MA201 / CS609 / MA8XX

Michael Baron

$> \text{rand}(\ )$

25% -

3-1-0-:4

35% - mid

45% - end

25% - quizzes

Prof
V P Sinha

Real-World

- Events occur.
  e.g., rain, temp^r,
  Stock market ↑↓
- We "perceive" events.
- Inductive.

Measurement

Engineering World

Numbers

$5 = 3 + 2$

$1+1+1 \quad 1+1$

- Number System
- Deductive.

$S = \{HHT, \quad , \cancel{HHT}$

Common Sense

Include ALL, once.

FFT

Euclid

No redundancy, no exclusion.

Q: Why probability is SO important?

$V = IR$

Catch

1 v.g. Charac.

- Stock Market fluctuations
- Temp. Variations over a period of time
- Ages of children in class-I of a School

Random (non-deterministic)

Catch the avg. char.:

$$\boxed{P(A) = \frac{\# N_A}{\# N}}$$

$N \to \infty$

Q: Can we write computer algorithms for large scale database?

Axiomatic definition of Probability:

$[0,1]$ 1. $p(A) \geq 0$

2. $p(S) = 1$

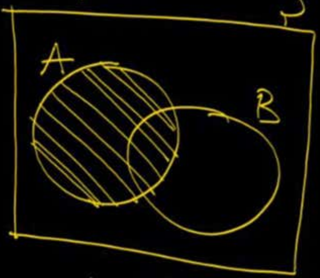
$-$ $\Rightarrow$ 3. $p(A+B) = p(A) + p(B) - p(AB)$

Venn diagram

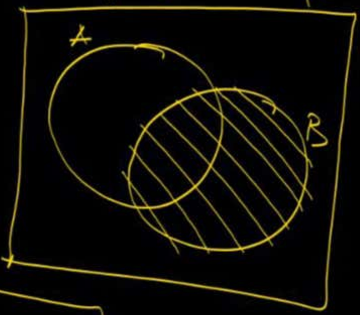
Set theory

To prove $p(A+B) = p(A) + p(B) - p(AB)$ Using Venn diag.

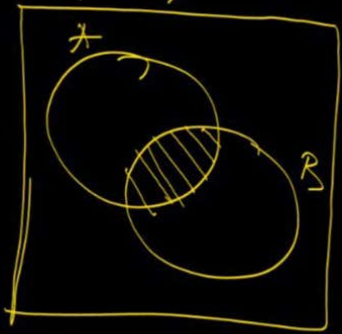
$p(A)$



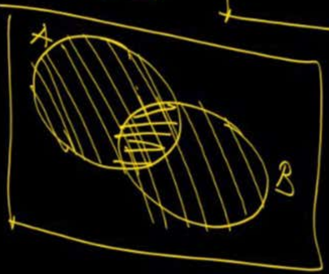
$p(B)$



$p(AB)$



$p(A+B)$



Mutually exclusive events

$$p(AB) = \{\phi\}$$

if M.E. events,

then $p(A+B) = p(A) + p(B)$

Homework

Consider two events, A and B.

s.t. $B \subset A$

$\therefore p(A) \geq p(B)$

prove Venn diag.

$A = B + \bar{B}A$

Recap

1. Real world / Eqq. world
2. prob. Space (S, A, p)
3. $p = \frac{N_A}{N}$ $[0,1]$

4. Axiomatic def
   - $p(A) \geq 0$
   - $p(S) = 1$
   - $p(A+B) = p(A) + p(B) - p(AB)$

5. M.E.

$$\boxed{p(A) = \frac{N_A}{N}}$$

TODAY

1. Conditional probability
2. Total probability
3. Bayes' theorem

Conditional probability of A

$$\boxed{p(A|M) = \frac{p(AM)}{p(M)}}$$

Joint prob.

given

occured

$A \subset M, \quad p(A|M) = \dfrac{p(A)}{p(M)}$

$M \subset A, \quad p(A|M) = 1$

if A & B are M.E.

$p((A+B)|M)$

$= p(A|M) + p(B|M)$

Example: fair-die expt. | Conditioning reduces Variance

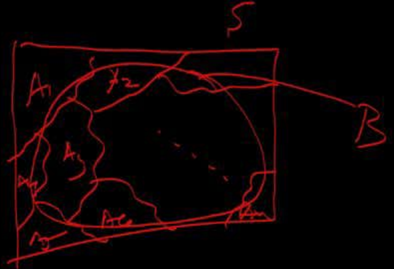
$S = \{1, 2, 3, 4, 5, 6\}$

$\boxed{GAN}$

- $A = \{f_2\} \Rightarrow p(A) = \frac{1}{6}$

- $M = \{f_2, f_4, f_6\} \Rightarrow p(A/M) = \frac{p(AM)}{p(M)} = \frac{1/6}{1/2} = 1/3$

Total probability



- $S = \{A_1, A_2, A_3, \ldots, A_n\}$ s.t. $A_i A_j = \{\phi\}, \forall i, j$

- $B = \cancel{A} \cdot S = B(A_1 + A_2 + A_3 \ldots + A_n) = BA_1 + BA_2 + \ldots + BA_n$

$p(B) = \underline{p(BA_1)} + p(BA_2) + \ldots$

Bayes' Theorem

~~Belief~~

Belief System

$$AB = BA$$

$$p(AB) = p(BA)$$

$$p(A|B)\,p(B) = p(B|A)\,p(A)$$

likelihood prior

$$\boxed{p(A|B) = \frac{p(B|A)\,p(A)}{p(B)}}$$

posterior evidence

Ex:

	defective	production
M-1	1%	40%
M-2	3%	20%
M-3	2%	40%

$$p(M_i|D) = \frac{p(D|M_k)\,p(M_k)}{\sum_{i=1}^{3} p(D|M_i)\,p(M_i)}$$

Sol

$p(D_1) = 0.01$

$p(D_2) = 0.03$

$p(D_3) = 0.02$

$$\boxed{p(M_3|D) = 4/9}$$

$p(M_1|D) = 2/9$

$p(M_2|D) = 1/3$

Recap

$$p(A) = \frac{N_A}{N}$$

frequentist

Conditional

$$p(A|M) = \frac{p(AM)}{p(M)}$$

Bayesian

Belief

prior

$$p(A|B) = \frac{p(B|A)\,p(A)}{p(B)}$$

TODAY

~~7~~

- Independent events.
- Random Variables.
  - → PDF
  - → CDF

M.E. $\Rightarrow$ $p(AB) = 0$

A B

Independent events

$$\boxed{p(AB) = p(A) \cdot p(B)}$$

$$p(AB) = p(A|B) \cdot p(B)$$

$$\downarrow$$

$$p(A)$$

A & B are DEPENDENT

$$p(A|B) = \frac{p(AB)}{p(B)}$$

$A = \{f_2\}$

$M = \{f_2, f_4, f_6\}$

$C =$

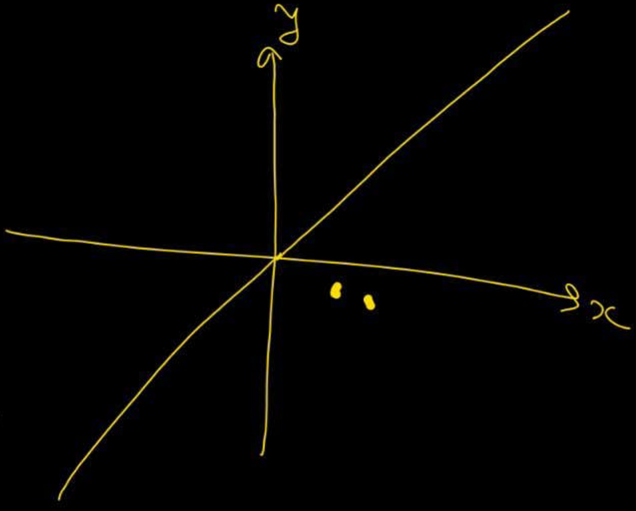
Random Variables

Function

$$f : X \xrightarrow{\text{maps}} Y$$

Domain

Co-domain
(range)



e.g., Toss two coins

<u>Random expt.</u>

outcome

S

Sample Space

HH
HT
TH TT

$\mathcal{A}$ = power set

Unions

Intersections

Compliments

Null

0

1

$P = [0,1]$

$(S, \mathcal{A}, P)$

Probability Space

$X = (-\infty, +\infty)$

Random Variable

$\mathbb{R}$

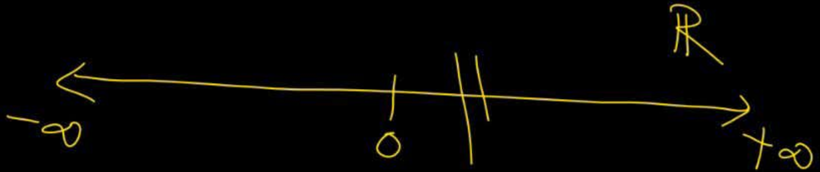
<u>No Subtraction</u>

Random Variable is a <u>real-valued function</u>.

$$\boxed{X : S \longrightarrow \mathbb{R}}$$

With conditions

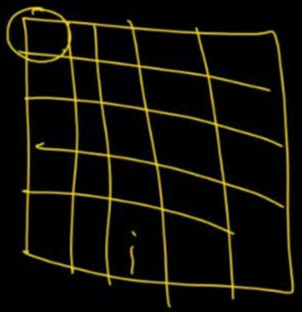
1) $\{X \le x\}$ is an event $-\infty < x < \infty$



2) $P(X = \pm\infty) = 0$.

Ex: throwing a fair die

$i = 1, 2, 3, \ldots, 6$



$$X(f_i) = 10i$$

$X(f_1) = 10 \longrightarrow 1/6$

$X(f_2) = 20 \longrightarrow 1/6$

$\vdots$

Full-view

Ex: A fair coin is tossed three times and faces shown up are observed.

