In The Name of God

Sharif University of Technology
Department of Aerospace Engineering

45-766: Optimal Control II

Instructor: Maryam Kiani

CH#1: Introduction

Probability

- How **likely** something is to happen.

- Axiomatic definition of Probability in probability space (Ω, E, P)

Sample space

event space

Probability measure

- Probability is a number that is assigned to each member of a collection of events from a random experiment that satisfies the following properties:

1) P(S) = 1 where S is the sample space.

2) $0 \leq P(A) \leq 1$ for any event A

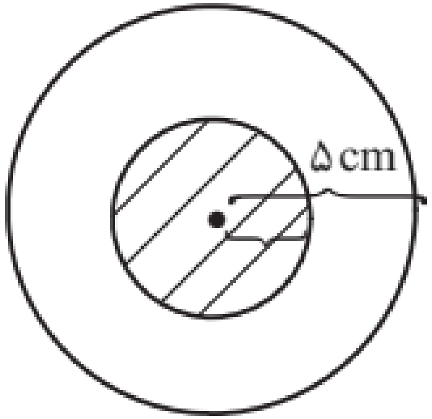
3) For two events A_1 and A_2 with $\underbrace{A_1 \cap A_2 = \varnothing}_{\text{mutually exclusive}}$ $: P(A_1 \cup A_2) = P(A_1) + P(A_2)$

مثال

- مطلوبست محاسبه احتمال آمدن عدد زوج در پرتاب یک تاس

- کمیته ای ۵ نفره از یک گروه شامل ۶ مرد و ۹ زن انتخاب می شود. اگر انتخاب به تصادف صورت گیرد، احتمال این که کمیته از ۳ مرد و ۲ زن تشکیل شود، چقدر است؟

- تیراندازی به هدفی شلیک می کند که قطر آن 10 cm و قطر دایره مرکزی هدف 2 cm است. احتمال این که تیر به مرکز هدف اصابت کند چقدر است؟



Conditional Probability

- Conditional probability that A occurs given that B has occurred is

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \qquad P(B)\neq 0$$

- Multiplication rule

$$P(A\cap B)=P(A|B)P(B)=P(B|A)P(A)$$

$$P(A_1A_2\ldots A_n)=P(A_1|A_2\ldots A_n)P(A_2|A_3\ldots A_n)\ldots P(A_{n-1}|A_n)P(A_n)$$

- Independence $P(A|B)=P(A)$

$$P(A\cap B)=P(A|B)P(B)=P(A)P(B)$$

مثال

- کیسه ای حاوی ۱۰ مهره سفید، ۵ مهره زرد و ۱۰ مهره سیاه است. مهره ای به تصادف از کیسه انتخاب شده است. مشاهده می شود که این مهره سیاه نیست. احتمال اینکه مهره زرد باشد، چقدر است؟

Random Variable

- Concept

$$S = \{HH, HT, TH, TT\} \xrightarrow{X} \{2,1,1,0\}$$

$$X(HH) = 2, \quad X(TH) = X(HT) = 1, \quad X(TT) = 0$$

- Random variable (RV): real-valued function defined on the sample space

$$X : S \to R$$

- Difference between X and x

- RV: Discrete, Continuous

Discrete Random Variables

x	0	1	2
$P(X=x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

x	1	2	3	4	5	6
$P(X=x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

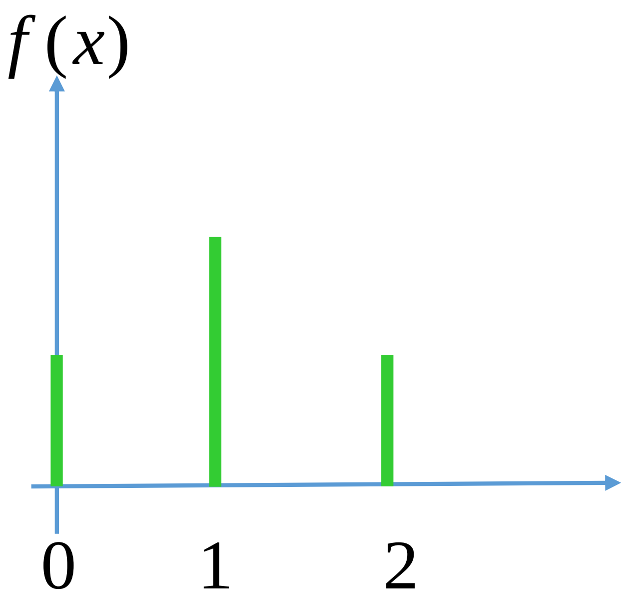
$$f_X(x) = \frac{1}{6}$$

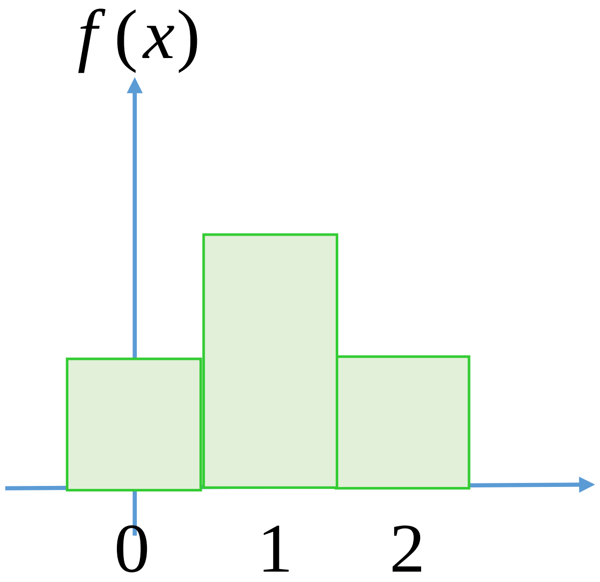
Discrete Random Variables

- Probability distribution function (probability mass function: pmf):

  For a discrete random variable X with possible values $x_1, x_2, ..., x_n$, a pmf is a function such that

  - $f(x_i) \geq 0$
  - $\sum_{i=1}^{n} f(x_i) = 1$
  - $f(x_i) = P(X = x_i)$



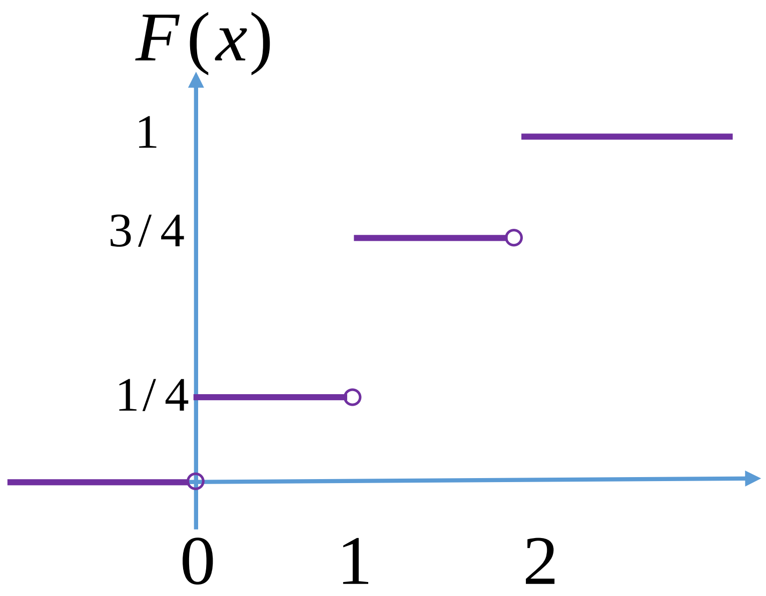


Discrete Random Variables

- Cumulative Distribution Function (CDF)

$$F_X(x): R \to [0,1]$$

$$F_X(x) = P(X \le x) = \sum_{t \le x} f(t)$$



مثال

- جعبه ای حاوی ۳ توپ قرمز و ۲ توپ زرد است. از این جعبه، دو توپ را به طور تصادفی و بدون جایگذاری خارج می کنیم. اگر متغیر تصادفی X نشان دهنده تعداد توپ قرمز در این آزمایش باشد، $F_X(x)$ را تعیین کنید.

Continuous Random Variables

- Probability density function (pdf)
  - $f(x) \geq 0$
  - $\int_{-\infty}^{+\infty} f(x)dx = 1$
  - $P(a < x < b) = \int_{a}^{b} f(x)dx$
- Cumulative density function (CDF): $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx$

$$f(x) = \frac{d}{dX} F_X(x)$$

$$P(a < x < b) = F(b) - F(a)$$

مثال

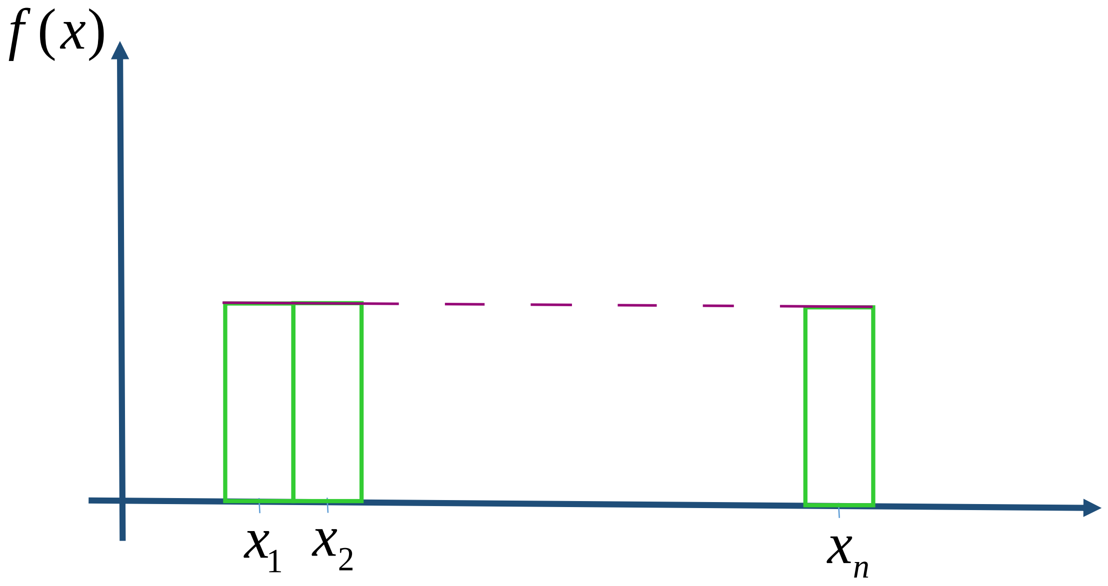
- Pdf متغیر تصادفی X به صورت زیر داده شده است:

$$f_X(x) = \begin{cases} axe^{-x^2} & x > 0 \\ 0 & x \le 0 \end{cases}$$

a را به گونه ای مشخص کنید که $f(x)$ معتبر باشد.

