Random Experiment:

Random experiments are experiments whose outcome is unpredictable.

Sample Space:

The set of all outcomes of a random experiment. It is denoted by 'Ω' or 'S'.

Event:

An event is a subset of sample space. The types of event are:

- Sure event
- Impossible event.

Sure event:

Whenever the experiment is performed, the occurence of this event is inevitable.

Impossible events:

Whenever the experiment is performed, the non-occurence of this event is inevitable.

When can we say an event has occured?

When the outcome of the experiment belongs to the event, then we say that the event has occurred.

* Ω is a sure trivial event.
* ϕ - null event is an impossible event.

Let A and B be two events.

Suppose $A \subseteq B$,

then, occurence of $A \Rightarrow$ Occurence of B.

but, occurence of $B \not\Rightarrow$ occurence of A.

Mutually Exclusive events:

Let A and B be mutually exclusive events.

then, $A \cap B = \phi$

Occurence of A prevents occurence of B and vice versa.

g. Throwing a die.

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

$O = \{1, 3, 5\}$

$E = \{2, 4, 6\}$

O and E are mutually exclusive events.

* In any experiments, an event and its complement are always mutually exclusive.

In general, $A_1, A_2, \ldots, A_n$ are mutually exclusive

if $A_1 \cap A_2 \cap \ldots \cap A_n = \phi$

$\bigcap_{i=1}^{n} A_i = \phi$

Collectively exhaustive set:

The sets $A_1, A_2, \ldots, A_n$ are said to be collectively exhaustive set, if their union is equal to the given ~~e~~ set.

Let $A_1, A_2, \ldots, A_n$ be events and S be the Sample Space.

$$A_1 \cup A_2 \cup \ldots \cup A_n = S$$

Mutually Exclusive and Collectively Exhaustive:

A collection of non-empty events which are mutually exclusive and collectively exhaustive forms a partition of the ~~sps~~ Sample Space.

Any collection of sets that forms a partition are known as covering (i.e said to cover the Sample Space)

Probability:

Probability is a measure of chance for an event to occur.

We can calculate probability in three ways,

* Classical approach
* Relative approach
* Arithmetic approach.
* Subjective approach

Classical approach:

E is ~~the~~ h no. of ways out of a total n possible ways, all of which are ~~li~~ equally likely to occur, then,

$$P(E) = \frac{\#(E)}{\#(S)}$$

Disadvantages of this approach:

* It is assumed that all events are equally likely to occur.

* When the sample space is infinite, the probability always zero.

Note:

Here probability is calculated apriori (i.e) ~~ex~~ without performing the experiment.

Relative Frequency approach:

If an experiment is performed n no. of times (n is large) and an event E occurs h no. of times, then the probability is estimated as

$$\boxed{P(E) = \lim_{n \to \infty} \frac{h}{n}}$$

Limitations:

* Estimate may be different at different times.
* Repetition may not be feasible for destructive, dangerous and expensive experiments.

* The limit may not exist.
* Large is vague.

Note:

* An approach based on experiment or observation is called empirical approach.
* This is a postpriori approach.

Subjective approach:

It is often not possible to assign probability based on empirical terms, in such situation probability is based on subjective judgement of an individual.

* It is based on expert's opinion.
* Assign probabilities based on subjective approach, if it is not possible ~~...~~ to assign probability based on logical or empirical approach.

Note:

An approach based on experiment or observation is called as empirical approach.

Probability Model:

It is a mathematical description of an uncertain situation.

2 main elements are,

* sample space
* probability law.

Probability law:

It assigns a number $P(A)$ to each event in the sample space. $P(A)$ is called the probability of A.

* Event space is the set of all events.
* When events have single outcomes it is called as atomic events.

The probability law satisfies the following axioms

Axiom 1: $P(A) \geq 0$ (non-negativity axiom)

Axiom 2: $P(S) = 1$ (normalization)

Axiom 3: For any list of pairwise disjoint or mutually exclusive events $A_1, A_2, \dots$

$$P[A_1 \cup A_2 \cup \dots] = P(A_1) + P(A_2) + \dots$$

(Additivity axiom)

Example 1: Toss a coin

$S = \{H, T\}$

Events : $\emptyset$, $\{H\}$, $\{T\}$, $\{HT\}$

Assume that its a fair coin, then the probabilit

$P(H) = P(T) = 0.5$ (probability law)

$P[A \cup B] = P(A) + P(B) = 0.5 + 0.5 = 1.$

Example 2: Roll a dice.

Sample Space $S = \{(1,1), (1,2), \ldots, (1,6),$
$(2,1), (2,2), \ldots, (2,6),$
$\vdots \quad \vdots \quad \vdots$
$(6,1), (6,2), \ldots, (6,6)\}$

$P[(i,j)] = \frac{1}{36}$ (probability law)

Results :

i) $0 \leq P(A) \leq 1$

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

iii) $P(\emptyset) = 0$

iv) If $A \subseteq B$, then $P(A) \leq P(B)$
and $P(B-A) = P(B) - P(A)$

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

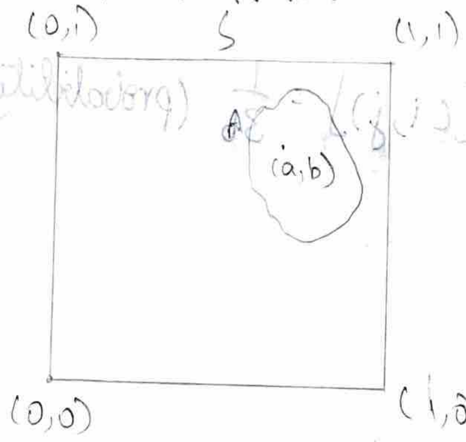
Discrete uniform probability law:

If the sample space has n outcomes, all of which are equally likely, then the probality law specifies probability for events containing of single outcomes, as

$$P(\{s_1\}) = P(\{s_2\}) = \dots = P(\{s_n\}) = \frac{1}{n}$$

Continuous uniform probability law:

Consider a unit square, the experiment is selecting a point in the square randomly. All the points in the square are equally likely to be selected.



$P(A) \propto$ Area of A

$$P(A) = \frac{\text{Area of A}}{\text{Area of S}}$$

$$= \frac{\text{Area of A}}{1}$$

Result 1: $P(\phi) = 0$

Proof:

A and ϕ are m.e for any event A.

$$A = A \cup \phi$$

$$P(A) = P(A \cup \phi)$$

$$P(A) = P(A) + P(\phi)$$

$$P(\phi) = 0$$

Result 2: $P(\bar{A}) = 1 - P(A)$

Proof:

$$A \cup \bar{A} = S$$

Here $A, \bar{A}$ are m.e by axiom 3 and 2

$$P(A) + P(\bar{A}) = P(S)$$

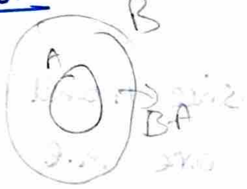
$$P(A) + P(\bar{A}) = 1$$

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

Result 3: If $A \subseteq B$ then $P(A) \leq P(B)$ — (i)

and $P(B-A) = P(B) - P(A)$ — (ii)

Proof:



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

A and B-A are m.e

$$P(B) = P[A \cup (B-A)]$$

$$P(B) = P(A) + P(B-A) \quad \text{— (1)}$$

$$P(B-A) = P(B) - P(A) \quad \text{— (ii)}$$

From $1 \Rightarrow P(A) \leq P(B)$

Result 4:

$$0 \leq P(A) \leq 1$$

Proof:

$P(A) \geq 0$ by axiom-1

W.K.T, $A \subseteq S$,

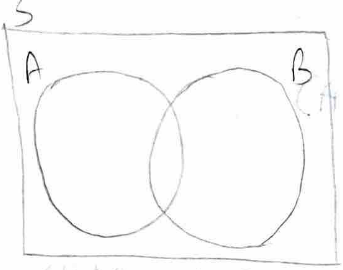
if $A \subseteq S$, then $P(A) \leq P(S)$

$P(A) \leq 1$

$$0 \leq P(A) \leq 1$$

Result 5: Addition rule

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



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

$P(A \cup B) = P[A \cup (B - A)]$

$= P(A) + P(B-A)$ — ① since A and B-A are m.e

Now, $B - A = B - A \cap B$ where $A \cap B \subseteq B$

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

$= P(B) - P(A \cap B)$ — ②

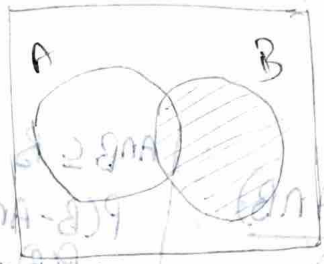
Sub ② in ①,

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

Conditional probability:

Let A, B be 2 events.

If we know that the outcome is in B, what is the probability that the outcome is in A?



By knowing that the outcome is in B, the sample space is reduced to B. That is, the new sample space is B. Now, we look for the favourable outcomes of A in B.

eg: Toss a die.

1) What is the probability that the outcome is <4 $= \frac{1}{2}$

2) If we know that the outcome is odd no. $B = \{1, 3, 5\}$ what is the probability of the number being < 4. $= \frac{2}{3}$

Probability of A given B is written as $P(A|B)$

$$P(A|B) = \frac{\#(A \cap B)}{\#(B)} = \frac{P(A \cap B)}{P(B)} \quad \text{provided } P(B) \neq 0$$

Note:

* Conditional probability is also a probability measure
* It also satisfies the three axioms,

- $P(A|B) \geq 0$
- $P(\Omega|B) = 1$; $P(B|B) = 1$
- $P[A_1 \cup A_2 \cup \ldots \cup A_n | B] = P(A_1|B) + \ldots + P(A_n|B)$

Prove $P(\bar{A}|B) = 1 - P(A|B)$

LHS:

$$P(\bar{A}|B) = \frac{P(\bar{A} \cap B)}{P(B)} = \frac{P(B - A \cap B)}{P(B)}$$

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

$$= \frac{P(B) - P(A \cap B)}{P(B)}$$

$$= \frac{\cancel{P(B)}}{\cancel{P(B)}} - \frac{P(A \cap B)}{P(B)}$$

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

Prove: $P(A \cup B \mid C) = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$

RHS :

$$= \frac{P(A \cap C)}{P(C)} + \frac{P(B \cap C)}{P(C)} - \frac{P(A \cap B \cap C)}{P(C)}$$

$$= \frac{P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)}{P(C)}$$

$$= \cancel{P(A) + P(C) - P(A \cup C)} + \cancel{P(B) + P(C)} - \cancel{P(C}$$

$$= \frac{1}{P(C)} \left[P(A \cap C) + P(B \cap C) - P(A \cap B \cap C) \right]$$

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

$$\boxed{= \frac{1}{P(C)} \left[P((A \cap C) \cup (B \cap C)) \right]}$$

$$= \frac{1}{P(C)} \left[P(A \cup B \cap C) \right]$$

$$= P(A \cup B \mid C)$$

Multiplication rule:

Think of an event E. We say E has occurred if both A and B have occurred.

$$E = A \cap B.$$

Suppose $E = A_1 \cap A_2 \cap \ldots \cap A_n$,

we say that E has occurred when $A_1, \ldots, A_n$ have occurred.

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

(or)

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

$$\boxed{P[A_1 \cap \ldots \cap A_n] = P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 . A_2) \times \ldots \times P(A_n|A_1 \ldots A_{n-1})}$$

We can write $P(A \cap B \cap C)$ as,

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

Independent events:

Two events A and B are independent, if the occurrence of one event has no influence on the occurrence of the other event.

A and B are independent events, if one of the following statements are true,

* $P(A/B) = P(A)$
* $P(B|A) = P(B)$
* $P(A \cap B) = P(A) . P(B)$

These three statements are equivalent statements.

Mutually exclusive vs independent events.

* Mutually exclusive events cannot happen simultaneously
  $A \cap B = \emptyset \Rightarrow P(A \cap B) = 0$

* Independent events can happen simultaneously.
  $P(A \cap B) = P(A) . P(B)$

Can two events A and B be both mutually exclusive as well as independent events?

No, ~~unless~~ ~~both~~ until both events have non zero probabilities, they cannot be both mutually exclusive as well as independent events.

(i.e) $P(A \cap B) = 0 = P(A \cap B) = P(A) . P(B)$

only if $P(A) . P(B) = 0$

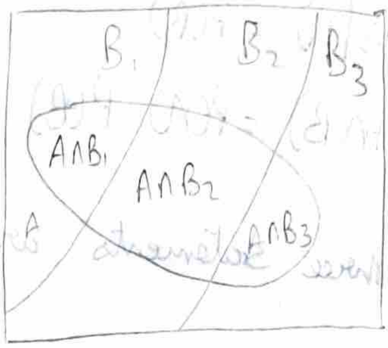
(i.e) $P(A) = 0$

or

$P(B) = 0$.

Total probility and Bayes' rule:

Let $B_1, \ldots, B_n$ be a list of M.e and exhaustive events and A be any events.



From the diagram,

$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_n)$

$A \cap B_1, A \cap B_2, \ldots, A \cap B_n$ are mutually exclusive events.

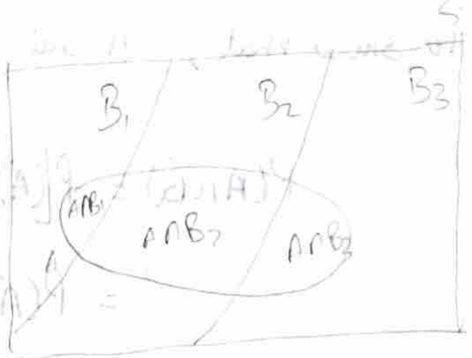
$P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_n)$ (axiom-3)

$$\boxed{P(A) = P(B_1) \cdot P(A|B_1) + \ldots + P(B_n) \cdot P(A|B_n)}$$

This is called the total probability of A.

Bayes' rule:

Let A be any event and $B_1, B_2, \ldots, B_n$ be mutually exclusive and exhaustive events.



Then, $P(B_k|A) = \dfrac{P(A\cap B_k)}{P(A)}$

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

$$\boxed{P(B_k|A) = \frac{P(B_k).P(A|B_k)}{P(A)}}$$

Result on independent events:

A and B are independent events iff

A and $\bar{B}$ are independent events;

$\bar{A}$ and B are independent events;

$\bar{A}$ and $\bar{B}$ are independent events;

That is, these 4 are equivalent statements.

Proof:

Assume A and B are independent events.

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

To show that, A and $\bar{B}$ are independent events.

$$P(A \cap \bar{B}) = P[(A) - (A \cap B)] \quad (A \cap B \subseteq A)$$

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

$$= P(A) - P(A).P(B)$$

$$= P(A)[1 - P(B)]$$

$$P(A \cap \bar{B}) = P(A).P(\bar{B}).$$

$\therefore$ proved that A and $\bar{B}$ are independent events

To show that, $\bar{A}$ and B are independent events

$$P(\bar{A} \cap B) = P[B - (A \cap B)] \quad (A \cap B \subseteq B)$$

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

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

$$= P(B)[1 - P(A)]$$

$$= P(B).P(\bar{A})]$$

$$P(\bar{A} \cap B) = P(\bar{A}).P(B)$$

$\therefore$ proved that $\bar{A}$ and B are independent events

To show that, $\bar{A}$ and $\bar{B}$ are independent events.

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

$\therefore$ proved that $\bar{A}$ and $\bar{B}$ are independent events.

Problems

1. An elevator with 2 passengers stops at 2^{nd}, 3^{rd} and 4^{th} floors. If it is equally likely that a passenger gets off at any of the three floors, find the probability that the two passengers get off at different floors.

Sol.

Let $a_i b_j$ denotes passenger 1 gets off at i^{th} floor and passenger 2 gets off at j^{th} floor for $i,j = 2,3,4$

Sample Space $S = \{ a_i b_j \mid i = j = 2, 3, 4 \}$

Since all outcomes are equally likely,

$$P(\{(a_i,b_j)\}) = \frac{1}{n} = \frac{1}{9}$$

Let A be the event that both passengers get off at different floors.

$$A = \{a_ib_j \mid i,j = 2,3,4,\ i \neq j\}$$

$$P(A) = \frac{6}{9} = \frac{2}{3}$$

2. If two dies are rolled, what is the probability that their sum is,

i) 8

ii) neither 7 nor 11

Sample Space $S = \{(i,j) \mid i,j = 1,2,\dots,6\}$

$$P(\{(i,j)\}) = \frac{1}{36}$$

i) $P(\text{sum}=8) = \dfrac{5}{36}$

m.e

ii) $P(\text{sum} \neq 7 \wedge \text{sum} \neq 11) = 1 - P(\text{sum}=7 \vee \text{sum}=11)$

$$= 1 - [P(\text{sum}=7) + P(\text{sum}=11)]$$

$$= 1 - \left[\frac{6}{36} + \frac{2}{36}\right]$$

$= 1 - \frac{8}{36}$

$= \frac{28}{36}$

$P(\text{sum} \neq 7 \cap \text{sum} \neq 11) = \frac{28}{36}$

3. Find the probability of a 4 turning up in two tosses of a fair die.

$P[\text{getting a } 4] = \frac{6+6-1}{36}$

$= \frac{11}{36}$

4. A card is drawn from a well shuffled deck of 52 cards. What is the probability that it is a spade or an ace?

$|S| = 52$

$P[\text{spade or ace}] = \frac{13+4-1}{52}$

$= \frac{16}{52} = \frac{4}{13}$

5. A box contains 6 red, 5 white and 4 black balls. 4 balls are taken at random. What is the probability that there is atleast one ball in each colour?

$|S| = {}^{15}C_4$

$\# S = 1365$

A : 2R 1W 1B

B : 1R 2W 1B

C : 1R 1W 2B

$$P(A) = \frac{6C_2 \times 5C_1 \times 4C_1}{1365} = \frac{300}{1365}$$

$$P(B) = \frac{6C_1 \times 5C_2 \times 4C_1}{1365} = \frac{246}{1365}$$

$$P(C) = \frac{6C_1 \times 5C_1 \times 4C_2}{1365} = \frac{180}{1365}$$

$$\therefore P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

$$= \frac{720}{1365}$$

6. The odds against A (person) solving a problem are 7 to 3 and odds in favour of B solving the same problem are 4 to 5. What is the probability that the problem is solved if A and B try independently?