(a) Write Sample Space.

(b) If X is a random variable representing number of heads in each outcome. Draw PDF/CDF.

Solution:

$S = \{HHH, HHT, HTH, THH, THT, TTH, TTT\}$

HTT

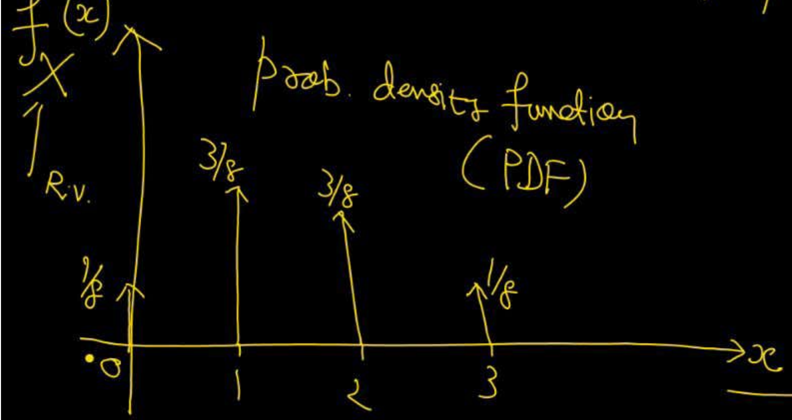
R.V. $X = \{3, 2, 2, 2, 1, 1, 0, 1\}$

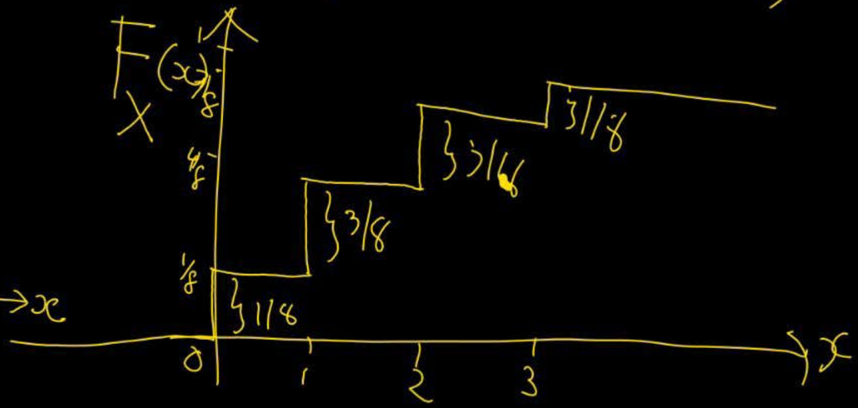
Cumulative distribution fun^{n}

(CDF)

prob. density function

(PDF)





TODAY

- Definition of CDF & PDF.
- Conditional distribution
- Well-known PDFs
-
-

x^-

CDF

$-\infty < x < +\infty$

$x \neq \pm\infty$

$$\triangleq \quad F_X(x) := P(X \leq x)$$

X: r.v.

x: Values taken by X

Ex: throw a fair die,

$S = \{f_1, f_2, \dots f_6\}$

let us define r.v. $X(f_i) = 10i$, $i = 1, 2, \dots, 6$

$\boxed{\text{Flexibity to define r.v.}}$

$X(f_1) = 10 \longrightarrow 1/6$

$X(f_2) = 20 \longrightarrow 1/6$

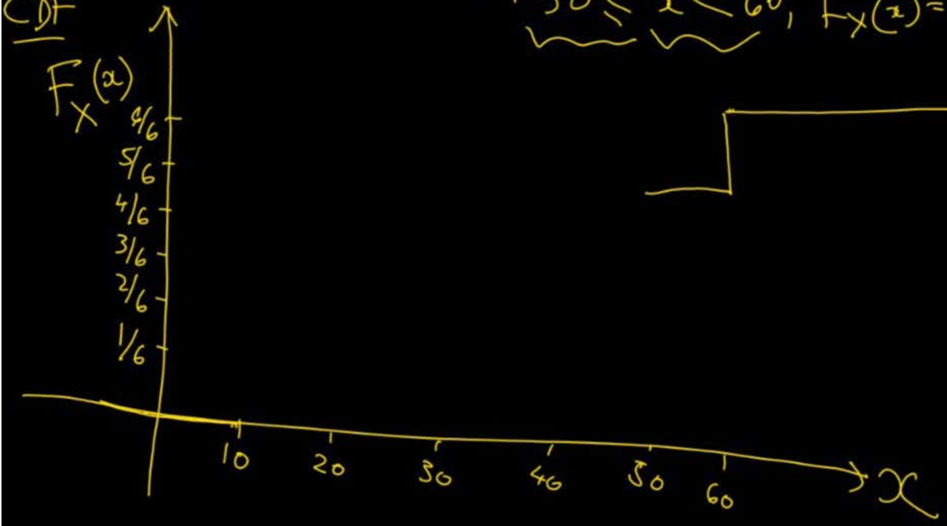
$\vdots$

für die expt.

let, $X > 60$, $F_X(60) = P(X \le 60) = 1$

let, $\underbrace{50 \le}\ \underbrace{x < 60}$, $F_X(x) = 5/6$

CDF

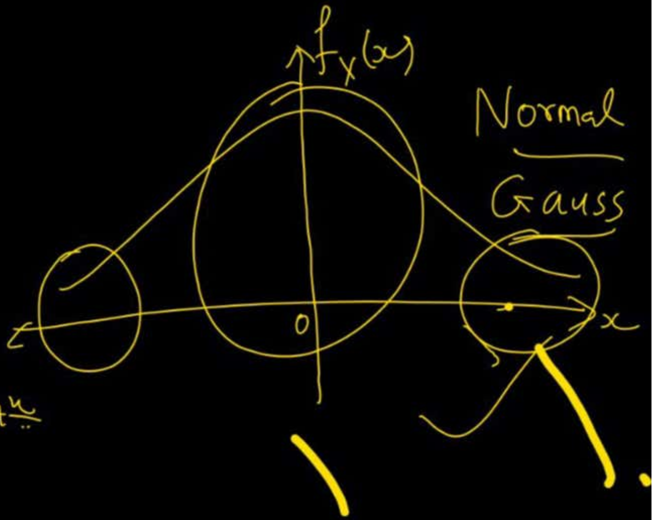


4/6

3/6

2/6

Some Well-Known distributions:



Discrete

- Binomial dist$^{\underline{n}}$
- Poission dist$^{\underline{n}}$

Continuous

- Normal or Gaussian dist$^{\underline{n}}$
- Uniform dist$^{\underline{n}}$
- Exponential dist$^{\underline{n}}$

① Conditional distribution

$$p(A|B) = \frac{p(AB)}{p(B)}$$

Conditional CDF

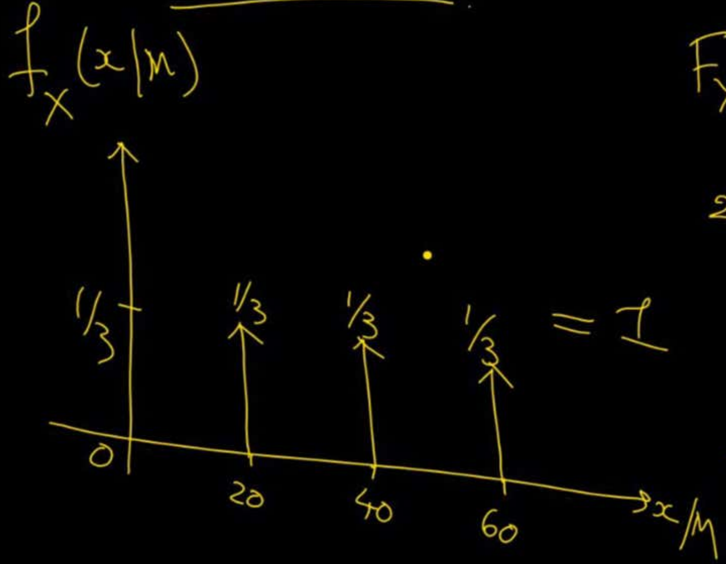
$$\boxed{F_X(x|M) := P(X \le x | M) = \frac{p(X \le x, M)}{p(M)}}$$

occured event

Conditional PDF

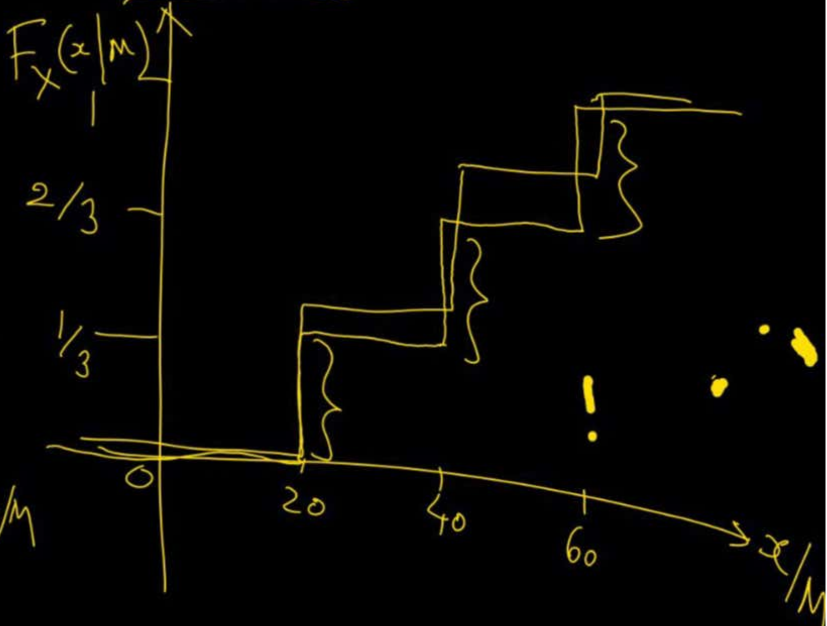
$$f_X(x|M) = \frac{d}{dx} F_X(x|M)$$

Conditional PDF



$= 1$

Cond. CDF

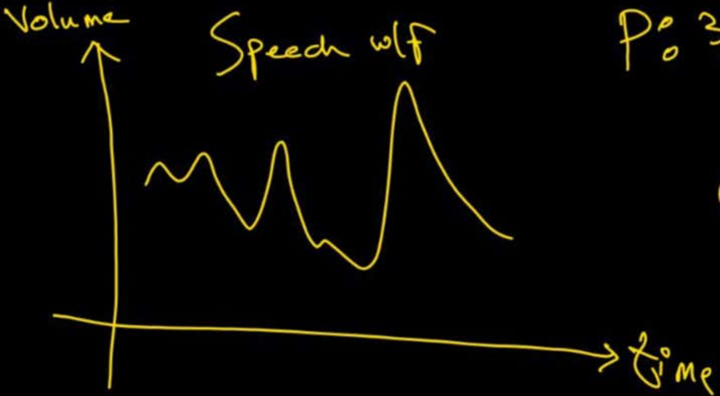


TODAY

"Functions describe the world." — 108

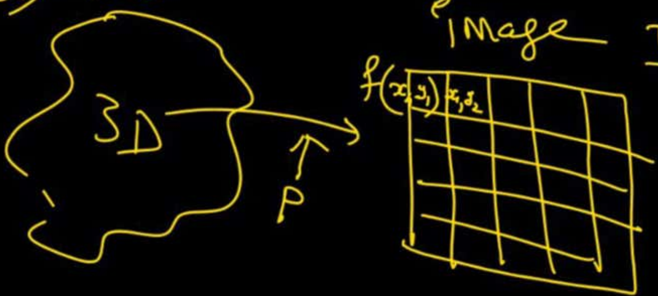
$\mathbb{R}\times\mathbb{R}$ $\quad \mathbb{R}^{\mathbb{R}\times\mathbb{R}}$

Speech w/f



$S: \text{Time} \to \text{pressure (volume)}$

$P: 3D \to 2D$



$f(x_1, y_1)$
$= 250$

$I : (x, y) \to \mathbb{R}^{x \times y}$

Function of a random Variable.

$\xi_e, X(\xi_e)$

- Consider a r.v., $X : S \to \boxed{\mathbb{R}}$
- let us define another r.v., $Y = g(X)$.

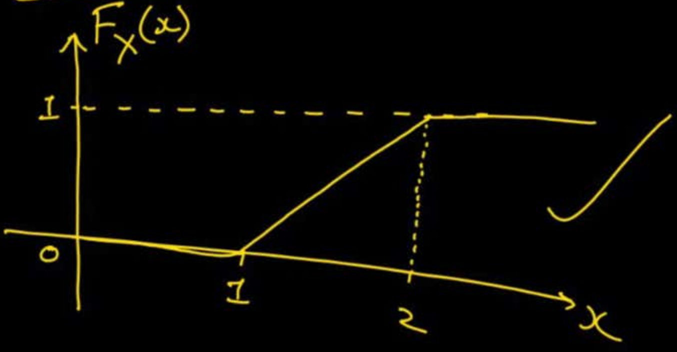
e.g., $\boxed{Y = X^2}$

$U(a, b)$

Fun$^{\underline{n}}$ of a random variable.

$\boxed{\text{A number map to another number}}$

Example:

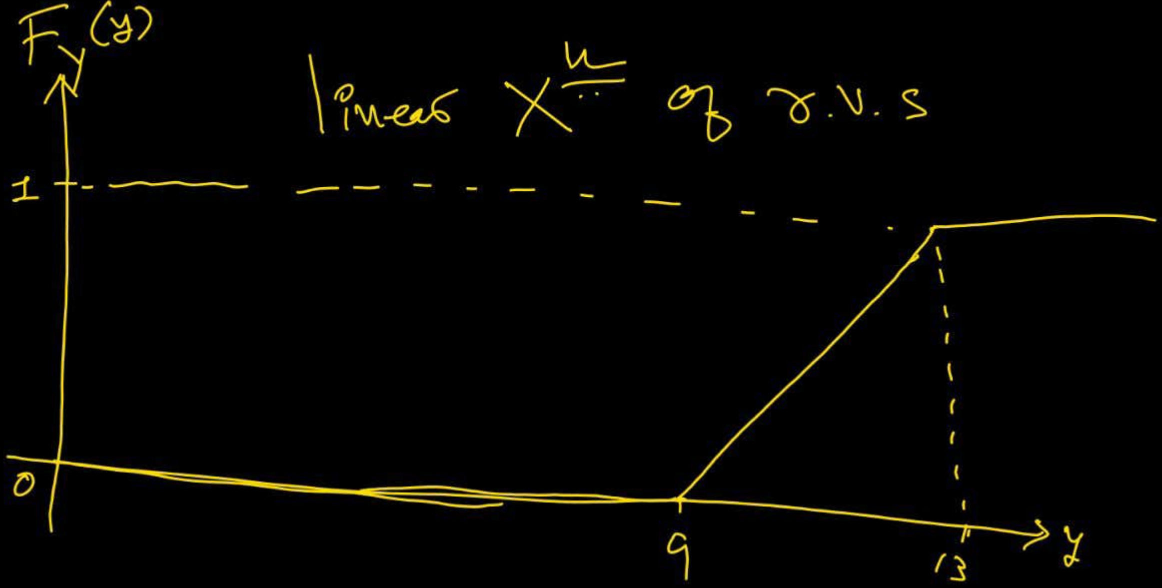


$$Y = g(X)$$
$$= aX + b \implies \frac{y-b}{a}$$
$$= 4X + \cancel{5} \quad \text{(e.g.,)}$$