توزیع احتمال متغیر تصادفی گسسته

- اگر متغیر تصادفی X تمام مقادیر خود را با احتمال مساوی اختیار کند، توزیع احتمال این متغیر تصادفی را **توزیع یکنواخت** می نامیم:



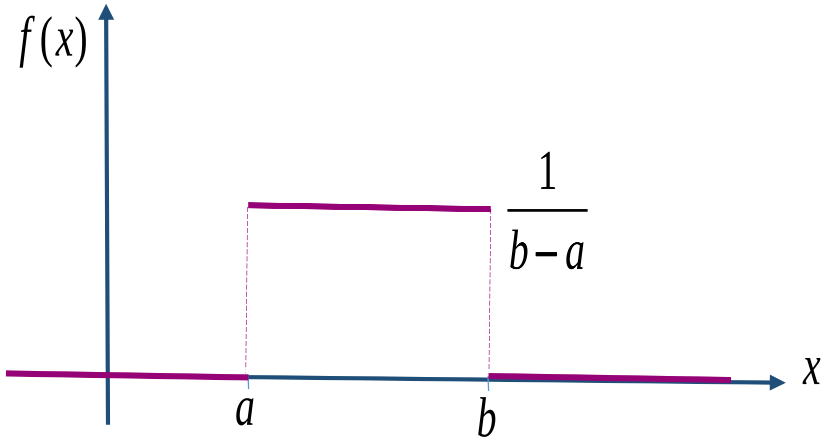
$$f(x;n)=\frac{1}{n},\quad x=x_1,x_2,\ldots,x_n$$

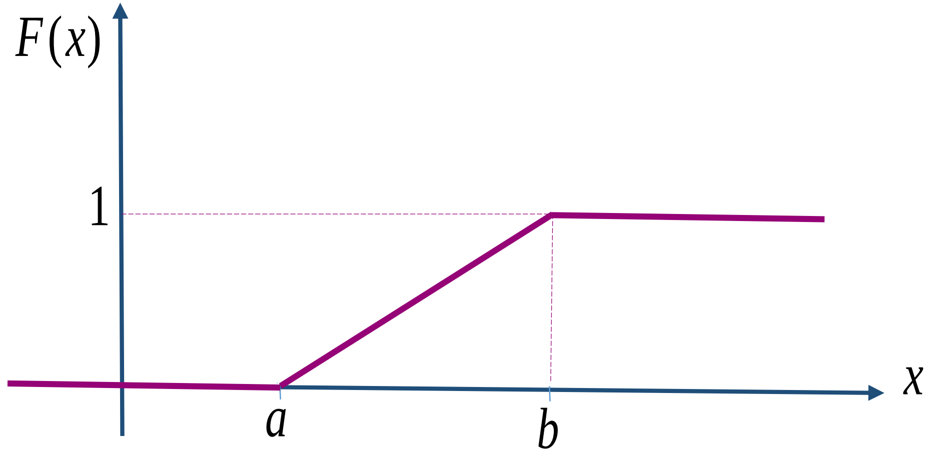
توزیع یکنواخت پیوسته

- متغیر تصادفی X در فاصله [a,b] دارای توزیع یکنواخت است اگر تابع چگالی احتمال آن به صورت زیر تعریف شود: $X \sim U[a,b]$

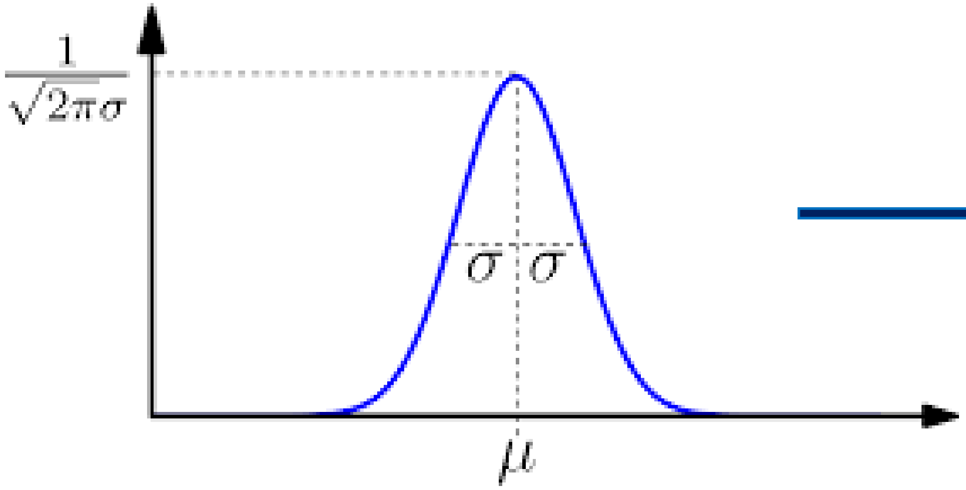
$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a\le x\le b\\ 0\ , & else\end{cases}$$

$$F(x)=\begin{cases}0 & x<a\\ \dfrac{x-a}{b-a}, & a\le x\le b\\ 1 & x>b\end{cases}$$





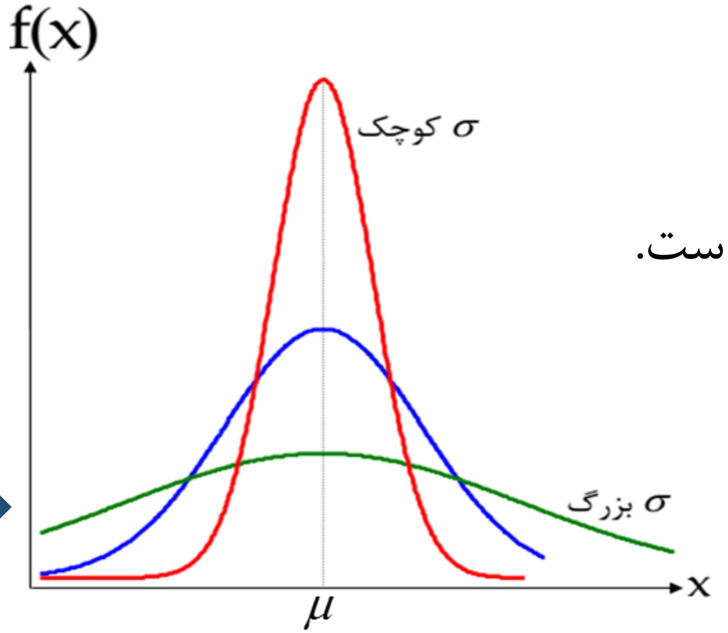
توزیع نرمال



- خصوصیات منحنی نرمال
  - منحنی نسبت به خط $x = \mu$ متقارن است

    (میانه، میانگین و مد بر هم منطبق هستند).
  - منحنی دارای دو نقطه عطف در نقاط $x = \mu \pm \sigma$ است.
  - $\lim_{x \to \pm\infty} f(x) = 0$ ، یعنی محور xها مجانب افقی منحنی است.
  - مساحت سطح واقع در زیر منحنی و بالای محورxها برابر یک است.

$$f(x) = N(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$



$N(x;\mu,\sigma^2)$

توزیع نرمال

- محاسبه سطح زیر منحنی نرمال

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} N(x;\mu,\sigma)$$

تغییر متغیر $z = \dfrac{x-\mu}{\sigma} \Rightarrow dz = \dfrac{1}{\sigma} dx$

$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{x_1}^{x_2} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

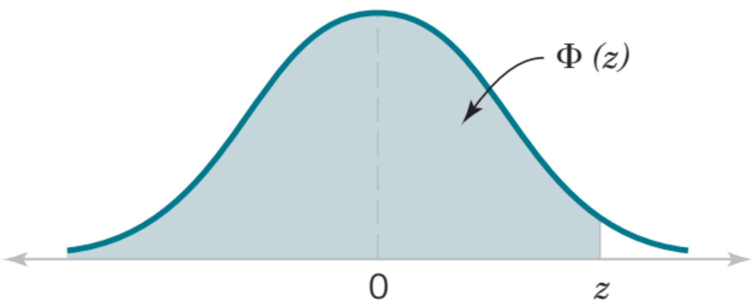
$$= \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{1}{2}z^2} dz = \int_{z_1}^{z_2} \boxed{N(z;0,1)}\, dz$$

متغیر تصادفی نرمال استاندارد (Z)

$$P(x_1 < X < x_2) = P(z_1 < Z < z_2)$$

- $$f(z) = \phi(z) = N(x;0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{u^2}{2}} du = G(z) = \Phi(z)$$



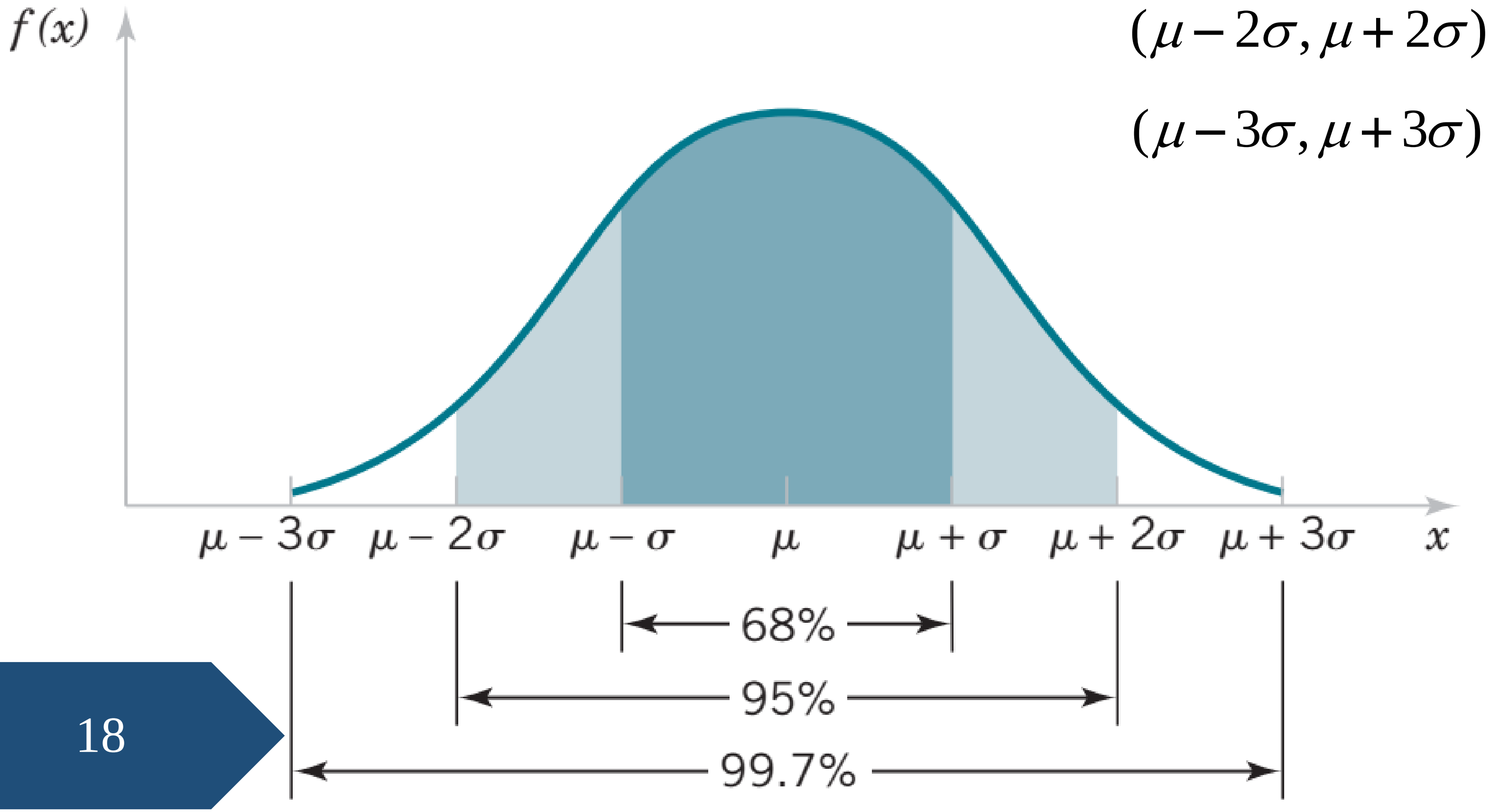
$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500000	0.503989	0.507978	0.511967	0.515953	0.519939	0.532922	0.527903	0.531881	0.535856
0.1	0.539828	0.543795	0.547758	0.551717	0.555760	0.559618	0.563559	0.567495	0.571424	0.575345
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.606420	0.610261	0.614092
0.3	0.617911	0.621719	0.625516	0.629300	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645	0.677242	0.680822	0.684386	0.687933
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840	0.712260	0.715661	0.719043	0.722405
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373	0.776373	0.779350	0.782305	0.785236
0.8	0.788145	0.791030	0.793892	0.796731	0.799546	0.802338	0.805106	0.807850	0.810570	0.813267
0.9	0.815940	0.818589	0.821214	0.823815	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913
1.0	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141	0.855428	0.857690	0.859929	0.862143

