Probability

Experiment - An experiment is a process from which we get a set of data.

Sample space - A set of all possible outcomes of an experiment is called sample space. It is denoted as S.

Ex: Exp - Rolling a die

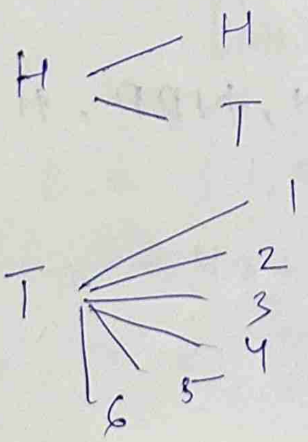
$S = \{1, 2, 3, 4, 5, 6\}$

Exp - Tossing two coins

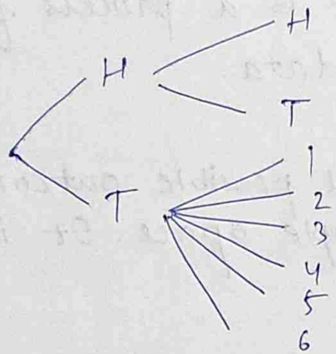
$S = \{HH, HT, TH, TT\}$

Q- Write the sample space of an experiment consisting of flipping a coin and then ~~flipping a coin twice and then~~ flipping a coin once if head appears & roll a die if tail appears.

Sol^n: $S = \{HH, HT, T1, T2, T3, T4, T5, T6\}$



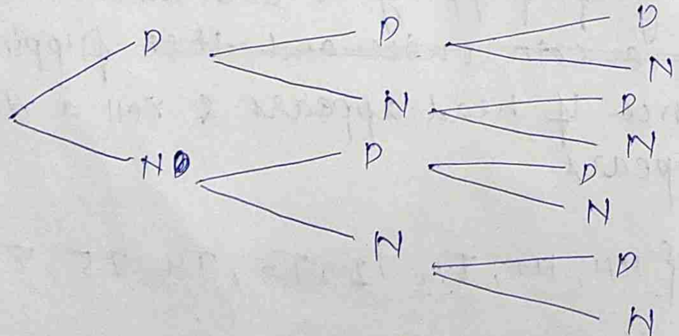
Tree diagram - It is a technique to list the elements of the sample space systematically



Q- Suppose that 3 items are selected at random from a manufacturing process each item is inspected and is classified as defective and non-defective. To list elements of the sample space using tree diagram.

Solⁿ

First Item | Second Item | Third Item



$\therefore$ S = { DDD, DDN, DND, DNN, NDD, NDN, NND, NNN }

Event - The event is a subset of the sample space of an experiment.

Experiment - Rolling a die

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

A, die shows even number

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

B, die shows odd number

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

C, die shows number less than 8

$C = \{1, 2, 3, 4, 5, 6, 7\}$

Q- Which of the following events are equal ?

$A = \{1, 3\}$

$B = \{x \mid x$ is a number on a die $\}$

$C = \{x \mid x^2 - 4x + 3 = 0\}$

$D = \{x \mid x$ is the number of heads when six coins flipped simultaneously $\}$

Solⁿ $B = \{1, 2, 3, 4, 5, 6\}$

$x^2 - 4x + 3 = 0$

$x = \frac{4 \pm \sqrt{16 - 12}}{2} = \frac{4 \pm 2}{2}$

$\Rightarrow x = 1, 3$

$\Rightarrow C = \{1, 3\}$

$D = \{0, 1, 2, 3, 4, 5, 6\}$

Event A and C are equal

Mutually Exclusive Event / Disjoint Events

Two events A & B are mutually exclusive ~~ex~~ events if there is no common ~~ele~~ event between A & B

$\Rightarrow A \cap B = \phi$

Ex: Suppose A is an event which shows even number by rolling a die and B is an event which shows odd number by rolling a die.

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

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

$\therefore A \cap B = \phi$

$\Rightarrow$ A & B are mutually exclusive event

union events

The union of two events A & B is an event in which an element lies in A or B ~~or both.~~

Intersection events

The intersection event of A & B is $A \cap B$ in which the element lies in A and B both or common element

Complement events

~~The~~ It is denoted by A', in which element in S or not in A. $A' = S - A$

Assignment - 3, 5, 7, 9, 11, 14

Probability of an event

The probability of an event A is the sum of weights of all sample points in A.

i.e. $P(A) = \dfrac{\text{Number of favourable cases}}{\text{Total no. of cases}}$

Rolling a fair die

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

A die shows even number

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

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

Q- Find the probability atmost two heads by tossing two coins simultaneously

Sol^n $S = \{HH, HT, TH, TT\}$

A is an event atmost two heads.

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

Properties

(1) For any event A of the sample space S, the probability of this event A always lies between $0 \le P(A) \le 1$.

(2) $P(\phi) = 0$, probability of null event,

(3) $P(S) = 1$,

Additive rule of Probability

If A and B are two events then ~~P(AUB~~

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

It is also written as:

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

Note: If A and B are mutually exclusive events then ~~prob~~ $P(A \cup B) = P(A) + P(B)$.

Q- Prove that $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + \cancel{P(A\cap} P(A \cap B \cap C)$

Sol^n Let $B \cup C = D$

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

$= P(A) + P(D) - P(A \cap D)$

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

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

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

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

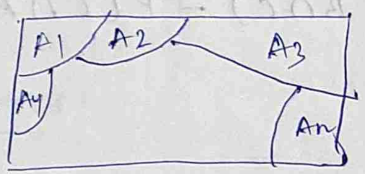
$= P(A) + P(B) + P(C) - P(B \cap C)$

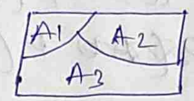
Note: If A, B and C are three mutually exclusive events then $P(A \cup B \cup C) = P(A) + P(B) + P(C)$.

If $A_1, A_2, A_3 \ldots\ldots A_n$ are mutually exclusive then $P(A_1 \cup A_2 \cup \ldots\ldots \cup A_n) = P(A_1) + P(A_2) + \ldots\ldots + P(A_n)$

Partition of an event

suppose $A_1, A_2, A_3 \ldots\ldots A_n$ are partitions of a sample space S then $P(A_1 \cup A_2 \cup \ldots\ldots \cup A_n) = 1$

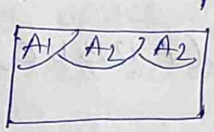




(partition) of S

$P(A_1 \cup A_2 \cup \ldots\ldots \cup A_n) = 1$

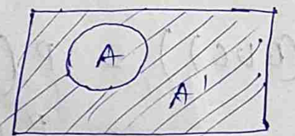
$\Rightarrow P(S) = 1$



Not a partition but mutually exclusive

Complement Rule

$A' = S - A$



$P(A') = 1 - P(A)$

[$A \cup A' = S$

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

$P(A \cup A') = P(S) = 1$

$P(A) + P(A') - P(A \cap A') = 1$]

$\therefore P(A) + P(A') = 1$

Q - prove that $P(A' \cap B') = 1 + P(A \cap B) - P(A) - P(B)$

Sol^n $P(A' \cap B') = P(A') + P(B') - P(A' \cup B')$

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

~~$P(A') + P(B') - P(A' \cup B') = P(A) + P(B) - P(A \cup B)$~~

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

$= 1 - P(A) + 1 - P(B) - P((A \cap B)')$

$= 1 - P(A) + 1 - P(B) - (1 - P(A \cap B))$

$= 1 + P(A \cap B) - P(A) - P(B)$

$=$ R.H.S

OR

L.H.S : $P(A' \cap B')$

$= P((A \cup B)')$

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

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

$= 1 + P(A \cap B) - P(A) - P(B)$

Q-50 An experiment involving tossing a pair of fair die one green and one red and recording the numbers turn up.