$P(A) = \frac{3}{10}$

$P(B) = \frac{5}{9}$

A, B are independent

W.K.T, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= P(A) + P(B) - P(A) \cdot P(B)$

$= \frac{3}{10} + \frac{4}{9} - \frac{3}{10} \cdot \frac{4}{9}$

$= \frac{27+40}{90} - \frac{12}{90}$

$= \frac{67-12}{90}$

$= \frac{\cancel{55}^{11}}{\cancel{90}_{18}}$

$P(A \cup B) = \frac{11}{18}$

7. 4 people participate in a chess tournament. In the first round, A will play against B and C will play against D. In the next round, the winners will play for the title and the losers will play for the third position. List all the outcomes. Find the probability that A wins the trophy? Find the probability that B gets in the final.

All outcomes

~~ABCD~~

$S = \{$ ACDB, ACBD, ADCB, ADBC, BCAD, BCDA, BDAC, BDCA, CABD, CADB, CBAD, CBDA, DABC, DACB, DBAC, DBCA $\}$

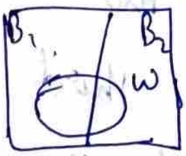
$P(\text{A winning}) = \frac{4}{16} = \frac{1}{4}$

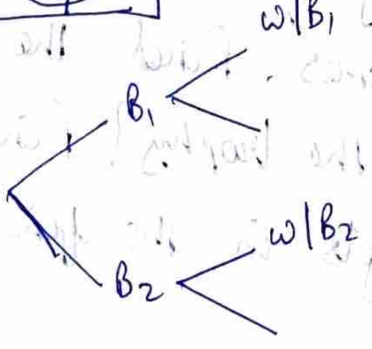
$P(\text{B getting in final}) = \frac{8}{16} = \frac{1}{2}$

8. Two boxes contain 3W, 4B and 4W, 3B balls. If a box is chosen and a ball is drawn from A, what is the probability that its a white ball

$P[\text{selecting white ball}] = \frac{1}{2} \cdot \frac{3}{7} + \frac{1}{2} \cdot \frac{4}{7}$

$= \frac{\cancel{7}\,1}{\cancel{14}\,2} = \frac{1}{2}$





$P(W) = \sum_{i=1}^{2} P(B_i) \cdot P(W|B_i)$

$= \frac{1}{2} \cdot \frac{3}{7} + \frac{1}{2} \cdot \frac{4}{7} = \frac{1}{2}$

9. Three horses A, B and C are in a race. A is twice as likely to win as B. B is twice as likely to win as C. What are their respective chances of winning

$P(B) = 2P(C)$

$P(A) = 2P(B) = 4P(C)$

A, B and C are M.E

$P(A \cup B \cup C) = 1$

$P(A) + P(B) + P(C) = 1$

$4P(C) + 2P(C) + P(C) = 1$

$7P(C) = 1$

$P(C) = \frac{1}{7}$

$P(A) = \frac{4}{7}$, $P(B) = \frac{2}{7}$, $P(C) = \frac{1}{7}$.

10. Among the workers in a factory, only 30% receive a bonus. Among those receiving bonus, only 20% are skilled. What is the probability that a randomly selected worker is skilled and receiving bonus.

Let B be the event of workers receiving bonus

Let S be the event that the worker is skilled

Given : $P(B) = 30\% = 0.3$

$P(S|B) = 20\% = 0.2$

$P(S \cap B) = P(B) \cdot P(S/B)$

$= 0.3 \times 0.2$

$P(S \cap B) = 0.06$

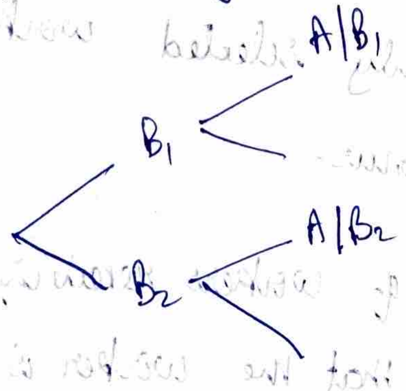
Probability that the worker receiving a bonus is skilled is 6%.

11. 2 boxes contains 5W, 3B and 4W, 5B balls. A box is chosen at random and two balls are drawn from it. What is the probability that one is white and other is black.

B_1 : Box 1 : 5W, 3B

B_2 : Box 2 : 4W, 5B

A : selecting 2 balls with 1W & 1B



$$P(A|B_1) = \frac{{}^5C_1 \times {}^3C_1}{{}^8C_2} = \frac{15}{28}$$

$$P(A|B_2) = \frac{{}^4C_1 \times {}^5C_1}{{}^9C_2} = \frac{20}{36}$$

$$P(A) = P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2)$$

$$= \frac{1}{2} \cdot \frac{15}{28} + \frac{1}{2} \cdot \frac{\cancel{20}\,10}{36}$$

$$= \frac{15}{56} + \frac{10}{36}$$

12. There are 10 urns. Each of three contains 1w and 9B, each of another three contains 9w and 1B, each of remaining four contains 5w and 5B balls respectively. One of the urns are selected at random and a ball is drawn from it turns out to be white. What is the probability that an urn containing 9w and 1B balls was selected.

$$P(A) = P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2) + P(B_3) \cdot P(A|B_3)$$

$$= \frac{\cancel{3}}{10} \times \frac{1}{\cancel{3}} \left(\frac{{}^1C_1}{{}^{10}C_1} \right) + \frac{\cancel{3}}{10} \times \frac{1}{\cancel{3}} \left(\frac{{}^9C_1}{{}^{10}C_1} \right) + \frac{\cancel{4}}{10} \times \frac{1}{\cancel{4}} \left(\frac{{}^5C_1}{{}^{10}C_1} \right)$$

13. In a certain assembly plant, three machines M_1, M_2 and M_3 make 30%, 45% and 25% respectively. It is known that 2%, 3% and 2% of the products made by each machine are defective.

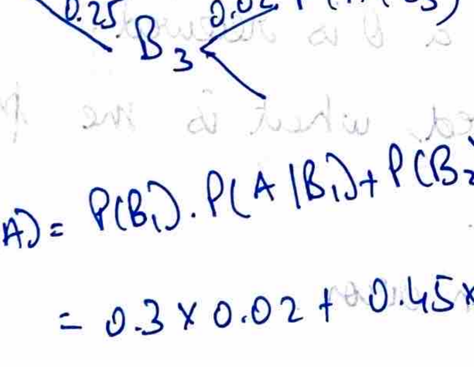
i) What is the probability that a randomly selected finished product is defective.

ii) If a product was found to be defective, what is the probability that it was made by machine M_3.

B_1 - product from M_1

B_2 - product from M_2

B_3 - product from M_3.



$$P(A) = P(B_1).P(A|B_1) + P(B_2).P(A|B_2) + P(B_3).P(A|B_3)$$

$$= 0.3 \times 0.02 + 0.45 \times 0.03 + 0.25 \times 0.02$$

$$P(A) = 0.0245$$

$$P(B_3|A) = \frac{P(B_3) \cdot P(A|B_3)}{P(A)}$$

$$= \frac{(0.25)(0.02)}{0.0245}$$

$$P(B_3|A) = 0.204\%$$

14. A binary communication channel carries data as one of two types of signals 0 and 1. Due to noise a transmitted 1 is sometimes received as 0 and vice versa. ~~The~~ Assume the probability of 0.95 that a transmitted 0 is correctly received as 0 and a probability of 0.85 that a transmitted 1 is correctly received as 1. Also assume that ~~0.45~~ with probability 0.45, a 0 is transmitted. Determine the following

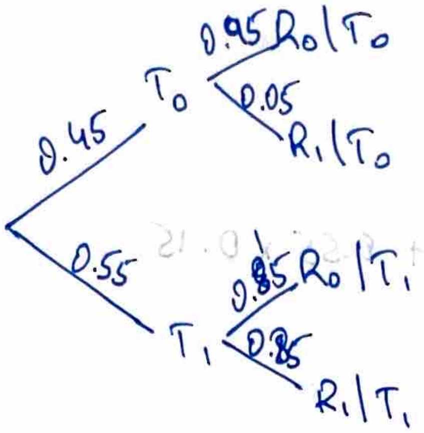
i) the probability that a 1 is received.

ii) the probability that a 0 is received.

iii) given a 0 is received, what is the probability that 1 was transmitted.

iv) the probability of an error.

For $i = 0, 1$ let T_i : Transmitting i
R_i : Receiving i.



i) $P(R_1) = P(T_0) \cdot P(R_1|T_0) + P(T_1) \cdot P(R_1|T_1)$

$$= \frac{45}{100}\left(\frac{5}{100}\right) + \frac{55}{100}\left(\frac{85}{100}\right)$$

$$= \frac{225}{10000} + \frac{\cancel{825}\ 4675}{10000}$$

$$= 0.0225 + \cancel{0.0825}\ 0.4675$$

$$P(R_1) = \cancel{0.105}\ 0.49$$

ii) $P(R_0) = P(T_0) \cdot P(R_0|T_0) + P(T_1) \cdot P(R_0|T_1)$

$$= \frac{\cancel{45}\ 9}{\cancel{100}\ 20}\left(\frac{\cancel{95}\ 19}{\cancel{100}\ 20}\right) + \frac{\cancel{55}\ 11}{\cancel{100}\ 20}\left(\frac{\cancel{185}\ 3}{\cancel{100}\ 20}\right)$$

$$= \frac{171}{400} + \frac{\cancel{187}\ 33}{400}$$

$$P(R_0) = \frac{\cancel{388}\ \cancel{204}\ 51}{\cancel{400}\ 100} = \cancel{0.1885}\ 0.51$$

iii) $P[T_1|R_0] = \dfrac{P(T_1) \cdot P(R_0|T_1)}{P(R_0)}$

$= \dfrac{0.55 \times 0.15}{0.51}$

$= 0.161$

iv) $P(\text{error}) = 0.45 \times 0.05 + 0.55 \times 0.15$

$= 0.105$

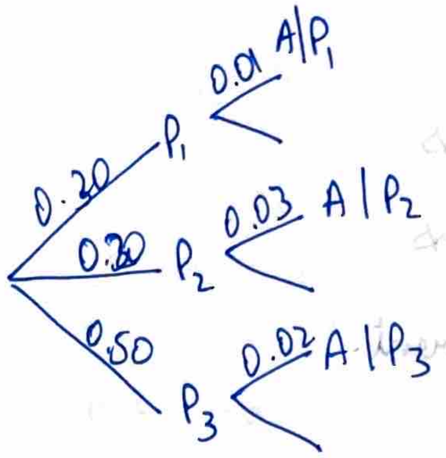
15. A manufacturing firm employees 3 analytical plants P_1, P_2, P_3 for design and development of its product. For cost reasons, all 3 are used in different times. Plants 1, 2 and 3 are used for 30%., 20%. and 50% of the products respectively. The defect rate for the three plants are 0.01, 0.03 and 0.02.

i) If a random product was observed, what is the probability that it is defective.

ii) If a product is defective, which plant was most likely used and thus responsible?

A - defective

P_1, P_2, P_3 - plants



i) $P[A] = 0.30 \times 0.01 + 0.30 \times 0.03 + 0.50 \times 0.02$

$= 0.003 + 0.006 + 0.010$

$P(A) = 0.019$

ii) $P[P_1|A] = \dfrac{P(P_1) \cdot P(A|P_1)}{P(A)}$

$= \dfrac{0.003}{0.09} = 0.157$

$P[P_2|A] = \dfrac{0.006}{0.019} = 0.315$

$P[P_3|A] = \dfrac{0.010}{0.019} = 0.526$

Since $P(P_3|A) > P(P_2|A) > P(P_1|A)$

P_3 is most likely to have produced the defective phone.

Pairwise Independence $\not\Rightarrow$ Independence.

Eg: Toss a coin twice

H_1 : Head in the 1st toss

H_2 : Head in the 2nd toss

D : The o/c s are different.

$S = \{HH, HT, TH, TT\}$

$P(H_1) = \frac{2}{4} = \frac{1}{2}$

$P(H_2) = \frac{2}{4} = \frac{1}{2}$

$P(D) = \frac{2}{4} = \frac{1}{2}$

$P(H_1 \cap H_2) = \frac{1}{4}$; $P(H_1).P(H_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

$P(H_2 \cap D) = \frac{1}{4}$; $P(H_2).P(D) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

$P(D \cap H_1) = \frac{1}{4}$; $P(D).P(H_1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

$\therefore$ H_1, H_2 and D are pairwise independent.

$P(H_1 \cap H_2 \cap D) = 0$ but $P(H_1).P(H_2).P(D) = \frac{1}{8}$

$\therefore$ H_1, H_2 and D are not independent but are pairwise independent.

$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C) \not\Rightarrow$ Pairwise Independence.

Eg: Roll a dice twice

A: O/C is 1, 2 or 3 in the 1st roll

B: O/C is 3, 4 or 5 in the 1st role

C: The sum is 9

$\# S = 36$

$P(A) = \frac{1}{2}$ $\quad P(B) = \frac{1}{2}$ $\quad P(C) = \frac{4}{36} = \frac{1}{9}$

$P(A \cap B) = \frac{9}{36} = \frac{1}{4}$; $P(A) \cdot P(B) = \frac{1}{4}$

$P(B \cap C) = \frac{3}{36} = \frac{1}{12}$; $P(B) \cdot P(C) = \frac{1}{18}$

$P(A \cap C) = \frac{1}{36}$; $P(A) \cdot P(C) = \frac{1}{18}$

$\therefore$ A, B and C are not pairwise independent

$P(A \cap B \cap C) = \frac{1}{36}$, $P(A) \cdot P(B) \cdot P(C) = \frac{1}{36}$

$\therefore$ A, B and C are not independent ~~but~~ and not pairwise independent.

Result:

A, B and C are independent if,

$$P(A \cap B) = P(A) \cdot P(B)$$
$$P(B \cap C) = P(B) \cdot P(C)$$
$$P(A \cap C) = P(A) \cdot P(C)$$
$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C).$$

Problems.

1. Is it possible for two events A and B where $P(A) = 0.6$, $P(B) = 0.7$ and $P(A \cap B) = 0.2$

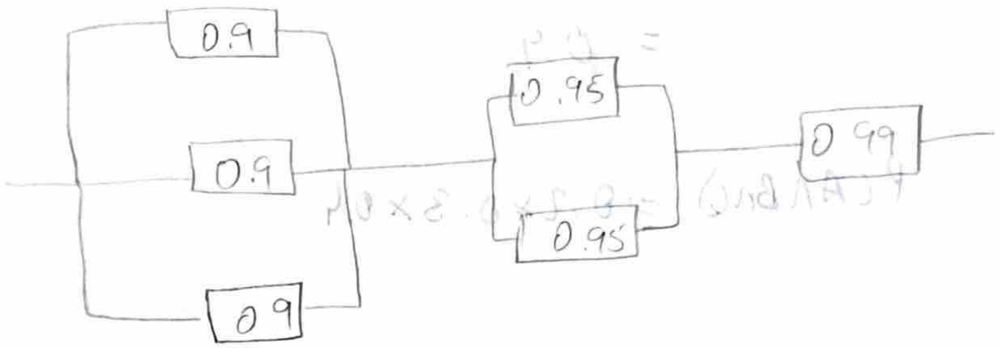
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.6 + 0.7 - 0.2$$
$$= 1.1$$
$$P(A \cup B) > 1$$

$\therefore$ It is not possible.

2. A fair coin is tossed twice and no. of heads obtained is observed. Since, the no of heads obtained can be 0, 1 or 2. The probability of obtaining no heads is $\frac{1}{3}$, comment on this statement.

Since the sample space contains 4 outcomes, the probability of no. of heads is $\frac{1}{4}$ not $\frac{1}{3}$.

3.



This circuit operates only if there's a path of functional devices from left to right. The probability that each device operates is shown. Assume that the device failed independently. What is the probability that the device operates?

$= [1 - P(\overline{C_1})] [1 - P(\overline{C_2})] [1 - P(\overline{C_3})]$

$= [1 - (1-0.9)(1-0.9)(1-0.9)] [1 - (1-0.95)(1-0.95)]$

[

4. A, B, C are m.e events with $P(A) = 0.2$, $P(B) = 0.3$ and $P(C) = 0.4$. Find $P(A \cup B \cup C)$ & $P(A \cap B \cap C)$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

$$= 0.9$$

$$P(A \cap B \cap C) = 0.2 \times 0.3 \times 0.4$$

5. A coin is tossed twice. Let A

A = atmost one head on the two tosses

B = one head and one tail in both tosses.

and A & B are independent.

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

$$P(A) = \frac{3}{4} \qquad P(B) = \frac{2}{4}$$

$$P(A) \times P(B) = \frac{3}{\cancel{4}_2} \times \frac{\cancel{2}}{4} \qquad P(A \cap B) = \frac{1}{2}$$

$$= \frac{3}{8}$$

$$P(A \cap B) \neq P(A) \cdot P(B)$$

6. Alice and Bob go target shooting. Both shoot at the target at same time. Suppose Alice hits the target at probability 0.7 and Bob hits with probability 0.4 independently.

i) Given that exactly one shot hit, find the probability that it was Bob's hit.

ii) Given that the target is hit, find the probability that Bob hit it.

$$P(A) = 0.7, \quad P(B) = 0.4$$

i)
$$P(\text{Bob's hit}) = \frac{0.4 \times 0.3}{0.12 + 0.42}$$

$$= \frac{0.12}{0.54}$$

$$= \frac{2}{9}$$

ii)
$$P(\text{Bob's hit}) = \frac{0.4 \times 0.3 + 0.7 \times 0.4}{0.12 + 0.42 + 0.28}$$

$$= \frac{0.4}{0.82}$$

$$= \frac{40}{82}$$

$$= \frac{20}{41}$$

6. <u>sol.</u>

$P(\text{Alice}) = 0.7$

$P(\text{Bob}) = 0.4$

$P(A \cap B) = P(A) \cdot P(B)$ (A, B independent)

then,

$\bar{A}, B$ independent

$A, \bar{B}$ independent

$\bar{A}, \bar{B}$ independent.

i) To find

$$P\Big[\underbrace{\text{Bob's hit}}_{E} \;\Big|\; \underbrace{\text{Exactly one shot hit}}_{F}\Big]$$

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

~~P(P(E)(F))/7~~

$F = (A \cap \bar{B}) \cup (\bar{A} \cap B)$ (exactly one shot either Alice's or Bob's)

$$P(F) = P(A \cap \bar{B}) + P(\bar{A} \cap B)$$
$$= P(A) \cdot P(\bar{B}) + P(\bar{A}) \cdot P(B)$$
$$= 0.7 \times 0.6 + 0.3 \times 0.4$$
$$= 0.54.$$

$$P(E|F) = \frac{P(\bar{A} \cap B)}{P(F)} = \frac{0.3 \times 0.4}{P(F)} = \frac{0.12}{0.54}$$

ii) To find P(~~Bob's~~ hit | ~~exactly one shot~~) P(Bob hit | Target B hit)

$$P(C|D) = \frac{P(C\cap D)}{P(D)}$$

Target is hit,

$A\cap B$, $\bar{A}\cap B$, $A\cap\bar{B}$ (bob didn't hit)

$$= \frac{P(A\cap B) + P(\bar{A}\cap B)}{P(A\cup B)}$$

$$= \frac{P(A\cap B)+P(\bar{A}\cap B)}{1-P(\bar{A}\cap\bar{B})}$$

$$= \frac{P(A\cap B)+P(\bar{A}\cap B)}{1-P(\bar{A}).P(\bar{B})}$$

$$= \frac{0.7\times 0.4+0.3\times 0.4}{1-0.3\times 0.6}$$

$$= \frac{0.40}{0.82}$$

Gambler's ruin problem.

