papergrid

Date: / /

CHAPTER 2

PROBABILITY CALCULUS

2.1: * Fundamental Concepts of Probability Calculus

→ A random ~~expe~~ experiment is a procedure with two or more possible outcomes where the outcome of one trial depends on chance.

		Possible outcomes
eg:	1. Tossing a coin	H / T
	2. Rolling a die	1/2/3/4/5/6
	3. Tossing a coin twice	HH / HT / TH / TT
	4. Rolling 2 ~~die~~ dice	(1,1) . . . . . (6,6) – 36 outcomes

• **Sample Space and Events**

→ possible outcomes of a random experiment are called simple or elementary events.

→ set of all elementary events of an experiment is called sample space.

$$\Omega = \{ \omega_1, \ldots \omega_n \}$$

where ω_n : outcome ~~of~~ or elementary events

Q Consider sample space Ω for the random experiment of rolling two dice where the 2 dice can be distinguished by their colour. Indicate the event $A \subseteq \Omega$

a) that the sum of 2 dice can be divided by 4.

b) product of 2 dice is larger than 8 but smaller than 20.

Ω =

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

a) $P(A) = \frac{|A|}{|\Omega|} = \frac{9}{36} = \frac{1}{4}$

$A = \{(1,3), (3,1), (2,2), (2,6), (6,2), (3,5), (5,3), (4,4), (6,6)\}$

$|A| = 9$ (cardinality)

b) $B = \{(2,5)\,(2,6)\,(3,3)\,(3,4)\,(3,5)\,(3,6)\,(4,3)\,(4,4)\,(5,2)\,(5,3)\,(6,2)\,(6,3)\}$

$|B| = 12$

$P(B) = \frac{|B|}{|\Omega|} = \frac{12}{36} = \frac{1}{3}$

- **Occurence and Complement of an Event**

→ $\bar{A} = \Omega \setminus A$ → difference

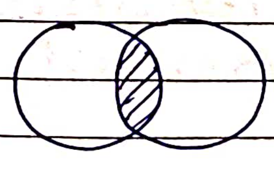
defines the complementary event or complement of A.

eg: $\Omega = \{1,2,3,4,5,6\}$ $\quad A = \{1,3,5\}$

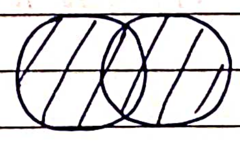
$\bar{A} = \{2,4,6\}$

- **Venn Diagrams**

→ intersection: $A \cap B$ (A and B)

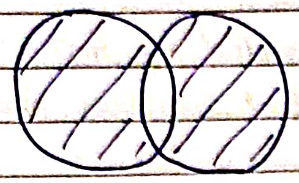


→ union: $A \cup B$ (A or B)

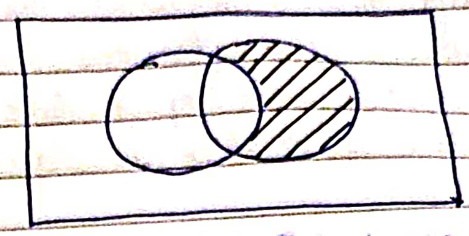


→ XOR : either A or B

$A \Delta B = (A \cup B) \setminus (A \cap B)$



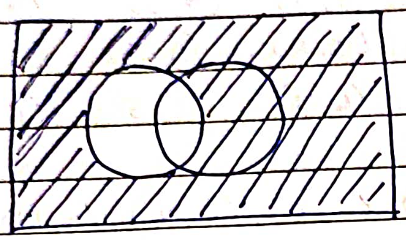
→ Relative Complement



$B \setminus A = \bar{A} \cap B$: B but not A

→ Complement of A

$\bar{A} = \Omega \setminus A$ = not A



→ Calculation Rules

- $(A \cap B) \cap C = A \cap (B \cap C)$
- $(A \cup B) \cup C = A \cup (B \cup C)$
- $A \cap B = B \cap A$
- $A \cup B = B \cup A$
- $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
- $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$

- Classical Approach to Probability

→ A laplace experiment is a random experiment with a finite number of outcomes where all possible outcomes of are equally probable.

→ $\underset{\text{probability}}{P(A)} = \dfrac{\text{number of outcomes in } A}{\text{number of all possible outcomes}} = \dfrac{|A|}{|\Omega|}$

Q Consider the random experiment of rolling two dice where the dice can be distinguished by their colour.

a) What is Probability of event A where the result of the red di[e] is smaller than result of the blue die?

b) what is P(A) that result of red die can be divided by result of blue die?

a) $\frac{15}{36} = P(A) = 0.41$ | $A = \{(2,1)(3,1)(3,2)(4,1)(4,2)(4,3)(5,1)(5,2)(5,3)(5,4)($ $(6,2)(6,3)(6,4)(6,5)\}$

b) $\frac{14}{36} = P(A) = 0.38$ | $A = \{(1,1)(1,2)(1,3)(1,4)(1,5)(1,6)(2,2)(2,4)(2,6)(3,$ $(3,6)(4,4)(5,5)(6,6)\}$

- **Empirical Approach to Probability**

→ empirical / frequentist approach, probabilities are interpreted as limits of relative frequencies.

→ to find $P(A)$, the random experiment is repeated n times. During n trials, event A occurs h_n times.
↳ absolute frequency

$\therefore$ relative frequency of A : $f_n = \dfrac{h_n}{n}$

$$\lim_{n \to \infty} f_n = P(A)$$

- **General Approach to Probability**

→ Goal: every event $A \subseteq \Omega$ shall be assigned a number $P(A)$ which indicates probability of occurrence of A.

→ The set of all subsets of Ω is represented by power set $P(\Omega)$

$|\Omega| = n \Rightarrow |P(\Omega)| = 2^n$

$P(\Omega) = \{ A \mid A \subseteq \Omega \}$

→ A probability measure P is a mapping

$P : P(\Omega) \to \mathbb{R} \qquad A \mapsto P(A)$

- **The Axioms of Kolmogorov**

With Ω the sample space of a random experiment, for arbitrary events $A, B \subseteq \Omega$, it holds:

(A1) $P(A) \geq 0$

(A2) $P(\Omega) = 1$

(A3) $P(A \cup B) = P(A) + P(B)$ if $A \cap B = \emptyset$

Rules derived from Axioms of Kolmogorov

a) $P(A) \leq 1$

b) $P(\emptyset) = 0$

c) $P(A \setminus B) = P(A) - P(A \cap B)$

d) $P(\bar{A}) = 1 - P(A)$

e) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

f) $A \subseteq B \Rightarrow P(A) \leq P(B)$

Q There are two local newspapers published in a country town, the Morning Post and the Evening Standard. The probability that a resident reads the

→ Morning post (event A) is 0.6 — $P(A) = 0.6$

→ evening standard (event B) is 0.5 — $P(B) = 0.5$

→ at least 1 newspaper is 0.9 — $P(A \cap B) = 0.9$

Determine probability that resident reads

a) both newspapers

b) no newspaper at all

c) exactly 1 of the 2 newspapers

a) $P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.6 + 0.5 - 0.9 = 0.2$

b) $P(\overline{A \cup B}) = P(\bar{A} \cap \bar{B}) = 1 - P(A \cup B) = 0.1$

excl. OR

c) $P(A \Delta B) = P((A \cup B) \setminus P(A \cap B))$

XOR

$= P(A \cup B) - P((A \cup B) \cap (A \cap B))$

$P(A \setminus B) = P(A) - P(A \cap B)$

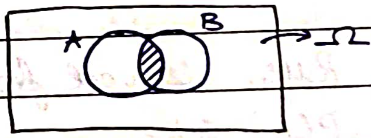
$= P(A \cup B) - P(A \cap B)$

$= 0.7$

• **Conditional Probability**

→ Probability of an event A under the condition that another event B has already occured.

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$ where $P(B) > 0$



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

→ If for two events A and B it holds that

$P(A|B) = P(A)$ or $P(B|A) = P(B)$,

the events A and B are called statistically independent.

For two independent events $P(A \cap B) = P(A) \cdot P(B)$

Events that are not independent are called dependent events

Q Consider the random experiment of rolling two dice, where the dice can be blue / red.

a) What is the conditional probability that red die shows 4 under the assumption that sum of two dice is 9.

b) Are two events "sum of two dice is 7" & "red die shows no. 1" statistically independent?

$\Omega = 36$

a) $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{1/36}{4/36} = \cancel{\tfrac{1}{2}}\ \dfrac{1}{4}$

$P(\overset{A\cap B}{\cancel{A}}) = \{ \cancel{(4,5)}\ (5,4) \}$

$P(A) = \{ (1,4)(2,4)(3,4)(4,4)(5,4)(6,4) \} \cancel{(4,1)(4,2)(4,3)(4,5)(4,6)}$

$P(B) = \{ (3,6)\ (4,5)\ (5,4)\ (6,3) \}$

b) $P(A|B) = \dfrac{1/36}{1/6} = \dfrac{1}{6}$

$P(B) = \{ (1,1)(1,2)(1,3)(1,4)(1,5)(1,6) \} = \dfrac{6}{36} = \dfrac{1}{6}$

$P(A) = \{ (1,6)(2,5)(3,4)(4,3)(5,2)(6,1) \} = \dfrac{6}{36} = \dfrac{1}{6}$

~~$P(B) = \{(1,6)(6,1)\} = \frac{2}{36} = \frac{1}{18}$~~

$P(A \cap B) = \dfrac{1}{36} = P(A) \cdot P(B) = \dfrac{1}{6} \cdot \dfrac{1}{6}$

~~$P(A \cap B) = \frac{2}{36} = \frac{1}{18} = \frac{2}{9}$~~

~~$P(A) \cdot P(B) = \frac{1}{18} \times \frac{1}{6} \neq \frac{2}{9}$~~

$\therefore$ they are independent ~~(not statistically independent)~~

- **Law of Total Probability**

→ Assume $B_1, \ldots B_n$ are a complete decomposition of the sample space Ω into mutually exclusive / disjoint events.

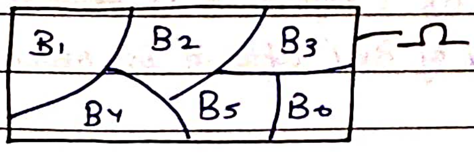
i) $B_1 \cup B_2 \cup \ldots \cup B_n = \Omega$ (complete decomposition)

ii) $B_i \cap B_j \neq \emptyset$ for $i \neq j$ (mutually exclusive)

iii) $P(B_i) > 0$ (no impossible events)

Under the above assumptions, for any event $A \subset \Omega$ it holds that

$\boxed{P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)}$: law of total probability



papergrid

Date: / /

Q A production facility order manufacturing components from three different suppliers S_1, S_2, S_3.

→ S_1 delivers 60% of the components. scrap rate = 9%

→ S_2 delivers 25% of components. scrap rate = 12%

→ S_3 delivers 15% of components. scrap rate = 4%

Out of the whole delivery, one manufacturing component is chosen randomly.

a) What is the probability that chosen component is scrap and what is probability that it is okay?

b) After choosing a component randomly it turns out to be scrap. What is the probability that it was delivered by supplier S_3

A = component is scrap

B_i = component was delivered by S_i $\quad i = 1,2,3$

a) $P(A) = \sum_{i=1}^{3} P(A|B_i) \, P(B_i) = (0.09 \times 0.6) + (0.12 \times 0.25) + (0.04 \times 0.15) = 0.09 = 9\%$

b) $P(B_3|A) = \dfrac{P(A \cap B_3)}{P(A)} = \dfrac{P(A|B_3) \cdot P(B_3)}{P(A)} \dfrac{\cancel{0.04 \times 0.15}}{0.09} = \dfrac{1}{15} \approx 0.07$

- **Bayes' Rule**

enables conclusions about probability of events that are not directly observable.

$$\boxed{P(B_k|A) = \frac{P(B_k) \cdot P(A|B_k)}{\sum_{i=1}^{n} P(B_i) \cdot P(A|B_i)}}$$

Deriv: $P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$

$$P(B_k|A) = \frac{P(B_k \cap A)}{P(A)} = \frac{P(B_k) \cdot P(A|B_k)}{\sum P(A|B_i) \cdot P(B_i)}$$

$P(B_k)$ = prior probability of B_k (before A is observed)

$P(B_k|A)$ = posterior probability of B_k (after A is observed)

Q During operation hours of a power plant : 3 operation modes – high, low, medium

→ op. mode high probability = 0.1
→ " " low " = 0.3
→ " " medium " = 0.6

$\Big\}$ B_i

diff. operation modes have diff. outage probabilities

→ during high, outage probability = 0.5
→ " low " " = 0.1
→ " medium " " = 0.2

There was an outage in the plant. What is the probability that plant was in op. mode high.