- $\Phi(-z) = 1 - \Phi(z)$

توزیع نرمال

- برای هر توزیع نرمال با میانگین μ و واریانس σ^2
  - تقریبا ۶۸.۲۸٪ مشاهدات در بازه $(\mu-\sigma, \mu+\sigma)$
  - تقریبا ۹۵.۴۴٪ مشاهدات در بازه $(\mu-2\sigma, \mu+2\sigma)$
  - تقریبا ۹۹.۷۴٪ مشاهدات در بازه $(\mu-3\sigma, \mu+3\sigma)$



$$P(\mu-\sigma < X < \mu+\sigma) = 0.6827$$

$$P(\mu-2\sigma < X < \mu+2\sigma) = 0.9545$$

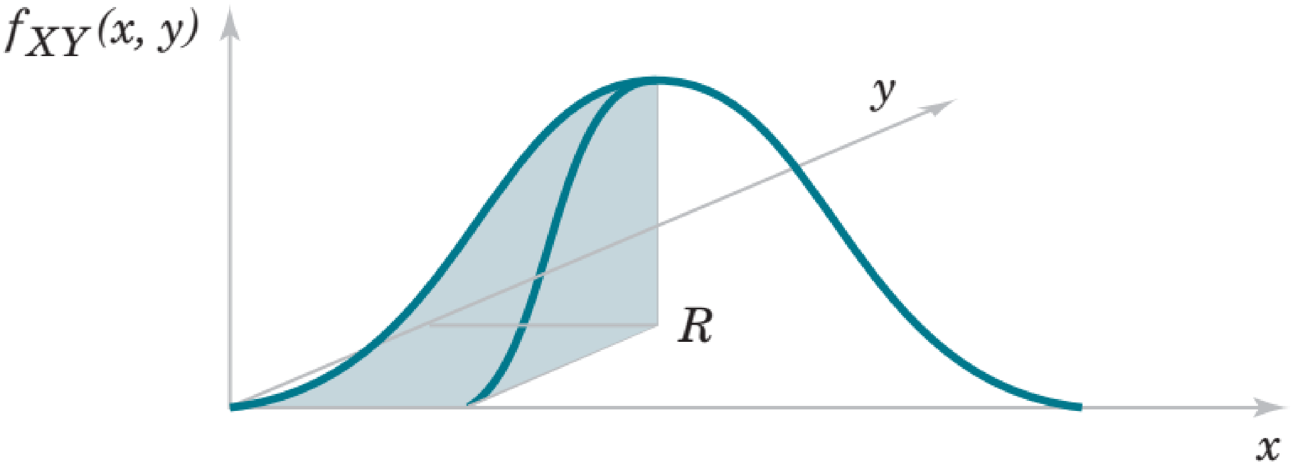
$$P(\mu-3\sigma < X < \mu+3\sigma) = 0.9973$$

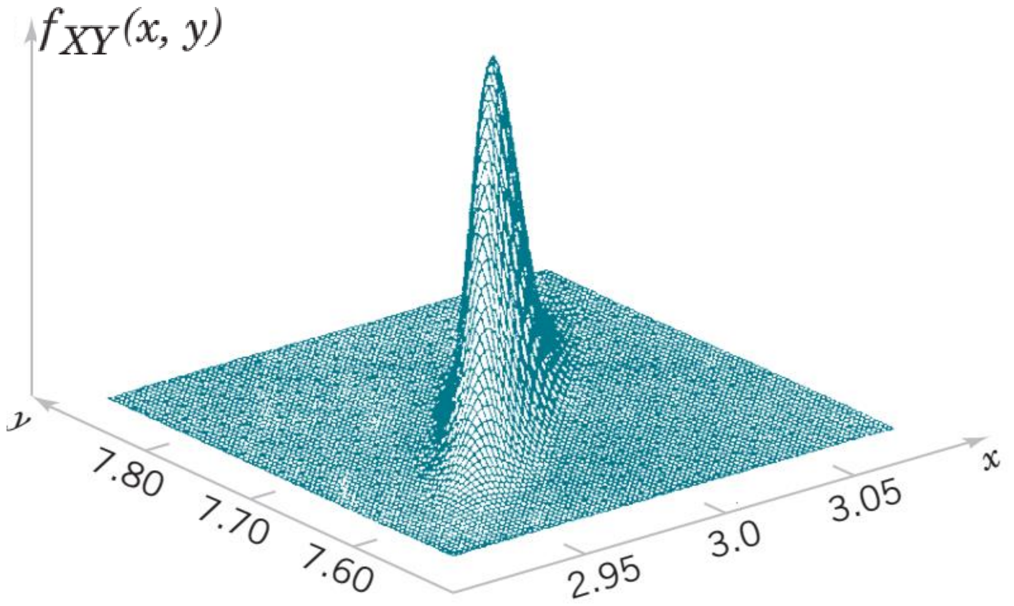
مثال

- اگر مقدار اشعه ای که کاربر کامپیوتر ممکن است در هر ساعت کاری دریافت کند دارای توزیع نرمال با میانگین ۴.۳۵ و واریانس ۰.۴۹ باشد، مطلوبست احتمال دریافت اشعه بین ۴ الی ۵ واحد.

Joint Probability Function

- Joint pmf
  - $f(x,y) \geq 0 \quad \forall x, y$
  - $\sum_x \sum_y f(x,y) = 1$
  - $P((X,Y) \in A) = \sum_A \sum f(x,y)$
- Joint pdf
  - $f(x,y) \geq 0 \qquad \forall x, y$
  - $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)dxdy = 1$
  - $P((X,Y) \in A) = \iint_A f(x,y)dxdy$





Joint Probability Distribution

- Joint CDF: $F(x,y) = P(X \le x, Y \le y)$

  - $F(x,y) = \sum_{X \le x} \sum_{Y \le y} f(x,y)$

  - $F(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(x,y)\,dxdy$

- Then

  - $F_X(a) = F(X \le a, Y < \infty) = F(a,\infty)$

    $F_Y(b) = F(X < \infty, Y \le b) = F(\infty, b)$

مثال

- در پرتاب ۳ سکه، اگر X را تعداد شیرها و Y را تعداد خط ها در نظر بگیریم:

$$\sum_x \sum_y f(x,y) = f(0,0) + f(0,1) + f(0,2) + f(0,3) + f(1,1) + \dots = 1$$

- به ازای چه مقداری از c تابع f(x,y) یک pdf است؟ احتمال پیشامد $A = \{(x,y) \mid 0 < x < 1/2,\ 0 < y < 1\}$ چیست؟

$$f(x,y) = c(xy + \frac{x^2}{2}) \qquad 0 < x < 1, \quad 0 < y < 2$$

Joint Probability Function

- Relationship between Joint pdf and CDF:

$$f(x,y)=\frac{\partial^2}{\partial x \partial y}F(x,y)$$

- Theorem: $P(X>a, Y>b)=1-F_X(a)-F_Y(b)+F_{X,Y}(a,b)$

- General case

$$P(a_1 \le X \le b_1, a_2 \le Y \le b_2)=F_{X,Y}(a_2,b_2)+F_{X,Y}(a_1,b_1)-F_{X,Y}(a_1,b_2)-F_{X,Y}(a_2,b_1)$$

Marginal Probability Function

- Discrete RV
  - $$f_X(x) = \sum_y f(x, y)$$
  - $$f_Y(y) = \sum_x f(x, y)$$
- Continuous Rv
  - $$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy$$
  - $$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)dx$$

مثال

- مثال: اگر
$$f(x,y)=\begin{cases} c(1-x^2-y^2) & 0<x^2+y^2\le 1 \\ 0 & x^2+y^2>1 \end{cases}$$
باشد. مطلوبست:

الف) مقدار ثابت c ؟

ب) $f_X(x)$ ؟

Conditional Probability Function

- $$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

- Baye's Theorem

$$f_{X,Y}(x,y) = f_{X|Y}(x|y) f_Y(y) = f_{Y|X}(y|x) f_X(x) \rightarrow f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)}$$

- $$P(a < X < b | Y = y) = \int_a^b f(x|y)dx$$

- Two (cont./disc) random variables of X and Y are independent iff

$$f(x|y) = f(x) \qquad or \qquad f(x,y) = f(x)f(y)$$

Distribution of a Function of Random Variables

$$Y = g(X_1, X_2, \ldots, X_n)$$

$$f(y) \leftarrow f(x_1, x_2, \ldots, x_n)$$

- CDF-based Method
- Variable Conversion Method
- Moment Generating Function

Variable Conversion Method

- g is a one-to-one transformation: $Y = g(X) \rightarrow X = u(Y)$

  - Discrete RV

$$f_Y(y) = P(Y = y) = P[x = u(y)] = f_X[u(y)]$$

  - Continuous RV

$$f_Y(y) = f_X[u(y)]\left|\frac{d}{dy}u(y)\right| \quad where \quad u(y) = g^{-1}(y)$$