Two gamblers A and B played a game in which a coin starts. If heads turn up, then A gets \$1 from B. If tails turns up, then A losses \$1 to B. Assume that the probability for heads and tails are p and $1-p$. Suppose A and B have \$a and \$b initially. If this game is played ~~successively~~ successfully, what is the probability

i) A will be ruined
ii) The game goes forever with nobody winning

Random Variables.

In most of the random experiments the outcomes are numerical. In other experiments, they are not.

A random variable X is a function that assigns numerical values to the outcomes of a random experiment

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

Example 1:

Three satellites are launched into space

$\Omega = \{SSS, SSF, SFS, SFF, FSS, FSF, FFS, FFF\}$

X: No of satellites that go into orbit.

$X = \{0, 1, 2, 3\}$

State Space (X) or
Space (X)

0, 1, 2, 3 are also called states.

$X = 0 \Rightarrow \{FFF\}$

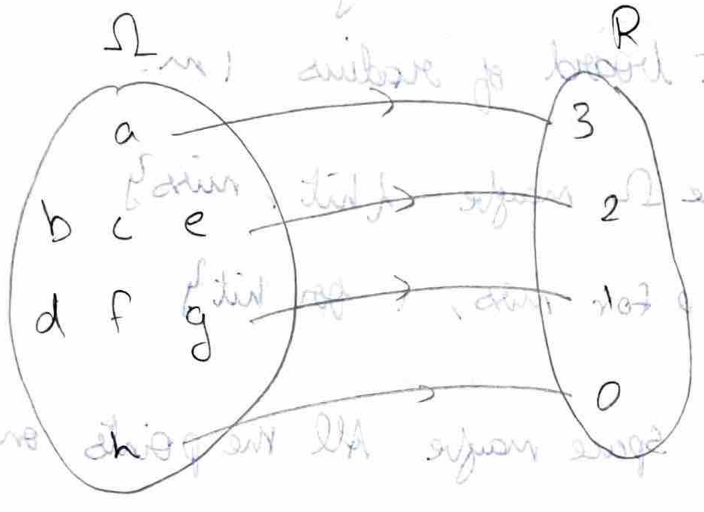
$X = 1 \Rightarrow \{SFF, FSF, FFS\}$

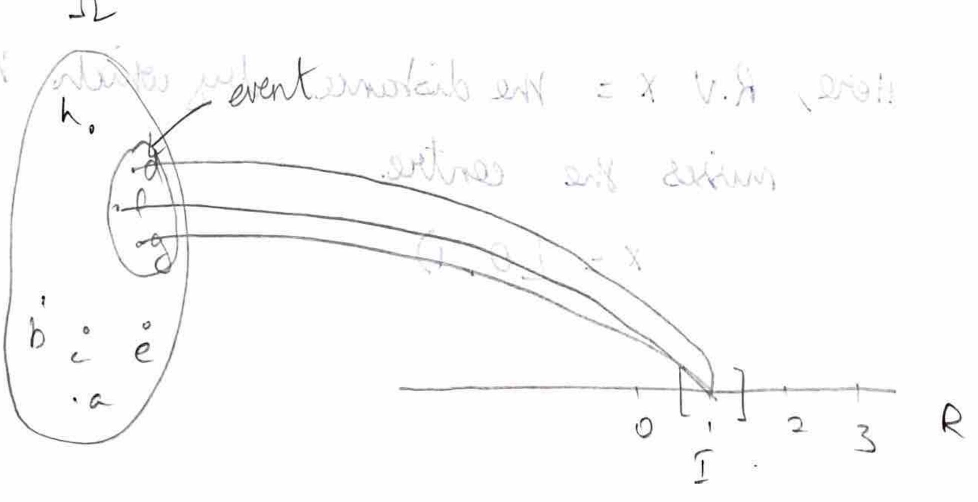
$X = 2 \Rightarrow \{SSF, SFS, FSS\}$

$X = 3 \Rightarrow \{SSS\}$

$[X = x]$ denotes the element containing the o/c s that are assigned to the value x.

A function $X : \Omega \to R$ is a random variable if for every interval $I \subseteq R$, $\{s : X(s) \in I\}$ is an event.





Example 2:

A person is tossing a coin repeatedly until head appears.

$\Omega = \{H, TH, TTH, TTTH, \dots\}$

R.V X: no. of trials for success

success : head appears.

$X = \{1, 2, 3, \dots\}$

Example 3:

A person is throwing a dart on a circular dart board of radius 1 m.

Sample Space Ω maybe {hit, miss}

R.V X = {0 for miss, 1 for hit}

or the Sample Space maybe All the points on the board.

Here, R.V X = the distance by which he/she misses the centre.

$X = [0, 1)$

Example 4:

A person is waiting for a message from his / her friend from 6:00 AM 4th Aug 2022

Ω = All time distances from 6:00 AM 4th Aug 2022

R.V X : Amount of time waiting.

$X = [0, \infty)$

Random Variables are classified into two types based on the nature of their state spaces

- Discrete Random Variables
- Continuous Random Variables

Discrete RV	Continuous RV
Takes discrete values or countable no. of values.	Takes any value in the interval of its range.
	Continuous or uncountable values.
eg: eg1, eg2, eg3 i)	eg: eg 3ii), eg4

Note:

A function of random variable defines another random variable

Example:

Toss 2 dice

X_1 = O/C on the 1^{st} dice

X_2 = O/C on the 2^{nd} dice.

$Y = X_1 + X_2$ → sum of the outcomes of the dice

$Z = X_1^2$

Z, Y are also random variables.

Advantages of using Random Variables

* Compactness and Compatability
* Dimensionality Reduction.

Probability distribution of a Random Variable:

A probability distribution specifies the probability of all events associated with the Random variable.

For a discrete random variable, the probability distribution is given by the probability mass function.

Probability mass function (pmf):

Probability mass function of a random variable X is given by,

$$\boxed{p(x) = P[X=x] \quad \forall \; x \in X}$$

Properties of pmf:

i) $p(x) \geq 0$

ii) $\sum_{x \in X} p(x) = 1$.

Example 1:

Toss two fair coins

Let X = ~~Head on the 1st toss~~ No. of heads

Y = ~~Head on the 2nd toss~~ No. of tails

Find pmf of X and Y

Ans

$X = \{0, 1, 2\}$

x	0	1	2
$p(x)$	1/4	1/2	1/4

$Y = \{0, 1, 2\}$

~~x~~	0	1	2
$p(y)$	1/4	1/2	1/4

Let I be the indicator of first toss landing heads

$$I = \begin{cases} 1 & \text{for } HH, HT \\ 0 & \text{for } TH, TT \end{cases}$$

pmf

I	0	1
$p_I(i)$	1/2	1/2

Example 2:

Roll 2 dice

Let X and Y denote the outcomes on the first and second die respectively.

$T = X + Y$ → sum of the dice.

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

X	1	2	3	4	5	6
$p_X(x)$	1/6	1/6	1/6	1/6	1/6	1/6

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

Y	1	2	3	4	5	6
$p_Y(y)$	1/6	1/6	1/6	1/6	1/6	

$T = \{2, 3, \dots, 12\}$

T	2	3	4	5	6	7	8	9	10	11	12
	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Cumulative distributive function. (cdf):

This function also specifies probabilities of events of a random variable.

→ While pmf is only for discrete rv, cdf is defined for any random variables.

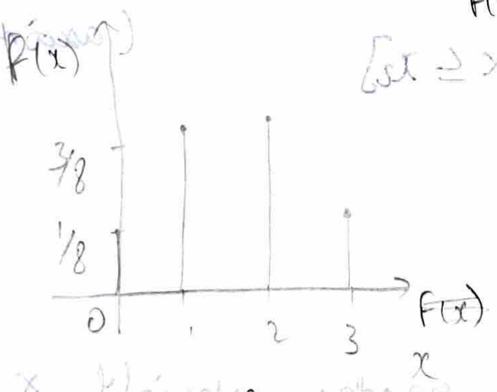
Definition:

The cumulative distributive function of a rv is given by

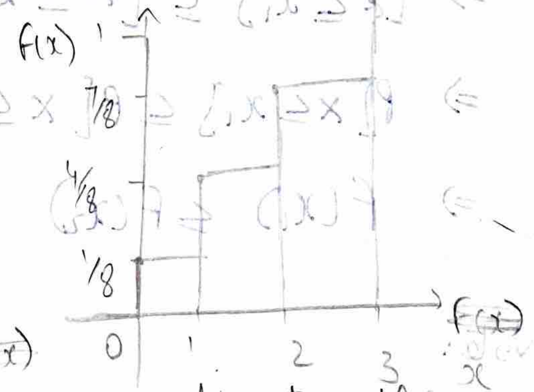
$$F(x) = P[X \leq x] \quad \forall \; x \in R$$

For the sattellite problem,

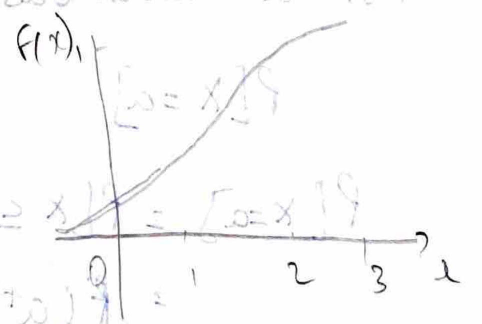
X	0	1	2	3
p(x)	1/8	3/8	3/8	1/8
F(x)	1/8	4/8	7/8	1



pmf



discrete cdf



continuous cdf

Properties of cdf:

* The graph is a continuous curve.

Three properties are,

* F is non decreasing
* F is right continuous, $F(a) = \lim_{x \to a^+} F(x)$
* $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$

Proof for property ①:

To prove F is non decreasing.

$$x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$$

LHS:

$$x_1 < x_2$$

$$\Rightarrow [X \leq x_1] \subseteq [X \leq x_2]$$

$$\Rightarrow P[X \leq x_1] \leq P[X \leq x_2] \quad \text{(axiom)}$$

$$\Rightarrow F(x_1) \leq F(x_2)$$

Note:

For a continuous random variable X,

$$P[X = a]$$

$$P[X=a] = P[X \leq a] - P[X < a]$$
$$= F(a^+) - F(a^-)$$

Since X is continuous random variable,

$$F(a^+) = F(a^-) = F(a)$$

$$P[X = a] = 0$$

Thus, for a continuous random value, probability that X takes a particular value is zero.

Probability computing using $F(x)$:

For any random variable X, we may be interested to compute the probability of the following events.

	Events	Probability
1.	$[X \leq a]$	$F(a)$
2.	$[X < a]$	$F(a^-)$
3.	$[X \geq a]$	$1 - F(a^-)$
4.	$[X > a]$	$1 - F(a)$
5.	$[X = a]$	$F(a) - F(a^-)$
6.	$[X \neq a]$	$1 - [F(a) - F(a^-)]$
7.	$[a < X \leq b]$	$F(b) - F(a)$
8.	$[a < X < b]$	$F(b^-) - F(a)$
9.	$[a \leq X \leq b]$	$F(b) - F(a^-)$
10.	$[a \leq X < b]$	$F(b^-) - F(a^-)$

Expected Value | Expectation | Mean of a random va

In a casino game, the probability of winning \$1, \$2 and \$3 are 0.3, 0.08 and 0.02 respectively. The probability of losing one dollar is 0.6. If u play this game n times, what is your expected gain

$= [(-1)(0.6) + (1)(0.3) + (2)(0.08) + 3(0.02)]n$

$= -0.08n$

The expected gain per game = -0.08 \$

r.v X = Gain

$X = \{-1, 1, 2, 3\}$

Pmf

X	-1	1	2	3
P(x)	0.6	0.3	0.08	0.02

Expected gain $E(x) = (-1)(0.6) + 1(0.3) + 2(0.08) + 3(0.02)$

$= -0.08$

Note:

* The expected value of a discrete random variable X is given by

$$E(x) = \sum_{x \in X} x \cdot P(x)$$

provided that the sum converges.

$$\text{Weighted Average} = \frac{\sum x_i f_i}{\sum f_i}$$ $f \to$ frequency of x

* $E(x)$ is the weighted average of all values of the random variable X in which the weight assigned to each $x \in X$ is $P(x)$.

Note:

Why pmf is called Probability Mass Function?

Total Mass (1) of an object is distributed among values of RV such that mass at x_i is $P(x_i)$

* In a physical sense, $E(x)$ is the centre of mass of an object distributed at all values of the random variable in which the mass at x_i is $p(x_i)$

Properties:

$* E(K) = K$

$* E(KX) = K.E(X)$

$* E(X+Y) = E(X)+E(Y)$

} Expectation is line[ar]

Example

A box contains 5 balls, each of two [balls] are marked 1\$, each of another two balls [are] marked 5\$ and the remaining one ball is marked 50\$. You select two balls at rando[m] from this box and win the sum of the amount marked on the balls. To play this ga[me] you have to pay 10\$. What is the expected g[ain]? Is this a fair game?

2	2	1
1\$	5\$	15\$

outcomes = $\{(1,1), (5,5), (1,5), (15,1), (5,15)\}$

2	10	6	16	20
-10	-10	-10	-10	-10

Gain X : -8 0 -4 6 10

$P(X)$ $\frac{{}^2C_2}{{}^5C_2}$ $\frac{{}^2C_2}{{}^5C_2}$ $\frac{{}^2C_1 \times {}^2C_1}{{}^5C_2}$ $\frac{{}^2C_1 \times {}^1C_1}{{}^5C_2}$ $\frac{{}^2C_1 \times {}^1C_1}{{}^5C_2}$

$1/10$ $1/10$ $4/10$ $2/10$ $2/10$

Expected gain, $E(X) = \sum_{x \in X} x P(x)$

$= -8(1/10) + 0 + -4(4/10) + 6(2/10) + 10(2/10)$

$= 0.8$

It is not a fair game, because,

$E(x) \neq 0$.

Both persons should not gain or lose anything

Variance of random variable:

Suppose you measure a quantity,

let 'x' denote the error. (true value - the values obtained in measurement)

$E(x)$ must be 0

$E(x) = 0$.

* Because, in measuring a quantity a large number of times, positive and negative errors occurs with equal probability.

Therefore, $E(X)$ only is not sufficient about to know about the nature of the random variable. It doesn't tell the variability in the values of the random variable, X.

Consider three random variables, X, Y, Z defined by $X = 0$ with probability 1.

$$Y = \begin{cases} -1 & \text{with probability } 1/2 \\ 1 & \text{with probability } 1/2 \end{cases}$$

$$Z = \begin{cases} -10 & \text{with probability } 1/4 \\ 0 & \text{with probability } 1/2 \\ 10 & \text{with probability } 1/4. \end{cases}$$

A drunken man has n keys. One of which opens the door of his house. If he tries to open the door using one key one and independently. What is the mean and variance of the no. of trials to open the door, If the wrong keys are i) eliminated ii) not eliminated

r.v X : No. of trials to open the door.

ii) not eliminated.

