some natural questions to ask about the sample sequences of a process $\{\mathbf{X}_i\}_{i \in \mathbf{N}}$, where a sample sequence is $\{\mathbf{X}(\omega)_i\}_{i \in \mathbf{N}}$.

Q 3 b) Given $Z = X+Y$ X & Y are r.v's.
Distribution function of r.v. Z will be

$$F_Z(z) = P\{Z \le z\}$$
$$= P\{X+Y \le z\}$$
$$= P\{(x,y) \in D_z\}$$ as shown below



The region D_z of the xy plane where $x+y \le z$ is shown above. So this probability will be

$$= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{z-y} f_{XY}(x,y)\,dx\,dy$$

i.e. integrating over the horizontal strip along the x axis first (inner integral) followed by sliding that

that strip along the y axis from $-\infty$ to $+\infty$ (outer integral) we cover the entire area.

$$\therefore \quad f_Z(z) = \frac{d}{dz} F_Z(z)$$

using Leibnitz rule i.e.

$$f_Z(z) = \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial z} \int_{-\infty}^{z-y} f_{XY}(x, y)\, dx \right) dy$$

$$= \int_{-\infty}^{\infty} \left(1 \cdot f_{XY}(z-y, y) - 0 + \int_{-\infty}^{z-y} \frac{\partial}{\partial z} (f_{XY}(x,y)) \right) dy$$

$$= \int_{-\infty}^{\infty} f_{XY}(z-y, y)\, dy$$

alternatively, the integration can be carried out first along the y axis followed by the x axis.

If X & Y are independent

$$f_{XY}(x, y) = f_X(x) f_Y(y)$$

so,

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\, dy$$

$$= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx$$

$$= f_X(x) \star f_Y(y).$$

Thus if two r.v's are independent then the density of their sum equals the convolution of their densities.

Q.4 a) The sample space has two elements (0 or 1)

let B_i, $i = 1, 2$ be the events "the symbol before the channel is 1", and "the symbol before the channel is 0" respectively.

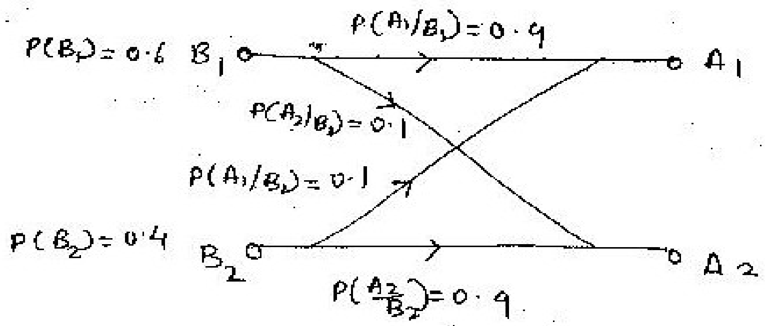
and A_i, $i = 1, 2$ be the events "the symbol after the channel is 1", and "the symbol after the channel is 0" respectively.

$P(B_1) = 0.6$ and $P(B_2) = 0.4$

also

$P(A_1/B_1) = 0.9$ and $P(A_2/B_1) = 0.1$

$P(A_1/B_2) = 0.1$ and $P(A_2/B_2) = 0.9$



$P(A_1/B_i) + P(A_2/B_i) = 1$ because A_1 & A_2 are mutually exclusive and are the only "receiver" events.

i) one is observed, probability of zero will be

$$P(B_2/A_1) = \frac{P(A_1/B_2)\,P(B_2)}{P(A_1)}$$

(law of total Probability)

$$= \frac{0.1 \times 0.4}{P(A_1)}$$

$$P(A_1) = P(A_1/B_1)\,P(B_1) + P(A_1/B_2)\,P(B_2)$$

(law of total Probability)

$$= 0.9 \times 0.6 + 0.1 \times 0.4$$

$$= 0.58$$

i.e. $$P(B_2/A_1) = \frac{0.1 \times 0.4}{P(A_1)}$$

$$= \frac{0.1 \times 0.4}{0.58} \approx 0.069$$

ii) $P(B_1/A_1)$

$$P(B_1/A_1) + P(B_2/A_1) = 1$$

$$P(B_1/A_1) = 1 - P(B_2/A_1)$$

$$= 1 - 0.069 = 0.931$$

Q5 a)Markov chains Transition probability matrix

The probability of going from state *i* to state *j* in *n* time steps is

$$p_{ij}^{(n)} = \Pr(X_n = j \mid X_0 = i)$$

and the single-step transition is

$$p_{ij} = \Pr(X_1 = j \mid X_0 = i).$$

For a time-homogeneous Markov chain:

$$p_{ij}^{(n)} = \Pr(X_{k+n} = j \mid X_k = i)$$

and

$$p_{ij} = \Pr(X_{k+1} = j \mid X_k = i).$$

The *n*-step transition probabilities satisfy the Chapman–Kolmogorov equation, that for any k such that $0 < k < n$,

$$p_{ij}^{(n)} = \sum_{r \in S} p_{ir}^{(k)} p_{rj}^{(n-k)}$$

where S is the state space of the Markov chain.