مثال

- $Y = (1 - X)^3$

$$f(x) = \begin{cases} 3(1-x)^2 & 0 < x < 1 \\ 0 & otherwise \end{cases}$$

$$f(y) = ?$$

Variable Conversion Method

- If g is not a one-to-one function, but it could be considered as k mutually exclusive random variables in different bounds such that $Y_1 = g_1(X),....,Y_k = g_k(X)$ produce one-to-one functions, then

$$f_Y(y) = \sum_{i=1}^{k} f_X(w_i(y))\left|J_i\right|$$

where $\quad J_i = w_i'(y) = \frac{d}{dy} w_i(y) \qquad i = 1,...,k$

$$w_i(y) = g_i^{-1}(y)$$

مثال

- X یک متغیر تصادفی $N(\mu,\sigma)$ است. مطلوبست تعیین تابع چگالی احتمال

$$Y=\left(\frac{X-\mu}{\sigma}\right)^2$$

$$Y=\left(\frac{X-\mu}{\sigma}\right)^2=Z^2\rightarrow z=\pm\sqrt{y}\rightarrow\begin{cases}z_1=\sqrt{y}\rightarrow J_1=\dfrac{1}{2\sqrt{y}}\\ z_2=-\sqrt{y}\rightarrow J_2=-\dfrac{1}{2\sqrt{y}}\end{cases}$$

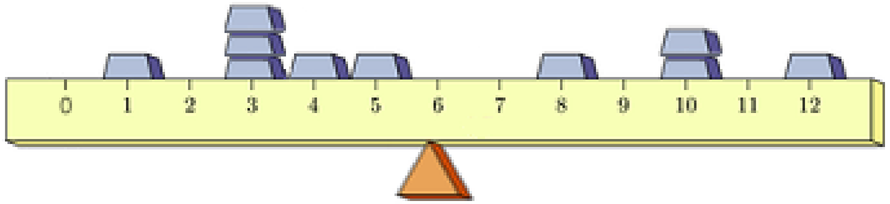
$$f_Z(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$$

$$f_Y(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y}{2}}\left|\frac{1}{2\sqrt{y}}\right|+\frac{1}{\sqrt{2\pi}}e^{-\frac{y}{2}}\left|-\frac{1}{2\sqrt{y}}\right|=\frac{e^{-\frac{y}{2}}}{\sqrt{2\pi y}}=\frac{1}{2^{\frac{1}{2}}\sqrt{\pi}}y^{\frac{1}{2}-1}e^{-y/2}$$

$$\sqrt{\pi} \rightarrow \Gamma(\frac{1}{2})$$

$$=\chi^2_{k=1}$$

Expected Value



- Concept

$$\bar{x} = \sum_{i=1}^{k} f_{p_i} x_i \qquad \textit{where} \qquad f_{p_i} = \frac{f_i}{\sum f_i} = \frac{n_i}{n}$$

- If $x_1, x_2, \ldots, x_n$ are realizations of X with probabilities of $f_X(x_i)$, then

$$E(X) = \mu_X = \sum_i f_X(x_i) x_i$$

probability-weighted average of outcomes

- For a continuous RV

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

مثال

- نقطه ای به تصادف در داخل دایره ای به شعاع ۲ و مرکز صفر انتخاب می شود. اگر فاصله این نقطه تا مرکز با متغیر تصادفی X نشان داده شود و

$$f(x)=\begin{cases}\dfrac{x}{2} & 0<x<2\\ 0 & otherwise\end{cases}$$

مطلوبست تعیین امید ریاضی X.

$$E(X)=\int_{-\infty}^{+\infty} xf(x)dx=\int_{0}^{2} x\frac{x}{2}dx=\frac{4}{3}$$

Expected Value

- Expected value of a function of X

$$Y = g(X)$$

- $E(Y) = \mu_Y = \sum_x g(x) f_X(x)$
- $E(Y) = \mu_Y = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$

مثال

- اگر X دارای توزیع یکنواخت در بازه $(0,1)$ باشد، مطلوبست $E(e^X)$

- $Y = e^X$

$$F_Y(y) = P(Y \le y) = P(e^X \le y) = P(X \le \ln(y)) = F_X(\ln(y))$$

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & else \end{cases} \qquad \Rightarrow F_X(x) = \begin{cases} x & 0 < x < 1 \\ 0 & else \end{cases}$$

$$\Rightarrow F_X(\ln(y)) = \ln(y) \Rightarrow f_Y(y) = F_Y'(y) = \frac{1}{y} \qquad 1 \le y \le e$$

$$E(y) = \int_1^e y \frac{1}{y} dy = e - 1$$

مثال

- اگر X دارای توزیع یکنواخت در بازه $(0,1)$ باشد، مطلوبست $E(e^{X})$

- $Y = g(X) = e^{X}$

$$E(y) = E[g(X)] = \int_0^1 e^{x} dx = e - 1$$

Expected Value

- Expected value can be defined similarly for joint pdf/pmf and conditional pdf/pmf

  - $$E[g(X,Y)] = \sum_{y}\sum_{x} g(x,y) f_{X,Y}(x,y)$$
    $$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y) f_{X,Y}(x,y)dxdy$$

  - $$E[X \mid Y=y] = \sum_{x} xf(x \mid Y=y)$$
    $$= \int_{-\infty}^{+\infty} xf(x \mid Y=y)dx$$

مثال

- اگر متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند:

$$f_{X,Y}(x,y)=\frac{3}{4}(xy+\frac{x^2}{2}) \qquad 0<x<1,\quad 0<y<2$$

الف) امید ریاضی $g_{X,Y}(x,y)=2XY+X+Y$ را بدست آورید.

ب) امید ریاضی $E[X|Y=1]$ را بدست آورید.

- $E[g(X,Y)]=\frac{3}{4}\int_0^1\int_0^2(2xy+x+y)(xy+\frac{x^2}{2})dydx=\frac{23}{6}$
- $E[X|Y=1]=\int_0^1 xf(x|Y=1)dx$

مثال

$$f_{X|Y}(x|Y=1)=\frac{f_{X,Y}(x,y=1)}{f_Y(y=1)}=\frac{f_{X,Y}(x,y)}{\int_0^1 f_{X,Y}(x,y)dx}=\frac{\frac{3}{4}(xy+\frac{x^2}{2})\Big|_{y=1}}{\frac{3}{16}y^2+\frac{1}{4}\Big|_{y=1}=\frac{7}{16}}=\frac{12}{7}(x+\frac{x^2}{2})$$

$$E[X\,|\,Y=1]=\frac{12}{7}\int_0^1 x(x+\frac{x^2}{2})dx=\frac{11}{14}$$

Expected Value Properties

- $E[C] = C$
- E[.] is a linear operator
  - $E[\alpha g(X)] = \alpha E[g(X)]$
  - $E[\alpha g(X) \pm \beta h(X)] = \alpha E[g(X)] \pm \beta E[h(X)]$
  - $E[\alpha g(X,Y) \pm \beta h(X,Y)] = \alpha E[g(X,Y)] \pm \beta E[h(X,Y)]$
- If X and Y are independent, Then $E[XY] = E[X]E[Y]$

$$E[XY] = \iint xyf(x)f(y)dxdy = \int xf(x)dx \int yf(y)dy = E(X)E(Y)$$

Moments

- We know that $E[g(x)] = \sum_x g(x) f_X(x)$

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

- If $g(X) = (X - a)^r$, then $E[g(x)]$ is called rth-order moment of X around a

$$\mu_r(a) = E[(X - a)^r]$$

- If $a = 0$

$$\mu_r' = E[X^r] = \sum_x x^r f_X(x)$$

$$= \int_{-\infty}^{+\infty} x^r f_X(x)$$

Moments

- If $a = \mu$, we have central moments: $\mu_r = E[(X-\mu)^r]$
- Variance: $\mu_2 = \sigma^2 = E[(X-\mu)^2] = E[X^2 + \mu^2 - 2X\mu] = E[X^2] - E^2[X]$
- Variance Properties:
  - $\mathrm{var}(C) = 0$
  - $\mathrm{var}(aX) = a^2\sigma_X^2$
  - $\mathrm{var}(aX+b) = a^2\sigma_X^2$
  - $\mathrm{var}(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 \sigma_{X_i}^2$

Conditional Moment

- $$\operatorname{var}(X \mid Y = y) = E[(X - \mu_{X|Y})^2 \mid Y = y]$$
$$= \sum_x (x - \mu_{X|Y})^2 f_{X|Y}(x \mid y)$$
$$= \int_{-\infty}^{+\infty} (x - \mu_{X|Y})^2 f_{X|Y}(x \mid y)dx$$

where $\mu_{X|Y} = E[X \mid Y = y]$

- $$E[(X - a)^r \mid Y = y] = \sum_x (x - a)^r f_{X|Y}(x \mid y)$$
$$= \int_{-\infty}^{+\infty} (x - a)^r f_{X|Y}(x \mid y)dx$$

مثال

$$f_{X,Y}(x,y)=\begin{cases}1 & 0<|y|<x<1\\ 0 & else\end{cases}$$

$$E[X\mid Y=y]=?,\quad E[Y\mid X=x]=?$$

$$f_X(x)=\int_{-x}^{x} f(x,y)dy=2x \qquad 0<x<1$$

$$f_Y(y)=\int_{|y|}^{1} f(x,y)dx=1-|y| \qquad |y|<1$$

$$f_{X|Y}(x\mid y)=\frac{f(x,y)}{f(y)}=\frac{1}{1-|y|} \qquad 0<|y|<x<1$$

$$f_{Y|X}(y\mid x)=\frac{f(x,y)}{f(x)}=\frac{1}{2x}$$

مثال

$$E[X \mid Y = y] = \int_{|y|}^{1} x f(x \mid y)\,dx = \frac{1 + |y|}{2}$$

$$E[Y \mid X = x] = \int_{-x}^{x} y f(y \mid x)\,dy = 0$$

Covariance

- X and Y are two RVs with defined $f_{X,Y}(x,y)$

$$\sigma_{XY} = \operatorname{cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

- $\sigma_{XY} = \sum_x \sum_y (x-\mu_X)(y-\mu_Y) f_{X,Y}(x,y)$
- $\sigma_{XY} = \int_{-\infty}^{+\infty} (x-\mu_X)(y-\mu_Y) f_{X,Y}(x,y)dxdy$

- $\sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)] = \underbrace{E[XY]}_{\text{Correlation}} - \mu_X\mu_Y = R_{XY} - \mu_X\mu_Y$

Covariance Properties

- $\mathrm{cov}(X, X) = \sigma_X^2$

- $\mathrm{cov}(aX + b, cY + d) = E\{[aX + b - (a\mu_X + b)][cY + d - (c\mu_Y + d)]\}$
$$= E\{ac(X - \mu_X)(Y - \mu_Y)\} = ac\,\mathrm{cov}(X, Y)$$

- $\mathrm{var}(aX + bY + C) = E\{[(aX + bY + C) - (a\mu_X + b\mu_Y + C)]^2\}$
$$= a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY}$$

- If X and Y are independent, then $\mathrm{var}(aX + bY + C) = a^2\sigma_X^2 + b^2\sigma_Y^2$

- $\sigma_{h(X)}^2 = E[(h(X) - \mu_{h(X)})^2]$

Correlation Coefficient

- $\rho = \rho(X, Y) = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}$

- Correlation Coef. Properties:
  - If X and Y are independent, $\rho = 0$
  - $\rho \in [-1, +1]$

- Proof

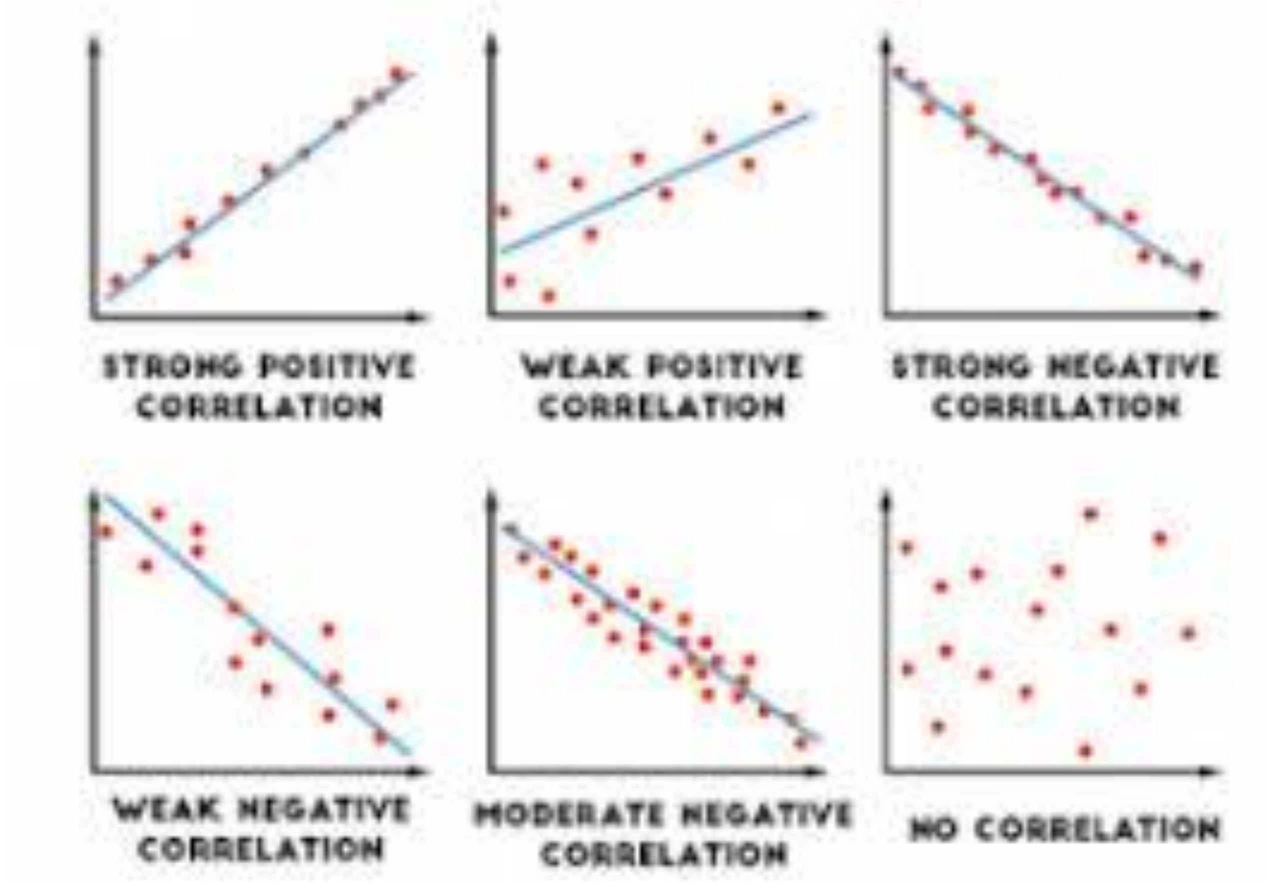
$$Z = Y - aX \rightarrow \mathrm{var}(Z) = \sigma_Y^2 + a^2\sigma_X^2 - 2a\sigma_{XY} = h(a) \geq 0$$

$$\frac{dh(a)}{da} = 2a\sigma_X^2 - 2\sigma_{XY} = 0 \rightarrow a = \frac{\sigma_{XY}}{\sigma_X^2} = \rho\frac{\sigma_Y}{\sigma_X}$$

$$\frac{d^2h(a)}{da^2} = 2\sigma_X^2 > 0 \rightarrow h_{\min} = \min(\mathrm{var}(Z)) = (1 - \rho^2)\sigma_Y^2 \geq 0 \rightarrow |\rho| \leq 1$$

Correlation Coefficient

- $\rho(aX+b, CY+d) = \rho(X,Y)$



مثال

- اگر متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند:

$$f_{X,Y}(x,y) = X + Y \qquad 0 < x < 1,\ 0 < y < 1$$

میانگین، واریانس، کواریانس و ضریب همبستگی X و Y را محاسبه کنید.

$$f_X(x) = \int_0^1 f(x,y)dy = x + \frac{1}{2} \qquad f_Y(y) = \int_0^1 f(x,y)dx = y + \frac{1}{2}$$

$$E(X) = \int_0^1 xf(x)dx = \frac{7}{12} \qquad E(Y) = \int_0^1 yf(y)dy = \frac{7}{12}$$

$$E(X^2) = \int_0^1 x^2 f(x)dx = \frac{5}{12} \qquad E(Y^2) = \int_0^1 y^2 f(y)dy = \frac{5}{12}$$

مثال

$$\sigma_X^2 = E(X^2) - E^2(X) = \frac{11}{144} \qquad \sigma_Y^2 = E(Y^2) - E^2(Y) = \frac{11}{144}$$

$$E[XY] = \int_0^1 \int_0^1 xyf(x,y)dxdy = \frac{1}{3} \to \sigma_{XY} = E[XY] - E[X]E[Y] = -\frac{1}{144}$$

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = -\frac{1}{11}$$

Some Definitions

- Two RVs of X and Y are called *Uncorrelated* if $\rho(X,Y)=0 \;\; or \;\; \sigma_{XY}=0$

- Two RVs of X and Y are called *Orthogonal* if $R_{XY}=E[XY]=0$

- Skewness

$$\alpha_3=\frac{\mu_3}{\sigma^3}=\frac{E[(X-\mu)^3]}{[\sigma^2]^{\frac{3}{2}}}$$

- Curtosis

$$\alpha_4=\frac{\mu_4}{\sigma^4}-3=\frac{E[(X-\mu)^4]}{[\sigma^2]^2}-3$$

Random Vectors Statistics

- Correlation between two **random vectors** of **X** and **Y**

$$
\begin{aligned}
R_{XY} &= E(XY^T) \\
&= \begin{bmatrix} E(X_1Y_1) & \cdots & E(X_1Y_m) \\ \vdots & & \vdots \\ E(X_nY_1) & \cdots & E(X_nY_m) \end{bmatrix}
\end{aligned}
$$

- Cross Covariance

$$
\begin{aligned}
C_{XY} &= E[(X-\bar{X})(Y-\bar{Y})^T] \\
&= E(XY^T) - \bar{X}\bar{Y}^T
\end{aligned}
$$

- Autocorrelation of the n-element RV **X**

$$
\begin{aligned}
R_X &= E[XX^T] \\
&= \begin{bmatrix} E[X_1^2] & \cdots & E[X_1X_n] \\ \vdots & & \vdots \\ E[X_nX_1] & \cdots & E[X_n^2] \end{bmatrix}
\end{aligned}
$$

Random Vectors Statistics

- Note
  - $R_X = R_X^T$
  - $z^T R_X z = z^T E[XX^T]z = E[z^T X X^T z] = E[(z^T X)^2] > 0$
- Autocovariance of **X**

$$
\begin{aligned}
C_X &= E[(X-\bar{X})(X-\bar{X})^T] \\
&= \begin{bmatrix} E[(X_1-\bar{X}_1)^2] & \cdots & E[(X_1-\bar{X}_1)(X_n-\bar{X}_n)] \\ \vdots & & \vdots \\ E[(X_n-\bar{X}_n)(X_1-\bar{X}_1)] & \cdots & E[(X_n-\bar{X}_n)^2] \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1n} \\ \vdots & & \vdots \\ \sigma_{n1} & \cdots & \sigma_n^2 \end{bmatrix}
\end{aligned}
$$

- $C_X = C_X^T \geq 0$

مثال

- تابع زیر را درنظر بگیرید:

$$f_{X,Y}(x,y)=\begin{cases} ae^{-2x-3y} & x>0, y>0 \\ 0 & otherwise \end{cases}$$

مطلوبست

الف) محاسبه واریانس σ_x^2 و σ_y^2 و کوواریانس C_{XY}

ب) محاسبه ماتریس کوواریانس بردار تصادفی $[X \quad Y]^T$

مثال

- $$\int_0^{\infty}\int_0^{\infty} f(x,y)dxdy = 1 \rightarrow a = 6$$

$$\sigma_X^2 = E(X^2) - \overline{x}^2 = \frac{1}{4}$$

$$\sigma_Y^2 = E(Y^2) - \overline{y}^2 = \frac{1}{9}$$

$$C_{XY} = E[XY] - \overline{xy} = 0$$

- $$C = \begin{bmatrix} E((X-\overline{x})^2) & E((X-\overline{x})(Y-\overline{y})) \\ E((X-\overline{x})(Y-\overline{y})) & E((Y-\overline{y})^2) \end{bmatrix} = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/9 \end{bmatrix}$$

Random Vectors Statistics

- Gaussian Random vector

$$\mathrm{pdf}(X) = \frac{1}{(2\pi)^{n/2}|C_X|^{1/2}} \exp\left[\frac{-1}{2}(X-\bar{X})^T C_X^{-1}(X-\bar{X})\right]$$

- Consider a Gaussian RV that undergoes a linear transformation:

$$\begin{aligned} Y &= AX + b \\ f_Y(y) &= |A^{-1}|\frac{1}{(2\pi)^{n/2}|C_X|^{1/2}} \exp\left[\frac{-1}{2}(A^{-1}y - A^{-1}b - \bar{x})^T C_X^{-1}(\cdots)\right] \\ &= \frac{1}{(2\pi)^{n/2}|AC_XA^T|^{1/2}} \exp\left[\frac{-1}{2}(y-\bar{y})^T (AC_XA^T)^{-1}(y-\bar{y})\right] \end{aligned}$$

$$\Rightarrow \quad y \sim N(A\bar{x}+b, AC_XA^T)$$

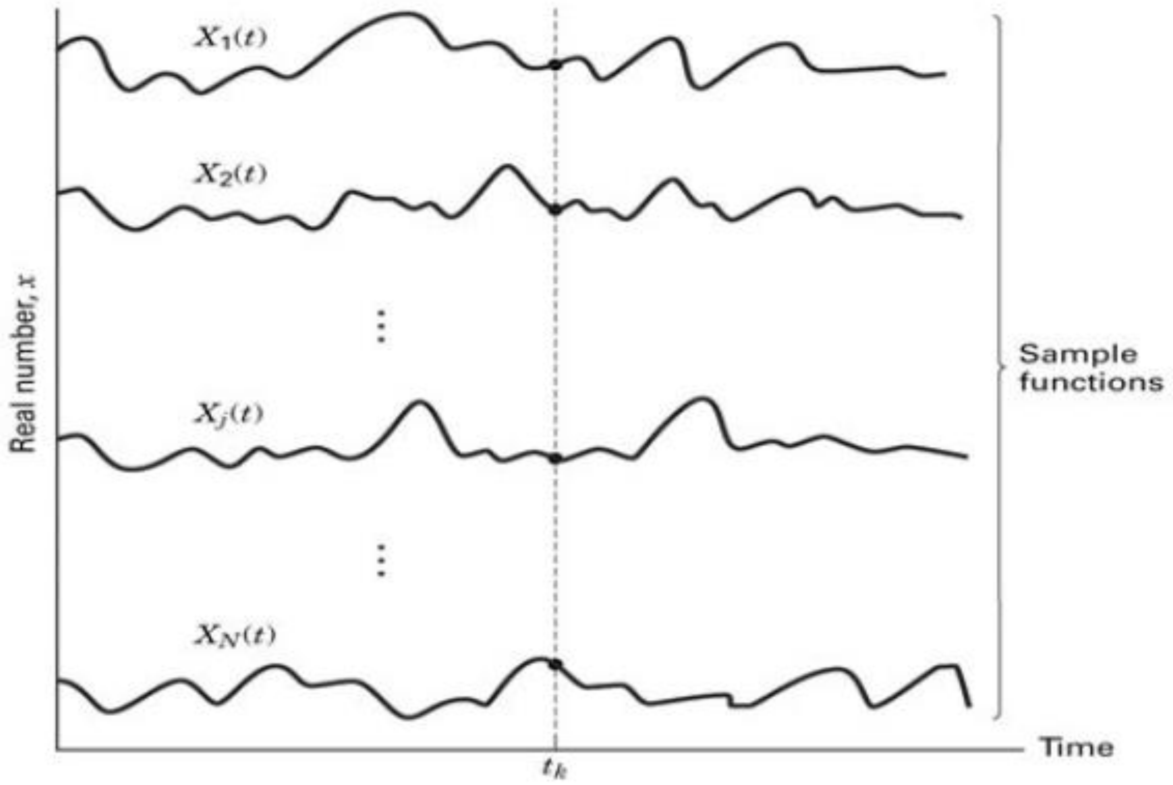
Normality is preserved in linear transformations.

Stochastic Process

- A stochastic process **X** (t) is a random vector **X** that changes with time/space:

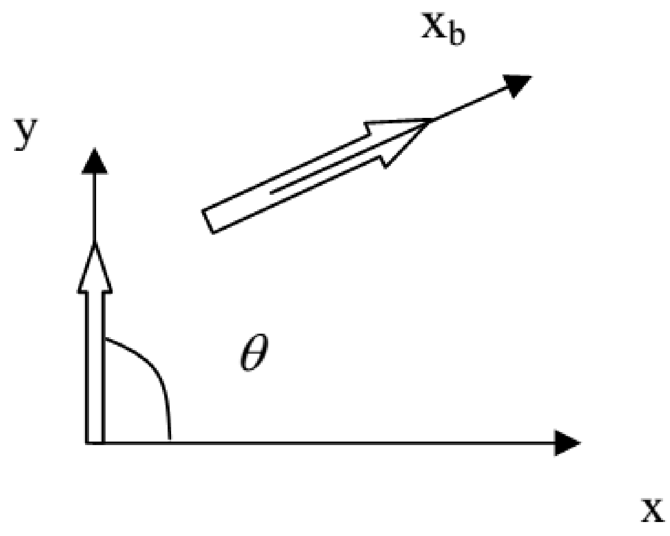
$$\mathbf{X}(t); \quad t_0 \le t \le t_f$$

- Time
  - ❑ Continuous
  - ❑ Discrete

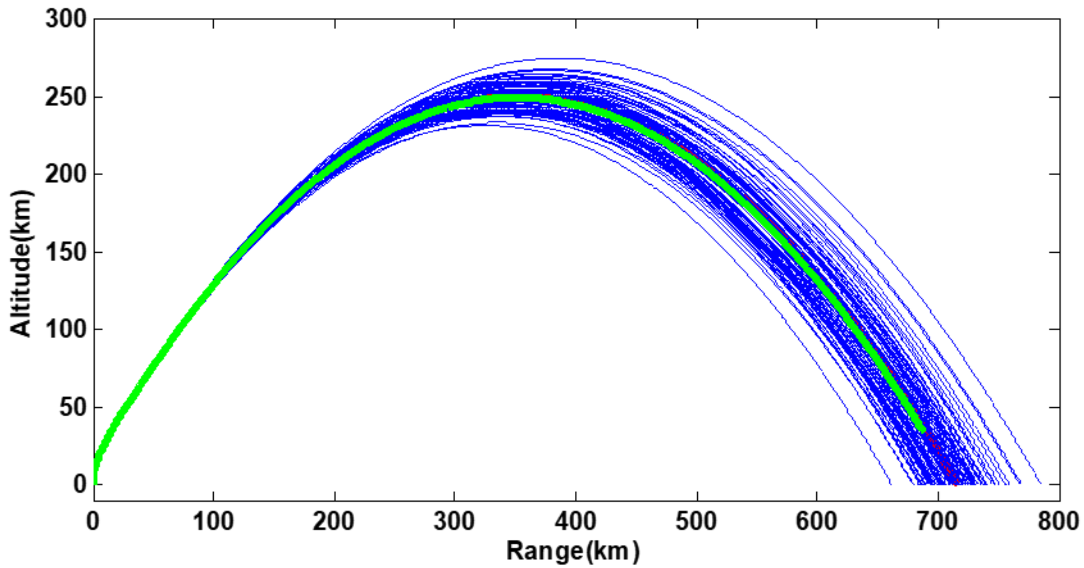


Stochastic Process

- Example



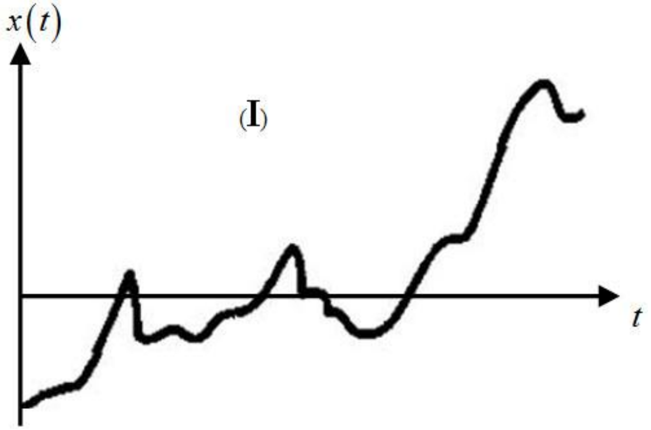
$$
Pitch\ Command\ \ \theta(t)^{\circ} = \begin{cases} 90 & t \le 5 \\ 45\left(1 + e^{\left(-\frac{10(t-5)^2}{t_b^2}\right)}\right) & 5 < t \le t_b \\ \theta(t_b) & t \ge t_b \end{cases}
$$



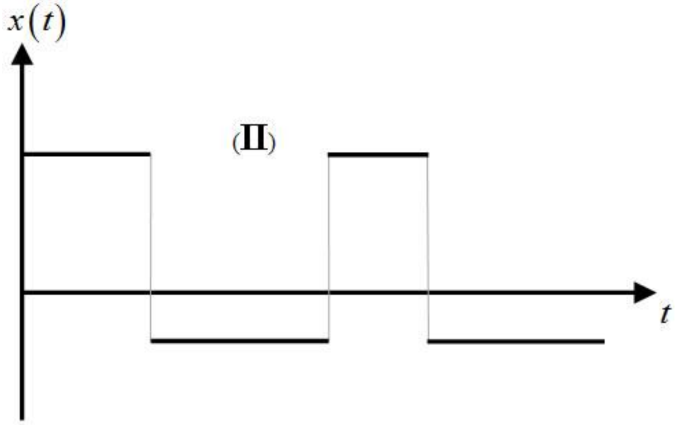
ضریب پراکندگی (δ)	نماد	میانگین (μ)	واحد	پارامتر
0.01	m_0	6000	kg	جرم اولیه
0.01	m_{p_0}	5000	kg	جرم سوخت اولیه
0.01	I_{sp}	240	sec	ایمپالس ویژه
0.01	ρ	1.25	kg/m^3	دانسیته هوا
0.01	g	9.81	m/s^2	شتاب جاذبه
0.01	S	0.1	m^2	مساحت مبناء
0.01	C_{D_α}	0.01	Rad^{-1}	شیب ضریب درگ
0.01	t_b	60	sec	زمان سوزش موتور

Stochastic Process

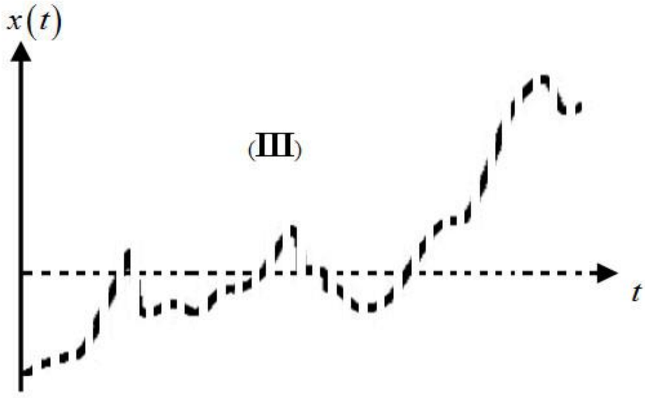
- A stochastic process can be one of four types:



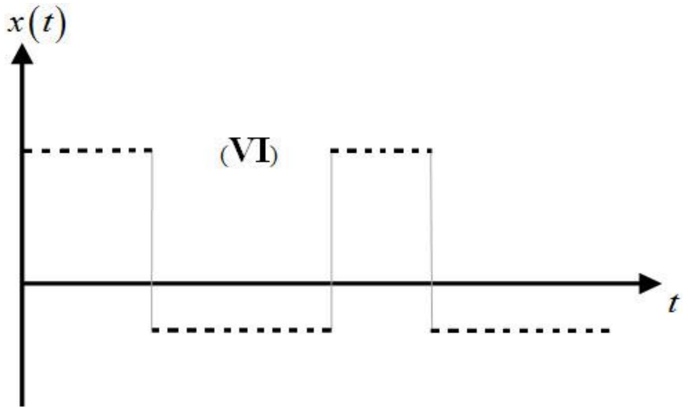
Continuous Random Process



Discrete Random Process



Continuous Random Sequence



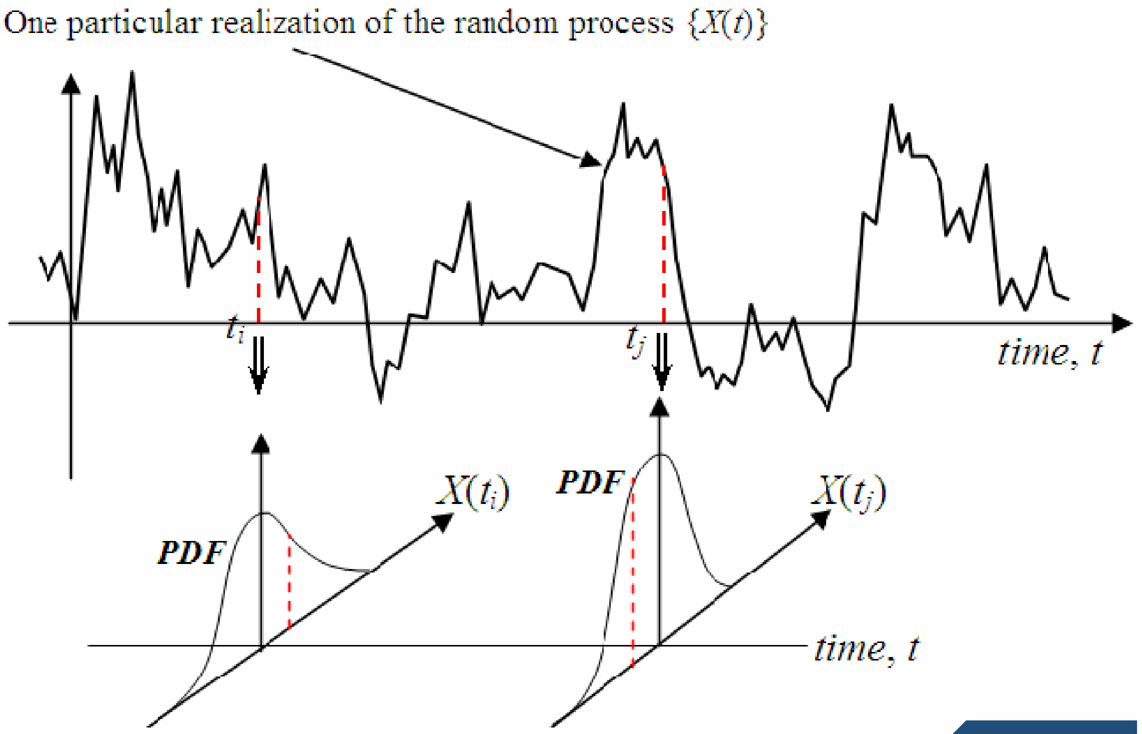
Discrete Random Sequence (Chain)

Stochastic Process

- CDF/pdf of a stochastic process **X** (t)

$$F_X(x,t) = P\left[X_1(t) \le x_1 \text{ and } \cdots X_n(t) \le x_n(t)\right]$$

$$f_X(x,t) = \frac{d^n F_X(x,t)}{dx_1 \cdots dx_n}$$



Stochastic Process

- Statistical properties of a random process
  - Mean: $$\bar{x}(t) = \int_{-\infty}^{\infty} x f(x,t)\, dx$$
  - Covariance: $$\begin{aligned} C_X(t) &= E\left\{[X(t) - \bar{x}(t)]\,[X(t) - \bar{x}(t)]^T\right\} \\ &= \int_{-\infty}^{\infty} [x - \bar{x}(t)]\,[x - \bar{x}(t)]^T f(x,t)\, dx \end{aligned}$$
- Second order CDF/pdf

$$\begin{aligned} F(x_1, x_2, t_1, t_2) &= P(X(t_1) \le x_1, X(t_2) \le x_2) \\ f(x_1, x_2, t_1, t_2) &= \frac{\partial^2 F(x_1, x_2, t_1, t_2)}{\partial x_1 \partial x_2} \end{aligned}$$

Stochastic Process

- Autocorrelation of the stochastic process

$$R_X(t_1, t_2) = E\left[X(t_1)X^T(t_2)\right]$$

- Autocovariance of a stochastic process

$$C_X(t_1, t_2) = E\left\{\left[X(t_1) - \bar{X}(t_1)\right]\left[X(t_2) - \bar{X}(t_2)\right]^T\right\}$$

- Strict-sense stationary vs. Wide-sense stationary

$$\begin{aligned} E[X(t)] &= \bar{x} \\ E[X(t_1)X^T(t_2)] &= R_X(t_2 - t_1) \end{aligned}$$

Stochastic Process

- For a WSS process

$$
\begin{aligned}
R_X(0) &= E[X(t)X^T(t)] \\
R_X(-\tau) &= R_X(\tau)
\end{aligned}
$$

- For a scalar process: $|R_X(\tau)| \leq R_X(0)$

- Cross correlation of **X** (t) and **Y** (t)

$$R_{XY}(t_1, t_2) = E[X(t_1)Y^T(t_2)]$$

- Cross covariance of **X**(t) and **Y**(t)

$$C_{XY}(t_1, t_2) = E\left\{[X(t_1) - \bar{X}(t_1)][Y(t_2) - \bar{Y}(t_2)]^T\right\}$$

Stochastic Process

- Time average & time autocorrelation of **X** (t)

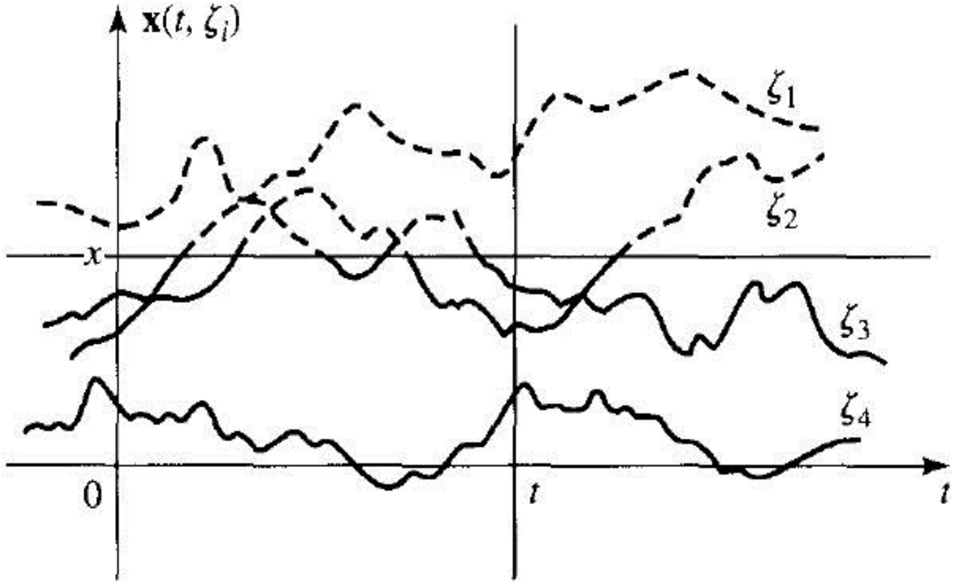
$$A[X(t)] = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, dt$$

$$R[X(t),\tau] = A[X(t)X^T(t+\tau)] = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)x(t+\tau)dt$$

- An ergodic process is a stationary random process for which

$$A[X(t)] = E(X)$$

$$R[X(t),\tau] = R_X(\tau)$$



مثال

- فرض کنید که X یک متغیر تصادفی و $Y(t) = X cos(t)$ یک فرایند تصادفی باشد.

الف) امید ریاضی Y(t) را بیابید.

ب) میانگین زمانی Y(t) را محاسبه کنید.

ج) تحت چه شرایطی $\overline{y}(t) = A[Y(t)]$

$$\overline{y}(t) = E[Y(t)] = E[X \cos t] = \overline{x} \cos t$$

$$A[Y(t)] = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} y(t)dt = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x \cos t dt = 0$$

$$\overline{y} = A[Y(t)] \to \overline{x} = 0$$

Moarkov process

- A stochastic process that satisfies Markov property; i.e. conditional pdf of future states of the process depends only upon the present state, not on the sequence of events that preceded it (memorylessness).

$$f_X(x(t), x(\tau)) = f_X(x(t) \mid x(\tau)) f_X(x(\tau)) \quad for \ \ all \ \ t, \tau \ \ in \ \ (t_0, t_f)$$

$$f_X(x_k \mid x_1, x_2, \ldots, x_{k-1}) = f_X(x_k \mid x_{k-1})$$

$$f_X(x_1, x_2, \ldots, x_{k-1}, x_k) = f_X(x_k \mid x_{k-1}) f_X(x_{k-1} \mid x_{k-2}) \ldots f_X(x_2 \mid x_1) f_X(x_1)$$

- Gauss-Markov random process: a Markov process in which both pdfs of $f_X(x(t) \mid x(\tau))$ and $f_X(x(\tau))$ are Gaussian.

Moarkov process

- Generalization of Markov models

$$x_1(k+1) = c_1(k)x_1(k) + c_2(k)x_1(k-1) + w(k)$$

$$X(k) = \begin{Bmatrix} x_1(k) \\ x_2(k) \end{Bmatrix}; \quad \text{where: } x_2(k+1) = x_1(k) \ or \ x_2(k) = x_1(k-1)$$

$$\{X(k+1)\} = \begin{Bmatrix} x_1(k+1) \\ x_2(k+1) \end{Bmatrix} = \begin{bmatrix} c_1(k) & \mathrm{c}_2(k) \\ 1 & 0 \end{bmatrix} \begin{Bmatrix} x_1(k) \\ x_2(k) \end{Bmatrix} + \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} w(k)$$

White Noise and Colored Noise

- If the RV $\mathbf{X}(t_1)$ is independent from the RV $\mathbf{X}(t_2)$ for all $t_1 \neq t_2$, then $\mathbf{X}(t)$ is called white noise. Otherwise, $\mathbf{X}(t)$ is called colored noise.

- Power spectral density of a WSS random process

$$\begin{aligned} S_X(\omega) &= \int_{-\infty}^{\infty} R_X(\tau) e^{-j\omega\tau}\, d\tau \\ R_X(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(\omega) e^{j\omega\tau}\, d\omega \end{aligned}$$

Wiener-Khintchine relations

- Power of a WSS random process

$$P_X = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(\omega)\, d\omega$$

مثال

- تابع خودهمبستگی یک فرایند تصادفی عبارت است از $R_X(\tau) = Ae^{-k|\tau|}$ که در آن A و k ثوابت مثبت هستند. مطلوبست

الف) محاسبه PSD و توان فرایند تصادفی

ب) به ازای چه مقداری از k نیمی از توان در فرکانس های کمتر از 1 Hz واقع می شود.

