PROBABILITY THEORY (CH.4)

Executive Summary

①

* **Deterministic signal**: signals that may be modeled as completely specified functions of time.

  Random signal: it is not possible to predict its precise value in advance.

* carrier → deterministic
  noise → random
  information (message) signal → random

* Then, how is the Fourier Transform performed, how is the bandwidth calculated?

* The mathematical discipline that deals with the statistical characterization of random signals is probability theory.

Relative Frequency: Let event A denote one of the possible outcomes of a random experiment. Suppose that in n trials of the experiment, event A occurs $N_n(A)$ times.

→ relative frequency of event A = $\frac{N_n(A)}{n}$.

* $0 \le \frac{N_n(A)}{n} \le 1$, $\quad P(A) = \lim_{n\to\infty} \frac{N_n(A)}{n}$

* A **Probability System**: consists of

  1. Sample space S of outcomes
  2. A probability measure $P(\cdot)$ assigned to each event (outcome) in S such that

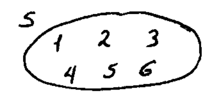
  (i) $P(S) = 1$ $\quad$ (ii) $0 \le P(A) \le 1$

  (iii) If $A+B$ is the union of two mutually exclusive events in S, then

  → $P(A+B) = P(A) + P(B)$

②

Ex: Die experiment



$P(1) = \cdots = P(6) = \frac{1}{6}$

$P(i) = \lim_{n\to\infty} \frac{N_n(i)}{n}$, $\quad i = 1, 2, \ldots, 6$

$P(1 \text{ or } 2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$

If A and B are not mutually exclusive, then

$P(A+B) = P(A) + P(B) - P(AB)$

$P(1 \text{ or } 2) = P(1) + P(2) - \underbrace{P(1 \text{ and } 2)}_{0}$

if A and B are mutually exclusive → $P(AB) = 0$

* conditional probability of B given A: $\quad P(B|A) = \frac{P(AB)}{P(A)}$

$P(AB) = P(B|A)P(A) = P(A|B)P(B)$

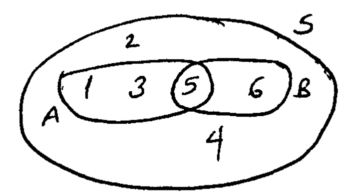
if A and B are statistically independent → $P(AB) = P(A)P(B)$

→ $P(B|A) = P(B)$

Bayes' Rule: $\quad P(B|A) = \frac{P(A|B)P(B)}{P(A)}$

Ex: A: odd

B: ≥ 5



$P(A) = \frac{1}{2}$

$P(B) = \frac{1}{3}$

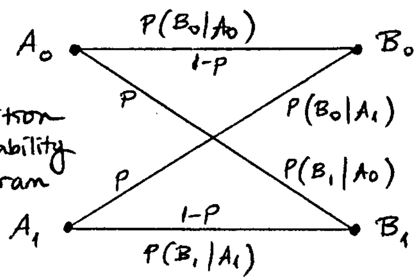
$P(AB) = \frac{1}{6}$

$P(A|B) = \frac{1}{2} = P(AB)/P(B) = \frac{1/6}{1/3}$

$P(B|A) = \frac{1}{3} = P(AB)/P(A) = \frac{1/6}{1/2} = 1/3$

Binary Symmetric Channel

Transition Probability Diagram



discrete memoryless channel

channel output at any time depends only on the channel input at that time

A priori probabilities of sending 0 and 1:

$P(A_0) = P_0$, $P(A_1) = P_1$ $\longrightarrow$ $P_0 + P_1 = 1$

$P(B_i|A_j)$: probability of receiving B_i given that A_j is sent, $i,j = 0, 1$.

conditional probability of error : $P(B_0|A_1) = P(B_1|A_0) = p$
$\longrightarrow$ symmetric channel

Obtain the $\underbrace{\text{a posteriori}}_{\text{after the fact}}$ probabilities: $P(A_0|B_0)$ and $P(A_1|B_1)$

B_0 and B_1 : mutually exclusive

$\rightarrow$ $P(B_0|A_0) + P(B_1|A_0) = 1$ $\rightarrow$ $P(B_0|A_0) = 1-p$
$P(B_0|A_1) + P(B_1|A_1) = 1$ $\rightarrow$ $P(B_1|A_1) = 1-p$