X	1	2	3
$P(x)$	$\frac{1}{n}$	$\left(1-\frac{1}{n}\right)\left(\frac{1}{n}\right)$	$\left(1-\frac{1}{n}\right)^2\left(\frac{1}{n}\right)$

$$E(x) = \sum x\,p(x), \quad x = 1, 2$$

$$= 1.p + 2.q.p + 3.q^2 p + \dots$$

$$= p\left[1 + 2q + 3q^2 + \dots\right]$$

$$= p\left[\frac{1}{(1-q)^2}\right] \qquad (p = 1-q)$$

$$= \frac{1}{p}$$

Note :

$$1 + x + x^2 + \dots = \frac{1}{1-x}$$

$$1 + 2x + 3x^2 + \dots = \frac{1}{(1-x)^2}$$

$$\frac{1}{2}\left[1.2 + 2.3x + 3.4x^2 + \dots\right] = \frac{1}{(1-x)^3}$$

$$E(x^2) = \sum x^2 p(x), \quad x = 1, 2, 3$$

$$= \sum x^2 q^{x-1} (p)$$

$$= \sum [(x-1)x + x] q^{x-1} p$$

$$= p \sum (x-1)x \, q^{x-1} + p \sum x \, q^{x-1}$$

$$= p [0 + 1.2q + 2.3q^2 + 3.4q^3 + \dots] + \frac{1}{p}$$

$$= pq [1.2 + 2.3q + 3.4q^2 + \dots] + \frac{1}{p}$$

$$= pq \frac{2}{(1-q)^3} + \frac{1}{p} \quad (p = 1-q)$$

$$= \frac{2pq}{p^{\cancel{3}2}} + \frac{1}{p}$$

$$E(x^2) = \frac{2q}{p^2} + \frac{1}{p}$$

$$\text{Var}(x) = E(x^2) - [E(x)]^2$$

$$= \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2}$$

$$= \frac{2q - 1}{p^2} + \frac{1}{p}$$

$$= \frac{2q + p - 1}{p^2} \quad [-q = p - 1]$$

$$= \frac{2q - q}{p^2}$$

$$= \frac{q}{p^2}$$

i) eliminated. X : no of trials to open the door

X	1	2	3 ...	n
$p(x)$	$\frac{1}{n}$	$\frac{\cancel{(n-1)}}{n} \frac{1}{\cancel{n-1}}$	$\frac{\cancel{(n-1)}}{n} \frac{\cancel{(n-2)}}{\cancel{(n-1)}} \frac{1}{\cancel{(n-2)}}$	$\frac{1}{n}$

$$E(x) = \sum x \, p(x) \quad x = 1, \ldots, n$$

$$= 1 \cdot \frac{1}{n} + 2 \cdot \frac{1}{n} + \ldots + n \frac{1}{n}$$

$$= \frac{1}{n} (1 + 2 + \ldots + n)$$

$$= \frac{1}{\cancel{n}} \left(\frac{\cancel{n}(n+1)}{2} \right)$$

$$E(x) = \frac{n+1}{2}$$

$$E(x^2) = \sum x^2 p(x)$$

$$= 1^2 \cdot \frac{1}{n} + 2^2 \cdot \frac{1}{n} + \ldots + n^2 \frac{1}{n}$$

$$= \frac{1}{n} (1^2 + 2^2 + \ldots + n^2)$$

$$= \frac{1}{\cancel{n}} \left(\frac{\cancel{n}(n+1)(2n+1)}{6} \right)$$

$$E(x^2) = \frac{(n+1)(2n+1)}{6}$$

$$Var(x) = E(x^2) - [E(x)]^2$$

$$= \frac{(n+1)(2n+1)}{6} - \left[\frac{n+1}{2}\right]^2$$

$$= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$

$$= \frac{n+1}{2}\left[\frac{2n+1}{3} - \frac{n+1}{2}\right]$$

$$= \frac{n+1}{2}\left[\frac{4n+2-3n-3}{6}\right]$$

$$= \frac{n+1}{2}\left[\frac{n-1}{6}\right]$$

$$var(x) = \frac{n^2-1}{12}$$

A box contains ten discs with radii 1, 2, ...

Find the expected value of the circumference of a randomly selected disc

Radii : 1, 2, ..., 10

RV X : circumference.

X	2π	4π	6π	...	20π
$P(x)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	...	$\frac{1}{10}$

$$E(x) = \sum x \, p(x) \qquad x = 2\pi, \ldots, 20\pi$$

$$= \frac{2\pi}{10} + \frac{4\pi}{10} + \frac{6\pi}{10} + \ldots + \frac{20\pi}{10}$$

$$= \frac{2\pi}{10}\left(1 + 2 + 3 + \ldots + 10\right)$$

$$= \frac{2\pi}{10}\left(\frac{10(10+1)}{2}\right)$$

$$= \frac{2\pi}{\cancel{10}\,\cancel{2}}\left(\overset{11}{\cancel{55}}\right)$$

$$E(x) = 11\pi$$

$$E(x^2) = \sum x^2 \, p(x)$$

$$= \frac{4\pi^2}{10} + \frac{16\pi^2}{10} + \ldots + \frac{400\pi^2}{10}$$

$$= \frac{4\pi^2}{10}\left(1^2 + 2^2 + \ldots + 10^2\right)$$

$$= \frac{4\pi^2}{10}\left(\frac{\overset{5}{\cancel{10}}(10+1)(\overset{7}{\cancel{20+1}})}{\cancel{6}\,\cancel{3}}\right)$$

$$= \frac{\overset{2}{\cancel{4}}\pi^2}{\underset{5}{\cancel{10}}}\left(\overset{77}{\cancel{385}}\right)$$

$$= 154\pi^2$$

$$Var(x) = E(x^2) - [E(x)]^2$$

$$= 154\pi^2 - \cancel{110}\ 121\pi^2$$

$$= 33\pi^2$$

$$Var(x) = 33\pi^2$$

3. Cdf of a random variable is given by

$$F(x) = \begin{cases} 0 & , x < 0 \\ \frac{x}{4} & , 0 \le x < 1 \\ \frac{1}{2} & , 1 \le x < 2 \\ \frac{x}{12} + \frac{1}{2} & , 2 \le x < 3 \\ 1 & , x \ge 3 \end{cases}$$

Find i) $P[x < 2]$

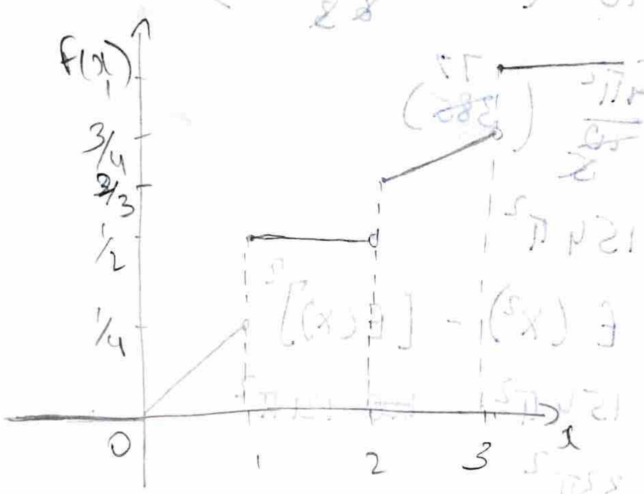
ii) $P[x = 2]$

iii) $P[1 \le x < 3]$

iv) $P[x > \frac{3}{2}]$

v) $P[x = \frac{5}{2}]$

vi) $P[2 < x \le 7]$



i) $P[x < 2] = F(2^-) = \frac{1}{2}$

ii) $P[x = 2] = F(2) - F(2^-) = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$

iii) $P[1 \le x < 3] = F(3^-) - F(1^-) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$

iv) $P[x > \frac{3}{2}] = 1 - F(\frac{3}{2}) = 1 - \frac{1}{2} = \frac{1}{2}$

v) $P[x = \frac{5}{2}] = F(\frac{5}{2}) - F(\frac{5}{2}^-) = \frac{17}{24} - \frac{17}{24} = 0$

vi) $P[2 < x \le 7] = F(7) - F(2) = 1 - \frac{2}{3} = \frac{1}{3}$

4. The sales x (in \$'000) of a continuous store on a randomly selected day is a random variable. It's cdf,

$$F(x) = \begin{cases} 0 & , x < 0 \\ \frac{x^2}{2} & , 0 \le x < 1 \\ k(4x - x^2) & , 1 \le x < 2 \\ 1 & , x \ge 2 \end{cases}$$

Suppose this store's total sales on any given day is less than \$2000, find the value of k.

i) Let A = Total sales \$ in any given day is b/w \$500 and \$1500

B = Total sales is over \$1000

Find P(A) and P(B)

ii) Are A and B independent.

Sol.

Given:

Sales on any given day is < \$2000

$$P[x<2]=1$$

$$F(2^-)=1$$

$$K(8-4)=1$$

$$4K=1$$

$$\boxed{K=\frac{1}{4}}$$

A = Total sales on any given day is b/w \$500 and \$1500

B = Total sales is over \$1000

i) To find P(A) and P(B)

$$P(A) = P[\text{sales b/w } \$500 \ \& \ \$1500]$$

$$= P[0.5<x<1.5]$$

$$= F(1.5^-)-F(0.5)$$

$$= \frac{1}{4}\left(4\left(\frac{3}{2}\right)-\left(\frac{3}{2}\right)^2\right)-\frac{\left(\frac{1}{2}\right)^2}{2}$$

$$= \frac{1}{4}\left(6-\frac{9}{4}\right)-\frac{1}{8}$$

$$= \frac{1}{4}\left(\frac{15}{4}\right)-\frac{1}{8}$$

$$P(A) = \frac{15}{16}-\frac{2}{16} = \frac{13}{16}$$

$P(B) = P[\text{sales over } \$1000]$

$= P[x > 1]$

$= 1 - F(1)$

$= 1 - \frac{1}{4}(4-1)$

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

$P(B) = \frac{1}{4}$

ii) A and B independent?

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

$P[A \cap B] = P[0.5 < x < 1.5,\ 1 < x]$

$= P[1 < x < 1.5]$

$= F(1.5^-) - F(1)$

$\cancel{= \frac{(\frac{1}{2})^2}{2} - \frac{1}{4}(4-1)}$

$\cancel{= \frac{1}{8} - \frac{3}{4}}$

$\cancel{= \frac{1}{8} - \frac{6}{8}}$

$= \frac{1}{4}\left(4\left(\frac{3}{2}\right) - \left(\frac{3}{2}\right)^2\right) - \frac{1}{4}(4-1)$

$= \frac{15}{16} - \frac{12}{16}$

$P[A \cap B] = \frac{3}{16}$

$$P(A).P(B) = \frac{13}{16} \cdot \frac{1}{4} = \frac{13}{64}$$

$$P(A \cap B) \neq P(A).P(B)$$

$\therefore$ A and B are not independent.

5. Two dice are rolled. Let X denote the absolute difference of the outcomes on the two dice. Find the pmf of X. Also find the mean and variance.

Special discrete Distribution:

Types of distribution:

* Bernoulli
* Binomial
* Geometric
* Poisson
* Discrete - uniform
* Negative Binomial.

Bernoulli distribution:

A random variable x is said to be a Bernoulli ~~distribution~~ random variable if it takes the values 0 and 1 with $P[x=1] = p$ and $P[x=0] = q$ where

$0 \le p \le 1$ and $p+q=1$

pmf:

$p \to$ success probability.

x	1	0
$P(x)$	p	q

We write $x \sim \text{Bernoulli}(p)$

Mean $E(x) = \sum x \cdot p(x) = 1 \cdot p + 0 \cdot q = p$

~~Variance~~ $E(x^2) = \sum x^2 p(x) = 1^2 \cdot p + 0^2 \cdot q = p$

Variance $(x) = E(x^2) - [E(x)]^2 = p - p^2 = p(1-p)$

Variance $(x) = pq$

Note:

For any event A, there is an associated random variable I_A defined by

$$I_A = \begin{cases} 1 & \text{if } A \text{ happens} \\ 0 & \text{if } A \text{ doesn't happen.} \end{cases}$$

$I_A \sim \text{Bernoulli}(p)$ where $p = P(A)$. This particular random variable is called an indicator random variable.

Bernoulli Trial:

Its an experiment with only two outcomes success and failure is called a bernoulli trial.

$$x = \begin{cases} 1 & \text{if o/c is success} \\ 0 & \text{if o/c is failure.} \end{cases}$$

A bernoulli random variable on a bernoulli trial can be thought of as an indicator random variable that indicates whether the outcome is success.

5.

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

X	0	1	2	3	4	5
$p(x)$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$

$$E(x) = 0 \times \frac{6}{36} + 1 \times \frac{10}{36} + 2 \times \frac{8}{36} + 3 \times \frac{6}{36} + 4 \times \frac{4}{36} + 5 \times \frac{2}{36}$$

$$= 0 + \frac{10}{36} + \frac{16}{36} + \frac{18}{36} + \frac{16}{36} + \frac{10}{36}$$

$$E(x) = \frac{70}{36} = 1.944$$

$$E(x^2) = \sum x^2 p(x)$$

$$= 0 + \frac{10}{36} + \frac{32}{36} + \frac{54}{36} + \frac{64}{36} + \frac{50}{36}$$

$$= \frac{210}{36}$$

$$= 5.833$$

$$Var(x) = E(x^2) - [E(x)]^2$$

$$= 5.833 - 3.7791$$

$$= 2.0539$$

Binomial Distribution:

Consider a sequence of n iid Bernoulli trials each with success probability p.

iid - independent and identical distribution.

R.V x = No. of success in n trials

$x = \{0, 1, 2, \ldots, n\}$

$x \sim Binomial(n, p)$

For example,

For passing exams,

$x \sim Binomial(5, 0.5)$

$$\boxed{P(x) = P[X=x] = {}^nC_x p^x q^{n-x}} \quad x = 0, 1, 2$$

$$(q+p)^n = q^n + {}^nC_1 p^1 q^{n-1} + \dots + p^n$$

Since it uses binomial theorem, it is called as binomial distribution.

A binomial random variable can be thought of as sum of n iid bernoulli random variables.

$$X = \sum_{i=1}^{n} x_i$$

where x_i is the bernoulli random variable of each ~~n~~ iid bernoulli trials.

Mean, $E(X) = E(x_1 + \dots + x_n)$

$= E(x_1) + \dots + E(x_n)$

$= p + \dots + p$

$= np$

$Var(X) = Var(x_1 + \dots + x_n)$

$= Var(x_1) + \dots + Var(x_n)$

$= pq + \dots + pq$

$= npq$

Geometric Distribution:

Consider a sequence of iid Bernoulli trials (similar to Binomial), but the no. of trials is not fixed, each with success probability p.

X = No of trials for success, $X = \{1, 2, 3, \ldots\}$

$\cancel{X}$ = No of failures before success, $\cancel{X} = \{0, 1, 2, \ldots\}$

* Geometric distribution is also known as first success indicator.

pmf, $P(x) = P[X=x] = q^{x-1} p$, $x = 1, 2, \ldots$

No. of trials	No. of failures	$P(x)$
1	0	p
2	1	qp
3	2	$q^2 p$
$\vdots$	$\vdots$	$\vdots$
x	$x-1$	$q^{x-1} p$

pmf for Y, $P(Y) = P[Y=y] = q^x p$, $x = 0, 1, 2, \ldots$

For Drunken Man problem ii), Mean $E(x) = 1/p$

Variance : $E(x^2) - [E(x)]^2 = q/p^2$

Memoryless property of GD

$$P[\{X > m+n\} | X > M] = P[X > n]$$

(RHS) The prob that someone needs more than n trials ~~to pass~~ for success is the same as (LHS) the prob that she needs further n trials for success having already made more than m trials.

RHS

$$P[X>n] = \sum q^{x-1} p \quad , \quad x = n+1, n+2, \ldots$$

$$= p\{q^n + q^{n+1} + q^{n+2} + \ldots\}$$

$$= p q^n \{1 + q + q^2 + \ldots\}$$

$$= p q^n \left(\frac{1}{1-q}\right)$$

$$P[X>n] = q^n.$$

LHS

$$P[X > m+n | X > m]$$

$$= \frac{P[X > m+n, X > m]}{P[X > m]}$$

$$= \frac{P[X > m+n]}{P[X > m]}$$

$$= \frac{q^{m+n}}{q^m}$$

$$P[X > m+n | X > m] = q^n. \qquad LHS = RHS$$

Negative Binomial Distribution:

Consider a sequence of iid Bernoulli trials (similar to experiment to geometric distribution.) with success probability p.

R.V X = No of trials for r^{th} success

$X = \{r, r+1, r+2, \dots\}$

(To have r success, there must atleast be r trials)

pmf, $P(n) = P[X=n] = \binom{n-1}{r-1} p^{r-1} q^{n-r} \times p$

$$P(n) = \binom{n-1}{r-1} p^r q^{n-r}$$

Among n trials there are r successes and the last one must be a success. So leaving out the last one, we get $r-1$ successes among $n-1$ trials.

Connection between Negative Binomial and Binomial

Here, we need $r-1$ successes during the first $n-1$ trials, applying this in Binomial distribution

$X \sim Bino(n-1, p)$

$P[X = r-1] = {}^{n-1}C_{r-1}\, p^{r-1} q^{(n-1)-(r-1)}$

$= \binom{n-1}{r-1} p^{r-1} q^{n-r} \times p$ (last trial success)

Relation between negative Binomial and geometric distribution.

X_1 - No. of trials for 1^{st} success.

X_2 - No. of trials for 2^{nd} success after 1^{st} success

X_3 - No. of trials for 3^{rd} success after 2^{nd} success

$\vdots$

X_r - No. of trials for r^{th} success after r^{n-1} th succ

Let $X \sim$ Negative Bino (r, p)

X can be thought of as the sum of r iid geometric distribution random variable.

$X = X_1 + X_2 + \ldots + X_r$ where each $X_i \sim$ Geometric

Mean, $E(X) = E(X_1 + X_2 + \ldots + X_r)$

$= E(X_1) + E(X_2) + \ldots + E(X_r)$

$= \frac{1}{p} + \frac{1}{p} + \ldots + \frac{1}{p}$

$= \frac{r}{p}$

Variance : $Var(x) = Var(x_1) + \ldots + Var(x_r)$

$$= \frac{q}{p^2} + \frac{q}{p^2} + \ldots + \frac{q}{p^2}$$

$$= \frac{rq}{p^2}$$

Poisson distribution:

Popular distribution for modeling discrete data.

RV x : No of occurences of an event in a specified interval of time / space.

A random variable x is said to follow poisson distribution with parameter $\lambda > 0$ if its pmf is $P(x) = P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}$, $x = 0, 1, 2, \ldots$

$x \sim Poisson(\lambda)$

λ is the mean.

eg: for time interval.

No of arrivals to a bank in an hour.

No of α particles emitted by a radioactive material per sec.

No of phones sold by an ecommerce portal per hour.

for space interval,

No of typos per page in a book published by a reputed company

- Poisson distribution is mostly commonly used to model the distribution of rare events.

Relation between poisson's and binomial distribution

Binomial distribution becomes poisson's distribution when $n \to \infty$ and $p \to 0$, so that mean $= np$ is finite

We can use poisson's distribution to compute binomial probabilities if $n \geq 20$ and $p \leq 0.05$, taking $\lambda = np$.