A is an event that ~~a number less than 8 occurs~~ sum is greater than 8 and C is an event a number greater than 4 comes up on the green die. Find P(A), P(C), P(A ∩ C)

Soln 1) A : sum > 8

C : 4 comes up on the green die.

$S = \{ (1,1) \ldots (1,6)$
$(2,1) \ldots (2,6)$
$\ldots$
$(6,1) \ldots (6,6) \}$

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

$P(A) = \frac{10}{36} \frac{5}{18} = \frac{5}{18}$

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

$P(C) = \frac{12}{36} = \frac{1}{3}$

~~$A \cap C = \{ (5,4), (6,4) \}$~~

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

$A \cap C = \{ (4,5), (5,5), (6,5), (6,4), (6,3), (6,6), (5,6) \}$

$P(A \cap C) = \frac{7}{36}$

Q 53 - The probability that an American Industry will locate in Shangai is 0.7. The probability that it will locate in Beijing is 0.4 and the probability it will locate in either the Shanghai or Beijing or both is 0.8. What is the probability the industry will locate (i) in both cities, (ii) in neither cities

Solⁿ (i) A is the event that industry will locate in shanghai

B event locate at beijing

$P(A \cap B) = P(A) + P(B) - P(A \cup B)$
$= 0.7 + 0.4 - 0.8$
$= 0.3$

~~(ii) $P(A \cap B) = 0.7 + 0.4 - 0.3$~~
~~$= 0.8$~~

(iii) $P(A' \cap B') = \cancel{P(A') + P(B')}\ 1 - P(A \cup B)$
$= P((A \cup B)') = 1 - 0.8$
$= 0.2$

Q 59 – In a poker hand consists of 5 cards. Find the probability of holding
a) 3 aces, b) 4 hearts and 1 club

Sol^n $^{52}C_5 = \frac{52!}{47!\,5!} = \frac{52 \times 51 \times 50 \times 49 \times 48}{120}$

$= 2598960$

(i) A = event of 3 aces

$P(A) = \frac{^4C_3 \times {}^{48}C_2}{^{52}C_5}$

$= \frac{4! \times 48!}{3! \times 46! \times 2! \times 2598960}$

$= \cancel{0.4512}\ 0.0017$

(ii) $P(B) = \frac{^{13}C_4 \times {}^{13}C_1}{^{52}C_5}$

$= \frac{13! \times 13!}{9!\,4!\,12! \times 2598960} = 0.0035$

Q 68 – Consider the decisions made by 6 customers purchase either a gas or electric oven. Suppose that the probability is 0.40 that at most 2 of these purchase an electric oven.

a) What is the probability that atleast 3 purchase the electric oven?

b) Suppose it is known that the probability of all 6 purchase the electric oven is 0.07. While 0.104 is all 6 purchase gas oven. What is the probability that at least one of each type is purchased?

Sol^n A : atmost 2 purchase electric oven
A' : atleast 3 purchase electric oven

C_1	C_2	C_3	C_4	C_5	C_6	
E	E	E	E	E	E	0.07
G	G	G	G	G	G	0.104

(a) $P(A') = 1 - P(A)$
$= 1 - 0.40$
$= 0.60$

(b) P (atleast one of each type)
$= 1 - (0.007 + 0.104)$
$= 0.889$

Assignment – 58, 58, 59, 65, 68, 72

Conditional Probability

The probability of an event B occuring when it is known that some event A has occured is the conditional probability.

The conditional probability is denoted as $P(B|A)$ reads as The probability of B given A

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

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

Ex Suppose that the sample space is the population of a small town according to gender and employment status which is given in the following table.

(i) What is the probability a person is randomly selected from the population who is male given that already he is employed.

	Employed	unemployed	Total
Male	460	40	500
Female	140	260	400
Total	600	300	900

(ii) Find the probability a person is selected randomly from the population who is employed person given that he is a male.

Solⁿ (i) Suppose M is an event

M : a person is male

and E is an event

E : employed person

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

$$P(M \cap E) = \frac{460}{900}$$

$$P(E) = \frac{600}{900} = \frac{2}{3}$$

$$\therefore P(M|E) = \frac{P(M \cap E)}{P(E)} = \frac{46}{90} \times \frac{3}{2} = \frac{23}{30}$$

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

$$P(E \cap M) = \frac{460}{900} = \frac{46}{90}$$

$$P(M) = \frac{500}{900} = \frac{5}{9}$$

$$\therefore P(E|M) = \frac{46}{90} \times \frac{9}{5} = \frac{23}{25}$$

(iii) Find the probability a person is selected at random is unemployed given that it is a femall person.

Sol^n: F : a person is femall

U : unemployed

P(U|F)

$\cancel{P(F|U)} = \cancel{\frac{P(F \cap U)}{P(U)}} \quad \frac{P(U \cap F)}{P(F)}$

P(U∩F)

$\cancel{P(F \cap U)} = \frac{260}{900}$

P(F)

$\cancel{P(U)} = \frac{\cancel{800}}{900}$

$\therefore P(U|F) = \frac{26}{90} \times \frac{9}{4} = \frac{13}{20}$

(iv) Find the probability a person is selected is femall and employed

Sol^n: $P(F|E) = \frac{P(F \cap E)}{P(E)}$

$P(F \cap E) = \frac{140}{900}$

$P(E) = \frac{600}{900}$

$P(F|E) = \frac{14}{90} \times \frac{9}{6} = \frac{7}{30}$

Independent Event

Two events A and B are independent iff

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

Q- Suppose an experiment in which two cards are drawn in succession from a pack of playing cards with replacement.

A : first card is an Ace

B : second card is club.

$P(B|A) = \frac{13}{52} = \frac{1}{4}$

$P(B) = \frac{13}{52} = \frac{1}{4}$

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

They are independent events.

Theorem If two events A and B are independent iff

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

Proof. $P(A|B) = \frac{P(A \cap B)}{P(B)}$

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

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

Theorem-2 If in an experiment the events $A_1, A_2, \ldots, A_K$ can occur then

$$P(A_1 \cap A_2 \cap \cdots \cap A_K) = P(A_1)\, P(A_2 | A_1)\, P(A_3 | (A_1 \cap A_2)) \cdots P(A_K | (A_1 \cap A_2 \cdots \cap A_{K-1}))$$

Note: If all the events $A_1, A_2, \ldots, A_K$ are independent then

$$P(A_1 \cap A_2 \cdots \cap A_K) = P(A_1) \cdot P(A_2) \cdots P(K)$$

Example-40 Three cards are drawn in succession without replacement from an ordinary deck of playing cards. Find the probability that the event $A_1 \cap A_2 \cap A_3$ occurs, where A_1 is the event that the first card is a red ace, A_2 is the event that the second card is a 10 or a jack, and A_3 is the event that the third card is greater than 3 but less than 7.

Soln.

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)\, P(A_2 | A_1)\, P(A_3 | (A_1 \cap A_2))$$

$$P(A_1) = \frac{2}{52} = \frac{1}{26}, \quad P(A_2) = \frac{8}{51}$$

$$P(A_3) = \frac{12}{50} = \frac{6}{25}$$

$$\therefore P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3) = \frac{1}{26} \times \frac{8}{51} \times \frac{6}{25}$$

2.75 A random sample 200 adults are classified below by sex and their level of education attained

Education	Male	Female
Elementary	38	45
secondary	28	50
college	22	17

If a person is picked at random from this group find the probability that

(a) the person is a male, given that the person has a secondary education.

(b) the person does not have a college degree, given that the person is a female.

Soln.

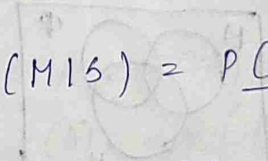
E : a person having elementary education

S : a person having secondary education

C : a person having college education.

M : a person is male.

F : a person is female.



(a) $P(M|S) = \dfrac{P(M \cap S)}{P(S)}$

$$P(M \cap S) = \frac{28}{200}, \quad P(S) = \frac{78}{200}$$

$$\therefore P(M|S) = \frac{28}{200} \times \frac{200}{78} = \frac{28}{78}$$

(b) $P(C'|F) = \dfrac{P(C' \cap F)}{P(F)}$

$P(C' \cap F) = \dfrac{95}{200}$, $P(F) = \dfrac{112}{200}$

$P(C'|F) = \dfrac{95}{200} \times \dfrac{200}{112} = \dfrac{95}{112}$

2.77 In the senior year of a high school graduating class of 100 students, 42 studied mathematics, 68 studied psychology, 54 studied history, 22 studied both mathematics and history, 7 studied history but neither mathematics nor psychology, 10 studied all the subject and 8 did not take any of the three. Randomly select a student from the class and find the probabilities of the following events

(a) A person enrolled in psychology takes all three subjects

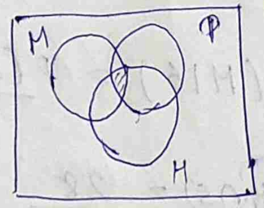
(b) A person not taking psychology is taking both history and mathematics.

Sol^n:

M : mathematics

P : psychology

H : History



$M \cap H = 22$, $M \cap P = 25$, $H \cap M' \cap P' = 7$

$M \cap P \cap H = 10$, $M' \cap H' \cap P' = 8$

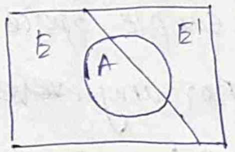
(a) $\dfrac{P(M \cap P \cap H)}{P(P)} = \dfrac{10}{68}$

(b) $\dfrac{P(M \cap H) - P(M \cap H \cap P)}{100 - P(P)}$

$= \dfrac{22 - 10}{100 - 68} = \dfrac{12}{32}$

Bayes Theorem

Total probability



$P(A) = P(E \cap A) + P(E' \cap A)$

or $P(A) = P(E)\,P(A|E) + P(E')\,P(A|E')$

Suppose S is the sample space of an experiment. Here E and E' are two partitions of S.

For any event A of S

$P(A) = P(E \cap A) + P(E' \cap A)$

Proof : $A = (E \cap A) \cup (E' \cap A)$

$P(A) = P\big((E \cap A) \cup (E' \cap A)\big)$

$P(A) = P(E \cap A) + P(E' \cap A) - \cancel{P((E \cap A) \cap (E' \cap A))} \to \phi$

$P(A) = P(E \cap A) + P(E' \cap A)$

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

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

$$P(E\cap A) = P(E)\,P(A|E)$$

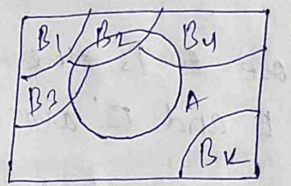
$$P(E'\cap A) = P(E')\,P(A|E')$$

General Formula

$B_1, B_2, B_3, \ldots B_K$ are the partitions of the sample space S. $P(B_i) \neq 0$, for $i = 1 \ldots K$

For any ~~evevent~~ event A of S

$$P(A) = P(B_1\cap A) + P(B_2\cap A) + \ldots\ldots + P(B_K\cap A)$$
$$= \sum_{i=1}^{K} P(B_i\cap A)$$



$$P(A) = P(B_1)\,P(A|B_1) + P(B_2)\,P(A|B_2) \ldots\ldots + P(B_K)\,P(A|B_K)$$
$$= \sum_{i=1}^{K} P(B_i)\,P(A|B_i)$$

Ex 2.41 In a certain assembly plant, three machines B_1, B_2, B_3 make 30%, 45% and 25% resp. of the products. It is known from past experience that 2%, 3% and 2% of the products made by each machine, respectively are defective. Now, suppose that a finished product is randomly selected. What is the probability that it is defective?

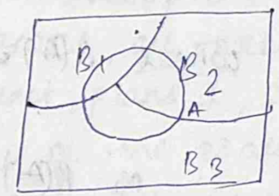
2%, 3%, 5%

5.1% A : the product is defective

30% B_1 : the product is made by machine B_1

45% B_2 : the product is made by machine B_2

25% B_3 : the product is made by machine B_3



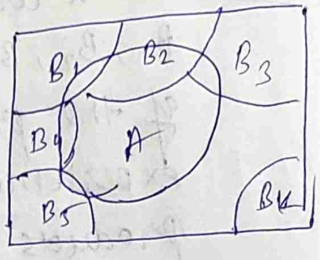
$P(B_1) = 0.30$, $P(B_2) = 0.45$, $P(B_3) = 0.25$

$P(A|B_1) = 0.02$, $P(A|B_2) = 0.03$, $P(A|B_3) = 0.02$

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)$$
$$= 0.024$$

Bayes' Theorem :-

If the events $B_1, B_2, \ldots, B_K$ constitute a partition ~~the~~ of the sample space S such that $P(B_i) \neq 0$, $\forall i = 1, 2, \ldots, K$ then for any event A such that $P(A) \neq 0$



$$P(B_r|A) = \frac{P(B_r \cap A)}{P(A)}$$

where $P(A) = \sum_{i=1}^{K} P(B_i \cap A)$

or $P(A) = \sum_{i=1}^{K} P(B_i)\, P(A|B_i)$

$r = 1, 2, \ldots, K$

Ex 2.42 With reference to example 2.41, if a product was chosen randomly and found to be defective what is the probability that it was made by machine B_3?

Sol^n $P(B_3|A) = \dfrac{P(B_3 \cap A)}{P(A)}$

where $P(A) = \sum_{i=1}^{3} P(B_i)\, P(A|B_i)$

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

$P(B_3 \cap A) = P(B_3)\, P(A|B_3)$

$= 0.25 \times 0.02$

$= 0.0050$

$\therefore P(B_3|A) = \dfrac{0.0050}{0.0245} = \dfrac{10}{49}$

Ex 2.43 A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varying times. In fact, plans 1, 2 and 3 are used for 30%, 20% and 50% of the products respectively. The defect rate is different for the three procedures as follows:

$P(D|P_1) = 0.01$, $P(D|P_2) = 0.03$, $P(D|P_3) = 0.02$

where $P(D|P_j)$ is the probability of a defective product, given plan j. If a random product was observed and found to be defective, which plan was most likely used and thus responsible?

Sol^n $P(P_1|D) = \dfrac{P(P_1 \cap D)}{P(D)}$

where

$P(D) = \sum_{i=1}^{3} P(P_i)\, P(D|P_i)$

$= 0.019$

$$\therefore P(P_1|D) = \frac{P(P_1 \cap D)}{P(D)}$$

$$= \frac{0.003}{0.019} = 0.158$$

$$P(P_2|D) = \frac{P(P_2 \cap D)}{P(D)}$$

$$P(P_2 \cap D) = P(P_1)P(D|P_2)$$
$$= 0.03 \times 0.02$$
$$= 0.0006$$

$$\therefore P(P_2|D) = \frac{0.006}{0.019} = 0.316$$

$$P(P_3|D) = \frac{P(P_3 \cap D)}{P(D)}$$

$$\cancel{P(P_3 \cap D) = P(P_2}$$

$$P(P_3|D) = \frac{(0.02)(0.50)}{0.019} = 0.526$$

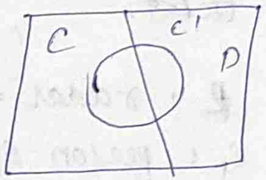
2.95 On a certain region of the country it is known from past experience that the probability of selecting an adult over 40 years of age with cancer is 0.05. If the probability of a doctor correctly diagnosing a person with cancer as having the disease is 0.78 and the probability of incorrectly diagnosing a person without cancer as having the disease is 0.06, what is the probability that an adult over 40 years of age is diagnosed as having cancer

Sol^n^ C : Adult selecting has cancer
D : Adult diagnosed as having cancer

$P(C) = 0.05$ $\quad P(C') = 0.95$

$P(D|C) = 0.78$ $\quad P(D|C') = 0.06$

(a) $P(D) = P(C)\,P(D|C) + P(C')\,P(D|C')$



$$= 0.05 \times 0.78 + 0.95 \times 0.06$$

$$= 0.096$$

2.97 What is the probability that a person diagnosed as having cancer actually has the disease?

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

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

$$= \frac{0.05 \times 0.78}{0.096}$$

$$= 0.40625$$

2.96 Police plan to enforce speed limits by using radar traps at four different locations within the city limits. The radar traps at each of the locations L_1, L_2, L_3, L_4 will be operated 40%, 30%, 20% and 30% of the time. If a person who is speeding on her way to work has probabilities 0.2, 0.1, 0.5 and 0.2 respectively, of passing through these locations, what is the probability that she will receive a speeding ticket.

Soln

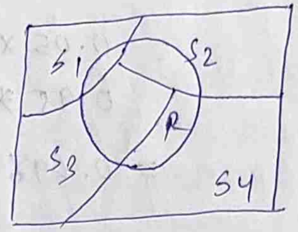
R : radar trap

S_1 : person speeding at L_1

S_2 : person speeding at L_2

S_3 : person speeding at L_3

S_4 : person speeding at L_4



$P(S_1) = 0.2$, $P(S_2) = 0.1$, $P(S_3) = 0.5$, $P(S_4) = 0.2$

$P(R|S_1) = 0.40$, $P(R|S_2) = 0.30$, $P(R|S_3) = 0.20$

$P(R|S_4) = 0.30$.

$P(R) = P(S_1)P(R|S_1) + P(S_2)P(R|S_2) + P(S_3)P(R|S_3) + P(S_4)P(R|S_4)$

$= 0.2 \times 0.40 + 0.1 \times 0.30 + 0.5 \times 0.20 + 0.2 \times 0.30$

$= 0.08 + 0.03 + 0.10 + 0.06$

$= 0.27$

2.98 If a person in exercise 96 received a speeding ticket on her way to work, what is the probability that she passed through radar trap located at L_2

Soln

$$P(S_2|R) = \frac{P(S_2 \cap R)}{P(R)}$$

$P(R) = 0.27$

$P(S_2 \cap R) = P(S_2) \cdot P(R|S_2)$

$= 0.1 \times 0.3$

$= 0.03$

$$P(S_2|R) = \frac{0.03}{0.27}$$

$= 0.11$

Random variables and probability distribution

(CHAPTER - 3)

Concept of random variables

Def^n A random variable is a function that associates a real number with each element of a sample space.

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

Generally a random variable is denoted by a capital letter.

Suppose an experiment 2 balls are drawn is succession without replacement from a bag containing 4 red ball, 3 blue balls and Y is a random variable that represent the number of red balls.

$S = \{RR, RB, BR, BB\}$

Y : No. of red balls

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

Discrete Sample space

A sample space if it contains a finite no. of possibilities or countable infinite no. of possiblities then that sample space is known as discrete sample space.

Continuous sample space

A sample space is said to be continuous sample space if it contains an infinite no. of possibilities equal to the no. of points on a line segment.

The random variables are of two types -

1) Discrete random variables
2) continuous random variables

Discrete random variables - A random variable is called discrete random variables if its set of possible outcomes are countable.

ex: Tossing 2 coins simultaneously and X is a random variable representing no. of heads.

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

Continuous random variables - A random variable ~~can~~ take some values on a continuous scale is called continuous random variables.

Bernoulli Random variables. A random variable is said to be Bernoulli random variable if it assigns either 0 or 1.

ex: Testing of a device ; for working → 1
for not working → 0

Discrete probability distribution

If X is a discrete random variable and $f(x)$ is the probability distribution ~~x~~ then for each x

① $f(x) \geq 0$

② $\sum_{x} f(x) = 1$

③ $P(X=x) = f(x)$

① $S = \{HH, HT, TH, TT\}$

X : No. of tails

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

X	0	1	2
$f(x)$	1/4	2/4	1/4

Cumulative distribution function

The cumulative distribution function ~~of a~~ $F(x)$ of a discrete random variable X with probability distribution function $f(x)$ is

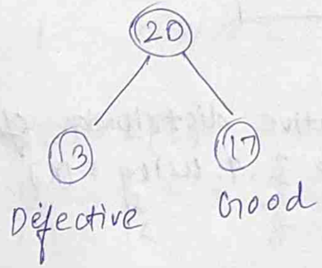
$$F(x) = P(X \leq x) = \sum_{x \leq x} f(x), \quad -\infty < x < \infty$$

X	0	1	2
$f(x)$	1/4	2/4	1/4

$F(0) = \frac{1}{4}$, $F(1) = \frac{3}{4}$, $F(2) = 1$

Ex-2.8 A shipment of 20 similar laptop computers to a retail outlet contains 3 that are defective. If a school makes a random purchase of 2 of these computers find the probability distribution for the no. of defectives.

Soln



X is a random variable representing no. of defective laptops.

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

$$f(0) = P(X=0) = \frac{^{17}C_2 \times {^3C_0}}{^{20}C_2}$$

$$= \frac{17!\ \cancel{2!}\ 18!\ 2!}{\cancel{2!}\ 15!\ 2!\ 20!}$$

$$= \frac{68}{95}$$

$$f(1) = P(X=1) = \frac{^3C_1 \times {^{17}C_1}}{^{20}C_2} = \frac{3!\ 17!\ 18!\ 2!}{2!\ 16!\ 20!}$$

$$= \frac{51}{190}$$

$f(2) = P(X=2) = \frac{{}^3C_2 \times {}^{17}C_0}{{}^{20}C_2}$

$= \frac{3! \times 17! \times \cancel{2!} \times 18!}{\cancel{2!} \times 20!}$

$= \frac{3 \times 17! \times 18!}{20 \times 19 \times 18 \times 17!} = \frac{3}{190}$

X	0	1	2
$f(x)$	$\frac{68}{95}$	$\frac{51}{190}$	$\frac{3}{190}$

Ex-3.10 Find the cumulative distribution of the random variable X in example 3.9. Using $F(x)$, verify that $f(2) = 3/8$

Soln

$f(x) = \frac{1}{16} {}^4C_x \quad , x = 0, 1, 2, 3, 4$

X : models with side airbags

$F(0) = f(0) = \frac{1}{16}$

$F(1) = f(0) + f(1) = \frac{5}{16}$

$F(2)$

$f(0) = \frac{1}{16} \qquad f(4) = \frac{1}{16}$

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

$f(2) = \frac{3}{8}$

$f(3) = \frac{1}{4}$

$F(0) = \frac{1}{16}, \quad F(1) = \frac{5}{16}, \quad F(2) = \frac{11}{16},$

$F(3) = \frac{15}{16}, \quad F(4) = 1$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{16} & 0 \le x < 1 \\ \frac{5}{16} & 1 \le x < 2 \\ \frac{11}{16} & 2 \le x < 3 \\ \frac{15}{16} & 3 \le x < 4 \\ 1 & x \ge 4 \end{cases}$$

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

$= \frac{11}{16} - \frac{5}{16} = \frac{3}{8}$

Continuous Probability distribution function

X is a continuous random variables

The function $f(x)$ is the probability density function if

① $f(x) \ge 0$ for all $x \in R$

② $\int_{-\infty}^{\infty} f(x)\, dx = 1$

③ $P(a < x < b) = \int_a^b f(x)\, dx$

$(a \le x < b)$

$(a < x \le b)$

$(a \le x \le b)$

Note: If X is a cts. random variable, the probability x lies between a to b

$$P(a < x < b) = P(a \le x < b) = P(a < x \le b) = P(a \le x \le b)$$

Cumulative distribution function

X $f(x)$

$F(x)$

The cumulative distribution function $F(x)$ of a continuous random variable X with the probability mass function $f(x)$ is

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = P(X \le x)$$

① $P(a < x < b) = F(b) - F(a)$

② $f(x) = \dfrac{d(F(x))}{dx}$

Ex 3.11 Suppose that the error in the reaction temperature in °C for a controlled laboratory experiment is a continuous random variable X having the probability density function

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

a) verify $f(x)$ is a density function

b) Find $P(0 \le x \le 1)$

Sol^n (a) $f(x) \ge 0$, for $x \in R$

(b) $\int_{-\infty}^{\infty} f(x)\,dx$

$$= \int_{-1}^{2} \frac{x^2}{3}\,dx \qquad \int_{-\infty}^{-1} \cancel{f(x)}\,dx + \int_{-1}^{2} f(x)\,dx + \int_{2}^{\infty} \cancel{f(x)}\,dx$$

$$= \int_{-1}^{2} \frac{x^2}{3}\,dx$$

$$= \left[\frac{x^3}{9}\right]_{-1}^{2} = 1$$

(C) $P(0 < x \le 1)$

$$= \int_{0}^{1} f(x)\,dx$$

$$= \int_{0}^{1} \frac{x^2}{3}\,dx$$

$$= \left[\frac{x^3}{9}\right]_{0}^{1} = \frac{1}{9}$$

Ex 3.12 For the density function of example 3.11. Find $F(x)$ and use it to evaluate $P(0 < X \le 1)$

Sol^n $F(x) = \int_{-\infty}^{x} f(t)\,dt$

$$= \int_{-\infty}^{-1} \cancel{f(t)\,dt} + \int_{-1}^{x} f(t)\,dt$$

$$= \int_{-1}^{x} \frac{t^2}{3}\,dt$$

$$= \left[\frac{t^3}{9}\right]_{-1}^{x} = \frac{x^3}{9} + \frac{1}{9} = \frac{x^3+1}{9}$$

$$F(x) = \begin{cases} 0 & , x < -1 \\ \frac{x^3+1}{9} & , -1 \le x < 2 \\ 1 & , x \ge 2 \end{cases}$$

$$P(0 < x \le 1) = F(1) - F(0)$$

$$= \frac{2}{9} - \frac{1}{9}$$

$$= \frac{1}{9}$$

3.3 Let w be a random variable giving the number of heads minus the number of tails in three tosses of a coin. List the elements of the sample space S for the three tosses of the coin and to each sample point assign a value w of W

Sol^n $S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$

$w = \{3, 1, -1, -3\}$

w	3	1	-1	-3
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

3.7 The total no. of hrs, measured in units of 100 hrs that a family runs a vacuum cleaner over period of one year is a continuous random variable X that has the density function

$$f(x) = \begin{cases} x, & 0 < x < 1 \\ 2-x, & 1 \le x < 2 \\ 0, & \text{elsewhere} \end{cases}$$

Find the probability that over a period of one year a family runs their vacuum cleaner

(a) less than 120 hrs.

(b) between 50 and 100 hrs.

Sol^n (a) $P(X < 1.2) = \int_{-\infty}^{1.2} f(x)\,dx$

$$= \int_{-\infty}^{0} \cancel{f(x)\,dx} + \int_{0}^{1} f(x)\,dx + \int_{1}^{1.2} f(x)\,dx$$

$$= \int_{0}^{1} x\,dx + \int_{1}^{1.2} (2-x)\,dx$$

$$= \left[\frac{x^2}{2}\right]_0^1 + \left[2x - \frac{x^2}{2}\right]_1^{1.2}$$

$$= \frac{1}{2} + \left[\left(2.4 - \frac{1.44}{2}\right) - \left(2 - \frac{1}{2}\right)\right]$$

$= \frac{1}{2} + 2.4 - 0.72 - 2 + \frac{1}{2}$

$= 0.68$

(b) $P(\frac{1}{2} < x < 1) = \int_{1/2}^{1} f(x)\,dx$

$= \int_{1/2}^{1} x \cdot dx$

$= \left[\frac{x^2}{2}\right]_{1/2}^{1}$

$= \frac{1}{2} - \frac{1}{8} = 3/8$

3.12 An investment firm offers its customers municipal bonds that mature after varying numbers of years. Given that the cumulative distribution function of T, the no of years to maturity for a randomly selected bond is

$$F(t) = \begin{cases} 0, & t < 1 \\ \frac{1}{4}, & 1 \le t < 3 \\ \frac{1}{2}, & 3 \le t < 5 \\ \frac{3}{4}, & 5 \le t < 7 \\ 1, & t \ge 7 \end{cases}$$

find

(a) $P(T=5)$;

(b) $P(T>3)$;

(c) $P(1.4 < T < 6)$;

(d) $P(T \le 5 \mid T \ge 2)$

Sol^n (a) $P(T=5) = F(5) - F(4)$

$= \frac{3}{4} - \frac{1}{2}$

$= \frac{1}{4}$.

(b) $P(T>3) = 1 - P(T \le 3)$

$= 1 - F(3)$

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

$= \frac{1}{2}$

(c) $P(1.4 < T < 6) = F(6) - F(1.4)$

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

$= \frac{1}{2}$

(d) $P(T \le 5 \mid T \ge 2)$

$= \frac{P[(T \le 5) \cap (T \ge 2)]}{P(T \ge 2)}$

$= \frac{P(2 \le T \le 5)}{P(T \ge 2)} = \frac{F(5) - F(2)}{1 - P(T < 2)}$

$= \frac{\frac{3}{4} - \frac{1}{4}}{1 - F(2)}$

$= \frac{\frac{2}{4}}{1 - \frac{1}{4}} = \frac{2}{3}$

3.21 Consider the density function

$$f(x) = \begin{cases} K\sqrt{x}, & 0 < x < 1, \\ 0, & \text{elsewhere} \end{cases}$$

(a) Evaluate K

(b) Find F(x) and use it to evaluate $P(0.3 < X < 0.6)$

Sol^n (a) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

$$= \int_{-\infty}^{0} \cancel{f(x)}^{0} dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{\infty} \cancel{f(x)}^{0} dx = 1$$

$$= \int_0^1 K\sqrt{x}\,dx = 1$$

$$= K\left[\frac{x^{3/2}}{3/2}\right]_0^1 = 1$$

$$\Rightarrow K = 3/2$$

(b) $P(0.3 < X < 0.6)$

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

$$= \int_{-\infty}^{0} \cancel{f(t)}^{0} dt + \int_0^x f(t)\cdot dt$$

$$= \int_0^x \frac{3}{2}\sqrt{t}\cdot dt$$

$$= \frac{3}{2}\left[\frac{t^{3/2}}{3/2}\right]_0^x$$

$$= x^{3/2}$$

$$F(x) = \begin{cases} 0, & x < 0 \\ x^{3/2}, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$$

$$P(0.3 < X < 0.6) = F(0.6) - F(0.3)$$

$$\cancel{= 0.6^{3/2} - (0.3)}$$

$$= (0.6)^{3/2} - (0.3)^{3/2}$$

$$= 0.3004$$

Joint Probability distribution

Let us consider X and Y are two discrete random variables. The function $f(x,y)$ is a joint probability distribution of X and Y if

① $f(x,y) \geq 0$ for all (x,y)

② $\sum_x \sum_y f(x,y) = 1$

③ $P(X = x, Y = y) = f(x,y)$

For any region A in ~~xy~~ x-y plane.

$$P((X,Y) \in A) = \sum\sum_A f(x,y)$$

For continuous case

X, Y continuous random variable

$f(x,y)$

① $f(x,y) \geq 0$ for all (x,y)

② $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$

③ $P((X,Y)\in A) = \iint_A f(x,y)\,dx\,dy$

where A is any region in my plane.

Marginal distributions

Marginal distributions of the random variables X ~~along~~ and Y ~~along~~ alone.

$g(x) = \sum_y f(x,y)$ (X alone) Discrete random variable

$h(y) = \sum_x f(x,y)$ (Y alone)

For cts. random variable

$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$ (X alone)

$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$ (Y alone)

Ex 3.14 Two ballpoint pens are selected at random from a box that contains 3 blue pens, 2 red pens, 3 green pens. If X is the number of blue pens selected and Y is the number of red pens selected find

(a) the joint probability function $f(x,y)$

(b) $P[(X,Y)\in A]$, where A is the region $\{(x,y) \mid x+y \leq 1\}$.

Sol^n $^8C_2 = \frac{8!}{2!\,6!} = 28$

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

$f(x,y)$ Y \ X	0	1	2
0	$f(0,0)$	$f(1,0)$	$f(2,0)$
1	$f(0,1)$	$f(1,1)$	$f(2,1)$
2	$f(0,2)$	$f(1,2)$	$f(2,2)$

$f(0,0) = \frac{^3C_2 \times {^3C_0} \times {^2C_0}}{28}$

$= \frac{3}{28}$

$f(1,0) = \frac{^3C_1 \times {^2C_0} \times {^3C_1}}{28}$

$= \frac{9}{28}$

$f(2,0) = \frac{{}^3C_2 \times {}^2C_0 \times {}^3C_0}{28} = \frac{3}{28}$

$f(0,1) = \frac{{}^3C_0 \times {}^2C_1 \times {}^3C_1}{28} = \frac{6}{28}$

$f(1,1) = \frac{{}^3C_1 \times {}^2C_1 \times {}^3C_0}{28} = \frac{6}{28}$

$f(2,1) = 0$

$f(0,2) = \frac{{}^3C_0 \times {}^2C_2 \times {}^3C_0}{28} = \frac{1}{28}$

$f(1,2) = 0$

$f(2,2) = 0$

General formula: $f(x,y) = \frac{{}^3C_x \, {}^2C_y \, {}^3C_{2-(x+y)}}{{}^8C_2}$

$x = 0,1,2$

$y = 0,1,2$

(b) $P((x,y) \in A) = f(0,0) + f(1,0) + f(0,1)$

$= \frac{3}{28} + \frac{6}{28} + \frac{9}{28}$

$= \frac{18}{28} = \frac{9}{14}$

(c) $P((x,y) \in A)$

$A = \{(x,y) \mid x+y \geq 1\}$

$P((x,y) \in A) = 1 - \frac{3}{28} = \frac{25}{28}$

Ex 3.16 (CBX-141) Show that the column and row totals of Table 3.1 give the marginal distribution of X alone and of Y alone respectively.

Soln The marginal distribution of X alone is

$g(x) = \sum_y f(x,y)$

$x=0$, $g(0) = \sum_y f(0,y) = f(0,0) + f(0,1) + f(0,2)$

$= \frac{3}{28} + \frac{6}{28} + \frac{1}{28}$

$= \frac{10}{28}$

$x=1$, $g(1) = \sum_y f(1,y) = f(1,0) + f(1,1) + f(1,2)$

$= \frac{15}{28}$

$x=2$, $g(2) = \sum_y f(2,y) = \frac{3}{28}$

x	0	1	2
$g(x)$	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

(column sum)

Ans The marginal distribution of Y alone is

$h(y) = \sum_x f(x,y)$

$y=0$, $h(0) = \sum_x f(x,0) = \frac{15}{28}$

$y=1$, $h(1) = \frac{12}{28}$

$y=2$, $h(2) = \frac{1}{28}$

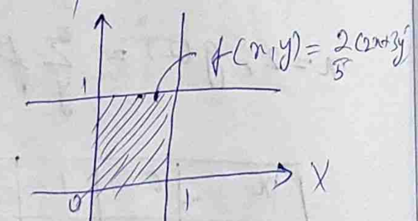
Ex 3.15 A privately owned business operates both a drive-in facility and a walk-in facility. On a randomly selected day let X and Y, respectively be the proportions of the time that the drive-in and the walk-in facilities are in use, and suppose that the joint density function of these random variable is

$$f(x,y) = \begin{cases} \frac{2}{5}(2x+3y), & 0 \le x \le 1,\ 0 \le y \le 1, \\ 0, & \text{elsewhere} \end{cases}$$

(a) verify condition 2 of Definition 3.9

(b) Find $P[(X,Y) \in A]$, where $A = \{(x,y) \mid 0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\}$

Soln (a) ① $f(x,y) \geq 0$



② $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy$

$= \int_0^1\int_0^1 \frac{2}{5}(2x+3y)\,dx\,dy$

$= \frac{2}{5}\int_0^1\int_0^1 (2x+3y)\,dx \cdot dy$

$= \frac{2}{5}\left[\int_0^1 2x\,dx + \int_0^1 3y\,dy\right]$

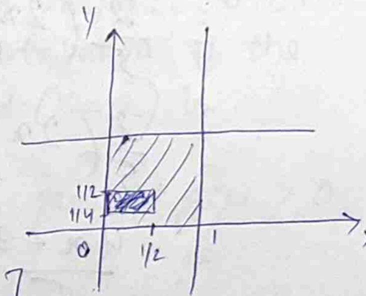
$= \frac{2}{5}\left([x^2]_0^1 + \left[\frac{3y^2}{2}\right]_0^1\right)$

$= \frac{2}{5}\left[1 + \frac{3}{2}\right]$

$= \frac{2}{5} \times \frac{5}{2} = 1$

$\Rightarrow f(x,y)$ is a mass density function

(b) $P[(x,y) \in A]$



$= \int_{1/4}^{1/2}\int_0^{1/2} \frac{2}{5}(2x+3y)\,dx\,dy$

$= \frac{2}{5}\left[\int_0^{1/2} 2x\,dx + \int_{1/4}^{1/2} 3y\,dy\right]$

$= \frac{2}{5}\left[[x^2]_0^{1/2} + \left[\frac{3y^2}{2}\right]_{1/4}^{1/2}\right]$

$= \frac{2}{5}\left[\frac{1}{4} + \left[\frac{3 \times 1}{4 \times 2} - \frac{3 \times 1}{16 \times 2}\right]\right]$

$= \frac{2}{5}\left[\frac{1}{4} + \left[\frac{3}{8} - \frac{3}{32}\right]\right]$

$= \frac{2}{5}\left[\frac{1}{4} + 3\left[\frac{3}{32}\right]\right]$

$= \frac{2}{5}\left[\frac{9}{32} + \frac{1.8}{4.8}\right]$

$= \frac{2}{5}\left[\frac{17}{32}\right] = \frac{34}{160}$ -13

Ex 3.17 Find $g(x)$ and $h(y)$ for the joint density function of Example 3.15

Sol^n The marginal distribution of X alone

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$= \int_0^1 \frac{2}{5}(2x+3y)\,dy$$

$$= \frac{2}{5}\left(2xy + \frac{3}{2}y^2\right)_0^1$$

$$= \frac{2}{5}\left(2x + \frac{3}{2}\right)$$

$$= \frac{4x+3}{5}$$

$$g(x) = \begin{cases} \frac{4x+3}{5}, & 0 \le x \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

The marginal distribution of Y alone

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

$$= \int_0^1 \frac{2}{5}(2x+3y)\,dx$$

$$= \frac{2}{5}\int_0^1 (2x+3y)\,dx$$

$$= \frac{2}{5}\left(x^2 + 3xy\right)_0^1$$

$$= \frac{2}{5}\left[1 + 3y\right]$$

$$h(y) = \begin{cases} \frac{2}{5}[1+3y], & 0 \le y \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

<u>Conditional distribution</u>

<u>Continuous distribution</u>

Let x and y be two random variables discrete or continuous. The conditional distribution of the random variable Y given that $(X=x)$ is

$$f(y|x) = \frac{f(x,y)}{g(x)}, \text{ provided } g(x) > 0$$

Similarly, the conditional distribution of X given that $Y=y$ is

$$f(x|y) = \frac{f(x,y)}{h(y)}, \quad h(y) > 0$$

$$P(a < x < b \mid Y=y) = \sum_{a<x<b} f(x|y)$$ (for discrete random variable)

~~Rcaxcb~~

$$P(a<x<b \mid Y=y) = \int_a^b f(x|y)\,dx \quad \text{(for cts. random variable)}$$

$$P(c<Y<d \mid X=x) = \sum_{c<y<d} f(y|x) \quad \text{(for discrete random variable)}$$

$$P(c<Y<d \mid X=x) = \int_c^d f(y|x)\,dy \quad \text{(for cts. random variable)}$$

Independent random variables

$X, Y \quad f(x,y) \quad g(x) \quad h(y)$

If $f(x,y) = g(x)\,h(y)$

3.38 If joint probability distribution of X and Y is given by

$$f(x,y) = \frac{x+y}{30} \quad \text{for } x = 0,1,2,3; \; y = 0,1,2$$

find

(a) $P(X \le 2, Y = 1)$;

(b) $P(X > 2, Y \le 1)$;

(c) $P(X > Y)$;

(d) $P(X + Y = 4)$

Sol^n (a) $P(X \le 2, Y = 1) = f(0,1) + f(1,1) + f(2,1)$

$$= \frac{1}{30} + \frac{2}{30} + \frac{3}{30} = \frac{6}{30} = \frac{1}{5}$$

(b) $P(X > 2, Y \le 1) = f(3,0) + f(3,1)$

$$= \frac{3}{30} + \frac{4}{30} = \frac{7}{30}$$

(c) $P(X > Y) = f(1,0) + f(2,0) + f(2,1) + f(3,0) + f(3,1) + f(3,2)$

$$= \frac{1}{30} + \frac{2}{30} + \frac{3}{30} + \frac{3}{30} + \frac{4}{30} + \frac{5}{30} = \frac{18}{30}$$

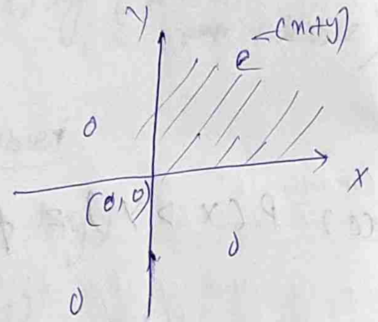
(d) $P(X + Y = 4) = f(2,2) + f(3,1)$

$$= \frac{4}{30} + \frac{4}{30} = \frac{8}{30}$$

3.42 Let X and Y denote the lengths of life, in years of two components in an electronic system. If the joint density function of these variables is

$$f(x,y) = \begin{cases} e^{-(x+y)}, & x>0,\ y>0 \\ 0, & \text{elsewhere,} \end{cases}$$

find $P(0<X<1 \mid Y=2)$



Sol^n $P(a<X<b \mid Y=y)$

$$= \int_a^b f(x|y)\,dx$$

$$= \int_0^1 e^{-x}\,dx$$

$$= -e^{-x}\Big]_0^1$$

$$= -e^{-1}+1$$

$$= 1-e^{-1}$$

$$= 1-\frac{1}{e}$$

$$f(x|y) = \frac{f(x,y)}{h(y)}$$

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

$$= \int_0^{\infty} e^{-(x+y)}\,dx$$

$$= \int_0^{\infty} e^{-x}.e^{-y}\,dx$$

$$= e^{-y}\left[e^{-x}\right]_0^{\infty}$$

$$= e^{-y}[1] = e^{-y}$$

$$f(x|y) = \frac{e^{-(x+y)}}{e^{-y}} = e^{-x}$$

3.49 Let X denote the number of times a certain numerical control machine will malfunction: 1, 2, or 3 times on any given day. Let Y denote the no. of times a technician is called on an emergency call. Their joint probability distribution is given as

$f(x,y)$		x = 1	2	3
	1	0.05	0.05	0.10
y	3	0.05	0.10	0.35
	5	0.00	0.20	0.10

(a) Evaluate the marginal distribution of X.
(b) Evaluate the marginal distribution of Y.
(c) Find $P(Y=3 \mid X=2)$

Sol^n (a) $g(x) = \sum_y f(x,y)$

X	1	2	3
$g(x)$	0.10	0.25	0.55

~~g(1) = Σ f(1,y)~~
(column sum)

$g(1) = \sum_y f(1,y) = f(1,1) + f(1,3) + f(1,5)$

(b)

Y	1	2	3
$h(y)$	0.20	0.50	0.30

(row sum)

(c) $P(Y=3 \mid X=2) = \dfrac{f(x,y)}{g(x)}$

$= \dfrac{f(2,3)}{g(2)} = \dfrac{0.10}{0.35}$

$= \dfrac{2}{7}$

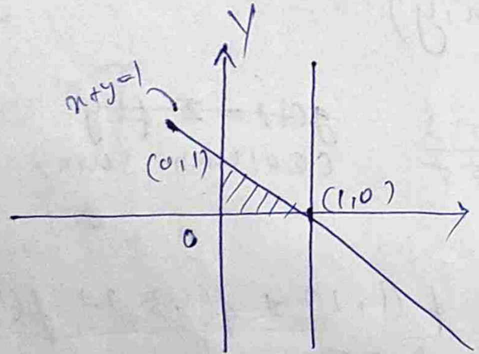
3.56 The joint density function of the random variables X and Y is

$$f(x,y) = \begin{cases} 6x, & 0<x<1,\ 0<y<1-x \\ 0, & \text{elsewhere} \end{cases}$$

(a) Show that X and Y are not independent

(b) Find $P(X>0.3 \mid Y=0.5)$

Sol^n



(a) $g(x) = \int_0^{1-x} f(x,y)\,dy = \int_0^{1-x} 6x\,dy$

$h(y) = \int_0^{1-y} f(x,y)\,dx = \int_0^{1-y} 6x\,dx$

(b) $P(X>0.3 \mid Y=0.5) = \int_{0.3}^{0.5} f(x|y)\,dx$

$f(x|y) = \dfrac{f(x,y)}{h(0.5)}$

Mathematical expectation
Ch - 4

Def^n Mean of a random variable

$x_1, x_2, x_3 \ldots\ldots\ldots, x_n$

$$\text{Avg / Mean} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}$$

$$= \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

Tossing a coin twice

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

Let X be a random variable that represents the no. of tails

$X = \{0, 1, 2\}$ (Discrete)

$f(x)$ probability distribution

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

The mean of random ~~var~~ variable X is known as expected value of X

* $E(X) = \mu$

$$E(X) = 0 \times \frac{1}{4} + 1 \times \frac{2}{4} + 2 \times \frac{1}{4}$$

$$= 0 + \frac{2}{4} + \frac{2}{4} = 1$$

* $E(X) = \sum_x x f(x)$

Note: If X is a cts. random variable with probability density $f(x)$. Then

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

Theorem 1 Let X be a random variable with probability distribution function $f(x)$

The expected value of $g(x)$ is

$$E(g(x)) = \mu_{g(x)} = \sum_x g(x) f(x) \quad \text{(discrete)}$$

$$E(g(x)) = \mu_{g(x)} = \int_{-\infty}^{\infty} g(x) f(x)\, dx$$

Theorem 3 : Let X and Y be two random variables with probability distribution function $f(x,y)$ then expected value of $g(X,Y)$ is

$$E(g(X,Y)) = \sum_x \sum_y g(x,y) f(x,y) \quad \text{(discrete)}$$

$$E(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f(x,y)\, dx\, dy \quad \text{(continuous)}$$

Note: The expected value of X alone is $E(X)$

$$E(X) = \mu_X = \sum_x \sum_y x f(x,y) = \sum_x x\, g(x) \quad \text{(discrete)}$$

~~The expected value of Y alone is E(Y)~~

$$E(X) = \mu_X = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f(x,y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} x\, g(x)\, dx \quad \text{(continuous)}$$

The expected value of Y alone is $E(Y)$

$$E(Y) = \mu_Y = \sum_x \sum_y y f(x,y) = \sum_y y\, g(y) \quad \text{(discrete)}$$

$$E(Y) = \mu_Y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, f(x,y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} y\, g(y)\, dy \quad \text{(continuous)}$$

4.4 A coin is biased such that a head is three times ~~as likely~~ as likely to occur as a tail. