1 Random Processes I

Why bother?

1. I like probability

2. The world is an interesting place

    - unpredicable disturbances (channel)

    - unpredicable information

1.1 Probability space

Ω sample space

Machinery

$\beta \equiv \sigma$ - field on Ω

(β is a subset of Ω and has special properties)

Set *Def*s

- $\Omega \in \beta$

- if event $E = \{w : w \in A_E\} \in \beta$ then so is $\bar{E}$ (E^c)

  A_E is event set

- if $E_i \in \beta \forall i \to \bigcup_{i=1}^{\infty} E_i \in \beta$

Measure *Def*s

Probability measure P

1. $0 \leq P(E) \leq 1 \qquad \forall E \in \beta$

2. $P(\Omega) = 1$

3. With $E_i \bigcap E_j = \phi \; \forall i \neq j \to P(\bigcup_{i=1}^{\infty}) = \sum_{i=1}^{\infty} P(E_i)$

   ie. probs. of disjoint events add

Fallout

1. $P(\bar{E}) = 1 - P(E)$

2. $P(\phi) = 0$

3. $P(E_1 \bigcup E_2) = P(E_1) + P(E_2) - P(E_1 \bigcap E_2)$

   growup notation

   $P(E_1 + E_2) = P(E_1) + P(E_2) - P(E_1 E_2)$

4. if $E_1 \subset E_2$ then $P(E_2) \geq P(E_1)$

conditional prob

$$P(E_1|E_2) = \begin{cases} \dfrac{P(E_1 E_2)}{P(E_2)} \text{provided} E_2 \neq 0 \\ 0 \qquad\qquad E_2 = 0 \end{cases}$$

If E_1 and E_2 indept. above implies $P(E_1 E_2) = P(E_1)P(E_2)$

Partition

$\{E_i\}_{i=1}^n$ is a finite partition of sample space

$$\sum_{i=1}^{n} E_i = \Omega \qquad E_i E_j = \phi \;\; i \neq j \;\; i,j = 1,2,...,n$$

can project on outcome A onto E_i!

$$P(A) = \sum_{i=1}^{n} P(E_i) P(A|E_i)$$

known as total probability theorem, which leads to

$$\begin{aligned} P(E_i|A) &= \frac{P(E_i)P(A|E_i)}{P(A)} \\ &= \frac{P(E_i)P(A|E_i)}{\sum_{i=1}^{n} P(E_i)P(A|E_i)} \end{aligned}$$

which is known as Bayes Thm.

1.2 Random Variables

CDF:

$Prob(X(w) \leq x) = F_X(x) = P(X \leq x)$

props

1. $0 \leq F_X(x) \leq 1$

2. $F_X(x)$ is non decreasing

3. $\lim_{x \to -\infty} F_X(x) = 0 \quad \lim_{x \to \infty} F_X(x) = 1$

4. $F_X(x)$ is continuous from right

$$\lim_{\epsilon \to 0} F_X(x + \epsilon) = F_X(x)$$

5. $P(a < X \leq b) = F_X(b) - F_X(a)$

6. $P(X = a) = F_X(a) - F_X(a^-)$

   pdf:

7. $f_X(x) = \frac{dF_X(x)}{dx}$, if discrete we get impulses props

   (a) $f_X(x) \geq 0$

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

   (c) $\int_{a^+}^{b^+} f_X(x) dx = P(a < X \leq b)$

   (d) $P(X \in A) = \int_A f_X(x) dx$

   (e) $F_X(x) = \int_{-\infty}^{x^+} f_X(x) dx$

Useful RV distributions

1. Bernoulli: biasd coin flip

   p - prob of success

   $1-p$ - prob of failure

$$p_K(k) = \begin{cases} p & k=1 \\ 1-p & k=0 \end{cases}$$

2. Binomial (a sequence of N indept Bernoulli trials)

$$P_K(k) = \binom{N}{k} p^k (1-p)^{N-k} \qquad 0 \le k \le N$$

3. uniform

4. Gaussian

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-m_X)^2/2\sigma^2}$$

5. Poisson

$$P_K(k) = \frac{(\lambda T)^k}{k!} e^{-\lambda T} \quad k = 0, 1, ...$$

6. Exponential

$$f_X(x) = \lambda e^{-\lambda x} \quad x \ge 0$$

Functions of RV

no time to review here

But $y = g(x)$ and $f_X(x)$ known,

first find $F_Y(y) = \int_{g(x) \le y} f_X(x) dx$

then differentiate

THIS ALWAYS WORKS! (book method doesn't always work)

More *Def*s

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

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

nth moment

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

nth central moment

$$\sigma_x^{(n)} = \int_{-\infty}^{\infty} (x - m_X)^n f_X(x) dx$$

$n = 2 \rightarrow$ variance σ_x^2

1. $E(cX) = cE(X)$
2. $E(c) = c$
3. $E(X + c) = E(X) + c$
4. $var(cX) = c^2 \sigma_x^2$
5. $var(c) = 0$
6. $var(X + c) = var(X) = \sigma_x^2$

Characteristic functions

$$\phi_X(s) = \int_{-\infty}^{\infty} f_X(x) e^{jsx} dx \quad \text{look like Laplace}$$

props

1. $$m_X^n = \frac{1}{j^n} \frac{d^n \phi_X(s)}{ds^n}\bigg|_{s=0}$$

2. X $N(m_X, \sigma)$ Gaussian

$$\phi_X(s) = e^{jsm_X - \frac{s^2\sigma^2}{2}}$$

More than 1 RV, some basic idea

$$F_{X,Y}(x,y) = P(X \leq x, Y \leq y)$$
$$f_{X,Y}(x,y) = \frac{\partial}{\partial x \partial y} F_{X,Y}(x,y)$$

Look at props (top p.155)

Covariance

$$E((x - m_X)(y - m_y)) = cov(X,Y)$$

correlation coefficient

$$\rho = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

Turns out that $-1 \leq \rho \leq 1$ so its a nice measure of how correlated $X + Y$ are.

no corr. $\rho = 0$ (not indept. necessarily!!!)

same RV $\rho = 1$ (only if Gaussian)

liar $\rho = -1$

prop

1. $$E(\sum_i c_i X_i) = \sum_i c_i E(X_i)$$
2. $$\begin{aligned} var(\sum_i c_i X_i) &= \sum_{i,j} cov(X_i X_j) c_i c_j \\ &= \sum_i c_i^2 var(X_i) + \sum_i \sum_{j \neq i} cov(X_i X_j) c_i c_j \end{aligned}$$

Jointly Gaussian

2 vars

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{(2\pi)^2\sigma_X^2\sigma_Y^2(1-\rho^2)}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-m_X)^2}{\sigma_X^2}+\frac{(y-m_y)^2}{\sigma_Y^2}-\frac{2\rho(x-m_X)(y-m_y)}{\sigma_X\sigma_Y}\right)}$$

general

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^2|C|}} e^{-\frac{1}{2}(\underline{x}-\underline{m_X})^T C^{-1}(\underline{x}-\underline{m_X})}$$

where

$$x = [x_1 x_2 ... x_n]^T$$

C = cov matrix, $|C| = det(C)$

Props (Important since will mostly use gaussians)

1. any subset of jointly gaussian RV's is gaussian as well
2. completely characterized by their first and second moments
3. conditioning still makes them jointly gaussian
   i.e. $f_{X_1,X_2|X_3,X_4}$ is jointly gaussian
4. $\sum a_i X_i$ is gaussian. set of linear combos is gaussian
5. uncorrelated jointly gaussian RV's are INDEPENDENT

RV sums

1. Indept. RV's $X + Y$ w/ $f_X(x), f_Y(y)$

$$Z = X + Y \to f_Z(z) = f_X(x) * f_Y(y)$$

   Prove it!

2. weak law of large numbers, $X_1, X_2, ...$ uncorrelated each w/ mean m_X and $\sigma_X^2 < \infty$ then

$$\epsilon > 0, \lim_{n\to\infty} Prob\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m_X\right| > \epsilon\right] = 0$$

   The empirical average converges in probability to the mean

3. Central limit thm. $X_1, ..., X_n$ indept. w/ means $(m_1, ..., m_n)$ and $(\sigma_1^2, ..., \sigma_n^2)$

$$Y = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i - m_i}{\sigma_i} \qquad n \to \infty \Rightarrow f_Y(y) = N(0,1)$$

   IID?$\to$ $Z = \frac{1}{N}\sum_{i=1}^{n} X_i \Rightarrow f_Z(z) = N(m_X, \sigma_x^2/n)$ when $n \to \infty$

   i.e. Large sums approach a gaussian!

   Very useful since physical world is often a superposition of uncorrelated events!

1.3 Random Processes

Each w gives a function. Random only because we don't know w ahead of time.

View

$x(t)$ random waveform

$x(t_1), x(t_2)$ random variables

$x(t, w)$ deterministic waveform

$x(t, w)$ constant

Some definitions

Def

Statistics of a Random Process $x(t)$ is the set of joint distributions on $(t_1, t_2, ...t_n)$ for any $n > 0$ and any values of $t_i \in R$, i.e. we have available
$f_{X(t_1)X(t_2)...X(t_N)}(x(t_1)...x(t_N))$ for any $n > 0$ and $it_i \in R$
Def
mth order statistics $\rightarrow$ we only know up to joint distribution or M instancesa. We sing and dance when $m = 2$ (gaussian random process has this property that it is completely described by $m = 2$ order stats)

Def Mean

$$E(X(t)) = m_X(t) = \int_{-\infty}^{\infty} x f_{X(t)}(x) dx$$

Where does it hang out?

Def Autocorr

$$R_{XX}(t_1 t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1),X(t_2)}(x_1, x_2) dx_1 dx_2$$

Def Autocov

$$C_{XX}(t_1 t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X(t_1))(x_2 - m_X(t_2)) f_{X(t_1),X(t_2)}(x_1, x_2) dx_1 dx_2$$

Stationarity

Def Strict sense stationarity (SSS)

$$f_{X(t_1)...X(t_N)}(x_1, ...x_N) = f_{X(t+\Delta)...X(t_N+\Delta)}(x_1, ...x_N)$$

$\forall D \in R \qquad \forall N > 0 \quad t_i \in R$
i.e. relative statistics look the same no matter what absolute time you look.
mth order stationarity $\Rightarrow$ above, only $N \leq M$

Def Wide sense stationarity (WSS)

$m_X(t) =$ constant

$R_{XX}(t_1 t_2) = R_{XX}(\tau)$ only time difference matters

($R_X(\tau)$ for short)

Def Cyclostationarity

$m_X(t) = m_X(t + kT)$ for some $T \in R$

$$R_X(t+\tau+kT, t+kT) = R_X(t+\tau, t)$$

where $t, \tau, T \in R$

Autocorr of Stationary Process

1. $R_X(\tau) = R_X(-\tau)$ (process must be real)

   pf. $R_x(\tau) = E(X(t)X(t+\tau)) = E(X(t-\tau)X(t)) = R_X(-\tau)$

2. $R_X(0) \geq |R_X(\tau)|$ pf. $E[(X(t) \pm X(t-\tau))^2] \geq 0$

   Therefore,

   $$\underbrace{E(X^2(t))}_{R_X(0)} + \underbrace{E(X^2(t-\tau))}_{R_X(0)} \pm 2R_X(\tau) \geq 0$$

   $R_X(0) \geq -R_X(\tau)$

   $R_X(0) \geq +R_X(\tau)$

   $R_X(0) \geq |R_X(\tau)|$

3. If $R_X(0) = R_X(T)$ then $R_X(kT) = R_X(0)$

   pf. more involved than its worth for review of a factoid

   Book use induction + Cauchy - Schwartz

   $$\left| \int_{-\infty}^{\infty} x(t) y^*(t) dt \right|^2 = \int |x(t)|^2 dt \int |y(t)|^2 dt$$

Ergodicity

(Strictly stationary) (simple argument)

Define averages first

1. $E(gX(t))) = \int_{-\infty}^{\infty} g(x) f_{X(t)}(x) dx$ ensemble average of possible values
2. $$\overline{g(x(t,w))} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} g(x(t,w))dt$$

   time average for a given sample function

   if $g(x(t,w)) = E(g(X(t)))$ $\omega \in \Omega$,

   then $x(t)$ is ergodic, $\rightarrow$ only need to look at a simple $x(t,w)$ to determine the props.

Problem: We don't live forever, so can't test ergodicity for arbitrary process.

Power and Energy (general)

$$E_i = \int_{-\infty}^{\infty} x^2(t,\omega_i)dt$$

energy in $x(t,\omega_i)$

$$P_i = \lim_{T\to\infty} \int_{-T/2}^{T/2} x^2(t,\omega_i)dt$$

power in $x(t,\omega_i)$

Def P_X, E_X random variables

CDFs:

$$F_{P_X}(P_X) = Prob(\bigcup \omega_i : P_i \leq \mathcal{P}_X)$$

$$F_{E_X}(E_X) = Prob(\bigcup \omega_i : E_i \leq \mathcal{E}_X)$$

Def

$P_X = E(\mathcal{P}_X) \qquad E_X = E(\mathcal{E}_X)$

$$E_X = \int_{-\infty}^{\infty} E(X^2(t))dt = \int_{-\infty}^{\infty} R_X(t,t)dt$$

$$P_X = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} E(X^2(t))dt = \lim_{T\to\infty} \int_{-T/2}^{T/2} R_X(t,t)dt$$

Stationary $\Rightarrow R_X(t,t) = R_X(0)$

$\mathcal{P}_X = R_X(0 \Rightarrow E_X$infinite for a stationary process

Ergodic? $\Rightarrow P_i = P_X$

More than one process

$x(t,\omega_i) \qquad y(t,\omega_i), \omega_i \in \Omega$ (need to be defined on the same sample space)

Def

1. $X(t), Y(t)$ indept. if

$$Prob(X(t_1) \le x, T(t_2) \le y) = Prob(X(t_1) \le x)Prob(Y(t_2) \le y) \quad \forall t_1, t_2$$

2. $X(t_1), Y(t_2)$ uncorrelated if

$$E(X(t_1)Y(t_2)) = E(X(t_1))E(Y(t_2)) \quad \forall t_1, t_2$$

3. Cross correlation

$$R_{XY}(t_1,t_2) = E(X(t_1)Y(t_2)) = R_{YX}(t_2,t_1)$$

4. Jointly WSS if $X(t), Y(t)$ each WSS and

$$R_{XY}(t_1,t_2) = R_{XY}(t_1 - t_2) = R_{XY}(\tau)$$

END