مثال

$$a)S_X(\omega)=\int_{-\infty}^{\infty} R_X(\tau)e^{-j\omega\tau}d\tau=\int_{-\infty}^{0} Ae^{k\tau}e^{-j\omega\tau}d\tau+\int_{0}^{\infty} Ae^{-k\tau}e^{-j\omega\tau}d\tau$$

$$=\frac{2Ak}{k^2+\omega^2}$$

$$b)P_X=\frac{1}{2\pi}\int_{-\infty}^{\infty} S_X(\omega)d\omega=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{2Ak}{k^2+\omega^2}d\omega=\frac{A}{\pi}\tan^{-1}(\frac{\omega}{k})\Big|_{-\infty}^{\infty}=A$$

$$c)P_1=\frac{A}{\pi}\tan^{-1}(\frac{\omega}{k})\Big|_{-2\pi}^{2\pi}=\frac{2A}{\pi}\tan^{-1}(\frac{2\pi}{k})=\frac{A}{2}\rightarrow\frac{2\pi}{k}=\tan(\frac{\pi}{4})\rightarrow k=2\pi$$

$$1Hz=2\pi\ rad/s$$

Simulating Correlated Noise

- Suppose we want to generate an n-element random vector $\boldsymbol{\omega}$ that has zero mean and covariance $\boldsymbol{Q}$

$$Q = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1n} \\ \vdots & & \vdots \\ \sigma_{1n} & \cdots & \sigma_n^2 \end{bmatrix}$$

- All eigenvalues are real and non-negative.

$$\lambda(Q) = \mu_k^2 \quad (k = 1, \ldots, n)$$

- Jordan form decomposition of $\boldsymbol{Q}$

$$Q = D\hat{Q}D^T$$

Orthogonal matrix

$$\hat{Q} = \mathrm{diag}(\mu_1^2, \cdots, \mu_n^2)$$

Simulating Correlated Noise

- Now we define the random vector $\boldsymbol{\nu}$ as $\boldsymbol{\nu} = \mathbf{D}^{-1}\boldsymbol{\omega}$, so that $\boldsymbol{\omega} = \mathbf{D}\boldsymbol{\nu}$, therefore

$$
\begin{aligned}
E(vv^T) &= E(D^T w w^T D) \\
&= D^T Q D \\
&= \hat{Q} \\
&= \mathrm{diag}(\mu_1^2, \cdots, \mu_n^2)
\end{aligned}
$$

- $\nu_i = \mu_i r_i$, where $r_i \sim (0,1)$ independent random number

Bayesian Vs. Non-Bayesian Inference

- The estimation procedure can follow one of two models:
  - The first model assumes that the parameters to be estimated are nonrandom and constant during the observation window but the observations are noisy and thus have random components.
  - The second model assumes that the parameters are random variables that have a prior probability and the observations are noisy as well.
- When the first model is used for state/parameter estimation, the procedure is called non-Baysian or Fisher estimation. State/parameter estimation using the second model is called Bayesian estimation.

Some basic definitions

- **Consistency**: An estimator is consistent if the estimator converges to the true value almost surely as the number of observations approaches infinity.

- **Unbiasedness**: An estimator is unbiased if its expected value is equal to the true value.

- **Efficiency**: An estimator is efficient if it produces the smallest error covariance matrix among all unbiased estimators, it is also regarded optimally using the information in the measurements. A well-known efficiency criterion is the Cramer-Rao bound.

Some basic definitions

- **Robustness**: An estimator is robust if it is insensitive to the gross measurement errors and the uncertainties of the model.
- **Minimal variance**: Variance reduction is the central issue of various Monte Carlo approximation methods, most improvement techniques are variance-reduction oriented.

Observability

- How well we can observe a system (determine the initial conditions after measuring the outputs).

- A *continuous-time system* is observable if for any initial state $X(0)$ and any final time $t > 0$, the initial state $X(0)$ can be uniquely determined by knowledge of the input $u(\tau)$ and output $z(\tau)$ for all $\tau \in [0, t]$.

- A *discrete-time system* is observable if for any initial state X_0 and some final time k, the initial state X_0 can be uniquely determined by knowledge of the input u_i and output z_i for all $i \in [0, k]$.

Observability

Continuous LTI System $\dot{x} = Ax + Bu$ $z = Cx$	**Discrete LTI System** $x_k = Fx_{k-1} + Gu_{k-1}$ $z_k = Hx_k$
$rank\left(\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}\right) = n$	$rank\left(\begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix}\right) = n$
Observability gramian is PD for some $t \in (0, \infty)$ $W_o(t) = \int_0^t e^{A^T\tau} C^T C e^{A\tau} d\tau$	Observability gramian is PD for some $k \in (0, \infty)$ $W_o(k) = \sum_{i=0}^{k} (F^T)^i H^T H\, F^i$
If A is stable, the unique solution of the following Lyapunov equation is PD. $A^T W_o(\infty) + W_o(\infty) A = -C^T C$	If A is stable, the unique solution of the following Lyapunov equation is PD. $W_o = F^T W_o F + H^T H$

Observability

- Popov-Belevitch-Haustus (PBH) tests:
  - For every eigenvalue λ_i of matrix A, the $(n+m)\times n$ matrix $\begin{pmatrix}\lambda I - A \\ C\end{pmatrix}$ has rank n.
  - (A,C) is un-observable IFF the eigenvector v_i of matrix A causes $Cv_i = 0$
- Condition number of an observability matrix denotes degree of observability.
- Observability concerns the sensor package (variety, location, etc.) , while the controllability concept returns to the actuators (arrangement, variety, etc.).

Example

- $$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$
  $$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$rank\left(\begin{bmatrix} C \\ CA \end{bmatrix}\right) = rank\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right) = 2$$

- $$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 4 & 5 & 1 \end{bmatrix}$$

$$rank\left(\begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix}\right) = rank\left(\begin{bmatrix} 4 & -6 & 6 \\ 5 & -7 & 5 \\ 1 & -1 & -1 \end{bmatrix}\right) = 2$$

Controllability

- How well we can control a system (drive the state to a desired value).
- Controllability of a linear system
  - A *continuous-time system* is controllable if for any initial state X(0) and any final time $t > 0$ there exists a control that transfers the state to any desired value at time t.
  - A *discrete-time system* is controllable if for any initial state X0 and some final time *k* there exists a control that transfers the state to any desired value at time *k*.
  - Controllability is independent of the output equation.

Controllability

Continuous LTI System $\dot{x} = Ax + Bu$	Discrete LTI System $x_k = Fx_{k-1} + Gu_{k-1}$
$rank([B \quad AB \quad \cdots \quad A^{n-1}B]) = n$	$rank([G \quad FG \quad \cdots \quad F^{n-1}G]) = n$
Controllability gramian is PD for any $t > 0$ $$W_c(t) = \int_0^t e^{A\tau}BB^Te^{A^T\tau}d\tau$$	Controllability gramian is PD for some $k \in (0, \infty)$ $$W_c(k) = \sum_{i=0}^{k} F^iGG^T(F^T)^i$$
If A is stable, the unique solution of the following Lyapunov equation is PD. $$AW_c(\infty) + W_c(\infty)A^T = -BB^T$$	If A is stable, the unique solution of the following Lyapunov equation is PD. $$W_c = FW_cF^T + GG^T$$

Example

- $$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

$$[\mathbf{B} \mid \mathbf{AB}] = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \text{singular}$$

- $$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} [u]$$

$$[\mathbf{B} \mid \mathbf{AB}] = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = \text{nonsingular}$$

Stabilizability & Detectability

- If a system is controllable or stable, then it is also stabilizable. If a system is uncontrollable or unstable, then it is stabilizable if its uncontrollable modes are stable.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

- If a system is observable, then it is also detectable. If a system is unobservable, then it is detectable if its unobservable modes are stable.

- Jordan-canonical forms can be utilized to these aims.

Controllability & Observability in Jordan form

- A system is completely controllable IFF 1) for every distinct eigenvalue, there is just one Jordan block. 2) the elements of the last row of each Jordan block are not all zero. 3) the elements of each row of transformed B that correspond to distinct eigenvalues are not all zero.

- A system is completely observable IFF 1) for every distinct eigenvalue, there is just one Jordan block. 2) no columns of the transformed matrix C that correspond to the first row of each Jordan block consist of zero elements. 3) no columns of the transformed matrix C that correspond distinct eigenvalues consist of zero elements.

Example

- Completely observable

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad y = \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \boxed{\begin{matrix} 3 \\ 4 \end{matrix}} & \begin{matrix} 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \end{matrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \left[\begin{array}{ccc:cc} 2 & 1 & 0 & & 0 \\ 0 & 2 & 1 & & \\ 0 & 0 & 2 & & \\ \hdashline & & & -3 & 1 \\ 0 & & & 0 & -3 \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}, \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \boxed{\begin{matrix} 1 \\ 0 \end{matrix}} & \begin{matrix} 1 \\ 1 \end{matrix} & \begin{matrix} 1 \\ 1 \end{matrix} & \boxed{\begin{matrix} 0 \\ 1 \end{matrix}} & \begin{matrix} 0 \\ 0 \end{matrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

Example

- Unobservable

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad y = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \boxed{\begin{matrix} 0 \\ 0 \end{matrix}} & \begin{matrix} 1 \\ 2 \end{matrix} & \begin{matrix} 3 \\ 4 \end{matrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \left[\begin{array}{ccc|cc} 2 & 1 & 0 & & 0 \\ 0 & 2 & 1 & & \\ 0 & 0 & 2 & & \\ \hline & & & -3 & 1 \\ 0 & & & 0 & -3 \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}, \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \boxed{\begin{matrix} 1 \\ 0 \end{matrix}} & \begin{matrix} 1 \\ 1 \end{matrix} & \begin{matrix} 1 \\ 1 \end{matrix} & \boxed{\begin{matrix} 0 \\ 0 \end{matrix}} & \begin{matrix} 0 \\ 0 \end{matrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

Example

- Completely controllable

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} u$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix} u$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \left[\begin{array}{ccc|cc} -2 & 1 & 0 & & 0 \\ 0 & -2 & 1 & & \\ 0 & 0 & -2 & & \\ \hline & & & -5 & 1 \\ 0 & & & 0 & -5 \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 3 & 0 \\ 0 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Example

- Uncontrollable

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix} u$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 4 & 2 \\ 0 & 0 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \left[\begin{array}{ccc:cc} -2 & 1 & 0 & & 0 \\ 0 & -2 & 1 & & \\ 0 & 0 & -2 & & \\ \hdashline & & & -5 & 1 \\ 0 & & & 0 & -5 \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \\ 1 \\ 3 \\ 0 \end{bmatrix} u$$

Principle of Duality

$E: \dot{X} = AX(t) + Bu(t)$

$Z(t) = CX(t)$

$E': \dot{X}' = A^{*}X'(t) + C^{*}u'(t)$

$Z'(t) = B^{*}X'(t)$

- System E is completely controllable, IFF its dual is completely observable.
- System E is completely observable, IFF its dual is completely controllable.
- System E is stabilizable, IFF its dual is detectable.
- System E is detectable, IFF its dual is stabilizable.