B_1 = high B_2 = low B_3 = medium

A = "outage" not A = "no outage"

$P(A|B_1) = 0.5$ $P(B_1) = 0.1$

$P(A|B_2) = 0.1$ $P(B_2) = 0.3$

$P(A|B_3) = 0.2$ $P(B_3) = 0.6$

$$P(B_1|A) = \frac{0.1 \times 0.5}{(0.1 \times 0.5) + (0.3 \times 0.1) + (0.6 \times 0.2)}$$

$$= \frac{0.05}{0.05 + 0.03 + 0.12} = \frac{0.05}{0.20} = \frac{5}{20} = \frac{1}{4}$$

$$= \boxed{0.25}$$

$P(B_1) < P(B_1|A)$

2.2 * DISCRETE RANDOM VARIABLES

→ Sample space of a random experiment = Ω

→ A random variable X is a real valued ~~function~~ function which assigns a numerical value $X(\omega)$ to every possible outcome $(\omega \in \Omega)$ of the random experiment.

→ The values $X(\omega) \in \mathbb{R}$ are called realizations of X.

→ The probability of a value $x \in \mathbb{R}$ is equal to the probability of the corresponding event $A \subseteq \Omega$ consisting of all outcomes $\omega \in \Omega$ which lead to the realization $X(\omega) = x$

$$P(X = x) = P(\{\omega \in \Omega \mid X(\omega) = x\}) = P(A)$$

Examples:

→ Scenario: randomly picking a THWS student.

Sample space $\Omega = \{\omega_1, \omega_2 \dots \omega_n\}$ → collection of all students

ω_j: random student $\qquad \omega_j \longmapsto X(\omega_j)$: height of student ω_j

probability of measuring height $x \in \mathbb{R}$ = probability of selecting any student with height $X(\omega_j) = x$

$$P(X = x) = P(\{\omega_j \in \Omega \mid X(\omega_j) = x\})$$

→ Random Variable

- Discrete
- Continuous

- **Discrete Random Variables**

A random value X is called discrete if set of all possible realizations $X(\omega)$ are isolated points along the real line.

eg: • tossing a ~~die~~ die and considering sum of two results, we have discrete random var. with realizations 2, 3, ..., 12

eg: • possible that X can have infinite values like if X indicates no. of coin flips to get "Head".

- **Probability Mass Function and Cumulative Distribution Function.**

→ p.m.f :

probability mass function

$$\boxed{p(x_k) = P(X = x_k)}$$

→ c.d.f :

cumulative distribution function

$$\boxed{F(x) = P(X \le x) = \sum_{x_k \le x} p(x_k)}$$

eg: Tossing two dice and considering the sum

x_k	Corresponding event	pmf	cdf
2	(1,1)	1/36	1/36
3	(1,2)(2,1)	2/36	3/36
4	(1,3)(2,2)(3,1)	3/36	6/36
5	⋮	4/36	10/36
6	⋮	5/36	15/36
7		6/36	⋮
8		5/36	⋮
9		4/36	⋮
10		3/36	⋮
11		2/36	⋮
12		1/36	36/36

papergrid

Date: / /

Q Exercise : Consider the random experiment of tossing two dice and interpret the sum of the numbers as a discrete random variable X.

a) What is the probability that X lies b/w 5 and 9?

b) What is the probability that X is larger than 6?

a) $P(5 \leq X \leq 9) = \frac{24}{36}$ $\left[P(X=5) + P(X=6) + \dots P(X=9) \right.$ (OR) $\left. F(9) - F(4) \right]$

b) $P(X > 6) = \frac{21}{36}$ $[1 - F(6)]$

• **Expectation Value, Variance and Standard deviation**

for a discrete random variable

→ expectation value / mean :

$$\boxed{\mu = E(x) = \sum_k x_k \cdot p(x_k)}$$

→ the variance

$$\boxed{\begin{aligned}\sigma^2 = V(X) &= \sum_k (x_k - \mu)^2 \cdot p(x_k) \\ &= \left(\sum_k x_k^2 \cdot p(x_k)\right) - \mu^2\end{aligned}}$$

→ standard deviation

$$\boxed{\sigma = \sqrt{V(X)}}$$

Q Exercise : Consider again prev exp. X - sum of numbers

Find μ and σ^2 of X.

- **Probability Distributions** - for discrete random variables can be characterized by specifying -
  1. pmf : probability mass function
  2. cdf : cumulative distribution function

The most important discrete probability distributions are
1. Hypergeometric distribution
2. Binomial distribution
3. Poisson distribution

- **Hypergeometric distribution**

→ sampling without replacement

→ $X \sim Hyp(n, n_1, m)$: n balls ; n_1 black balls
$n - n_1$ ~~t~~ white balls

X: total black balls taken out ; m balls taken out w/o replacement

→ $$p(k) = P(X=k) = \frac{\binom{n_1}{k} \cdot \binom{n_2}{m-k}}{\binom{n}{m}} \; ; \; k = 0, \ldots, m$$

↓ probability mass function

→ expectation value : $\mu = E(X) = m \cdot \frac{n_1}{n}$ $\qquad \mu = \sum_{k=0}^{m} k \cdot p(k)$

variance $\sigma^2 = V(X) = m \cdot \frac{n_1}{n}\left(1 - \frac{n_1}{n}\right) \cdot \frac{n-m}{n-1}$

$$\sigma^2 = V(X) = \sum_{k=0}^{m} \big(k - \underbrace{m \cdot \tfrac{n_1}{n}}_{\mu}\big)^2 \cdot p(k)$$

example : $X \sim Hyp(30, 20, 8)$

$$p(k) = P(X=k) = \frac{\binom{20}{k} \cdot \binom{10}{8-k}}{\binom{30}{8}} \; ; \; k = 0, \ldots, 8$$

k	0	1	2	3	4	5	6	7	8
p(k)	0	0	0.01	0.05					
F(k)	0	0	0.01	0.06					

$\mu = m \cdot \frac{n_1}{n}$

$\mu = \frac{8 \cdot 20}{30} = 5.33$

$$\sigma^2 = m \cdot \frac{n_1}{n}\left(1 - \frac{n_1}{n}\right) \cdot \frac{n-m}{n-1} = 8 \cdot \frac{20}{30}\left(1 - \frac{20}{30}\right) \cdot \frac{30-8}{30-1} = 1.34$$

Q After the legal warranty period is over, 20 out of 30 airbag triggers are still functional. What is the probability that from 8 randomly selected airbags atleast 5 trigger?

$X \sim Hyp.(30, 20, 8)$

$P(X \geq 5) = P(X=5) + P(X=6) + P(X=7) + P(X=8)$

$= 1 - P(X \leq 4)$

$= 1 - F(4) = 1 - 0.23$

(from table on pg 51)

- Binomial distributions

→ sampling with replacement

→ m repetitions with replacement performed

If $m = 1$, a.k.a Bernoulli distribution

→ probability of selecting black ball in one trial is $p \in [0,1]$

∴ probability of selecting white ball is $1-p$.