$$F_Y(y) = F_X\left(\frac{y-b}{a}\right). \quad F_X(2.5)$$

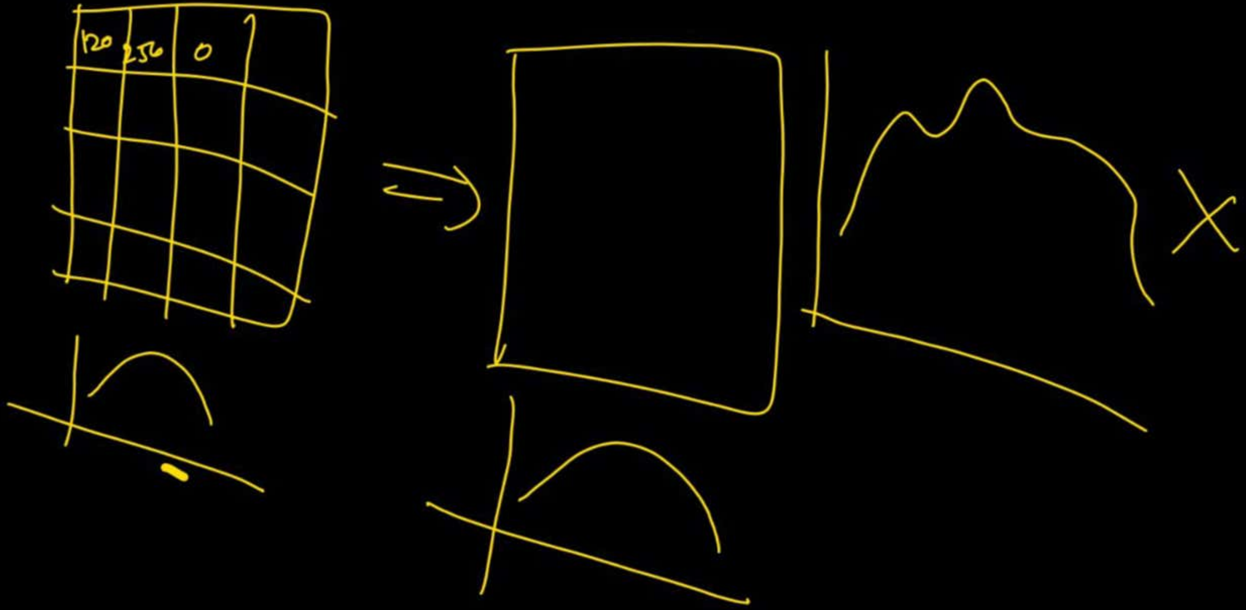
$$y = 15$$

linear X^{n} of r.v.s



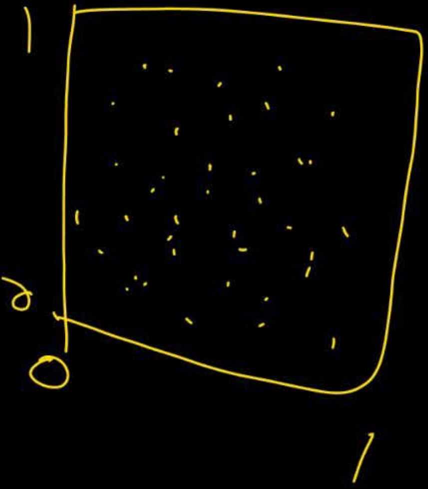
$Y = x^2$

$aX + b$



2D RN

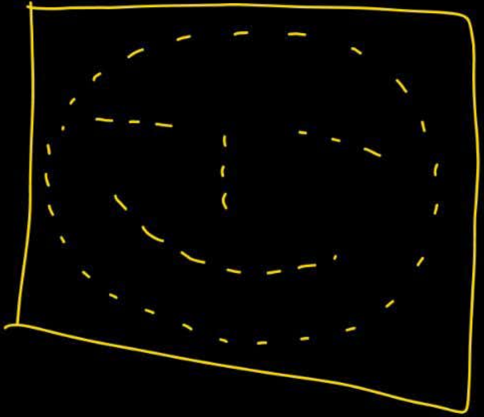
X



eg circle

$f(x) \Rightarrow$

"learn" ⇒ ML

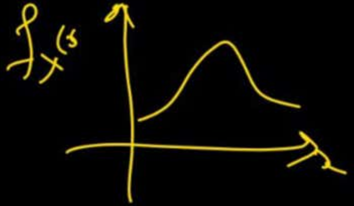


Joint distribution

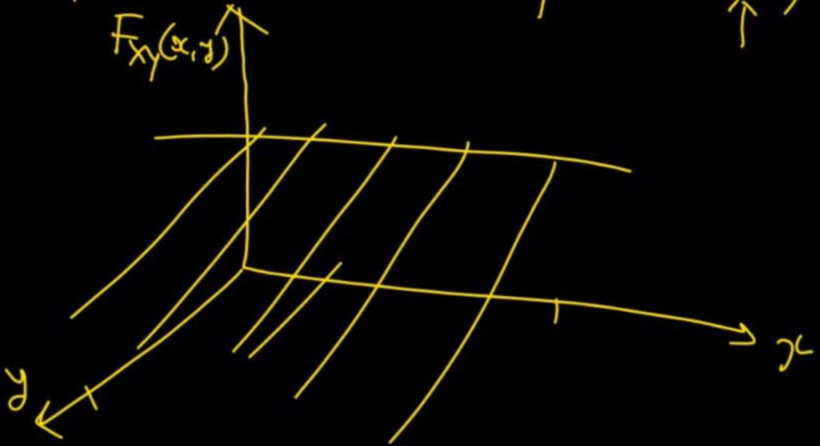
- Consider two r.v.s, X and Y.
- Now, interested in joint behaviour of the X and Y.

Illustration: Speech w/fs, images, Web pages, etc...

Characterization



$$F_{XY}(x,y) = P(X \le x,\ Y \le y)$$



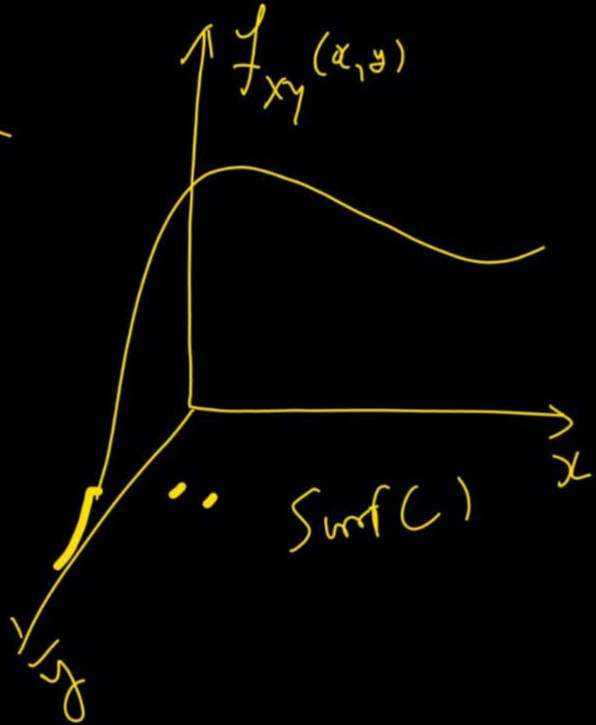


Illustration: Given joint dist^n, find marginal dist^n

Tossing two coins $\Rightarrow$ $S = \{HH, HT, TH, TT\}$

$\Rightarrow \frac{1}{2} \quad \frac{1}{8} \quad \frac{1}{4} \quad \frac{1}{8}$

Y \ X	X=0	X=1	
Y=0	1/8	1/8	1/4
Y=1	1/4	1/2	3/4
	3/8	5/8	

Marginal dist^n

$\frac{1}{4} + \frac{3}{4} = 1$

Illustration:

Given marginals, find joint

Y \ X	X=0	X=1	
Y=0	?	?	___
Y=1	?	?	___
	___	___	

answer is NOT unique

$\therefore \infty$ solutions

$$\boxed{\begin{array}{lll} \text{Joint dist}^{n} \longrightarrow \text{Marginal dist}^{n} & \Rightarrow & \text{Unique sol}^{n} \\ \text{Marginal dist}^{n} \longrightarrow \text{joint dist}^{n} & \Rightarrow & \infty \text{ Solutions} \end{array}}$$

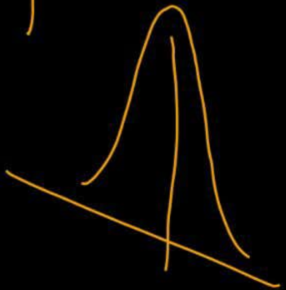
$x+y=2$

Working formulae

$$\mathcal{N}(\mu, \sigma^2)$$

- Set of observations, $O_1, O_2, ---, O_n$.

$$\text{Mean} = \bar{O} = \mu_O = \frac{1}{n}\sum_{i=1}^{n} O_i$$



$$\text{Variance}(O) = \frac{1}{n}\sum_{i=1}^{n} (O_i - \bar{O})^2$$

- Now, let's consider another set of $obs^{\underline{n}}$:

$$p_1, p_2, \ldots, p_n.$$

$\mu_p =$

$Var(p) =$

$\Rightarrow$ - Now, for <u>Covariance</u>, consider pairs ✓

$$(O_1, P_1), (O_2, P_2), \ldots, (O_n, P_n)$$

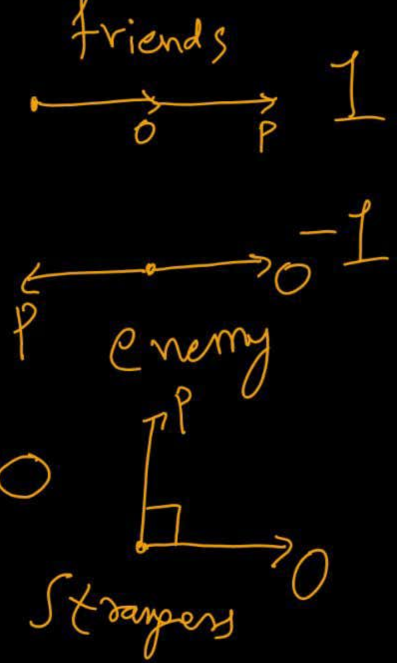
$$COV(O, P) = \frac{1}{n} \sum_{i=1}^{n} O_i P_i - \overline{O}\,\overline{P}$$

· **Correlation** bet^n sets of obs^{ns}

$$\rho_{OP} = \frac{Cov(O,P)}{\sqrt{Var(O)}\,\sqrt{Var(P)}}$$

$$-1 \le \rho_{OP} \le 1$$

Standard deviation



Concept of EXPECTATION

$O_1, O_2, \dots\dots, O_n \longrightarrow$ obsⁿ

$f_1, f_2, \dots\dots, f_n \longrightarrow$ frequency of appearing

$$\bar{O} = \frac{1}{\sum_{i=1}^{n} f_i} \sum_{i=1}^{n} O_i \cdot f_i$$

$$\bar{O} = \sum_{i=1}^{n} O_i \cdot \frac{f_i}{\sum_{i=1}^{n} f_i}$$

EXPECTATION

Value

Relative freq.

Prob.

Remarks:

$$E[aX+b] = aE(x)+b$$

1) Expectation is a linear operator.

2) $$\mathrm{Var}(x) = E(X-\mu)^2$$

$$= \sum_x (x-\mu)^2 \cdot p(x)$$

$$= \sum_x (x^2 - 2\mu x + \mu^2) \cdot p(x)$$

$$= \sum_x x^2 \cdot p(x)$$

$$-2\mu \sum_x x \cdot p(x)$$

$$+\mu^2 \sum_x p(x)$$

$$Var(x) = E(x^2) - 2\mu^2 + \mu^2$$

$$= E(x^2) - \mu^2$$

$$\therefore \boxed{Var(x) = E(x^2) - \left(E(x)\right)^2}$$

Useful identity

Variance is NOT a linear op^{r}:

$$\boxed{\begin{aligned}&Var(aX+b)\\ &= a^2 Var(X)\end{aligned}}$$

$$= E\left[(aX\cancel{+b} - a\mu \cancel{-b})^2\right]$$

$$= E\left[a^2 (X-\mu)^2\right]$$

$$= a^2 E\left[(X-\mu)^2\right]$$

Expectation of funⁿ of two r.v.s : $g(X, Y)$ and Variance

from defⁿ: $$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\cdot f_{XY}(x,y)\cdot dx\cdot dy$$

for example, $$g(X,Y) = XY$$

$$g(X,Y) = X+Y$$

$$E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, \underline{f_{XY}(x,y)\cdot dx\, dy} + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y \cdot \underline{f_{XY}(x,y)\cdot dx \cdot dy}$$

$$= \int_{-\infty}^{\infty} x \left(\underbrace{\left(\int_{-\infty}^{\infty} f_{XY}(x,y)\cdot dy\right)}_{\text{Marginal density of } x} \cdot dx \right) + \int_{-\infty}^{\infty} y \underbrace{\left(\int_{-\infty}^{\infty} f_{XY}(x,y)\, dx\right)}_{y} dy$$

$$\int_{-\infty}^{\infty} x\, f(x)\cdot dx + \int_{-\infty}^{\infty} y \cdot f(y)\, dy$$

$$\cdot \ \mathrm{Var}(\check{X}+\check{Y}) \triangleq E\left[(X+Y)^2\right] - \left[E(X+Y)\right]^2$$

$$= E\left[X^2+Y^2+2XY\right] - (m_X+m_Y)^2$$

$$= E(X^2)+E(Y^2)+2E(XY) - m_X^2 - m_Y^2 - 2m_X m_Y$$

$$= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\left[\underbrace{E(XY) - E(X)\cdot E(Y)}_{\text{Covariance}}\right]$$

In matrix form,

$$\mathrm{Var}(X+Y) = \begin{array}{cc} & \begin{array}{cc} \quad X & \qquad\qquad Y \end{array} \\ \begin{array}{c} X \\ Y \end{array} & \begin{bmatrix} \mathrm{Var}(X) & \underline{\mathrm{Cov}(X,Y)} \\ \underline{\mathrm{Cov}(Y,X)} & \mathrm{Var}(Y) \end{bmatrix} \end{array}$$

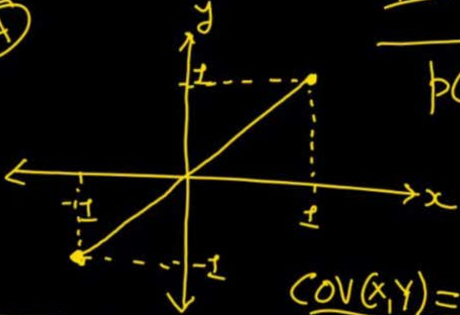
- PSD
- Square
- Symmetric ✓
- diagonals are non-neg.