$P(B_0) = P(B_0|A_0)P(A_0) + P(B_0|A_1)P(A_1) = (1-p)P_0 + pP_1$

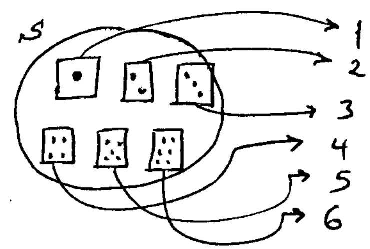
$P(B_1) = P(B_1|A_0)P(A_0) + P(B_1|A_1)P(A_1) = pP_0 + (1-p)P_1$

$$P(A_0|B_0) = \frac{P(B_0|A_0)P(A_0)}{P(B_0)} = \frac{(1-p)P_0}{(1-p)P_0 + pP_1}$$

$$P(A_1|B_1) = \frac{P(B_1|A_1)P(A_1)}{P(B_1)} = \frac{(1-p)P_1}{pP_0 + (1-p)P_1}$$

Random Variables

A function whose domain is a sample space and whose range is some set of real numbers is called a random variable of the experiment.



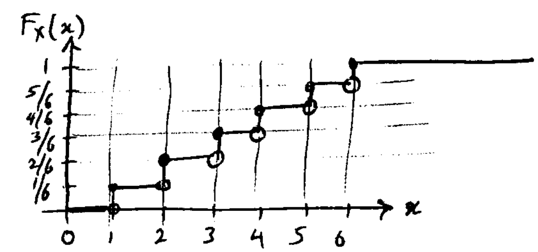
outcome : s $\longrightarrow$ random variable X

X: discrete values $\longrightarrow$ discrete rv

X : continuous values $\longrightarrow$ continuous rv

$F_X(x) = P(X \le x)$ ... cumulative distribution function (cdf)

$F_X(4) = P(X \le 4) = \frac{4}{6}$



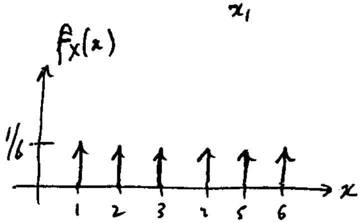
properties of $F_X(x)$ $\left\{ \begin{array}{l} \bullet\ 0 \le F_X(x) \le 1 \\ \bullet\ \text{monotone-nondecreasing function of } x \end{array} \right.$

$x_1 < x_2 \rightarrow P(X \le x_1) \le P(X \le x_2)$

$\rightarrow F_X(x_1) \le F_X(x_2)$

$f_X(x) = \frac{d}{dx} F_X(x)$ ... probability density function (pdf)

$$P(x_1 < X \le x_2) = P(X \le x_2) - P(X \le x_1) = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x)\,dx$$



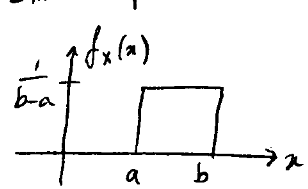
$P(X=3) = 1/6$

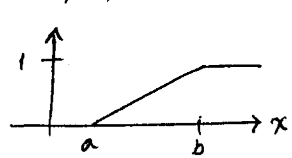
$P(.5 < X \le 3.5) = 1/2$

$$F_X(x) = \int_{-\infty}^{x} f_X(\mu)\,d\mu$$

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1 \quad \rightarrow \quad f_X(x): \begin{cases} \text{nonnegative function} \\ \text{area} = 1 \end{cases}$$

EX: uniform r.v





Several R.V.s

$F_{X,Y}(x,y) = P(X \le x, Y \le y)$ ... joint distribution function

$f_{X,Y}(x,y) = \dfrac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$ ... joint density function

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$$

$$F_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{X,Y}(\alpha,\beta)\,d\alpha\,d\beta$$

$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,\eta)\,d\eta$ ... marginal density

The conditional probability density function of Y given that $X=x$:

$f_Y(y|X=x) = f_Y(y|x) = \dfrac{f_{XY}(x,y)}{f_X(x)}$ provided $f_X(x) > 0$

$$= \frac{f_{XY}(X=x, Y=y)}{f_X(X=x)}$$

$f_Y(y|x) \ge 0$

$\int_{-\infty}^{\infty} f_Y(y|x)\,dy = 1$

If the rvs X and Y are statistically independent, then knowledge of the outcome of X does not affect the distribution of Y.