→ $X \sim Bin(m, p)$

↳ total no. of black balls taken out of the urn.

→ $p(k) = P(X=k) = \binom{m}{k} \cdot p^k \cdot (1-p)^{m-k}$ $k = 0, \ldots,$

$\mu = E(X) = m \cdot p$

$\sigma^2 = V(X) = m \cdot p \cdot (1-p)$

~~Approximation of Hypergeometric distribution by Binomial~~

papergrid

Date: / /

eg: $X \sim Bin(8, 2/3)$

$$p(k) = P(X=k) = \binom{8}{k} \cdot \left(\frac{2}{3}\right)^k \left(\frac{1}{3}\right)^{8-k} \qquad k=0,\dots,8$$

$$\mu = m.p = \frac{8 \times 2}{3} = 5.33$$

$$\sigma^2 = m.p.(1-p) = 8 \times \frac{2}{3} \times \frac{1}{3} = 1.77$$

Q Donald misses the 8 pm news each day with a probability of 0.3. What is the probability that Donald misses the 8 pm news during ~~eight~~ 7 consecutive days at most twice.

$$\underbrace{X}_{\text{number of missed news}} \sim Bin(\underbrace{7}_{\text{days of the week}}, \underbrace{0.3}_{\text{daily miss rate}})$$

$$P(X \leq 2) = F(2) = P(X=0) + P(X=1) + P(X=2)$$

$$= \binom{7}{0} \cdot 0.3^0 \cdot 0.7^7 + \binom{7}{1} \cdot 0.3^1 \, 0.7^6 + \binom{7}{2} \, 0.3^2 \, 0.7^5$$

$$= 1 \times 1 \times 0.7^7 + 7 \times 0.3 \times 0.7^6 + 21 \times 0.09 \times 0.7^5$$

→ Approximation of Hypergeometric Distribution by the Binomial Distribution

Consider $Hyp(n, n_1, m)$ for $m << n$; $\left(\frac{m}{n} \leq 0.1\right)$

$$Hyp(n, n_1, m) \approx Bin\left(m, \frac{n_1}{n}\right)$$

- **POISSON DISTRIBUTION**

→ it is limit of binomial distribution when p gets very small and m gets very large.

→ binomial dist. can be replaced by Poisson when $m \geq 1($ and $\lambda = mp \leq 10$

→ Let X be a random variable $\in N_0 = \{0, 1, 2 \dots\}$, then

$$X \sim Ps(\lambda) \quad (\lambda > 0) : \boxed{p(k) = P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}} \quad k = 0, 1, 2 \dots$$

$$Bin(m, p) \xrightarrow[mp = \lambda]{m \to \infty,\ p \to 0} Ps(\lambda)$$

→ for $X \sim Ps(\lambda)$: $\boxed{\mu = E(X) = \lambda \text{ and } \sigma^2 = V(X) = \lambda}$

Q On avg. 5% of output of a production facility do not posses required quality for further processing. Determine the prob. that out of a ~~s~~ sample of 100 output items, exactly 10 do not meet quality requirements.
a) using Binomial
b) using Poisson

$X \sim Bin(100, 0.05)$

a) $p(k) = P(X = 10) = \binom{m}{k} \cdot p^k \cdot (1-p)^{m-k}$

$$= \binom{100}{10} \cdot (0.05)^{10} \cdot (0.95)^{90}$$

$$\binom{100}{10} = \frac{100!}{10!\,90!}$$

$$= 0.0167$$

b) $p(k) = P(x=10) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$

$\boxed{\lambda = mp = 100 \times 0.05}$

$$= \frac{5^{10}}{10!} \cdot e^{-5} = 2.6911 \times e^{-5}$$

$$= 0.0181$$

23 * CONTINUOUS RANDOM VARIABLES

- Discrete vs. Continuous

→ for discrete random variables, all possible values can be ennumerated : $x_1, x_2, x_3 \cdots$

→ for continuous random variables, it can be any real no. in a whole interval.

- Cumulative Distribution Function

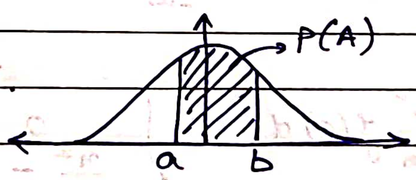
→ The probability $P(X \le x)$ to wait at most x minutes is given by c.d.f $F(x)$ as follows :

→ $$F(x) = P(X \le x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{x}{5} & \text{if } 0 \le x \le 5 \\ 1 & \text{if } x > 5 \end{cases}$$

- Probability Distribution Function - pdf

→ The p.d.f $f(x)$ of continuous random variables X takes on positive values on whole intervals of the real line and not on isolated points.

→ specific probability will be represented by area under the p.d.f.



$$P(A) = P(A \in A) = P(a \le X \le b) = \int_a^b f(x)dx = F(b) - F(a)$$

- Relationship b/w cdf and pdf

→ The cdf $F(x)$ is now integral of pdf $f(x)$ upto point x

$$\boxed{F(x) = P(X \le x) = \int_{-\infty}^{x} f(s)ds}$$

→ vice versa

$$\boxed{f(x) = \frac{d}{dx} F(X)}$$

papergrid
Date: / /

- **Probabilities for Continuous Random Variables**

1. $P(X=a) = P(a \le X \le a) = \int_a^a f(x)\,dx = F(a) - F(a) = 0$

2. $P(X<a) = P(X \le a) - P(X=a) = P(X \le a) = F(a)$

3. $P(a<X<b) = P(a<X\le b) = P(a \le X < b) = P(a \le X \le b)$
$= F(b) - F(a)$

- **Properties of Probability Density function (pdf)**

for pdf of cont. rand. var. X, the following MUST HOLD

i) $f(x) \ge 0$ for all $x \in \mathbb{R}$ — non negativity

ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$ — normalization

Q $f(x) = \begin{cases} x/8 & \text{if } 3 \le x \le 5 \\ 0 & \text{otherwise} \end{cases}$

a) Check $f(x)$ is pdf of cont rand var. X

b) Determine cdf $F(x)$ and compute $P(X \le 4)$ and $P(4<X<5)$.

a) i) Non negativity ✓

ii) Normalization : $\int_{-\infty}^{\infty} f(x)\,dx = \int_3^5 \frac{x}{8}\,dx = \frac{x^2}{16}\Big|_3^5 = \frac{25-9}{16} = 1$

continuous

b) $F(x) = \int_{-\infty}^{x} f(s)\,ds = \begin{cases} 0 & x<3 \\ \frac{x^2}{16} - \frac{9}{16} & 3 \le x \le 5 \\ 1 & x>5 \end{cases}$

$\|$

subtract this so that the function is continuous at both its end limits 3 & 5

$P(X \le x) \rightarrow P(X \le 4) = \frac{16}{16} - \frac{9}{16} = \frac{7}{16}$

$P(4<X<5) = F(5) - F(4)$
$= \frac{9}{16}$

$\frac{12-10}{40} = \frac{2}{40}$

papergrid

Date: /

- **Continuous Random Variable – Remarks**

→ $f(x)$ can have jumps.