The marginal distribution $\Pr(X_n = x)$ is the distribution over states at time n. The initial distribution is $\Pr(X_0 = x)$. The evolution of the process through one time step is described by

$$\Pr(X_n = j) = \sum_{r \in S} p_{rj} \Pr(X_{n-1} = r) = \sum_{r \in S} p_{rj}^{(n)} \Pr(X_0 = r).$$

The superscript (n) is an index and not an exponent.

Reducibility

A state j is said to be **accessible** from a state i (written $i \rightarrow j$) if a system started in state i has a non-zero probability of transitioning into state j at some point. Formally, state j is accessible from state i if there exists an integer $n \geq 0$ such that

$$\Pr(X_n = j | X_0 = i) = p_{ij}^{(n)} > 0.$$

Allowing n to be zero means that every state is defined to be accessible from itself.

A state i is said to **communicate** with state j (written $i \leftrightarrow j$) if both $i \rightarrow j$ and $j \rightarrow i$. A set of states C is a **communicating class** if every pair of states in C communicates with each other, and no state in *C* communicates with any state not in *C*. It can be shown that communication in this sense is an equivalence relation and thus that communicating classes are the equivalence classes of this relation. A communicating class is **closed** if the probability of leaving the class is zero, namely that if *i* is in *C* but *j* is not, then *j* is not accessible from *i*.

That said, communicating classes need not be commutative, in that classes achieving greater periodic frequencies that encompass 100% of the phases of smaller periodic frequencies, may still be communicating classes provided a form of either diminished, downgraded, or multiplexed cooperation exists within the higher frequency class.

Finally, a Markov chain is said to be **irreducible** if its state space is a single communicating class; in other words, if it is possible to get to any state from any state.

Q5 (b) $X(t) = A\sin(w_0 t + \Theta)$

r.v.s A & Θ are independent &

$\Theta \sim [-\pi, \pi]$

$$\mu_X(t) = E\{A\sin(w_0 t + \Theta)\}$$

$$= E\{A\}\,E\{\sin w_0 t + \Theta\}$$ $\because$ A & Θ are independent

$$= E\{A\}\,E\{\sin w_0 t + \Theta\}$$

$$E\{\sin w_0 t + \Theta\} = \int_{-\infty}^{\infty} f_\Theta(\theta)\sin(w_0 t + \theta)\,d\theta$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sin(w_0 t + \theta)\,d\theta$$

$$= \frac{1}{2\pi}\left[-\cos(w_0 t + \theta)\right]_{-\pi}^{\pi}$$

$$= 0$$

$$\mu_x(t) = E\{A\} \cdot 0 = 0, \text{ for all } t.$$

$$R_{xx}(t_1, t_2) = E\{X(t_1) . X(t_2)\}$$

$$= E\{A \sin(w_0 t_1 + \theta) \cdot A \sin(w_0 t_2 + \theta)\}$$

$$= E\{A^2\} \cdot E\{\sin(w_0 t_1 + \theta) \sin(w_0 t_2 + \theta)\}$$

$$E\{\sin(w_0 t_1 + \theta) \sin(w_0 t_2 + \theta)\}$$

$$= E\frac{1}{2}\{\cos(w_0 t_1 + \theta - w_0 t_2 - \theta) - \cos(w_0 t_1 + \theta + w_0 t_2 + \theta)\}$$

$$= E\frac{1}{2}\{\cos(w_0(t_1 - t_2)) - \cos(w_0 t_1 + w_0 t_2 + 2\theta)\}$$

$$= \frac{1}{2}\cos[w_0(t_1 - t_2)] - \underbrace{E\frac{1}{2}\{\cos(w_0 t_1 + w_0 t_2 + 2\theta)\}}_{0}$$

$$= \frac{1}{2}\cos[w_0(t_1 - t_2)]$$

$$R_{xx}(t_1, t_2) = \frac{1}{2}E\{A^2\}\cos[w_0(t_1 - t_2)]$$

Q 6 a) WSS r.p with

$$R_{xx}(\tau) = A e^{-a|\tau|}$$

$$h(t) = e^{-bt} u(t)$$

$$H(\omega) = \int_{-\infty}^{\infty} h(t) e^{-j\omega t} dt = \int_{0}^{\infty} e^{-bt} \cdot e^{-j\omega t} dt$$