$\rightarrow f_Y(y|x) = f_Y(y)$

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

$f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots x_n)$ joint pdf of n rv.s

6a

A bag contains 3 balls: 1, 2, 3

Take one ball, put it back in the bag, then take a second ball. $\boxed{X|Y}$

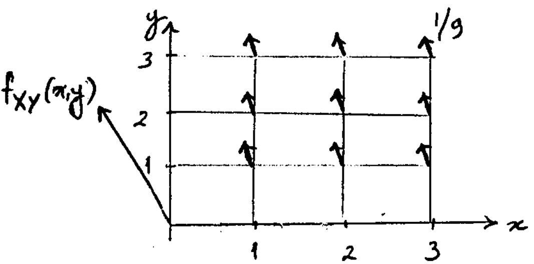
$$P(2,3) = \frac{1}{3}\cdot\frac{1}{3} = \frac{1}{9}$$

$$f_X(x) = \frac{1}{3}\delta(x-1) + \frac{1}{3}\delta(x-2) + \frac{1}{3}\delta(x-3)$$

$$f_Y(y) = \frac{1}{3}\delta(y-1) + \frac{1}{3}\delta(y-2) + \frac{1}{3}\delta(y-3)$$

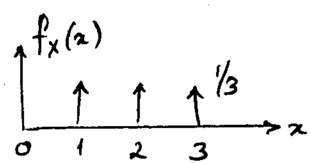
$$P(X=i, Y=j) = P(X=i)\,P(Y=j) = \frac{1}{9} \qquad i,j = 1,2,3$$

$$f_{XY}(x,y) = \frac{1}{9}\delta(x-1)\delta(y-1) + \frac{1}{9}\delta(x-1)\delta(y-2) + \frac{1}{9}\delta(x-1)\delta(y-3) + \cdots + \cdots + \frac{1}{9}\delta(x-3)\delta(y-3)$$



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

X and Y are independent



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

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

6b

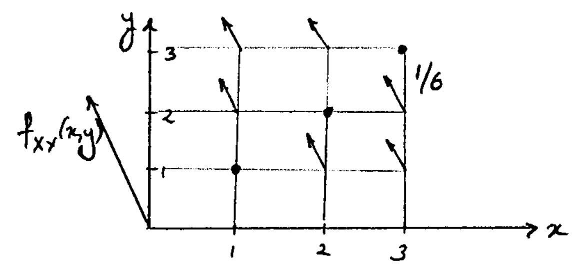
Take one ball, do not put it back in the bag, then take a second ball. $\boxed{X|Y}$

$$P(2,3) = \frac{1}{3}\cdot\frac{1}{2} = \frac{1}{6}$$

$$f_X(x) = \frac{1}{3}\delta(x-1) + \frac{1}{3}\delta(x-2) + \frac{1}{3}\delta(x-3)$$

$$f_Y(y) = \frac{1}{3}\delta(y-1) + \frac{1}{3}\delta(y-2) + \frac{1}{3}\delta(y-3)$$

$$P(X=i, Y=j) = \begin{cases} 0 & \text{if } i=j \\ \frac{1}{6} & \text{if } i\neq j \end{cases}$$



$$f_{XY}(x,y) = \frac{1}{6}\delta(x-1)\delta(y-2) + \frac{1}{6}\delta(x-1)\delta(y-3) + \frac{1}{6}\delta(x-2)\delta(y-1) + \frac{1}{6}\delta(x-2)\delta(y-3) + \frac{1}{6}\delta(x-3)\delta(y-1) + \frac{1}{6}\delta(x-3)\delta(y-2)$$

$\neq f_X(x)f_Y(y) \longrightarrow$ X an Y are not independent

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,z)\,dz \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(z,y)\,dz$$

$$f_Y(y|X=1) = \frac{1}{2}\delta(y-2) + \frac{1}{2}\delta(y-3) \neq f_Y(y)$$

$$= \frac{f_{XY}(X=1,y)}{f_X(X=1)} = \frac{\frac{1}{6}\delta(0)\delta(y-2) + \frac{1}{6}\delta(0)\delta(y-3)}{\frac{1}{3}\delta(0)}$$

$$= \frac{1}{2}\delta(y-2) + \frac{1}{2}\delta(y-3)$$

Statistical Averages

Expected value (mean)

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

$\hookrightarrow$ statistical expectation operator

* μ_X locates the center of gravity of the area under the pdf curve.

Die example: $\text{average} = \dfrac{1+2+3+4+5+6}{6} = 3.5$

$$f_X(x) = \frac{1}{6}\delta(x-1) + \frac{1}{6}\delta(x-2) + \cdots + \frac{1}{6}\delta(x-6)$$

$$\mu_X = \int_{-\infty}^{\infty} x f_X(x)\,dx = \frac{1}{6}(1+\cdots+6) = 3.5$$

For discrete r.v.: $\mu_X = \sum_i x_i P(X=x_i)$

$\rightarrow$ sample average

Function of a Random Variable

$Y = g(X)$, Find μ_Y.

* Brute Force method: $E[Y] = \int_{-\infty}^{\infty} y \underbrace{f_Y(y)}\,dy$

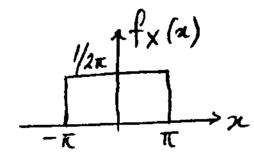
have to obtain $f_Y(y)$ from $f_X(x)$

* Simpler method:

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

Ex: X is uniform in $(-\pi, \pi)$. Find $E[\cos X]$.

$Y = g(X) = \cos(X)$



$$E[Y] = \int_{-\pi}^{\pi} \cos(x)\frac{1}{2\pi}\,dx = \frac{-1}{2\pi}\sin x\Big|_{-\pi}^{\pi} = 0$$

* Mean-Square Value: $E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx$

$\sigma_X^2 = \text{var}[X] = E[(X-\mu_X)^2] = \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x)\,dx$: variance

σ_X : standard deviation

σ_X^2 (or σ_X): a measure of randomness, measures how wide is the pdf around the mean.

Note: if $f_X(x) = \delta(x-\mu_X)$ (no randomness)

$$\rightarrow \sigma_X^2 = \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x)\,dx = \int_{-\infty}^{\infty} (x-\mu_X)^2 \delta(x-\mu_X)\,dx$$

$$= (x-\mu_X)^2\Big|_{x=\mu_X} = 0$$

$$\sigma_X^2 = E(X^2 - 2\mu_X X + \mu_X^2)$$

$$= E(X^2) - 2\mu_X^2 + \mu_X^2$$

$$= E[X^2] - \mu_X^2$$

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

correlation between r.v.s X and Y: $E[XY]$

* covariance: correlation between centered r.v.s

$$cov[XY] = E[(X-\mu_X)(Y-\mu_Y)]$$
$$= E[XY] - \mu_X\mu_Y$$

* X and Y are uncorrelated <u>iff</u> $cov[XY]=0$

  " orthogonal <u>iff</u> $E[XY]=0$

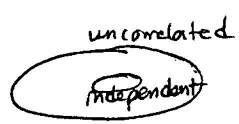
* $\left(\mu_X=0 \vee \mu_Y=0 \vee (\mu_X=0 \wedge \mu_Y=0)\right)$ and X,Y: orthogonal

  $\rightarrow$ X,Y: uncorrelated

* if X, Y : statistically independent

  $\rightarrow$ $E[XY] = \int x f_X(x)\,dx \int y f_Y(y)\,dy = E[X]E[Y]$

  $\therefore$ statistical independence $\underset{\not\leftarrow}{\rightarrow}$ uncorrelatedness



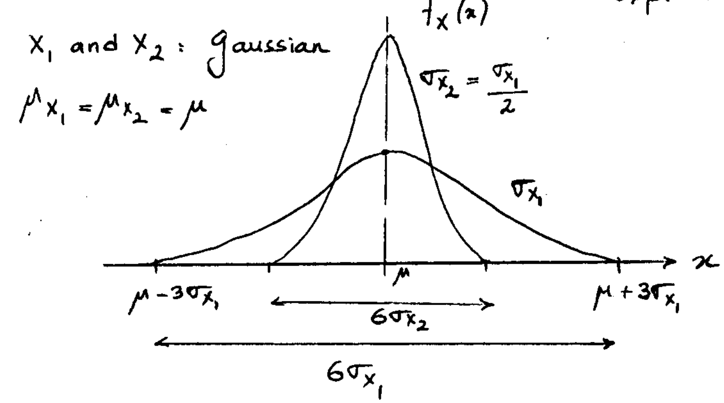
Gaussian (Normal) Random Variable

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}}\, e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}}$$

mean and variance are explicit

X_1 and X_2 : Gaussian

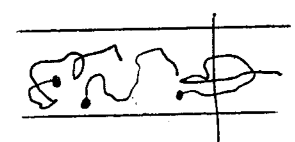
$\mu_{X_1} = \mu_{X_2} = \mu$



Central Limit Theorem

If a rv X is the sum of a large number of "small" r.v.s, then under general conditions, the pdf of X approaches that of a Gaussian rv.

Ex: thermal noise is the electrical noise arising from the random motion of electrons in a conductor.



current: $i = \frac{q}{t}$: $\frac{\text{charge}}{\text{time}}$

X: the (net) # of electrons passing through a crosssectional area in 1 sec.

$X = X_1 + \cdots + X_n$ $\quad$ n: very large

$f_X(x)$: Gaussian