→ pdf $f(x)$ can take on values larger than 1 (since probabilities are now expressed as area below pdf $f(x)$ and not $f(x)$ directly)

→ for a given pdf $f(x)$, corresponding cdf $F(x)$ is always a continuous function differentiable at every point where pdf $f(x)$ is continuous. $\frac{d\,F(x)}{dx} = f(x)$

- **Expectation value, variance and Standard Deviation**

for a cont. random variable X with pdf $f(x)$,

→ expectation value or mean

$$\boxed{\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx}$$

→ variance

$$\boxed{\begin{aligned}\sigma^2 = V(X) &= \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\,dx \\ &= \left(\int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx\right) - \mu^2\end{aligned}}$$

→ Std. deviation

$$\boxed{\sigma = \sqrt{V(X)}}$$

Q $f(x) = \begin{cases} 6(x-x^2) & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$

a) Check that $f(x)$ is pdf of cont. rand. var. X.

b) Find μ & σ^2

~~$(x^2 < x$ for $0 \le x \le 1)$~~

a) non negativity ✓ normalization : $\int_{-\infty}^{\infty} 6x - 6x^2 = \left.\frac{6x^2}{2} - \frac{6x^3}{3}\right|_0^1 = 3-2=1$ ✓

b) $\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \int_0^1 6x^2 - 6x^3\,dx$

$= \left.\frac{{}^2\cancel{6}x^3}{\cancel{3}} - \frac{{}^3\cancel{6}x^4}{{}_2\cancel{4}}\right|_0^1 = 2 - \frac{3}{2} = \frac{1}{2}$ //

$\sigma^2 = \int_0^1 (6x^3 - 6x^4) - \mu^2 = \left.\frac{6x^4}{4} - \frac{6x^5}{5}\right|_0^1 - \frac{1}{4} = \frac{3}{10} - \frac{1}{4} = \frac{1}{20}$ //

Probability distributions for continuous random variables can be characterized by specifying 1. pdf
2. cdf

The most important continuous probability distributions are
1. uniform distribution
2. exponential distribution
3. normal distribution

- **Uniform Distribution**

→ for uniform distribution, one specifies
1. the smallest possible value a
2. the largest possible value b

→ on the obtained interval $[a, b]$, the pdf $f(x)$ is assumed to be constant.

→ for a uniformly distributed random variable X on $[a, b]$, we write $X \sim U(a, b)$

$$f(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{if } x > b \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } x > b \end{cases}$$

$$\mu = E(X) = \frac{a+b}{2} \qquad \sigma^2 = V(X) = \frac{(b-a)^2}{12}$$

$\frac{1}{\frac{1}{10}}, \frac{1}{10}$

Exponential Distribution

→ applied for random variables measuring the length of a time span.

eg: • duration of telephone conversations
• life span of wireless equipment
• interarrival time

→ for an exponentially distributed random variable X, the conditional distribution of the remaining time is independent of the already consumed time. Thus it is called distribution without memory.

$$P(X \le t+s \mid X \ge t) = P(X \le s)$$

→ pdf $f(x) = \begin{cases} 0 & x<0 \\ \lambda e^{-\lambda x} & x \ge 0 \end{cases}$

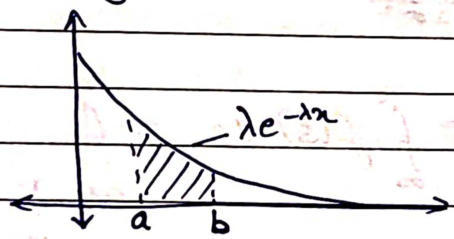
parameter $\lambda > 0$

→ cdf $F(x) = \begin{cases} 0 & x<0 \\ 1-e^{-\lambda x} & x \ge 0 \end{cases}$

→ $\mu = E(X) = \frac{1}{\lambda}$ $\sigma^2 = V(X) = \frac{1}{\lambda^2}$

→ short form: $X \sim Exp(\lambda)$

→ For probability $P(a \le X \le b) = \int_a^b \lambda e^{-\lambda x} dx = F(b) - F(a)$
$= e^{-\lambda a} - e^{-\lambda b}$



Q time span b/w arrival of trucks at a gate in minutes is assumed to be an exponentially distributed rand. var. $X \sim Exp(\lambda)$. After t minutes have passed since the last arrival, the probability that the next truck...

a) $X \sim Exp : P(X \leq t+1 \mid x \geq t) = \frac{1}{2} = P(X \leq 1) = F(1)$

↳ memoryless property

$F(1) = 1 - e^{-\lambda} = \frac{1}{2}$

$\Rightarrow e^{-\lambda} = \frac{1}{2} \quad \Rightarrow \ln e^{-\lambda} = \ln \frac{1}{2}$

$\Rightarrow +\lambda = -\ln\left(\frac{1}{2}\right) = \ln 2 \approx 0.7$ // $\left(\ln \frac{1}{2} = -\ln 2\right)$

b) $P(1 \leq X \leq 3) = e^{-\lambda} - e^{-3\lambda}$

↳ $F(3) - F(1)$

$\Rightarrow P(1 \leq X \leq 3) = e^{-\ln 2} - e^{-3\ln 2}$

$\Rightarrow P = e^{\ln \frac{1}{2}} - e^{\ln \frac{1}{8}} = \frac{1}{2} - \frac{1}{8} = \frac{3}{8}$ //

- **Normal Distribution / Gaussian distribution**

→ most imp. continuous probability distribution

→ can take on values on the whole real line.

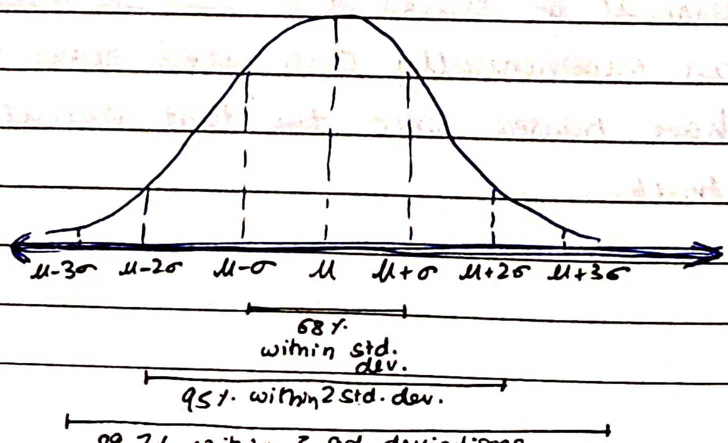
→ pdf $f(x)$ is called Gaussian bell curve / normal curve

→ this distribution plays an imp role in inductive / inferential stat as well

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

→ cdf $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(s)\,ds = \int_{-\infty}^{x} \underbrace{\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(s-\mu)^2}{2\sigma^2}}}_{\text{not directly computable}}$

→ we write shortly $X \sim N(\mu, \sigma^2)$



$F_0(-z) = P(X \le -z)$
$F_0(-z) = 1 - F_0(z)$ ~~$P(X \ge z)$~~

$F_0(z) = P(X \le z)$

→ PROPERTIES :

1. $f(x) > 0$ for all $x \in R$
2. maximum of $f(x)$ is attained at $x = \mu$ → second derivative = 0
3. $f(x)$ has inflection points at $x = \mu - \sigma$ and $x = \mu + \sigma$
4. $f(x)$ is symmetric around the mean

$$f(\mu - x) = f(\mu + x) \text{ and } F(\mu - x) = 1 - F(\mu + x)$$

• **Standard Normal Distribution**

→ since cdf $F(x)$ of $X \sim N(\mu, \sigma^2)$ cannot be computed directly, values are transformed to the std. norm. dist. $N(0,1)$ which is supported by table values.

→ $$f_0(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad F_0(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds$$

→ transformation rule for $X \sim N(\mu, \sigma^2)$ to $N(0,1)$ is

$$F(x) = P(X \le x) = F_0\left(\left(\frac{x-\mu}{\sigma}\right)\right) = F_0(z) \text{ where } z = \frac{x-\mu}{\sigma}$$

Q Past experience has shown that the time needed for the successful deployment of an SAP ERP system is normally distributed with the mean $\mu = 39$ weeks. & std dev. $\sigma = 2$ weeks. What is probability that time for deployment of a system lies b/w 37 and 41 weeks.

$$X \sim N(\underset{\mu}{39}, \underset{\sigma^2}{4}) = P(\underset{\mu-\sigma}{37} \le X \le \underset{\mu+\sigma}{41}) = F(41) - F(37) = 0.68$$

$$= F_0\left(\frac{41-39}{2}\right) - F_0\left(\frac{37-39}{2}\right)$$

$$= F_0(1) - F_0(-1)$$

$$= F_0(1) - (1 - F_0(1))$$

$$= 2F_0(1) - 1 = 2 \times 0.84134 - 1$$

$$= 0.68$$

2.4 * MULTIVARIATE DISTRIBUTIONS

→ The joint realization of two or more variables is described by multivariate distributions.

→ Two / more simultaneously observed random variables is ~~called~~ combined to give a random vector.

- Random Vectors

→ $X : \Omega \rightarrow \mathbb{R}^D \quad , \quad \omega \mapsto X(\omega) = (X_1(\omega), \ldots, X_D(\omega))^T$

↳ D dimensional random vector
a.k.a multivariate random variables

→ The joint probability distribution of X is denoted by P_X and we write $X \sim P_X$

eg: Bivariate data sample : joint measurement of height and weight of a random THWS student.

$$X : \underset{\text{all students}}{\Omega} - \mathbb{R}^2 \; ; \; \omega_j \mapsto X(\omega_j) = (\underbrace{X_1(\omega_j)}_{\text{height of student } \omega_j}, \underbrace{X_2(\omega_j)}_{\text{weight of student } \omega_j})$$

→ F_X : joint cumulative distribution function (joint cdf)

- Joint Distribution of Random Vectors → (Joint cdf)

→ (Joint Probability Density function)

f_X : joint probability density function (joint pdf) of X.

$$F_X(x_1, \ldots, x_D) = \int_{-\infty}^{x_D} \cdots \int_{-\infty}^{x_1} f_X(s_1, \ldots, s_D)\, ds_1 \ldots ds_D$$

$$F_X(x_1, \ldots x_D) = P(\{\omega \in \Omega \mid X_1(\omega) \leq x_1, \ldots, X_D(\omega) \leq x_D\})$$
$$= P(X_1 \leq x_1, \ldots X_D \leq x_D)$$

$$\underbrace{P_X(B)}_{\text{probability}} = P(X \in B) = \int_B f_X(x_1, \ldots, x_D)\, dx_1 \ldots dx_D$$

→ Properties of Joint Cdf and Pdf

1. The joint cdf $F_X(x) = F_X(x_1, \ldots, x_D)$ is monotonically increasing for all x_i $(i = 1, 2, 3 \ldots D)$

2. $$\lim_{x_i \to -\infty} F_X(x_1, \ldots, x_D) = 0$$

3. $$\lim_{x_1, \ldots x_D \to +\infty} F_X(x_1, \ldots, x_D) = 1$$

4. $$f_X(x_1, \ldots, x_D) = \frac{\partial^D F_X(x_1, \ldots, x_D)}{\partial x_1 \ldots \partial x_D}$$

- Marginal Cdf and Pdf

→ for continuous random vector $\mathbf{X} = (X_1, \ldots, X_D)^T$ with joint pdf $f_X(x)$, the marginal pdf is

$$f_{X_i}(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, \ldots, x_D)\, dx_1 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_D$$

→ marginal cdf $F_{X_i}(x_i) = \int_{-\infty}^{x_i} f_{x_i}(s_i)\, ds_i$

- Statistical Independence of Random Variables

→ The components $X_1, \ldots, X_D$ of $\mathbf{X}$ are called statistically independent if

$$F_X(x_1, \ldots, x_D) = F_{X_1}(x_1) \cdot \ldots \cdot F_{X_D}(x_D)$$

→ In case $X_1, \ldots, X_D$ are statistically independent and have same marginal distribution, they are called independent & identically distributed

- **Mean of a Random vector**

for a random vector $X = (X_1, \ldots, X_D)^T$,

$$\text{meanvector or mean } X = \underbrace{\mu}_{\text{measure of central tendency}} = E(X) = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_D) \end{pmatrix}$$