$\mathrm{Cov}(X,Y) = E(XY) - E(X)\cdot E(Y)$

$\mathrm{Cov}(Y,X) = E(YX) - E(Y)\cdot E(X)$

Illustration: Covariance

Fig. A



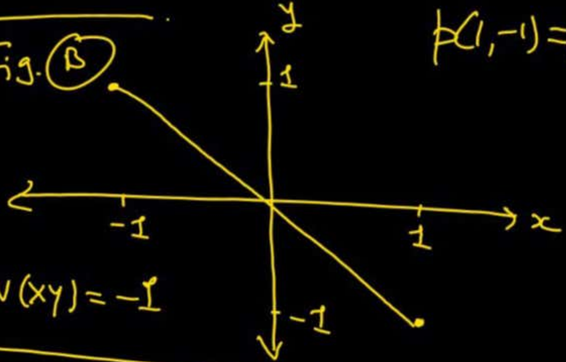
$p(1,1) = p(-1,-1) = \frac{1}{2}$

Relation

$X = Y$

$COV(X,Y) = 1$

Fig. B



$p(1,-1) = p(-1,1) = \frac{1}{2}$

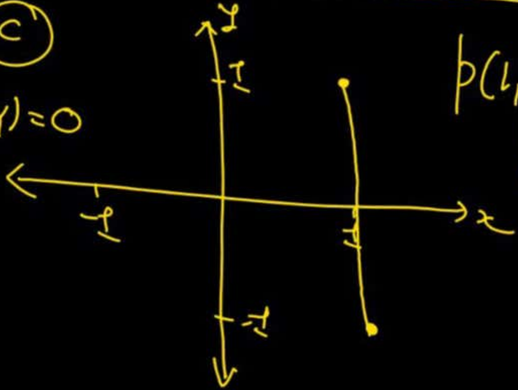
Relation

$X = -Y$

$COV(XY) = -1$

Fig. C

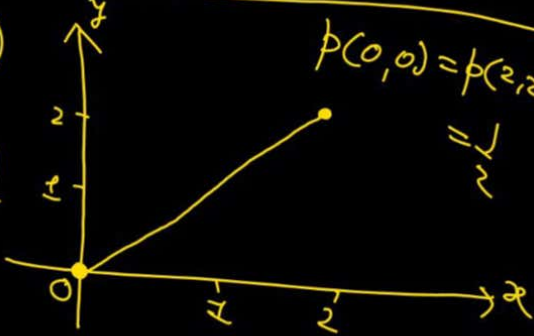
$COV(XY) = 0$



$p(1,-1) = p(1,1) = \frac{1}{2}$

Fig. D

$COV(XY) = 1$



$p(0,0) = p(2,2) = \frac{1}{2}$

Fig. A

$E(x) = 0$

$E(Y) = 0$

$$E(XY) = \sum_i \sum_j x_i y_j \cdot p(x_i, y_j)$$

$$= \left(1 \times 1 \times \tfrac{1}{2}\right) + \left(-1 \times -1 \times \tfrac{1}{2}\right)$$

$E(XY) = 1$

Fig. B

$E(x) = 0$

$E(Y) = 0$

$E(XY) = -1$

Fig. C

$E(x) = 1$

$E(Y) = 0$

$E(XY) = 0$

Fig. D

$E(x) = 1$

$E(Y) = 1$

$E(XY) = 2$

The measure will be,

$$\boxed{E\left[\underbrace{(\overset{1}{X-m_x})(\overset{2}{Y-m_y})}_{3}\right] \triangleq \operatorname{cov}(X,Y)}$$

PCA

$$\operatorname{cov}(X,Y) = E\left[XY - Xm_y - m_xY + m_xm_y\right]$$

$$= E(XY) - m_yE(X) - m_x\cancel{E(Y)} + \cancel{m_x}m_y$$

$$\boxed{\operatorname{cov}(X,Y) = E(\underbrace{\overset{1}{X}\overset{2}{Y}}_{3}) - \underline{E(X)}\,\underline{E(Y)}}$$

$$
\mathrm{Cov}(X,Y) = \begin{cases} +ve & \Rightarrow X, Y \text{ in same direction} \\ 0 & \Rightarrow \text{NOT related.} \\ -ve & \Rightarrow X, Y \text{ are in opposite direction.} \end{cases}
$$

Imp: If $\text{cov}(X, Y) = 0$

$$E(XY) = E(X)\,E(Y)$$

Such X, Y are called "UNCORRELATED" r.v.s

Remarks:

1) $$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$$

if X, Y are Uncorrelated,

$$\therefore \mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

2) Independent of two r.v.s

$$f_{XY}(x,y) \triangleq f_X(x)\cdot f_Y(y)$$

$$p(x,y) = p(x)\cdot p(y)$$

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

Uncorrelate.

If X, Y are Statistically independent $\Longrightarrow$ Uncorrelated. ✓

If X, Y are Uncorrelated $\not\Longrightarrow$ Independent. (HW)

- For, Gaussian r.v.s Independent $\Longleftrightarrow$ Uncorrelated.

-

X

$$\begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix}_{n\times 1}$$

Cov

Cov

$$\begin{bmatrix} \# & \# & \# \\ \# & & \\ \# & & \# \end{bmatrix}_{n\times n}$$

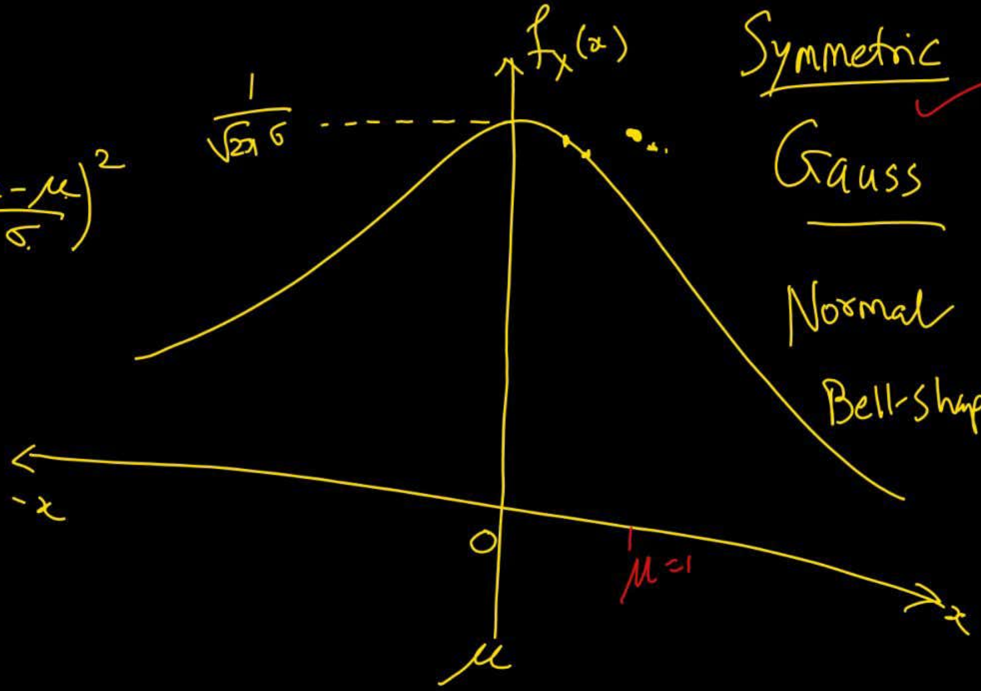
A Gaussian PDF:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

μ = mean

σ^2 = Variance

σ = Std. deviation



Symmetric

Gauss

Normal

Bell-Shaped

n-Variate Gaussian PDF

$n=1$, $n=2$

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma|}}\, e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu})}$$

$\cancel{1\times n}$ $\cancel{n\times n}$ $\cancel{n\times 1}$

determinant of Σ Matrix

Variance-Covariance Matrix $= \Sigma =$

$$\begin{bmatrix} Var(X_1) & cov(X_1, X_2) & \cdots & cov(X_1, X_n) \\ cov(X_2, X_1) & Var(X_2) & \cdots & \\ \vdots & & \ddots & \\ cov(X_n, X_1) & \cdots & \cdots & Var(X_n) \end{bmatrix}_{n\times n}$$

Vector-Valued random variable

$$X : S \to \mathbb{R}$$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$\underline{X} = \begin{pmatrix} \underline{X}_1 \\ \underline{X}_2 \\ \vdots \\ \underline{X}_n \end{pmatrix}$$

$$\underline{X}_1 = \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{pmatrix}$$

$$\underline{X}_2 = \begin{pmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{pmatrix}$$

120×2^{10}

$$\underline{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}$$

Put, $n=1$

Univariate

$$f_X(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Scalar

$$(x-\mu)^T (\sigma^2)^{-1} (x-\mu)$$

$$\left(\frac{x-\mu}{\sigma}\right)^2$$

pls. provide prob. value of X at x

$$X := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

put, $\underline{n=2}$ $\underline{(\text{Bivariate})}$

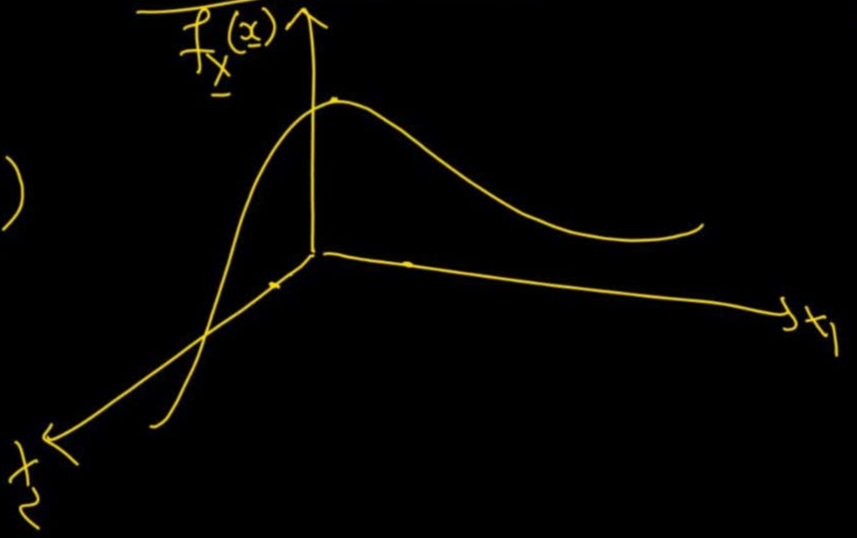
$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)\,|\Sigma|} \cdot e^{-\frac{1}{2}\underset{1\times\cancel{p}}{(\underline{x}_1-\underline{\mu})^T}\underset{\cancel{p}\times\cancel{p}}{\Sigma^{-1}}\underset{2\times 1}{(\underline{x}-\underline{\mu})}}$$

$\underline{X}$ → $\underline{\text{Scalar}}$

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{2\times 1} \qquad \underline{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}_{2\times 1} \qquad \Sigma = \begin{pmatrix} V(x_1) & \mathrm{cov}(x_1,x_2) \\ \mathrm{cov}(x_2,x_1) & V(x_2) \end{pmatrix}_{2\times 2}$$

Bivariate Gaussian

$$\underline{x} = \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix}$$



n-Variate Gaussian

Properties of Gaussian dist$^{\underline{n}}$

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

(I) Normalization of X and Standard normal.

$$\boxed{Y = \frac{X - \mu}{\sigma}}$$

$$f_Y(y) \sim \mathcal{N}(0, 1)$$

(2) Linear X^{n} of $X \sim \mathcal{N}(\mu, \sigma^2)$

$$Y = \alpha X + \beta \Rightarrow f_Y(y) \sim \mathcal{N}(\alpha\mu + \beta, \alpha^2\sigma^2)$$

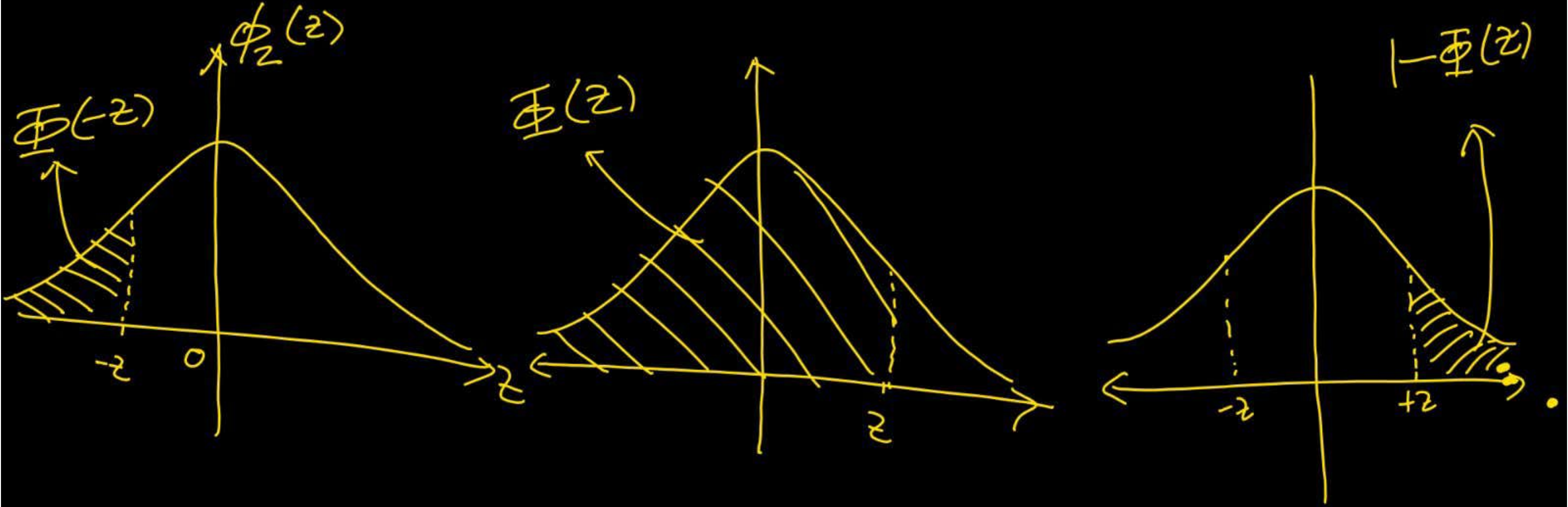
(3) PDF of Standard normal, $\phi_Z(z) \sim \mathcal{N}(0,1)$

$$\phi_Z(z) = \frac{1}{\sqrt{2\pi}\cdot\cancel{\sigma}} e^{-z^2/2} \Leftrightarrow$$

(4) $\phi_Z(z) = \phi_Z(-z)$ Symmetry of PDF.

(5) <u>CDF</u>: $\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \phi_Z(z) \cdot dz$

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



$\phi_Z(z)$
$\Phi(-z)$
$-z$
0
z
$\Phi(z)$
z
$1-\Phi(z)$
$-z$
$+z$

(7) For any α, $\quad 1-\alpha = \Phi(\tau_\alpha)$

Std. normal

	0.00	0.01	- - - - - -	0.09
0	0.5			
.1				
.2				
⋮				

Que: Grading of marks on 4-pointer scale.

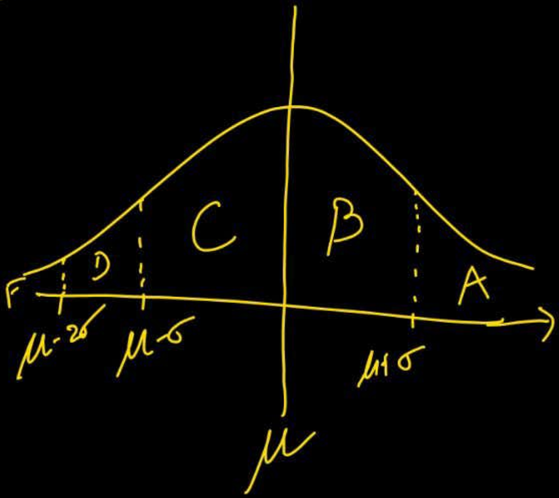
$X > \mu + \sigma$: A grade.

$\mu \leq X < \mu + \sigma$: B grade

$\mu - \sigma \leq X < \mu$: C grade

$\mu - 2\sigma \leq X < \mu - \sigma$: D grade.

F



(a) Find % of students getting A grade.

$$P(X > \mu + \sigma)$$

$$= P\left(\frac{X-\mu}{\sigma} > \frac{\mu+\sigma-\mu}{\sigma}\right)$$

$$= P(Z > 1)$$

$$= 1 - P(Z \leq 1)$$

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

$$= 1 - 0.8413$$

$$= 0.1587$$

$$\approx 15\%$$

(b) Find % of Students getting grade - B.

$$P(\mu \le X < \mu+\sigma)$$

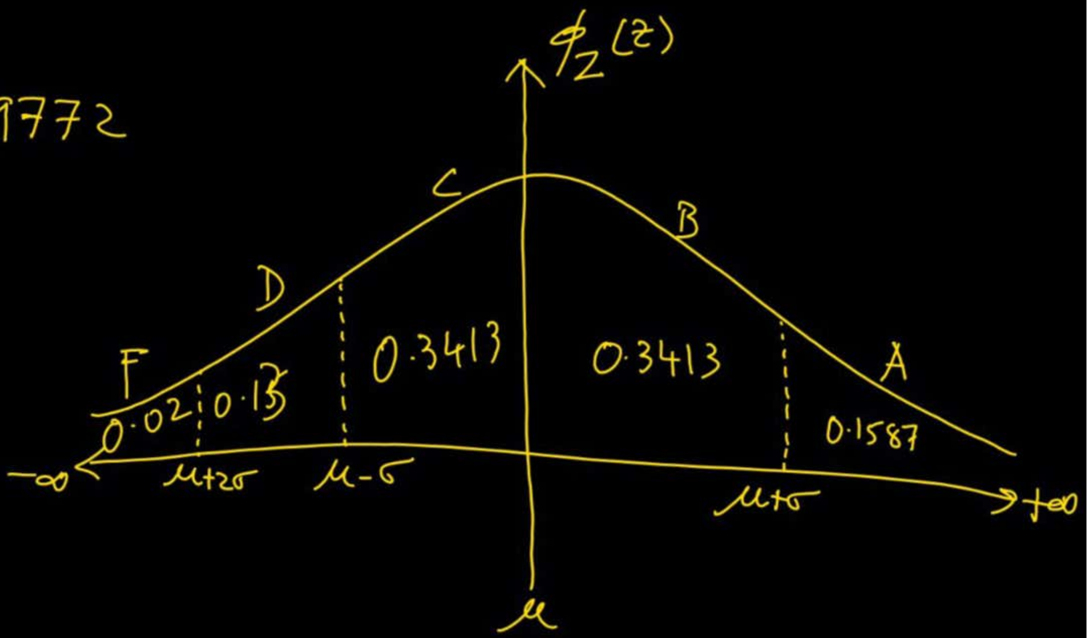
$$= P\left(\frac{\mu-\mu}{\sigma} \le \frac{X-\mu}{\sigma} < \frac{\mu+\sigma-\mu}{\sigma}\right)$$

$$= \Phi(1) - \Phi(0)$$

$$= 0.3413$$

$$\approx 34\%$$

$$\Phi(2) = 0.9772$$



$\phi_Z(z)$
C
B
D
F
A
0.3413
0.3413
0.02
0.13
0.1587
$-\infty$
$\mu+2\sigma$
$\mu-\sigma$
$\mu+\sigma$
$+\infty$
μ

Master Recap

inv()

pinv()

$\underline{b} \notin \text{col}(A)$

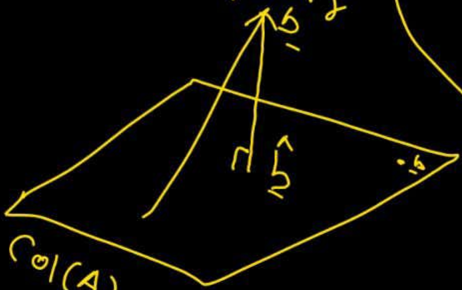
data

$A\underline{x} = \underline{b} + \underline{n}$

"Random" expt. $\}$ → Uncertainty

$\underline{x} = A^{-1}\underline{b}$

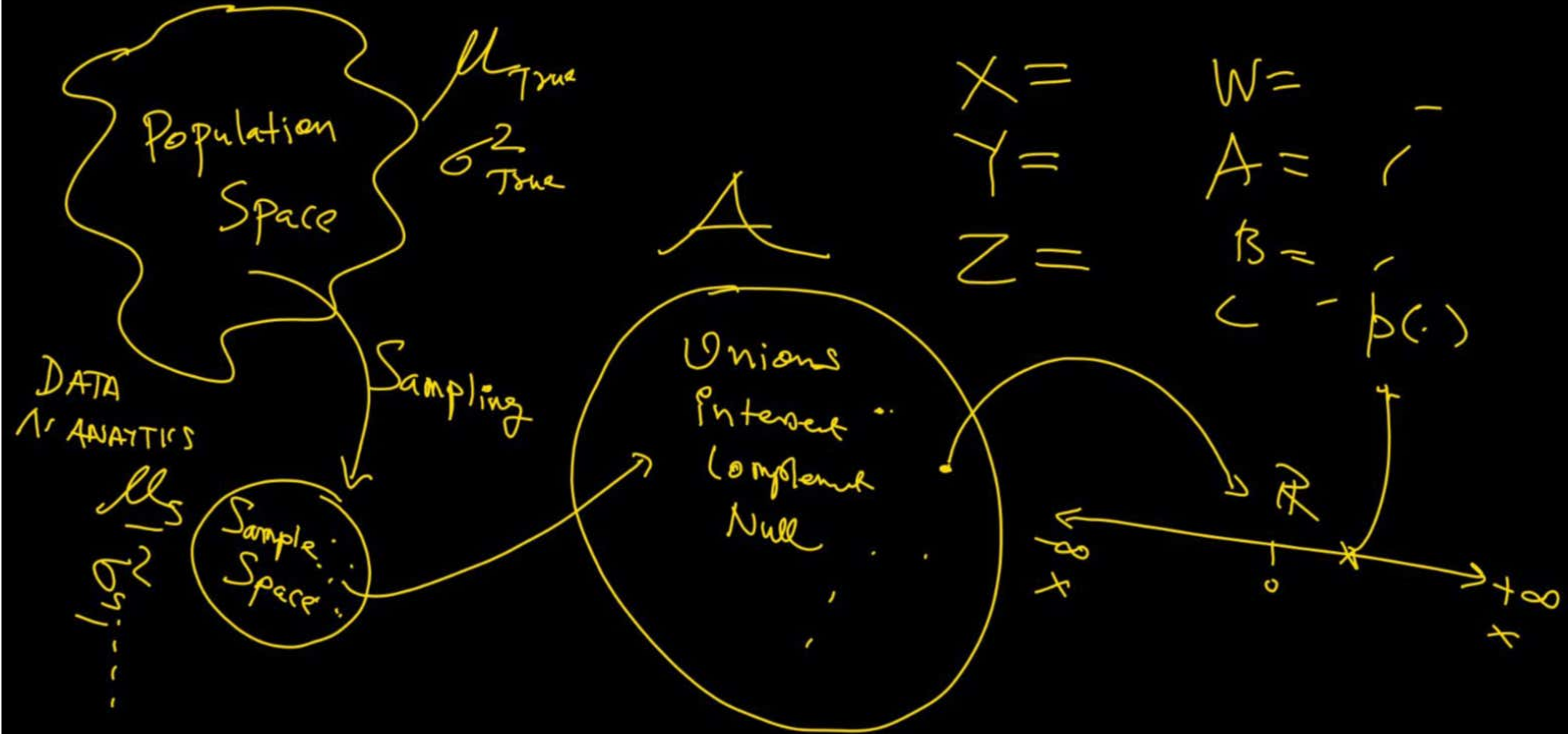
Probability Space



Col(A)

$A^T A \underline{x} = A^T \underline{b}$

$\underline{x}_{pinv} = (A^T A)^{-1} A^T \underline{b}$


Population Space
μ_{True}
σ^2_{True}
Sampling
DATA ANAYTICS
μ_s
σ^2_s
Sample Space
A
Unions
Intersect
Complement
Null
X =
Y =
Z =
W =
A =
B =
C -
p(·)
R
-∞
0
+∞

$$X : S \longrightarrow \mathbb{R} \quad , \quad -\infty < x < \infty$$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \begin{matrix} p(x_1) \\ p(x_2) \\ \vdots \end{matrix}$$

$$Y : S \longrightarrow \mathbb{R}$$

$$X \quad Y$$

$$p(A) = \frac{n_A}{N}$$

1.
2.
3.

$$p(A|B) = \frac{p(AB)}{p(B)}$$

$$p(A|B) = \frac{p(B|A)\, p(A)}{p(B)}$$

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

$$\frac{\partial}{\partial x} F_X(x) = f_X(x)$$

$$F_{X|m}(x|m) =$$

$$E(X) = \sum_i x_i \cdot p(x_i)$$

$$Var(X) = E(X-\mu)^2$$

$$= E(X^2) - (E(X))^2$$

$$Cov(X, Y)$$

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\{\mathrm{cov}(X,Y)\}$$

$$\begin{bmatrix} \mathrm{Var}(X) & \mathrm{cov}(X,Y) \\ \mathrm{cov}(Y,X) & \mathrm{Var}(Y) \end{bmatrix}_{2\times 2}$$

$$E(XY) - E(X)\cdot E(Y)$$

$$E\left[(X-m_x)(Y-m_y)\right]^2$$

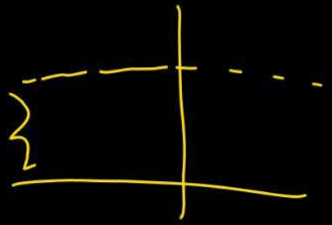
$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\,\sigma_Y} \qquad -1 \le \rho \le +1$$

Statistical indep. $\Longrightarrow$ Uncorrelated.

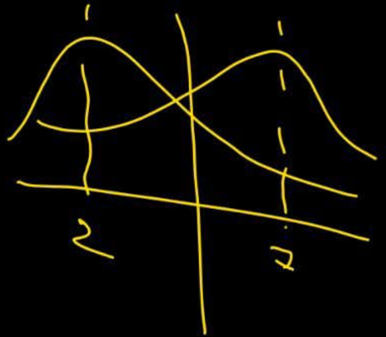
$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$$

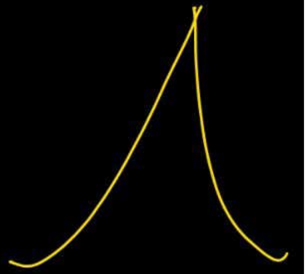
unique

Joint dist^{n} $\xrightarrow{\checkmark}$ Marginal



Marginals $\xrightarrow{\infty}$ joint





Today:

1. Cauchy-Schwarz inequality

2. Chebyshev inequality

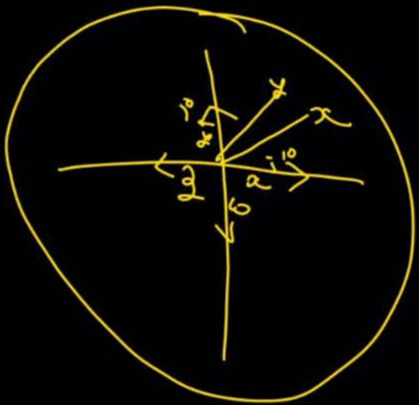
3. Markov inequality

$$4x^2 + 2x + 5 = 10 \rightarrow$$

$$4x^2 + 2x + 5 \leq 10$$

Cauchy-Schwarz inequality

- For vectors, $\underline{v}$ and $\underline{w}$ ✓ $\quad |\underline{v}\cdot\underline{w}| \leq \|\underline{v}\|\,\|\underline{w}\|$



$\mathbb{R}^2$

equality holds, at $\underline{w} = c\,\underline{v}$

constant

- Random variables, V and W

equality holds if $V = cW$

Cauchy-Schwarz inequality

$$|E(VW)| \leq \sqrt{E(V^2)}\sqrt{E(W^2)}$$

· Let, $V = X - m_x$ and $W = Y - m_y$

then, $$\left| \underbrace{E(X-m_x)(Y-m_y)}_{Cov(X,Y)} \right| \leq \sqrt{\underbrace{E(X-m_x)^2}_{\sigma_X^2}} \sqrt{\underbrace{E(Y-m_y)^2}_{\sigma_Y^2}}$$

$$\frac{Cov(X,Y)}{\sqrt{\sigma_X^2}\sqrt{\sigma_Y^2}} \leq \left| \rho_{XY} \right|$$

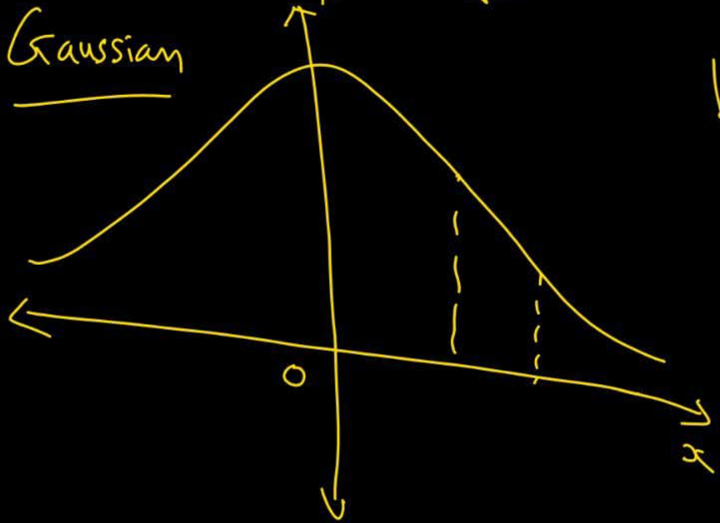
Chebyshev inequality

V.v. practical

- When no PDF is given or required.
- finds probability that random Variable lies between certain interval (bound)

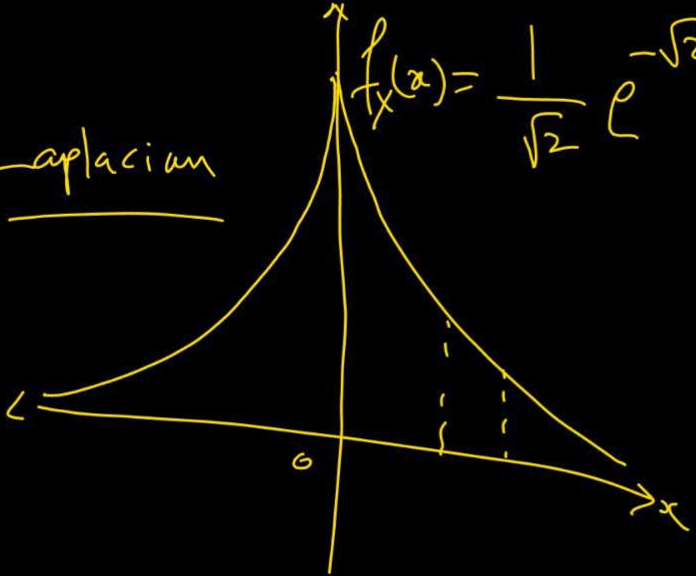
$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Gaussian



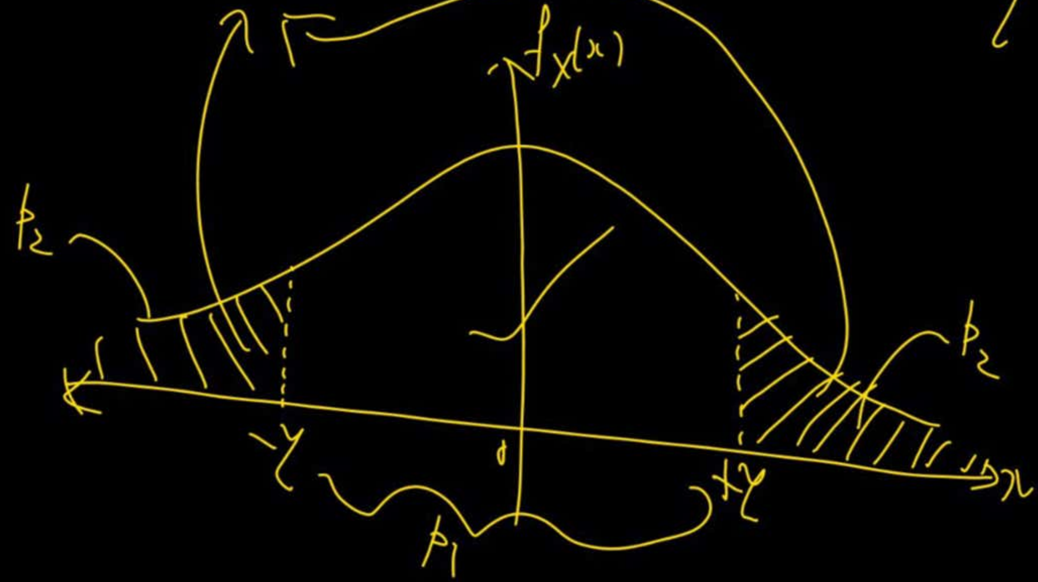
Laplacian

$$f_X(x) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|}$$



Defn: $P\left[|X - m_X| > \gamma\right] < \frac{Var(X)}{\gamma^2}$

let, $m_X = 0$



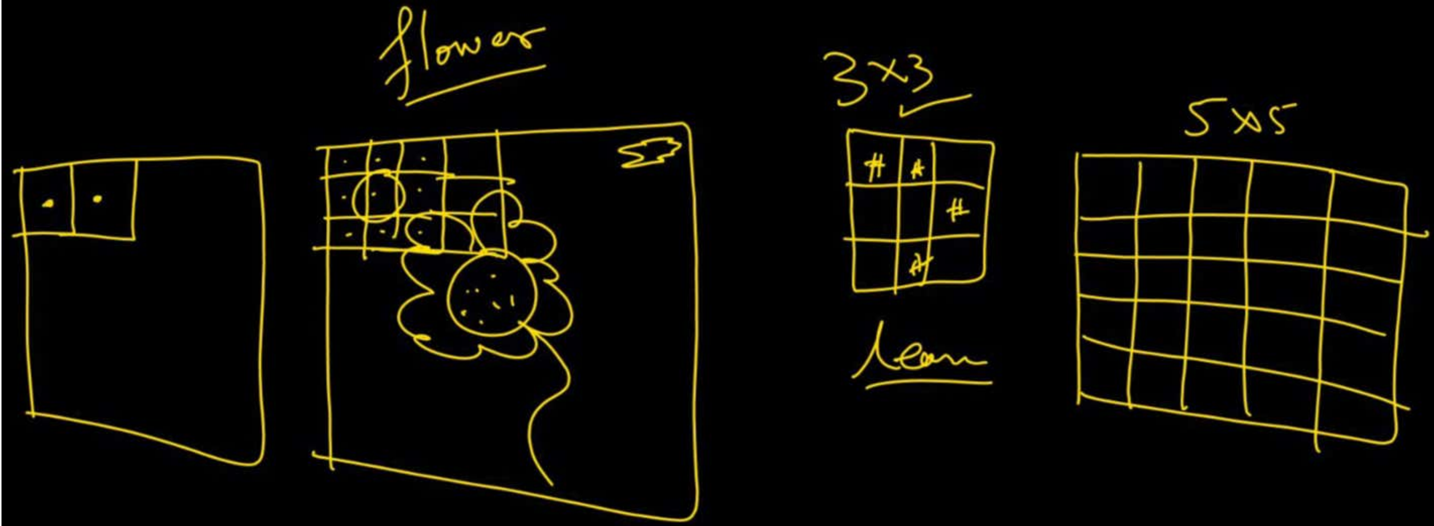
Remark

Exp:

Put $y = k\sigma_x$

then, $p[|x - m_x| > k\sigma_x] \leq \frac{\sigma_x^2}{k^2\sigma_x^2} \leq \frac{1}{k^2}$

$$\left.\begin{array}{l} k=2 \\ \sigma_x=1 \\ m_x=0 \end{array}\right\} \quad p[|x| > 2] \leq \frac{1}{4}$$


flower
3x3
learn
5x5

Consider, $\mathrm{Var}(x) \triangleq \int_{-\infty}^{\infty} (x-m_x)^2 \cdot f_X(x) \cdot dx$

$$\therefore \mathrm{Var}(x) = \underbrace{\int_{|x-m_x| \leq \gamma} (x-m_x)^2 \cdot f_X(x) \cdot dx}_{P_1} + \underbrace{\int_{|x-m_x| > \gamma} (x-m_x)^2 \cdot f_X(x)\, dx}_{P_2}$$

$$\mathrm{Var}(x) \geq \int_{|x-m_x|\geq \gamma} (x-m_x)^2 \cdot f_X(x) \cdot dx$$

replace, $(x-m_x)^2 = \gamma^2$

$$\mathrm{Var}(x) \geq \int_{|x-m_x|\geq \gamma} \gamma^2 \cdot f_X(x) \cdot dx \geq \gamma^2 \cdot P\left[|x-m_x|\geq \gamma\right]$$

Markov's inequality

For any non-negative r.v. X, with $\alpha > 0$

$$P[X \geq \alpha] \leq \frac{m_X}{\alpha}$$

Proof:

$$E[X] = \int_0^\infty x \cdot f_X(x) \cdot dx$$

$$\geq \int_\alpha^\infty x \cdot f_X(x) \cdot dx$$

$$\geq \alpha \int_\alpha^\infty f_X(x) \cdot dx \geq \alpha \, P(X \geq \alpha)$$