Mean, $E(x) = \sum x . p(x) = \sum x \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum \frac{\lambda^x}{(x-1)!}$

$= \lambda e^{-\lambda} \sum \frac{\lambda^{(x-1)}}{(x-1)!} = \lambda e^{-\lambda} \{ 1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots$

$= \lambda \cancel{e^{-\lambda}} \cancel{e^{\lambda}}$

$E(x) = \lambda$

Variance, $E(x^2) = \sum x^2 p(x) = \sum x^2 \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum \frac{x \lambda^x}{(x-1)!}$

$= \lambda e^{-\lambda} \sum \frac{(x-1+1)}{(x-1)!} \lambda^{x-1}$

$$= \lambda e^{-\lambda} \left\{ \sum \frac{\lambda^{x-2}}{(x-2)!} + \sum \frac{\lambda^{x-1}}{(x-1)!} \right\}$$

$$= \lambda e^{-\lambda} \{ \lambda e^{\lambda} + e^{\lambda} \}$$

$$= \lambda^2 + \lambda$$

Variance, $Var(x) = E(x^2) - [E(x)]^2$

$$= \cancel{\lambda^2} + \lambda - \cancel{\lambda^2}$$

$$Var(x) = \lambda$$

	Ber dist	Bin dist	Geo dist	N Bin dist	Poisson
Mean	p	np	$1/p$	r/p	λ
Variance	pq	npq	q/p^2	rq/p^2	λ

Discrete Uniform distribution:

Let x be a number randomly selected from a finite $c = \{a, a+1, \dots, b\}$ of n elements. Then we says x follows uniform discrete distribution on the set c.

$x \sim$ discrete-uniform (c)

pmf, $P[x] = P[X = x] = \frac{1}{n}$, $a \le x \le b$, where

$$n = b - (a-1)$$
$$= b - a + 1$$

Mean, $E(x) = \sum x \cdot p(x) = a \cdot \frac{1}{n} + (a+1) \cdot \frac{1}{n} + \dots + b \cdot \frac{1}{n}$

$$= \left(\frac{b}{2}(b+1) - \frac{(a-1)a}{2} \right) \frac{1}{n}$$

$$= \frac{1}{n} \left(\frac{b(b+1)}{2} - \frac{(a-1)a}{2} \right) = \frac{b^2 + b - a^2 + a}{2}$$

$$= \left(\frac{b^2 - a^2 + a + b}{2}\right)\frac{1}{n}$$

$$= \frac{(b+a)(b-a) + (b+a)}{2n}$$

$$= \frac{(b+a)(b-a+1)}{2n}$$

$$= \frac{(b+a)\cancel{n}}{2\cancel{n}}$$

$$E(x) = \frac{a+b}{2}$$

$$E(x^2) = \sum x^2 px = a^2\frac{1}{n} + (a+1)^2\frac{1}{n} + \dots + b^2\frac{1}{n}$$

$$= \left(\frac{b(b+1)(2b+1)}{6} - \frac{(a-1)(a)(2a-1)}{6}\right)\frac{1}{n}$$

Example 1:

Four dice are rolled. What is the probability that atleast one six appears?

The random variable x can be,

$x \sim Bin(n, p)$ - No. of success in n trials

$x \sim Geo(p)$ - No. of trials for success.

$x \sim$ Negative $Bin(r, p)$ - No. of trials for r^{th} success.

$x \sim Poisson(\lambda)$ - No. of occurences of an event in a specified boundary.

$x \sim$ Discrete Uniform(c) - A number selected randomly from c.

Here, the random variable x follows,

x = No. of times 6 appears in 4 rolls.

$x \sim Bino(n, p)$

Here $n = 4$ and $p = \frac{1}{6}$

pmf, $p(x) = P[X = x] = {}^4C_x \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{4-x}, \quad x = 0, 1, \ldots, 4.$

To find P[atleast one six]

$= P[x \geq 1]$

$= 1 - P[x < 1]$

$= 1 - P[x = 0]$

$= 1 - q^n$

$= 1 - \left(\frac{5}{6}\right)^4$

In Bin,

$P[x = 0] = q^n$

$P[x = n] = p^n$

P[atleast one six] $= 1 - \left(\frac{5}{6}\right)^4$

...ple 2:

The probability of an item produced by a machine is defective is 0.05. What is the probability that the machine will produce a defective item in its 6^{th} run.

Random variable, X = No. of trials for first defective item.

$X \sim$ Geometric (p)

Here, $p = 0.05$

pmf, $P(x) = P[X = x] = q^{x-1} p$, $x = 1, 2, \ldots$

To find, $P[\text{defective item on } 6^{th} \text{ run}] = P[X = 6]$

$= q^5 p$

$= (0.95)^5 (0.05)$

Example 3:

The person has two taxis. The number of demands for a taxi in any given day follows poisson distribution with the mean of 1.5. What is the probability that on a given day

i) no taxi is used

ii) Some demands are refused.

Random variable, X = No. of demands for a taxi on a day

$X \sim$ Poisson's distribution (λ)

Here, $\lambda = 1.5$

pmf, $P(x) = P[X=x] = \frac{e^{-\lambda}\lambda^{x}}{x!}$, $x = 0, 1, 2, \ldots$

To find,

i) $P[\text{no taxi is used}] = P[X=0]$ In poisson's $P[X=0] = e^{-\lambda}$

$= e^{-1.5}$

ii) $P\left[\begin{matrix}\text{some demands}\\ \text{are refused}\end{matrix}\right] = P[X>2]$

$= 1 - P[X \le 2]$

$= 1 - P[X=0] - P[X=1] - P[X=2]$

$= 1 - e^{-1.5} - (e^{-1.5})(1.5) - \frac{e^{-1.5}(1.5)^2}{2!}$

Example 4:

In a factory manufacturing razor blades, there is a small chance of $\frac{1}{500}$ for any blade is defective. Use poisson's distribution to find the number of packes (each with 10 blades) containing

i) no defective blades.

ii) one defective blades

iii) two defective blades.

in a consignment of 10,000 packets.

Random Variable, x = No. of defective blades in a packet (of 10).

$x \sim$ Poisson (λ) where $\lambda = np = 10 \times \frac{1}{500} = \frac{1}{50} = 0.$

i) P[no defective] $= P[x=0] = e^{-\lambda} = e^{-0.02} = 0.980$

ii) P[1 defective] $= P[x=1] = e^{-0.02}(0.02) = 0.019$

iii) P[2 defectives] $= P[x=2] = \dfrac{e^{-0.02}(0.02)^2}{2!} = 0.000$

no defective blade $= .9801 \times 10000 = 9801$

one defective blade $= .0.196 \times 10000 = 196$

two defective blades $= 0.000196 \times 10000 = 1.96$

Example 5:

An oil company conducts a geological study that indicates that an exploratory oil well should have a 20% chance of striking oil. What is the probability that the first strike comes from the third well drill. What is the probability that the third strike comes on the seventh well drill.

i) X = No of drills for first strike

$X \sim$ Geometric (p) where $p = 20\% = 0.2$

$$p(x) = q^{x-1} p \qquad x = 1, 2, 3, \ldots$$

$$P[X=3] = (0.8)^2 (0.2) = 0.128$$

ii) X = No of drills for 3rd strike.

$X \sim NB(r, p)$ where $r = 3$, $p = 0.2$

$$P[X=n] = \binom{n-1}{r-1} p^r q^{n-r}$$

$$= \binom{6}{2} (0.2)^3 (0.8)^4$$

$$= 0.049$$

Continuous Probability Distributions:

pdf : probability density function:

For a continuous random variable with cdf F, the pdf f is the derivative of F, given by,

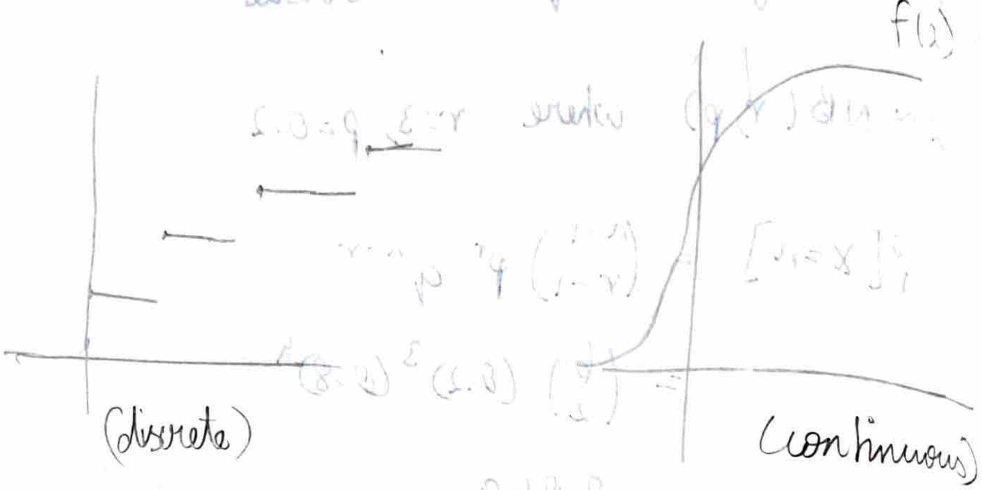
$$f(x) = F'(x)$$

Another definition.

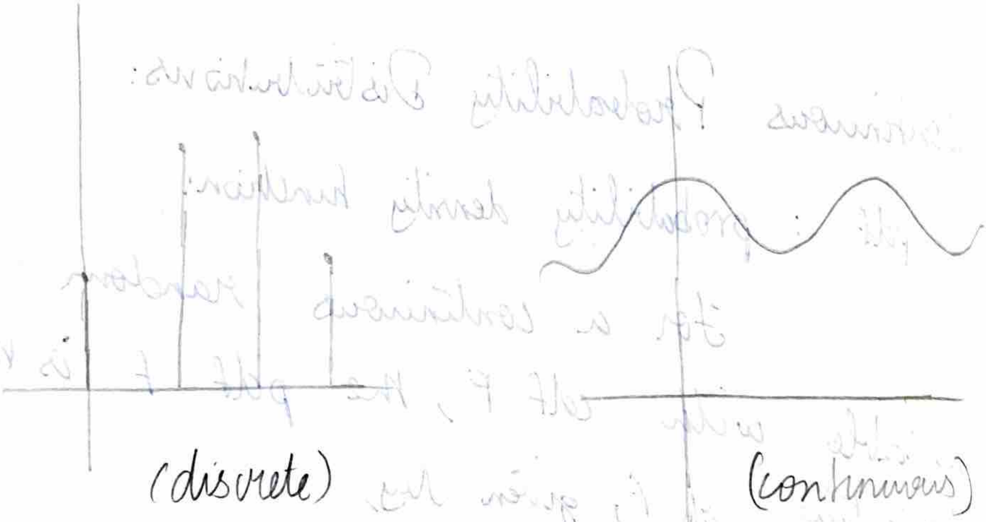
The function f is the pdf of continuous random variable X if,

$$\int_a^b f(x)\,dx = P[a < x < b]$$

cdf



pdf:



Note:

* $F(x)$ is a probability measure while $f(x)$ is not, unless it is multiplied by

an infinitesimal Δx

$$f(x)\,\Delta x = P[x < X < x+\Delta x]$$

$$F(x) = P[X \leq x]$$

$$f(x) = \lim_{\Delta x \to 0} \frac{P[x < X < x+\Delta x]}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x}$$

$$f(x) = F'(x)$$

Special continuous distributions:

continuous uniform distributions:

If X is a point selected randomly in an interval (a, b), then X follows continuous uniform distribution over (a, b)

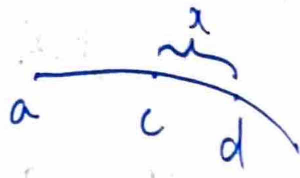
$$X \sim \text{uniform}(a, b)$$

pdf: $$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}$$

cdf: $$F(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & x \geq b \end{cases}$$

Note:

$$P[c < x < d] = \frac{d-c}{b-a}$$



Example.

1. Busses arrive at a bus stop at 15 mins intervals starting from 7:00 AM, (i.e) 7:00, 7:15, 7:30, 7:45 and so on. If a passenger arrives at the bus stop at a random time between 7:00 and 7:30 AM, find the probability that he waits

i) less than 5 minutes

ii) more than 10 minutes.

for the bus. What is the average amount of time spent waiting.

X: Arrival time of the psg

$X \sim$ Uniform $(7:00, 7:30)$

pdf : $f(x) = \frac{1}{7:30 - 7:00} = \frac{1}{30}$

i) P[waits < 5 mins

$= P[7:10 < X < 7:15] + P[7:25 < X < 7:30]$

$= \frac{5}{30} + \frac{5}{30} = \frac{1}{3}$

iii) P[waits > 10 mins]

$= P[7{:}00 < x < 7{:}05] + P[7{:}15 < x < 7{:}20]$

$= \frac{5}{30} + \frac{5}{30}$

$= \frac{1}{3}$

Note:

Mean and Variance of Continuous uniform dist

Mean : $E(x) = \frac{a+b}{2}$

Variance : $Var(x) = \frac{(b-a)^2}{12}$

To find the average time spent waiting

x: waiting time

x ~ uniform (0, 15)

Mean, $E(x) = \frac{15}{2} = 7.5$ mins.

Avg waiting time is 7.5 minutes.

2. It takes a random time between 20 and 27 minutes for a student to walk from hostel to college. If he has a class at 8:30 AM what is the probability that he will reach the class on time if he starts at 8:07 AM

x : Times to reach college from hostel

$x \sim$ Uniform $(20, 27)$

pdf =, $f(x) = \frac{1}{27-20} = \frac{1}{7}$, $20 < x < 27$

To find, ~~the~~

P[reaches class on time if he starts at 8:07 AM]

$= P[20 < x < 23] = \frac{23-20}{27-20} = \frac{3}{7}$

Note:

$E(X) =$ Mean $= \frac{a+b}{2}$

Variance $(X) = \frac{(b-a)^2}{12}$

Average Time spent waiting

X: Waiting time

$X \sim$ Uniform $(0, 15)$

Average time $= \frac{0+15}{2} = 7.5$ mins

2) It takes a random time between 20 and 27 mins for a student to walk from hostel to college. If he has a class at 8.30am, what is the probability that he will reach the class on time if he starts at 8:07 am

Ans. X: Time in minutes to reach the college from hostel

$X \sim$ Uniform $(20, 27)$

pdf: $f(x) = \frac{1}{27-20} = \frac{1}{7}$ $(\because 20 < X < 27)$

$$f(x) = \begin{cases} 1/7 & 20 < X < 27 \\ 0 & \text{otherwise} \end{cases}$$

To find:

P[he reaches the class on time if he starts at 8:07AM] $= P[20 < X < 23]$

$= \frac{23-20}{27-20}$

$= \frac{3}{7}$ //

10/08/2022

Exponential distribution.

R.V. X: Waiting time for the success

eg: (i) Time for the emission of α-particle.

(ii) Time for the next booking of ticket in IRCTC website

(iii) Time for the next emergency arrival in a hospital.

(iv) Time to failure of a component.

pdf: $f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$

Mean & variance

$$E(x^r) = \int_0^\infty x^r . f(x)\, dx = \int_0^\infty x^r . \lambda e^{-\lambda x} dx$$

$$= \lambda \int_0^\infty x^r e^{-\lambda x} dx$$

$$= \lambda \, L(x^r) \Big]_{s=\lambda}$$

$$= \lambda \frac{\Gamma(r+1)}{\lambda^{r+1}} = \frac{r!}{\lambda^r}$$

$r=1 \Rightarrow$ Mean $E(X) = \frac{1}{\lambda}$

$r=2 \Rightarrow$ $E(X^2) = \frac{2}{\lambda^2}$

$$\text{Variance } (X) = E(X^2) - E(X)^2$$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2}$$

$$= \frac{1}{\lambda^2}$$

Exponential and Poisson distribution.

If X is a Poisson R.V that counts the number of occurrences of an event ~~then the~~ per unit time with mean λ, then the time between successive occurences follows exponential distribution.

Memoryless property of exponential distribution

$$P[X \geq t+s \mid X \geq s] = P[X \geq t]$$

$$\text{RHS} = P[X > t] = \int_t^\infty f(x)\,dx = \int_t^\infty \lambda e^{-\lambda x}\,dx = \lambda\left(\frac{e^{-\lambda x}}{-\lambda}\right)_t^\infty = e^{-\lambda t}$$

$$\boxed{P[X > t] = e^{-\lambda t}}$$

$$\text{LHS} = P[X > t+s \mid X > s]$$

$$= \frac{P[X > t+s, X > s]}{P[X > s]}$$

$$= \frac{P[X > t+s]}{P[X > s]}$$

$$= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}}$$

$$= e^{-\lambda t}$$

1) Time between arrivals of electronic messages in your computer is exponentially distributed with a mean of 2 hours. Find the probability that

(i) You do not receive a message during a 2 hour period ?

(ii) If you have not had a message in last 4 hours what is the probability that you do not receive a message in next 2 hours.

(iii) What is the expected time between 5th and 6th message.

Ans. X : Time between arrivals of e-messages

$X \sim$ Exponential (λ)

Given: Mean $\frac{1}{\lambda} = 2$

$$\boxed{\lambda = \frac{1}{2}}$$

i) $P[X > 2] = e^{-2\lambda} = e^{-1} = \frac{1}{e}$

ii) $P[X \geq 6 | X \geq 4] = P[X \geq 2] = e^{-1}$.

iii) P[Expected time for the arrival of 5th and 6th messages = 2 hours .

6/9/2022

Weibull distribution.

It is the most popular distribution to model the distribution of failures in electronic components.

R.V. X: To failure (lifetime)

pdf: $f(x) = \alpha \beta x^{\beta - 1} e^{-\alpha x^{\beta}}, x > 0$

with parameters $\alpha > 0$ (scale parameter) and $\beta > 0$ (Shape parameter)

$X \sim Beta(\alpha, \beta)$

Mean & variance

$$E(X^r) = \int_0^\infty x^r . f(x)\, dx$$

$$= \int_0^\infty x^r \alpha \beta x^{\beta-1} e^{-\alpha x^\beta}\, dx$$

Take $y = \alpha x^\beta$

$\frac{dy}{dx} = \alpha \beta x^{\beta-1}$

$x^\beta = \frac{y}{\alpha}$

$x = \left(\frac{y}{\alpha}\right)^{1/\beta}$

$$= \int_0^\infty \left(\frac{y}{\alpha}\right)^{r/\beta} e^{-y}\, dy$$

$$= \frac{1}{\alpha^{r/\beta}} \int_0^\infty y^{r/\beta} e^{-y} dy$$

$$\boxed{E(X^r) = \frac{1}{\alpha^{r/\beta}} \Gamma\left(\frac{r}{\beta} + 1\right)}$$

$$\boxed{\Gamma(n) = \int_0^\infty e^{-x} x^{n-1} dx}$$

From this we can find mean and variance

Hazard rate function

$z(x) = \alpha \beta x^{\beta - 1}$

$$\boxed{H(t) = \frac{f(t)}{1 - F(t)}}$$

7/9/22

To find $P[X \geq t]$

$$P[X \geq t] = \int_t^\infty f(x)\,dx$$

$$= \int_t^\infty \alpha \beta x^{\beta-1} e^{-\alpha x^\beta}\,dx$$

Take $u = \alpha x^\beta$
$du = \alpha \beta x^{\beta-1}\,dx$

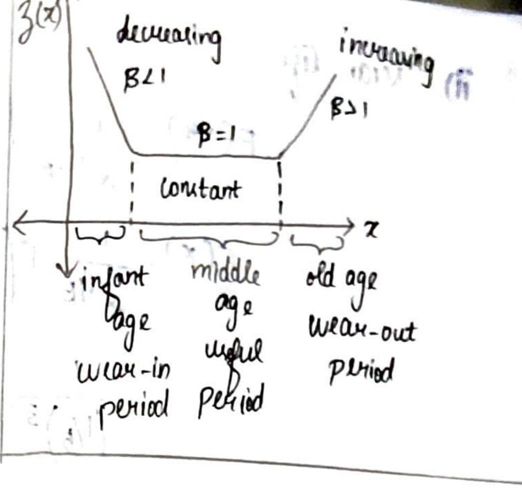
$x = t$, $u = \alpha t^\beta$

$x = \infty$, $u = \infty$

$$= \int_t^\infty e^{-u}\,du$$

$$= \left[-e^{-u}\right]_{\alpha t^\beta}^{\infty}$$

$$= e^{-\alpha t^\beta}$$



$$\boxed{P(X \geq t) = e^{-\alpha t^\beta}}$$

1) Suppose the time to failure in hours of a certain electronic component follows Weibull distribution with $\alpha = 1/5$, $\beta = 1/3$. What is the probability that such a component will fail before 27 hours. What is the mean life of the component

Ans.

$\alpha = 1/5$, $\beta = 1/3$

X: Life (in hours)

$X \sim$ Weibull (α, β)

where $\alpha = 1/5$, $\beta = 1/3$

i) $P[X < 27] = 1 - P[X \geq 27]$

$= 1 - e^{-\alpha t^\beta}$

$= 1 - e^{-\frac{1}{5}(27)^{1/3}}$

$= 1 - e^{-\frac{3}{5}}$

$= 1 - e^{-0.6}$

$= 1 - 0.548$

$= 0.452$

ii) Mean life

$$E(X^r) = \frac{1}{\alpha^{r/\beta}} \Gamma\left(\frac{r}{\beta}+1\right)$$

$$E(X) = \frac{1}{\alpha^{1/\beta}} \cdot \Gamma\left(\frac{1}{\beta}+1\right)$$

$$= \frac{1}{(1/5)^3} \Gamma(3+1)$$

$$= 125 \times 3!$$

$$= 750 \text{ hours}$$

2) The lifetime X in hours of a component has Weibull distribution with $\beta = 2$. Find the value of α, if it is known that 15% of the components that have lasted 90 hours failed before 100 hours.

Ans:

X: life (in hours)

$X \sim$ Weibull (α, β) with $\beta = 2$

To find α.

~~$[X<100 \mid X<900] = 0.15$~~

$$P[X<100 \mid X \geq 90] = 0.15$$

$$\frac{P[X<100, X \geq 90]}{P[X \geq 90]} = 0.15$$

$$\frac{P[90<X<100]}{P[X \geq 90]} = 0.15$$

$$\frac{P[X<100] - P[X<90]}{P[X \geq 90]} = 0.15$$

$$\frac{1-e^{-\alpha(100)^2} - 1 + e^{-\alpha(90)^2}}{e^{-\alpha(90)^2}} = 0.15$$

$$\frac{e^{-\alpha(10000)} + e^{-\alpha(8100)}}{e^{-\alpha(8100)}} = 0.15$$

$$\frac{e^{-\alpha(18100)}}{e^{-\alpha(8100)}} = 0.15$$

3) The life in years of a certain component has Weibull distribution with $\beta = 2$. Find the value of α given that the probability that the components life exceeds 5 years is $e^{-0.25}$ and also find mean and variance.

4) Suppose the lifetime of a certain battery follows Weibull distribution with $\alpha = 0.1$ & $\beta = 0.5$. Find the probability that such a battery will last more than 300 hours. Also find its mean life.

09/09/2022

Poisson process $N(t)$: No of arrivals by time t

An arrival process $\{N(t) : t \geq 0\}$ in continuous time is called a Poisson process with rate λ if

i) the number of arrivals in an interval of length t is Poisson(λt) i.e.,

$$P[N(t) = n] = \frac{e^{-\lambda t} (\lambda t)^n}{n!}$$

ii) the number of arrivals that occur in disjoint time intervals are independent.

r - Erlang distribution.

Let $\{N(t) : t \geq 0\}$ be a Poisson process where $N(t)$: No of arrivals by time t i.e. in the interval $(0, t)$

Let X_1: Time for the 1st arrival

X_2: Time between 1st and 2nd arrivals

X_3: Time between 2nd and 3rd arrivals

and so on.

$\Rightarrow$ Then $X_1, X_2, \ldots\ldots$ is a sequence of iid exponential random variables with rate λ.

$\Rightarrow$ $X = X_1 + X_2 + X_3 + \cdots\cdots + X_r$ is the time for the rth arrival with mean $1/\lambda$.

$X \sim \text{Erlang}(r, \lambda)$

$$\text{pdf } f(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{r-1}}{\Gamma(r)}, \quad x > 0$$

$$\text{Mean} = \frac{r}{\lambda}$$

$$\text{Variance} = \frac{r}{\lambda^2}$$

Note: The generalised r-Erlang distribution for any real number r is called the r-Gamma distribution

<u>cdf of r-Erlang distribution</u>

$F(x) = P[X \le x] = 1 - P[X > x]$

$= 1 - P[\text{Time for } r\text{th arrival is} > x]$

$= 1 - P[\text{no arrival or 1 arrival or} \ldots\ldots (r-1) \text{ arrivals by time } x]$

$= 1 - P[N(x) \le (r-1)]$

$$= 1 - \sum_{n=0}^{r-1} \frac{e^{-\lambda x} (\lambda x)^n}{n!}$$

$$\boxed{F(x) = 1 - \sum_{n=0}^{r-1} \frac{e^{-\lambda x} (\lambda x)^n}{n!}}$$

?/09/2022

1) The time between arrivals of customers in a bank is exponentially distributed with a mean of 10 mins. What is the probability that the time between the arrival of 2nd and 6th customer is more than 20 mins.

Ans. X: Time for 4th arrival $\Leftrightarrow$ Time between 2nd and 6th customer

$X \sim$ Erlang (r, λ) where $r = 4$ and $\lambda = \frac{1}{10}$

Time between arrivals $\sim$ Expo (λ) with Mean $= \frac{1}{\lambda} = 10$

To find: $P[X > 20]$

$$= \sum_{k=0}^{r-1} \frac{e^{-\lambda x} (\lambda x)^k}{k!}$$

$$= \sum_{k=0}^{3} \frac{e^{-1/10(20)} \cdot (2)^k}{k!}$$

$$= \sum_{k=0}^{3} \frac{e^{-2}\, 2^k}{k!}$$

$$= e^{-2}\left(1 + 2 + 2 + \frac{4}{3}\right)$$

$$= \frac{19}{3} e^{-2}$$

Normal distribution. (ND)

Most popular application in theory and application

$\Rightarrow$ It is used to model distribution of error in measurements and many other naturally occurring phenomenon.

Standard normal distribution. (SND)

* A ND with mean 0 and variance 1 is called SND. It is the simplest ND.

* A normal random variable Z is said to follow SND if its pdf is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < x < \infty$$

$\frac{1}{\sqrt{2\pi}}$ is normalizing constant

13/09/2022

We Know: $\boxed{\Gamma(n) = \int_0^\infty e^{-x} x^{n-1} dx}$ $\qquad \frac{\Gamma(n)}{a^n} = \int_0^\infty e^{-ax} x^{n-1} dx$

Take $x = y^2 \Rightarrow dx = 2y\,dy$

$$\Gamma(n) = \int_0^\infty e^{-y^2} (y^2)^{n-1} 2y\,dy$$

$$\boxed{\Gamma(n) = 2\int_0^\infty e^{-y^2} y^{2n-1} dy}$$

$$\int_{-a}^{a} f(x)\,dx = \begin{cases} 2\int_0^a f(x)\,dx & \text{if } f(-x) = f(x) \\ 0 & \text{otherwise} \end{cases}$$

To show: $\int_{-\infty}^{\infty} e^{-z^2/2} dz = \sqrt{2\pi}$

Proof:

LHS : $\int_{-\infty}^{\infty} e^{-z^2/2} dz = 2\int_0^\infty e^{-z^2/2} dz$ since $e^{-z^2/2}$ is an even function

Take $y = z/\sqrt{2} \Rightarrow dz = \sqrt{2}\,dy$ $\qquad \frac{dy}{dz} = \frac{1}{\sqrt{2}}$

$$= 2\int_0^\infty e^{-y^2} \sqrt{2}\,dy$$

$2n-1=0$

$\boxed{n=1/2}$

$$= 2\sqrt{2}\int_0^\infty e^{-y^2} y^0 dy$$

specs podala kaam ujuthathuu

$$= \sqrt{2}\,\Gamma(1/2) = \sqrt{2}\cdot\sqrt{\pi}$$

$$= \sqrt{2\pi} = \text{RHS}$$

Mean and variance.

$$E(z) = \int_{-\infty}^{\infty} z\, f(z)\,dz$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z\, e^{-z^2/2} dz \quad \text{- odd function}$$

$$= 0$$

$$E(z^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-z^2/2} dz$$

Take $y = z/\sqrt{2}$, $dz = \sqrt{2}\, dy$

$$= \frac{2\sqrt{2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2} dy$$

$$= \frac{4\sqrt{2}}{\sqrt{2\pi}} \int_{0}^{\infty} y^2 e^{-y^2} dy$$

$2n-1 = 2$
$2n = 3$
$n = 3/2$

$$= \frac{2\sqrt{2}}{\sqrt{2\pi}} \Gamma(3/2)$$

$$= \frac{2\sqrt{2}}{\sqrt{2\pi}} \cdot 1/2\, \Gamma(1/2)$$

$$= \frac{\sqrt{2\pi}}{\sqrt{2\pi}}$$

$$= 1$$

$$\text{Variance}(Z) = E(Z^2) - E(Z)^2$$
$$= 1$$

Normal distribution with Mean μ and Variance σ^2.

If $Z \sim N(0,1)$, then the r.v. $X = \mu + \sigma Z$ follows ND with mean μ and var σ^2.

Verify:

Mean $= E(X) = E(\mu + \sigma Z) = E(\mu) + \sigma E(Z)$
$= \mu + \sigma(0)$

$$\boxed{E(X) = \mu}$$

$\text{Var}(X) = \text{Var}(\mu + \sigma Z) = \text{Var}(\mu) + \sigma^2 \text{Var}(Z)$
$= \sigma^2(1)$

$$\boxed{\text{Var}(X) = \sigma^2}$$

To find probabilities in ND.

For a SND, the probabilities $P[Z \le z] = \phi(z)$ for any z are available in Normal table.

A normal random variable with mean μ and s.d. σ can be transformed into a Standard Normal Variable by

$$\boxed{Z = \frac{X-\mu}{\sigma}} \rightarrow Z\text{-score}$$

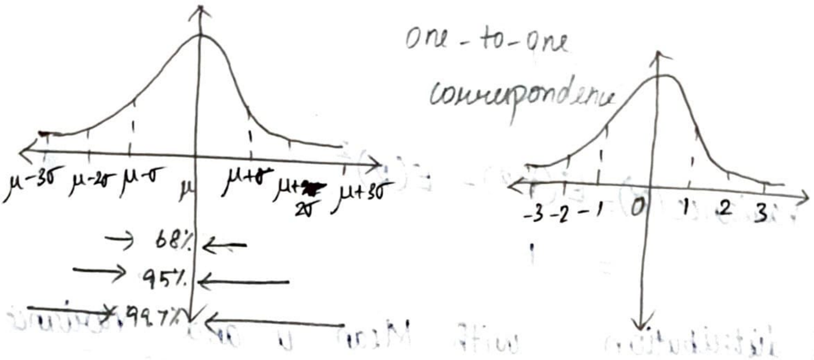
14/09/2022

$X \sim ND(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty$$

⇒ Bell shaped

⇒ Normal distribution is the only distribution in which Mean, median and mode coincide.



⇒ 68 – 95 – 99.7 rule

FOR ND:

$P[\mu-\sigma < X < \mu+\sigma] = 0.68$

$P[\mu-2\sigma < X < \mu+2\sigma] = 0.95$

$P[\mu-3\sigma < X < \mu+3\sigma] = 0.997$

FOR SND:

$P[-1 < Z < 1] = 0.68$

$P[-2 < Z < 2] = 0.95$

$P[-3 < Z < 3] = 0.997$

Gaussian distribution.

⇒ $P[Z \le z] = \phi(z)$

For any z

$P[Z < -2.5] = \phi(-2.5) = 0.0062$

$P[Z < -1.27] = \phi(-1.27) = 0.1020$

$P[Z \leq 1.9] = \phi(1.9) = 0.9713$

$P[Z < 2.32] = \phi(2.32) = 0.9898$

$P[Z < 3.49] = \phi(3.49) = 0.9998$

1) A survey indicates that for each trip to a super market a customer spends an average of 45 mins and a SD of 12 mins. The length of time spent X follows ND. Find the probability that a randomly selected customer will be the store for
(i) between 24 and 54 mins
(ii) more than 39 mins
(iii) between 33 and 60 mins.

Sol. X: Time spent in the store

$X \sim ND(\mu, \sigma)$

$\mu = 45, \sigma = 12$

Z score: $Z = \frac{X-\mu}{\sigma} = \frac{X-45}{12}$

(i) $P[24 < X < 54] = P\left[\frac{24-45}{12} < Z < \frac{54-45}{12}\right]$

$= P\left[\frac{-21}{12} < Z < \frac{9}{12}\right] = P\left[\frac{-7}{4} < Z < \frac{3}{4}\right]$

$= P[-1.75 < Z < 0.75]$

$= \phi(0.75) - \phi(-1.75)$

$= 0.7734 - 0.0401$

$= 0.7333$

ii) $P[X > 39] = P\left[Z > \frac{39-45}{12}\right] = P[Z > -0.5]$

$= 1 - P[Z \leq -0.5]$

$= 1 - 0.3085 = 0.6915$

$P[33 < X < 60] = P[-1 < Z < 1.25]$

$= \cancel{\phi(-1)} \; \phi(1.25) - \phi(-1)$

$= 0.8944 - 0.1587$

$= 0.7357$

15/09/2022

2. Amount of soft drink in a can is normally distributed. Suppose 7% of the bottles contain less than 155 ml and 10% of the bottles contain more than 163 ml. Find the mean and SD of the amount of soft drink in a randomly selected bottle.

Sol:

X: Amount of soft drink in a bottle

$X \sim ND(\mu, \sigma)$

To find μ & σ.

Given: 7% of bottles contain < 155 ml

$P[X < 155] = 0.07$

$P\left[Z < \frac{155-\mu}{\sigma}\right] = 0.07$

$\phi\left(\frac{155-\mu}{\sigma}\right) = 0.07$

From normal table, for prob $= 0.07$, corresponding Z score is $z=-1.48$

$\phi\left(\frac{155-\mu}{\sigma}\right) = \phi(-1.48)$

$\frac{155-\mu}{\sigma} = -1.48$

$\mu - 1.48\sigma = 155 \rightarrow (1)$

Given: 15% of bottles contain > 163 ml

$P[X > 163] = 0.1$

$P\left[Z > \frac{163-\mu}{\sigma}\right] = 0.1$

$$1-\phi\left[\frac{163-\mu}{\sigma}\right]=0.1$$

$$\phi\left[\frac{163-\mu}{\sigma}\right]=0.9$$

$$\phi\left[\frac{163-\mu}{\sigma}\right]=\phi[1.28]$$

$$\frac{163-\mu}{\sigma}=1.28$$

$$\mu+1.28\sigma=163 \rightarrow (2)$$

$$(1) \Rightarrow \cancel{\mu}-1.48\sigma=155$$
$$(2) \Rightarrow \cancel{\mu}+1.28\sigma=163$$
$$(-)\quad(-)$$
$$-2.76\sigma=-8$$

$$\boxed{\sigma=2.89}$$

$$\mu=-1.28(2.89)+163$$

$$\mu=-3.6992+163$$

$$\boxed{\mu=159.3}$$

3) In a normal distribution, 31% of the items are under 45 and 8% of the items are over 64. Find the mean and SD.

Sol:

X: Amount of items

$X \sim ND(\mu, \sigma)$

Given: 31% of items are under 45

$$P[X<45]=0.31$$

$$P\left[Z<\frac{45-\mu}{\sigma}\right]=0.31$$

$$\phi\left(\frac{45-\mu}{\sigma}\right)=0.31$$

$$\phi\left(\frac{45-\mu}{\sigma}\right)=\phi(-0.5) \Rightarrow \frac{45-\mu}{\sigma}=-0.5$$

$$\mu-0.5\sigma=45 \rightarrow (1)$$

Given: 8% of items are over 64

$$P[X>64]=0.08$$

$\Rightarrow 1 - P[X < 64] \Rightarrow 1 - \phi\left[\frac{64-\mu}{\sigma}\right] = 0.08$

$0.92 = \phi\left[\frac{64-\mu}{\sigma}\right]$

$\phi(\cancel{-2.05}\ 1.41) = \phi\left(\frac{64-\mu}{\sigma}\right)$

$\mu \cancel{-2.05\sigma}\ +1.41\sigma = 64 \rightarrow (2)$

~~$(1) \Rightarrow \mu - 0.5\sigma = 45$~~

~~$(2) \Rightarrow \mu - 2.05\sigma = 64$~~

~~$(-)\ (+)\ (-)$~~

~~$1.55\sigma = -19$~~

~~$\boxed{\sigma = 12.25}$~~

~~$(1) \Rightarrow \mu - 0.5(12.25) = 45$~~

~~$\mu - 6.125 = 45$~~

~~$\boxed{\mu = 51.125}$~~

$\mu - 0.5\sigma = 45$

$\mu + 1.41\sigma = 64$

$(-)\ (-)\ (-)$

$-1.91\sigma = -19$

$\boxed{\sigma = 9.947}$

$(1) \Rightarrow \mu - 0.5(9.947) = 45$

$\mu - 4.9735 = 45$

$\boxed{\mu = 49.9735}$

4) The weight of a randomly selected adult in a certain region has ND with mean 160 pounds and 85% of the adult population has a weight between 120 and 200 pounds. Find the variance of X.

Sol:

~~$P[120 < X < 200] = 0.85$~~

~~$P\left[\frac{120-\mu}{\sigma} < X < \frac{200-\mu}{\sigma}\right] = 0.85$~~

~~$P\left[\frac{-40}{\sigma} < X < \frac{40}{\sigma}\right] = 0.85$~~

~~$\phi\left(\frac{+40}{\sigma}\right) - \phi\left(\frac{-40}{\sigma}\right) = 0.85$~~

$P[120 < X < 200] = 0.85$

$2P[160 < X < 200] = 0.85$

$2\left(P\left[0 < X < \frac{40}{\sigma}\right]\right) = 0.85$

$\phi\left(\frac{40}{\sigma}\right) - \phi(0) = 0.425$

$$\phi\left(\frac{40}{\sigma}\right) - 0.5 = 0.425$$

$$\phi\left(\frac{40}{\sigma}\right) = 0.925$$

$$\frac{40}{\sigma} = 1.44$$

$$\sigma = \frac{40}{1.44}$$

16/09/2022

Joint Probability distribution.

We have to deal with 2 or more random variables defined on the same SS not only to study their individual behaviour but also to determine the degree of relationship between them.

eg: RV X: Height of a person
Y: Weight of the same person

Joint pmf / Joint pdf.

For 2 discrete random variables X and Y,
Joint pmf of X and Y,

$$P(x,y) = P[X=x, Y=y] \ \forall\ x \in X \text{ and } y \in Y$$

1) Roll 2 dice. Let X denotes maximum of the 2 dice and Y denotes number of times an even number appears. Find the joint pmf of (X, Y).

Sol:

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

$$Y = \{0, 1, 2\}$$

Space $(X, Y) \subseteq \mathbb{R}^2$

X \ Y	0	1	2
1	1/36	0	0
2	0	2/36	1/36
3	3/36	2/36	0
4	0	4/36	3/36
5	3/36	4/36	0
6	0	8/36	5/36

Marginal pmf of

X	1	2	3	4	5
$P_X(x)$	1/36	3/36	5/36	7/36	[illegible]

Marginal pmf of

Y	0	1	2
$P_Y(y)$	9/36	18/36	9/36

The pmfs of the individual random variables are called marginal pmfs.

Marginal pmf of X:

$$P_X(x) = \sum_{y \in Y} P(x, y)$$

Marginal pmf of Y:

$$P_Y(y) = \sum_{x \in X} P(x, y)$$

Joint pdf:

For 2 continuous random variables X and Y

$f(x, y)$ is the Joint pdf of X and Y if

$$\int_a^b \int_c^d f(x, y)\, dy\, dx = P[a < X < b, c < Y < d]$$

Marginal pdf of X:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

Marginal pdf of Y.

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

17/09/2022

1) 3 cards are drawn from a deck of 52 cards. Let X and Y denote the number of diamond and spade cards. Find the joint pmf of X and Y and also find the marginal pmf.

Sol:

X: No of diamonds

Y: No of spades

$X = Y = \{0, 1, 2, 3\}$

Joint pmf of (X, Y)

X\Y	0	1	2	3
0	$\frac{^{26}C_3}{^{52}C_3}$	$\frac{^{13}C_1 \cdot {^{26}C_2}}{^{52}C_3}$	$\frac{^{13}C_2 \cdot {^{26}C_1}}{^{52}C_3}$	$\frac{^{13}C_3}{^{52}C_3}$
1	$\frac{^{13}C_1 \cdot {^{26}C_2}}{^{52}C_3}$	$\frac{^{13}C_1 \cdot {^{13}C_1} \cdot {^{26}C_1}}{^{52}C_3}$	$\frac{^{13}C_1 \cdot {^{13}C_2}}{^{52}C_3}$	0
2	$\frac{^{13}C_2 \cdot {^{26}C_1}}{^{52}C_3}$	$\frac{^{13}C_2 \cdot {^{13}C_1}}{^{52}C_3}$	0	0
3	$\frac{^{13}C_3}{^{52}C_3}$	0	0	0

Marginal pmf of X 0 1 2 3

2) Roll a fair die, let the outcome be X. Then toss a fair coin X times and let Y denotes the no of tails. Find the joint pmf of X and Y.

3) A thief has stolen 4 animals from a farm containing ~~3~~ 2 horses, 3 cows and 5 sheep. Let X and Y denote the no of cows and sheep stolen. Obtain the joint pmf of X and Y.

$$f_X(x) = \int f(x,y)\,dy$$

$$f_Y(y) = \int f(x,y)\,dx$$

Conditional distribution of x given $Y=y$

Conditional density of X given $Y=y$:

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

Conditional density of Y given $X=x$

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

$$P_{X|Y=1}(1|1) = \frac{P(1,1)}{P_Y(1)} = 0$$

$$P_{X|Y=1}(2|1) = \frac{2/36}{18/36} = 2/18$$

Conditional expectation of x given $Y=y$

$E(X) = \sum x \cdot P_X(x)$

Expection of x given $Y=y$

$\int x \cdot f_X(x)\,dx$

$E(Y) = \int y \cdot f_Y(y)\,dy$

$$\frac{1}{36}(1+6+15+28+45+66) = \frac{161}{36} = 4.47$$

$$E[X|Y=y] = \int x\, f_{X|Y}(x|y)\,dx$$

$$E[Y|X=x] = \int y\, f_{Y|X}(y|x)\,dy$$

$E(XY) = \iint xy\, f(x,y)\, dy\, dx$

10/10/2022

1) 2 random variables x and y has the joint pdf

$$f(x,y) = xy^2 + \frac{x^2}{8}, \quad 0 \le x \le 2, \ 0 \le y \le 1$$

Find i) $P[X > 1]$

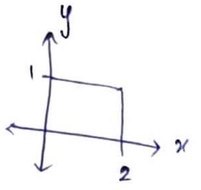
ii) $P[Y < 1/2]$

iii) $P[X > 1, Y < 1/2]$

iv) $P[X < Y]$

v) $P[Y < 1/2 \mid X > 1]$

vi) $P[X + Y \le 1]$



Marginal pdf of X.

$$f_X(x) = \int f(x,y)\, dy = \int_0^1 \left(xy^2 + \frac{x^2}{8}\right) dy$$

$$= \left[\frac{xy^3}{3} + \frac{x^2 y}{8}\right]_0^1$$

$$\boxed{f_X(x) = \frac{x}{3} + \frac{x^2}{8}}$$

$$f_Y(y) = \int f(x,y)\, dx = \int_0^2 \left(xy^2 + \frac{x^2}{8}\right) dx$$

$$= \left[\frac{x^2y^2}{2} + \frac{x^3}{24}\right]_0^2$$

$$= \frac{4y^2}{2} + \frac{8}{24}$$

$$\boxed{f_Y(y) = 2y^2 + \frac{1}{3}}$$

i) $P[X > 1] = \int_1^2 f_X(x)\, dx = \int_1^2 \left(\frac{x}{3} + \frac{x^2}{8}\right) dx = \left[\frac{x^2}{6} + \frac{x^3}{24}\right]_1^2$

$$= \frac{2}{3} + \frac{1}{3} - \frac{1}{6} + \frac{1}{24}$$

$$= \frac{24 - 4 + 1}{24}$$

$$= 19/24$$

ii) $P[Y < 1/2] = \int_{1/2}^{1} f_Y(y)\, dy = \int_{1/2}^{1} (2y^2 + 1/3)\, dy$

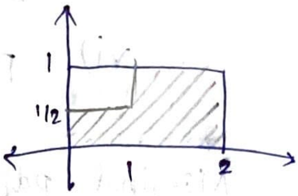
$= \left[\frac{2y^3}{3} + \frac{y}{3}\right]_{0}^{1/2}$

$= \frac{1}{12} + \frac{1}{6} \qquad \cancel{= \frac{2}{3} + \frac{1}{3} - \frac{1}{12} - \frac{1}{6}}$

$= \frac{3}{12} \qquad \cancel{= \frac{12 - 1 - 2}{12}}$

$= 1/4$ // $\qquad \cancel{= 9/12 = 3/4}$

iii) $P[X > 1, Y < 1/2]$



$= \int_{1}^{2} \int_{0}^{1/2} f(x,y)\, dy\, dx$

$= \int_{1}^{2} \int_{0}^{1/2} \left[xy^2 + \frac{x^2}{8}\right] dy\, dx$

$= \int_{1}^{2} \left[\frac{xy^3}{3} + \frac{x^2 y}{8}\right]_{0}^{1/2} dx$

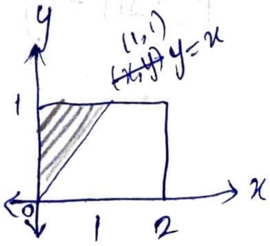
$= \int_{1}^{2} \left[\frac{x}{24} + \frac{x^2}{16}\right] dx$

$= \left[\frac{x^2}{48} + \frac{x^3}{48}\right]_{1}^{2}$

$= \frac{4 + 8 - 1 - 1}{48}$

$= 5/24$ //

iv) $P[X < Y]$



$= \int_{0}^{1} \int_{x}^{1} (xy^2 + x^2/8)\, dy\, dx$

$= \int_{0}^{1} \left[\frac{xy^3}{3} + \frac{x^2 y}{8}\right]_{x}^{1} dx$

$$= \int_0^1 \left[\frac{x}{3} + \frac{x^2}{8} - \frac{x^4}{3} - \frac{x^3}{8}\right] dx$$

$$= \left[\frac{x^2}{6} + \frac{x^3}{24} - \frac{x^5}{15} - \frac{x^4}{32}\right]_0^1$$

$$= \frac{1}{6} + \frac{1}{24} - \frac{1}{15} - \frac{1}{32}$$

$$= \frac{5}{24} - \frac{32+15}{480}$$

$$= \frac{100}{480} - \frac{47}{480}$$

$$= \frac{53}{480}$$

$3 \times 5 \times 8 \times 4$

$$\begin{array}{r} 32 \\ 15 \\ \hline 320 \\ 160 \\ \hline 480 \end{array}$$

v) $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$ $\quad P[Y < 1/2 \mid X \leq 1]$

$$P[Y < 1/2 \mid X \leq 1] = \frac{P[Y < 1/2, X \leq 1]}{P[X \leq 1]}$$

$$= \frac{5/24}{19/24}$$

$$= 5/19$$

vi) $P[X + Y \leq 1]$

7) The joint pdf r.v.s X and Y is

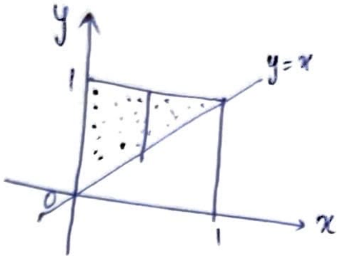
$$f(x,y) = \begin{cases} Kxy^2, & 0 \le x \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

i) Find the value of K

ii) Find the marginal densities of X and Y

iii) Find E(X), E(Y) and E(XY)

Sol:



$0 \le x \le 1$

$0 \le y \le 1$

$x \le y$

i) To find K:

For a pdf : $\int_0^1 \int_x^1 f(x,y)\, dy\, dx = 1$

$$\int_0^1 \int_x^1 Kxy^2\, dy\, dx = 1$$

$$\int_0^1 K\left(\frac{xy^3}{3}\right)_x^1 dx = 1$$

$$K \int_0^1 \left(\frac{x}{3} - \frac{x^4}{3}\right) dx = 1$$

$$K\left[\frac{x^2}{6} - \frac{x^5}{15}\right]_0^1 = 1$$

$$K\left[\frac{1}{6} - \frac{1}{15}\right] = 1$$

$\cancel{\boxed{K = 12}}$

$\boxed{K = 10}$

ii) Marginal pdf of X:

$$f_X(x) = \int_x^1 f(x,y)\, dy$$

$$= \int_x^1 10xy^2\, dy = \left[\frac{10y^3x}{3}\right]_x^1$$

$$= \frac{10x}{3} - \frac{10x^4}{3}$$

$\boxed{f_X(x) = \frac{10x(1-x^3)}{3}}$

Marginal pdf of Y:

$$f_Y(y) = \int_0^1 f(x,y)\,dx = \int_0^y 10xy^2\,dx = \left[\frac{10x^2y^2}{2}\right]_0^y$$

$$\boxed{f_Y(y) = 5y^4}$$

iii) To find $E(X)$, $E(Y)$ and $E(XY)$.

$$E(X) = \int_0^1 x\, f_X(x)\,dx$$

$$\cancel{\int_0^1 x\,(10xy^2)\,dx}$$

$$= \frac{10}{3}\int_0^1 x^2(1-x^3)\,dx$$

$$= \frac{10}{3}\left[\frac{x^3}{3} - \frac{x^6}{6}\right]_0^1$$

$$= \frac{10}{3}\cdot\frac{1}{6}$$

$$= \frac{5}{9}$$

$$E(Y) = \int_0^1 y\, f_Y(y)\,dy = \int_0^1 5y^5\,dy$$

$$= 5\left[\frac{y^6}{6}\right]_0^1$$

$$= 5/6$$

$$E(XY) = \int_0^1\int_x^1 xy\, f(x,y)\,dy\,dx$$

$$= \int_0^1\int_x^1 10x^2y^3\,dy\,dx$$

$$= 10\int_0^1 \left[\frac{x^2y^4}{4}\right]_x^1 dx$$

$$= 10\int_0^1 \left[\frac{x^2}{4} - \frac{x^6}{4}\right] dx$$

$$= 10\left[\frac{x^3}{12} - \frac{x^7}{28}\right]_0^1 = 10\left(\frac{1}{12} - \frac{1}{28}\right) = 10/21$$

$\therefore$ X and Y are not independent

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

To check if the Random variable is independent.

* $f(x,y) = f_X(x) \cdot f_Y(y)$

(or)

* $E(X \cdot Y) = E(X) \cdot E(Y)$

21/09/2022

2) Two r.vs X and Y have the joint pdf

$$f(x,y) = c(x+y), \quad 0 \le x \le 1, \quad 0 \le y \le 1$$

Find (i) value of c

(ii) Marginal densities of X and Y

(iii) $P[X \ge Y]$, $P[Y \le 1/2]$, $P[X \le Y^2]$

Sol:

To find c:

For a pdf: $\int_0^1 \int_0^1 f(x,y)\, dy\, dx = 0+1$

$$\int_0^1 \int_0^1 c(x+y)\, dy\, dx = 0+1$$

$$\int_0^1 c\left(xy + \frac{y^2}{2}\right)_0^1 dx = 0+1$$

$$\int_0^1 c\left(x + \frac{1}{2}\right) dx = 0+1$$

$$c\left(\frac{x^2}{2} + \frac{x}{2}\right)_0^1 = 0+1$$

$$c\left(\frac{1}{2} + \frac{1}{2}\right) = 0+1$$

$\cancel{c = 1}$ $\boxed{c=1}$

ii)

$$f_X(x) = \int_0^1 f(x,y)\, dy$$

$$= \int_0^1 (x+y)\, dy$$

$$= \left(xy + \frac{y^2}{2}\right)_0^1$$

$$= x + \frac{1}{2}$$

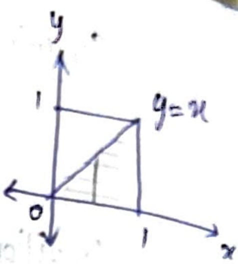
$$f_Y(y) = \int_0^1 f(x,y)\, dx$$

$$= \int_0^1 (x+y)\, dx$$

$$= \left(\frac{x^2}{2} + xy\right)_0^1$$

$$= \frac{1}{2} + y$$

c) i) $P[X<Y] = \int_0^1 \int_0^x (x+y)\,dy\,dx$



$= \int_0^1 \left[xy + \frac{y^2}{2}\right]_0^x dx$

$= \int_0^1 \left[x^2 + \frac{x^2}{2}\right] dx$

$= \left[\frac{x^3}{3} + \frac{x^3}{6}\right]_0^1$

$= \frac{1}{3} + \frac{1}{6}$

$= \frac{1}{2}$ //

ii) $P[Y<1/2]$

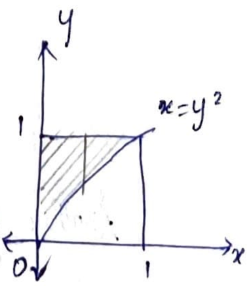
$P[Y<1/2] = \int_0^{1/2} f_Y(y)\,dy = \int_0^{1/2} \left(\frac{1}{2} + y\right) dy$

$= \left(\frac{y}{2} + \frac{y^2}{2}\right)_0^{1/2}$

$= \frac{1}{4} + \frac{1}{8}$

$= \frac{3}{8}$ //

iii) $P[X<Y^2] = \int_0^1 \int_{\sqrt{x}}^1 (x+y)\,dy\,dx$



$= \int_0^1 \left[xy + \frac{y^2}{2}\right]_{\sqrt{x}}^1 dx$

$= \int_0^1 \left[x + \frac{1}{2} - x^{3/2} - \frac{x}{2}\right] dx$

$= \left[\frac{x^2}{4} + \frac{x}{2} - \frac{x^{5/2}}{5/2}\right]_0^1$

$= \frac{1}{4} + \frac{1}{2} - \frac{1}{5/2}$

$= \frac{3}{4} - \frac{2}{5}$

$= \frac{7}{20}$ //

Conditional pdf of X given Y = y.