If X is a continuous random vector with joint pdf $f_X(x)$, then

$$\boxed{\mu_i = E(X_i) = \int_{\mathbb{R}^D} x_i \cdot f_X(x_1, \ldots, x_D)\, dx_1 \ldots dx_D}$$

- **Covariance Matrix of a Random Vector**

for a random vector X,

measure of dispersion & correlation

$$\Sigma = \mathrm{Cov}(X) = (\sigma_{ij})_{i,j} = (\mathrm{Cov}(X_i, X_j))_{i,j}$$

covariance matrix of X. → symmetric matrix of the pairwise covariances of X_i and X_j

$$\boxed{\sigma_{ij} = \mathrm{Cov}(X_i, X_j) = \int_{\mathbb{R}^D} (x_i - \mu_i)(x_j - \mu_j) f_X(x_1 \ldots x_D)\, dx_1 \ldots dx_D}$$

- **Properties of Covariance Matrix**

Σ is always positive semidefinite i.e

$$\forall a \in \mathbb{R}^D : \quad a^T \Sigma a \geq 0$$

Also
$$\begin{aligned} \Sigma &= E((X-\mu)(X-\mu)^T) \\ &= E(XX^T) - E(X)\mu^T - \mu E(X^T) + \mu\mu^T \end{aligned}$$

$$\boxed{\Sigma = \underbrace{E(XX^T)}_{\text{auto correlation matrix}} - \mu\mu^T}$$

papergrid

Date: / /

- **Linear Transformation of Random Vectors**

let $X = (X_1, \dots X_D)^T$ with mean μ & cov. matrix Σ

let A = arbitrary real matrix

b = " " vector

Then for $Y = AX + b$

1. $E(Y) = E(AX + b) = A\mu + b$
2. $Cov(Y) = Cov(AX+b) = A\Sigma A^T$

Q Exercise - The joint pdf $f_X(x)$ of $X = (X_1, X_2)^T$ is given by

$$f_X(x_1, x_2) = \begin{cases} 3x_1^2 + 3x_2^2 & \text{for } 0 \le x_1 \le 1,\ 0 \le x_2 \le x_1 \\ 0 & \text{otherwise} \end{cases}$$

Determine $\mu = E(X)$ and $\Sigma = Cov(X)$

$$\mu_1 = E(X_1) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_1 \cdot f_X(x_1, x_2)\, dx_1 dx_2$$

$$= \int_0^1 \int_0^{x_1} x_1 (3x_1^2 + 3x_2^2)\, dx_2 dx_1$$

$$= \int_0^1 \left[3x_1^3 x_2 + \frac{3x_1 x_2^3}{3} \right]_0^{x_1} dx_1$$

$$= \int_0^1 3x_1^4 + x_1^4\, dx_1 = \int_0^1 4x_1^4\, dx_1$$

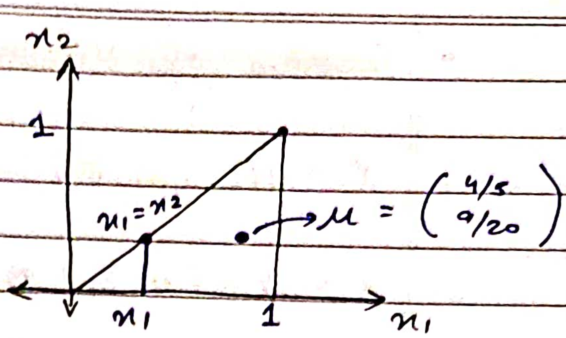
$$= \frac{4x_1^5}{5}\Big|_0^1 = \frac{4}{5}$$

$$\mu_2 = E(X_2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_2 \cdot f_X(x_1, x_2)\, dx_1 dx_2$$

$$= \int_0^1 \int_0^{x_1} x_2 \cdot (3x_1^2 + 3x_2^2)\, dx_2 dx_1$$

$$= \frac{9}{20}$$

$$\mu = \begin{pmatrix} 4/5 \\ 9/20 \end{pmatrix}$$



$$\Sigma = \begin{pmatrix} 2/75 & 3/200 \\ 3/200 & 77/1200 \end{pmatrix}$$

$$\Sigma = E(X \cdot X^T) - \mu \cdot \mu^T$$

$$= \begin{pmatrix} E(X_1^2) & E(X_1 \cdot X_2) \\ E(X_1 \cdot X_2) & E(X_2^2) \end{pmatrix} - \begin{pmatrix} \mu_1^2 & \mu_1 \mu_2 \\ \mu_1 \mu_2 & \mu_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

$$\underset{\downarrow\ >0}{\sigma_{11}} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_1^2 \cdot f_X(x_1, x_2)\, dx_1\, dx_2 - \frac{16}{25} =$$

$$\underset{\downarrow\ >0}{\sigma_{22}} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_2^2 \cdot f_X(x_1, x_2)\, dx_1\, dx_2 - \frac{81}{400} =$$

$$\sigma_{12} = \sigma_{21} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_1 \cdot x_2 \cdot f_X(x_1, x_2)\, dx_1\, dx_2 - \underbrace{\frac{4}{5} \cdot \frac{9}{20}}_{\frac{36}{100}} =$$

- **Multivariate Normal Distribution** a.k.a multivariate Gaussian distribution

$a_1 X_1 + a_2 X_2 + \dots a_D X_D \rightarrow \mathcal{N}(\mu, \sigma^2)$

$\mu \rightarrow$ real vector $\qquad \Sigma \rightarrow$ symmetric positive definite real matrix

The joint pdf of D-dimensional normal distribution is given by

$$\boxed{f_X(x_1, \dots, x_D) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)}$$

where $|\Sigma| = \det(\Sigma)$

joint pdf $f_X(x) > 0$ for all $x \in \mathbb{R}^D$

$X \sim N_D(\mu, \Sigma)$. for $D = 1$, we get univariate normal distribution

papergrid
Date: / /

- **Properties of Multivariate Normal Distribution**

1. $X \sim N_D(\mu, \Sigma) \Rightarrow X_i \sim N(\mu_i, \sigma_{ii})$ for $i = 1, \dots D$

2. $X \sim N_D(\mu, \Sigma)$, $X = \begin{pmatrix} \tilde{X}_1 \\ \tilde{X}_2 \end{pmatrix}$

$\Rightarrow \tilde{X}_1 \sim N_D(\tilde{\mu}_1, \Sigma_{11})$, $\tilde{X}_2 \sim N_D(\tilde{\mu}_2, \Sigma_{22})$

3. $\tilde{X}_1, \tilde{X}_2$ statistically independent $\Leftrightarrow \Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$

i.e $\tilde{X}_1, \tilde{X}_2$ uncorrelated.