$$= \frac{1}{b + j\omega}$$

Power density $S_{xx}(\omega)$ of $X(t)$ is the F.T of $R_{xx}(\tau)$

$$S_{xx}(\omega) = FT\{R_{xx}(\tau)\} = \int_{-\infty}^{\infty} A e^{-a|\tau|} e^{-j\omega\tau} d\tau$$

$$= \frac{2Aa}{a^2 + \omega^2}$$

Power spectral density of the output is given by

$$S_{YY}(\omega) = |H(\omega)|^2 S_{xx}(\omega)$$

$$= \frac{1}{b^2 + \omega^2} \cdot \frac{2Aa}{a^2 + \omega^2}$$

using Partial fraction expansion

$$S_{YY}(\omega) = \frac{1}{\omega^2 + b^2} \cdot \frac{2aA}{\omega^2 + b^2} = \frac{aA}{(a^2 - b^2)b}\left(\frac{2b}{\omega^2 + b^2}\right) - \frac{A}{a^2 - b^2}\left(\frac{2a}{\omega^2 + a^2}\right)$$

using PFE.

Taking inverse FT.

$$R_{YY}(\tau) = \frac{aA}{(a^2 - b^2)b} e^{-b|\tau|} - \frac{A}{a^2 - b^2} e^{-a|\tau|}$$

Q 6 b. since for $X = x$, Y is poisson, we can write

$$P\{Y = K / X = x\} = \frac{x^K}{K!} e^{-x} \qquad K = 0, 1, 2, \ldots$$

$$E\{Y / X = x\} = \sum_{K=0}^{\infty} K \cdot \frac{x^K}{K!} e^{-x}$$

$$= x \sum_{K=0}^{\infty} \frac{e^{-x}}{(K-1)!} x^{K-1}$$

$$= x \sum_{i=0}^{\infty} e^{-x} \frac{x^i}{i!}$$

$$= x e^{-x} e^{x} = x$$

* * students are not expected to prove above result.

Thus,

Expected value of Poisson r.v. will be parameter x

Expected value of Y

$$E\{Y\} = \int_{-\infty}^{\infty} E\{Y / X = x\} f_X(x)\, dx$$

$$= \int_{-\infty}^{\infty} x f_X(x)\, dx$$

$$= \int_{-\infty}^{\infty} x \left[\frac{1}{\mu_x} \exp\left(-\frac{x}{\mu_x}\right)\right] dx$$

we obtain, by integration by parts

$$\boxed{E\{Y\} = \mu_x}$$

* Alternate approaches should be considered & Marks should be awarded accordingly.

Q. 7 a) i)

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$$

i.e. $1 = 10K + 0.25 + 0.25$

$1 = 10K + 0.5$

$$\boxed{K = 0.05}$$

ii) $P\{X \le 5\} = P\{X < 5\} + P\{X = 5\}$

= impulse at $x = 5$ must be included

so $$P\{X \le 5\} = \int_0^{5+} (0.05 + 0.25\,\delta(n-5))\,dn$$

$$= \int_0^{5+} 0.05\,dn + \int_0^{5+} 0.25\,\delta(n-5)\,dn$$

$$= 0.25 + 0.25 = \underline{\underline{0.5}}$$

for

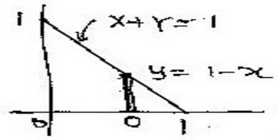
$P\{5 \le X < 10\}$ we leave out the impulse at $x = 10$ but include the impulse at $x = 5$.

Thus,

$$P\{5 \le X < 10\} = \int_{5-}^{10-} [0.05 + 0.25\,\delta(n-5)]\,dn$$

$$= \underline{\underline{0.5}}$$

1 (b) Given area is shown below.



The constant C is determined from the condition that the total probability is equal to 1.

i.e

$$\iint_{\Delta} f_{XY}(x,y)\,dx\,dy = 1$$

$$= \int_0^1 \left[\int_0^{1-x} [C(1-x-y)]\,dy \right] dx = 1.$$

$$= \frac{1}{2} C \int_0^1 (1-x)^2\,dx = 1$$

$$= \frac{1}{6} C = 1 \qquad \therefore \boxed{C = 6}$$

The marginal pdf of x

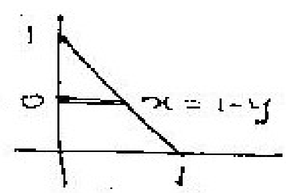
$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$$

$$= 6 \int_0^{1-x} (1-x-y)\,dy$$

$$f_X(x) = 3(1-x)^2, \qquad 0 \le x \le 1$$

$$= 0, \qquad x < 0,\ x \geq 1$$

Similarly,



$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx$$

$$= \int_0^{1-y} 6(1-x-y)\,dx$$

$$= \begin{cases} 3(1-y)^2, & 0 \le y < 1 \\ 0, & y < 0,\ y \geq 1. \end{cases}$$