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

Conditional pdf of Y given X = x

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

We can recover the joint pdf $f(x,y)$ if $f_{X|Y}$ and f_Y are known

as $\boxed{f(x,y) = f_{X|Y}(x|y) \, f_Y(y)}$

if $f_{Y|X}$ and f_X are known

as $\boxed{f(x,y) = f_{Y|X}(y|x) f_X(x)}$

Independence of random variables X and Y

2 r.v.s X and Y are independent if

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

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

$$\boxed{f(x,y) = f_X(x) \cdot f_Y(y)}$$

Another formula:

If X and Y are independent, then

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

$$\cancel{E(XY) =} \iint$$

$$E(X) = \int x \cdot f_X(x) \, dx$$

$$E(Y) = \int y \cdot f_Y(y) dy$$

$$\cancel{E(XY) = \int x \cdot f_X(x) dx} \cdot \int y \cdot f_Y(y) dy$$

$$= \iint xy \, f(x,y) \, dy \, dx$$

$$LHS = E(XY) = \iint xy\, f(x,y)\, dy\, dx$$

$$= \iint x\, f(x,y)\quad y . f(x,y)\, dy\, dx$$

$$= \int x\, f(x,y)\quad E(Y)\, dx$$

$$= E(X). E(Y) = RHS$$

Result:

$$E(E(X|Y)) = E(X)$$

$$E(E(X|Y)) = \int_{-\infty}^{\infty} E(X|Y)\, f_Y(y)\, dy$$

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

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

$$= \int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f(x,y)\, dy\right) dx$$

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

$$= E(X)$$

2022

Continuous version of Bayes rule.

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

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

$$\boxed{f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x) . f_X(x)}{\int_{x\in X} f_{Y|X}(y|x)\, f_X(x)\, dx}}$$

$$\boxed{f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y) . f_Y(y)}{\int_{y\in Y} f_{X|Y}(x|y) . f_Y(y)\, dy}}$$

1) A man invites his friend to KFC for a brunch. They decide to meet at KFC between 11:30 A.M. and 12:00 P.M. They arrive at a random time in this period. What is the probability that they meet within 10 mins of their arrival.

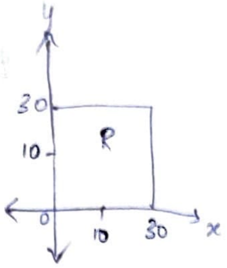
Sol:

Let X: Time past 11:30 A arrives

Y: Time past 11:30 B arrives

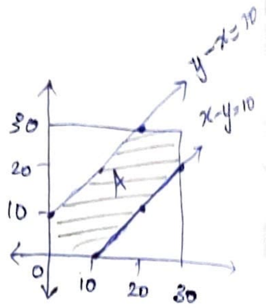
$X = Y = (0, 30)$

X & $Y \sim$ Uniform $(0, 30)$



Joint pdf of (X, Y): $f(x,y) = \frac{1}{\text{Area}(R)}$, $x, y \in (0, 30)$

$$f(x,y) = \frac{1}{900}$$



$$P[|X-Y| \le 10] = \frac{\text{Area of } A}{\text{Area of } R}$$

$$= \frac{900 - 400}{900}$$

$$= \frac{500}{900}$$

$$= 5/9$$

$x - y = 10$

x	10	30
y	0	20

$y - x = 10$

x	0	20
y	10	30

Note: $\boxed{f(x,y) = \frac{1}{(b-a)(d-c)}}$

23/09/2022

Sum of independent r.v.s

If X and Y are independent r.v.s. What distribution $T = X + Y$ follows?

Convolution:

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(u)\, g(t-u)\, du$$

Result:

If X and Y are independent r.vs, then the pdf of $T = X+Y$ is the convolution of the pdfs of X and Y.

$$\boxed{f_T(t) = \int_x f_X(x) f_Y(t-x)\, dx}$$

or

$$\boxed{f_T(t) = \int_y f_Y(y) \cdot f_X(t-y)\, dy}$$

If X and Y are independent discrete r.vs, then the pmf of $T = X+Y$ is the convolution of the pmfs of X and Y.

$$\boxed{P[T=t] = \Sigma\, P[X=x]\, P[Y=t-x]}$$

(or)

$$\boxed{P[T=t] = \Sigma\, P[Y=y]\, P[X=t-y]}$$

1) If X, Y ~~follows~~ ~ iid exponential (λ), then find the pdf of $T = X+Y$

Ans.

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$
$$f_Y(y) = \lambda e^{-\lambda y}, \quad y \geq 0$$

To find: pdf of $T = X+Y$

$$f_T(t) = \int_0^t f_X(x)\, f_Y(t-x)\, dx$$

$$= \int_0^t \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda(t-x)}\, dx$$

$$= \int_0^t \lambda^2 e^{-\lambda t}\, dx$$

$$= \left[x \lambda^2 e^{-\lambda t}\right]_0^t$$

$$\boxed{f_T(t) = \lambda^2 t e^{-\lambda t}}$$

This is the pdf of 2-Erlang distribution

$\therefore X+Y \sim$ Erlang $(2, \lambda)$

1) Let ~~If~~ $X \sim$ Binomial distribution (n,p) and $Y \sim$ Binomial (n,p)

if X and Y are independent ; find the pmf of $T = X+Y$.

2). Let ~~If~~ $X \sim$ Poisson (λ_1) and $Y \sim$ Poisson (λ_2).

If X and Y are independent, find the pmf of $T = X+Y$.

Ans:

1) $X \sim$ Binomial (n, p)

$Y \sim$ Binomial (m, p)

$P[X = x] = {}^nC_x\, p^x q^{n-x}$

$P[Y=y] = {}^mC_y\, p^y q^{m-y}$

To find:

$T = X+Y$

$$P[T=t] = \sum_{x\geq 0}^{t} P[X=x] \cdot P[Y=t-x \mid X=x]$$

$$= \sum_{x \geq 0}^{t} {}^nC_x\, p^x q^{n-x}\; {}^mC_{t-x}\, p^{t-x} q^{m-t+x}$$

(Vandermonde's formula)

$$P[T=t] = {}^{n+m}C_t\, p^t q^{m+n-t}$$

$T \sim$ Binomial $(m+n, p)$

2) $X \sim$ Poisson (λ_1)

$Y \sim$ Poisson (λ_2)

$P[X = x] = \dfrac{e^{-\lambda_1} \lambda_1^x}{x!}$ $\quad P[Y=y] = \dfrac{e^{-\lambda_2} \lambda_2^y}{y!}$

To find: $T = X+Y$

$$P[T=t] = \sum_{x\geq 0}^{t} P[X=x] \cdot P[Y=t-x]$$

$$= \sum_{x\geq 0}^{t} \cancel{\frac{e^{-2\lambda}\, \lambda^{x+t-x}}{x!\,(t-x)!}}\; \frac{e^{-\lambda_1} \lambda_1^x}{x!} \cdot \frac{e^{-\lambda_2} \lambda_2^{t-x}}{(t-x)!}$$

$$= e^{-(\lambda_1+\lambda_2)} \sum_{x} \lambda_1^x \cdot \lambda_2^{t-x} \cdot \frac{1}{x!} \cdot \frac{1}{(t-x)!}$$

$$P[T=t] = \sum_{x} P[Y=t-x]\, P[X=x]$$

Proof:

$$\underline{LHS} = P[T=t] = P[X+Y=t]$$

$$= \sum_{x} P[X+Y=t \mid X=x] \cdot P[X=x]$$

$$= \sum_{x} P[Y=t-x \mid X=x] \cdot P[X=x]$$

$$= \sum_{x} P[Y=t-x]\, P[X=x]$$

($\because$ X and Y are independent)

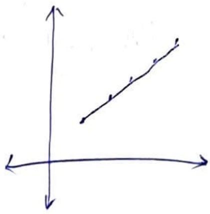
$$= \underline{RHS}$$

07/09/2022

Covariance.

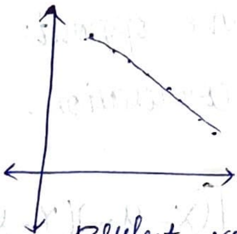
Mean and variance provide single number summaries of the distribution of a single random variable.

Covariance measures the tendency of ~~the~~ 2 random variables whether they go up or go down. Covariance is a measure of a linear association between 2 random variables X and Y.



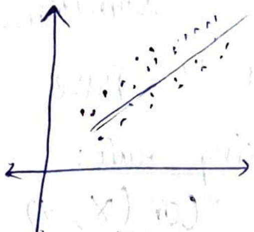
Perfect +ve correlation

(a)



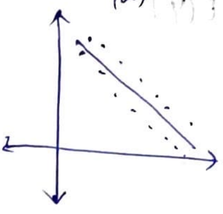
perfect -ve correlation

(b)



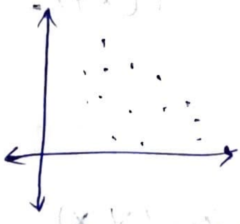
+ve correlation

(c) $cov(X,Y) > 0$



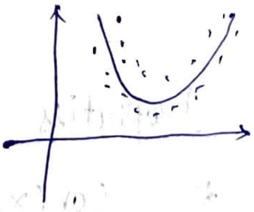
-ve correlation

(d) $cov(X,Y) < 0$



no correlation

(e) $cov(X,Y) = 0$



(f) $cov(X,Y) = 0$

a, b, c, d - linear correlation

e, f - non-linear correlation

Definition:

Covariance between 2 random variables X and Y is defined by

$$cov(X, Y) = E[(X-\mu_X)(Y-\mu_Y)] \text{ where } \mu_X = E(X),\ \mu_Y = E(Y)$$

* If $cov(X, Y) > 0$, it indicates X goes up and Y also tends to go up.
* If $cov(X, Y) < 0$, it indicates when X goes up, Y tends to go down
* When X and Y tend to go in the same direction then $(X-\mu_X)$ and $(Y-\mu_Y)$ have same signs (either both +ve or both -ve). We say, X and Y have positive correlation
* If X and Y tend to go in the opposite direction then $X-\mu_X$ and $Y-\mu_Y$ have opposite signs. We say, X and Y have -ve correlation.
* If $X-\mu_X$ and $Y-\mu_Y$ sometimes have same signs and sometimes have opposite signs, we say X and Y have no correlation.

Working rule:

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

Properties

* $cov(X, Y) = cov(Y, X)$
* $cov(X, X) = var(X)$
* $cov(X, c) = 0$
* $cov(kX, Y) = k\,cov(X, Y)$
* $cov(X+Y, Z) = cov(X, Z) + cov(Y, Z)$

* $var(aX+bY) = a^2 var(X) + b^2 var(Y) + 2ab\, cov(X,Y)$

$\underline{\underline{LHS}} =$ $var(aX+bY) = E[((aX+bY) - E(aX+bY))^2]$

$= \cancel{a^2E(X^2)+b^2E(Y^2) - a^2E(X^2)}$ $\mu_X = E(X)$

$= \cancel{E((aX+bY)^2) - [E(aX+bY)]^2}$ $\mu_Y = E(Y)$

$= \cancel{E(a^2X^2+b^2Y^2+2abXY)}$

$= E[((aX+bY) - a\mu_X - b\mu_Y)^2]$

$= E[(a(X-\mu_X) + b(Y-\mu_Y))^2]$

$= E[a^2(X-\mu_X)^2 + b^2(Y-\mu_Y)^2 + 2ab(X-\mu_X)(Y-\mu_Y)]$

$\therefore var(aX+bY) = a^2 var(X) + b^2 var(Y) + 2ab\, cov(X,Y)$

28/09/2022

Note: $cov(X,Y)$ is not independent of units

1) Suppose X and Y are measured in cm with $cov(X,Y) = 0.15$.

If X and Y are measured in mm let $X_1 = 10X$, $Y_1 = 10Y$

$cov(X_1, Y_1) = cov(10X, 10Y) = 10^2 cov(X,Y)$

$= \underline{\underline{15}}$

So to have the measure independent of units, we normalize the covariance formula to get coefficient of correlation.

$$\boxed{\rho_{XY} = \frac{cov(X,Y)}{\sigma_X \cdot \sigma_Y}}$$

OR

$$\boxed{\rho_{XY} = \frac{E(XY) - E(X)E(Y)}{\sigma_X \cdot \sigma_Y}}$$

X, Y independent $\Rightarrow cov(X,Y) = 0$

$cov(X,Y) \not\Rightarrow X$ and Y are independent

eg: Let $X \sim Uniform(-1,1)$ and $Y = X^2$

$E(X) = 0 \quad \because \left[E(X) = \frac{a+b}{2}\right]$

and $E(X^3) = 0$ $\quad \because x^3$ is odd function

$\Rightarrow \int x f(x) dx$

$\Rightarrow f(x) = \frac{1}{b-a}$

$cov(X,Y) = E(X,Y) - E(X).E(Y)$

$= E(X^3) - E(X).E(X^2)$

$= 0$

Properties of correlation.

theorem
lemma - short theorems
corollary - results following theorem

1) $-1 \leq \rho \leq 1$

cor 1: $var\left(\left(\frac{X}{\sigma X}\right) + \left(\frac{Y}{\sigma Y}\right)\right) = \frac{1}{\sigma_x^2} var(X) + \frac{1}{\sigma_y^2} var(X) + 2\left(\frac{1}{\sigma_{XY}^2}\right) cov(X,Y)$

$= \frac{1}{1} + \frac{1}{1} + 2\rho_{XY}$

$= 2 + 2\rho_{XY}$

cor 2: $var\left(\frac{X}{\sigma X} - \frac{Y}{\sigma Y}\right) = 2 - 2\rho_{XY}$

By cor1, $2 + 2\rho_{XY} \geq 0 \quad \because var$ is ≥ 0

$\rho_{XY} \geq -1$

By cor 2, $2 - 2\rho_{XY} \geq 0$

$\rho_{XY} \leq 1$

$\therefore -1 \leq \rho_{XY} \leq 1$.

2) If $\rho_{XY} = 1$, then $Y = aX + b$ for constants a and b and $a > 0$

3) If $\rho_{XY} = -1$, then $Y = aX + b$ with $a < 0$

2) Proof:

$$Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 2 - 2\rho_{XY} \quad \text{where } \rho \leq 1$$

$$Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 0$$

$$\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y} = K$$

$$\frac{Y}{\sigma_Y} = \frac{X}{\sigma_X} - K$$

$$Y = \frac{\sigma_Y X}{\sigma_X} - \sigma_Y K \quad (Y = aX + b)$$

where $a = \frac{\sigma_Y}{\sigma_X}$ ⇒ variance is +ve

∴ a is positive

∴ $a > 0$

3) Proof:

$$Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 0$$

$$\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y} = K$$

$$\frac{Y}{\sigma_Y} = -\frac{X}{\sigma_X} + K$$

$$\frac{Y}{\sigma_Y} = K - \frac{X}{\sigma_X}$$

$$Y = -\sigma_Y K - \frac{\sigma_Y X}{\sigma_X} \quad (Y = aX + b)$$

where $a = -\frac{\sigma_Y}{\sigma_X}$

∴ $a < 0$

29/09/2022

Inequalities and limit theorem

What should we do if we can't compute probability or expectation exactly?

Ans: Simulate it, bound it or approximate it (limit)

* Markov inequality
* Chebyshev inequality
* law of large nos (LLN)
* Central limit theorem

Markov's inequality.

Let x be a non-negative random variable with mean $E(x) = \mu$. Then

$$P[x \geq a] \leq \frac{\mu}{a}$$

Proof:

$$E(x) = \int_0^\infty x f(x) dx = \int_0^a x f(x) dx + \int_a^\infty x f(x) dx$$

$$\geq \int_a^\infty x f(x) dx$$

$$\geq \int_a^\infty a f(x) dx$$

$$\geq a \int_a^\infty f(x) dx$$

$$E(x) \geq a \, P[x \geq a]$$

$$P[x \geq a] \leq \frac{\mu}{a}$$

Interpretation

⇒ If a is large, it is very unlikely that $E(x) \geq a$.

⇒ If $x \geq 0$ and $E(x)$ is small then x is unlikely to be very large

⇒ Use a bit of information (μ is known) about the distribution to learn something about the probability of extreme events.

→ gives crude bound

The bound may not be close to the actual value.

→

Eg:1

Let $X \sim \text{Expo}(1)$. Then $P[X \geq a] \leq \frac{1}{a}$

$\lambda = 1$

$E(X) = \frac{1}{\lambda} = 1$

Actual probability: $P[X \geq a] = e^{-a}$

0 e^{-2} $1/2$

0.135 0.5

Actual probability is far away from the bound given by the Markov's inequality.

Eg: 2

Let $X \sim \text{Uniform}(-4, 4)$. Then $P[X \geq 3] \leq 1$

$|X| \sim \text{Uniform}(0, 4)$ Then $P[X \geq 3] = \frac{1}{2} P[|X| \geq 3]$

$$\leq \frac{1}{2} \frac{E[|X|]}{2}$$

$$\leq \frac{1}{2} \cdot \frac{2}{3}$$

$$P[X \geq 3] \leq \frac{1}{3}$$

$$P[X \geq 3] \leq P[|X| \geq 3]$$

$$\leq \frac{E(|X|)}{3}$$

$$\leq \frac{2}{3}$$

Actual probability: $P[X \geq 3] = \frac{1}{8}$ $\quad \frac{d-c}{b-a}$

0 1/8 1/3 2/3

The bounds are not close to actual probability 1/8.

Chebyshev's inequality.

Let X be a random variable with finite mean μ and finite var σ^2. Then $P[|X-\mu| \geq a] \leq \frac{\sigma^2}{a^2}$

Proof:

$$LHS = P[|X-\mu| \geq a] = P[\underbrace{(X-\mu)^2}_{\text{non-negative}} \geq a^2]$$

$$\leq \frac{E(X-\mu)^2}{a^2} \quad \text{by Markov's inequality}$$

$$\leq \frac{Var(X)}{a^2} = RHS$$

Eg 1:

Let $X \sim Expo(1)$

Use Chebyshev's inequality to find an upperbound for $P[X \geq a]$ if $X \sim Expo(1)$

Ans: To find: $P[X \geq a]$

$\lambda = 1$
$E(X) = \frac{1}{\lambda} = 1$
$Var(X) = \frac{1}{\lambda^2} = 1$

$$P[X \geq a] = P[X-1 \geq a-1] = \cancel{\leq} \; P[|X-1| \geq a-1] \quad (\because X \text{ is non-negative})$$

$$\leq \frac{Var(X)}{(a-1)^2}$$

$$\leq \frac{1}{(a-1)^2}$$

$$\cancel{\leq \frac{1}{a^2}} \approx \frac{1}{a^2} \quad \text{for large } a$$