most imp.
4. $X \sim N_D(\mu, \Sigma) \Rightarrow Y = AX + b \sim N_C(A\mu + b, A\Sigma A^T)$

Q Consider $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}^T \sim N_2(\mu, \Sigma)$

$\mu = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ $\Sigma = \begin{pmatrix} 4 & 1/2 \\ 1/2 & 3 \end{pmatrix}$

Determine pdf of $Y = X_1 + X_2$

$Y = AX + b$
$\downarrow$
$A = (1, 1)$ $b = 0$

Find $(A\mu + b)$ and $A\Sigma A^T$

$\cancel{\mu = \begin{pmatrix} 4/5 \\ 9/20 \end{pmatrix}}$ $\cancel{\Sigma = \begin{pmatrix} \frac{2}{75} & \frac{3}{200} \\ \frac{3}{200} & \frac{77}{1200} \end{pmatrix}}$

$\underbrace{Y = (1,1)}_{A \in \mathbb{R}^{1 \times 2}} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = X_1 + X_2$

$E(Y) = A \cdot \mu = 1 + 2 = 3$

$V(Y) =$ "$Cov(Y)$" $= A\Sigma A^T = (1,1) \begin{pmatrix} 4 & \frac{1}{2} \\ \frac{1}{2} & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 8$

$\boxed{Y \sim \mathcal{N}(3,8)}$

* 2.5 CONDITIONAL DISTRIBUTIONS

Recap: Conditional Probability $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$; $P(B) > 0$

- Conditional Density and Conditional Distribution
  conditional distribution of a random vector X given Y has occured.

→ Let $(X, Y) \in \mathbb{R}^{D+C}$ be a continuous random vector with the joint pdf $f_{(X,Y)}(x,y)$ and the marginal pdfs $f_X(x)$ and $f_Y(y)$.

Then conditional density of X under Y is given by

$$\boxed{f_{X|Y}(x) = f_{X|Y}(x|y) = \begin{cases} \dfrac{f_{(X,Y)}(x,y)}{f_Y(y)} & \text{if } f_Y(y) > 0 \\ f_X(x) & \text{if } f_Y(y) = 0 \end{cases}}$$

→ The corresponding distribution is the conditional distribution of X under $Y = y$, in short $P_{X|Y} = y$.

For events B, $P_{X|Y=y}(B) = P(X \in B \mid Y = y)$

$$= \int_B f_{X|Y}(x|y)\, dx_1 \ldots dx_D$$

→ the marginal pdf of X can be obtained as

law of total probability $P(A) = \sum_{i=1}^{n} P(A|B_i)\cdot P(B_i)$

$$f_X(x) = \int_{\mathbb{R}^C} f_{X|Y}(x|y) \cdot f_Y(y)\, dy_1 \ldots dy_C$$

The conditional expectation of $g(X)$ under $Y = y$ can be expressed as

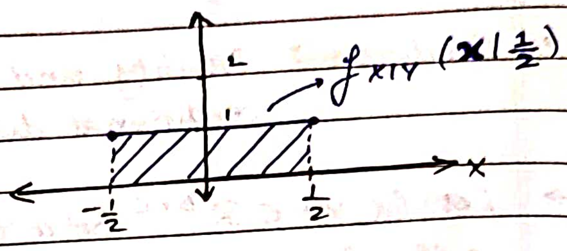
$$E(g(X) \mid Y = y) = \int_{\mathbb{R}^D} g(x) \cdot f_{X|Y}(x|y)\, dx_1 \ldots dx_D$$

Q Exercise – Consider $Y \sim U(0,1)$
conditional distribution of the random variable X under $Y=y$ is assumed to be characterized by conditional density

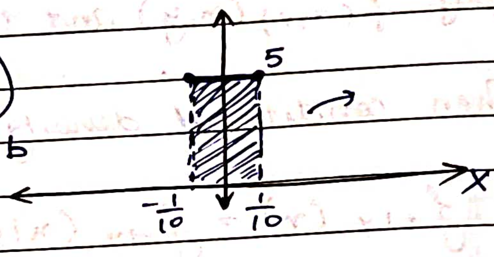
$$f_{X|Y}(x|y) = \frac{1}{2y} \cdot \mathbb{I}_{[-y,y]}(x)$$ i.e it should hold

that $X|Y=y \sim U(-y,y)$. Determine the marginal density $f_X(x)$ of a random variable X.

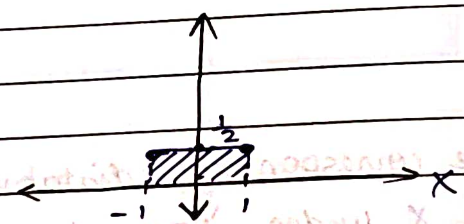
$$Y = \frac{1}{2} : X|Y = \frac{1}{2} \sim U\left(\underset{a}{\frac{-1}{2}}, \underset{b}{\frac{1}{2}}\right) \quad (-y,\ y)$$



$$Y = \frac{1}{10} : X|Y = \frac{1}{10} \sim U\left(\underset{a}{\frac{-1}{10}}, \underset{b}{\frac{1}{10}}\right)$$



$$Y = 1 : X|Y = 1 \sim U(-1,1)$$



marginal pdf of X : $f_X(x) = \int f_{X|Y}(x|y) . f_Y(y)\, dy , \ldots dy$

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X|Y}(x|y) . \underbrace{f_Y(y)}_{\text{pdf of uniform dist b/w } (0,1)} dy$$

$$f_Y(y) = \mathbb{I}_{[0,1]}\, y = \begin{cases} 1, & y \in [0,1] \\ 0, & \text{otherwise} \end{cases}$$

$$\therefore f_X(x) = \int_0^1 f_{X|Y}(x|y)\, dy = \int_0^1 \frac{1}{2y} \underbrace{\mathbb{I}_{[y,y]}(x)}_{\begin{cases} 1; & x \in [-y,y] \\ 0; & \text{otherwise} \end{cases}} dy$$

indicator function – returns 1 if x is element of set.
returns 0 " " " not " " "

papergrid

Date: / /

$$\therefore f_X(x) = \int_0^1 \frac{1}{2y} \mathbb{1}_{[-y,y]}(x)\, dy = \int_{|x|}^1 \frac{1}{2y}\, dy = \frac{1}{2} \ln(y) \Big|_{|x|}^{1}$$

$$= \frac{1}{2} \underbrace{\ln(1)}_{0} - \frac{1}{2} \ln|x| = -\frac{1}{2} \ln|x| \text{ for}$$

$$x \in [-1,1] \setminus \{0\}$$

excluding 0.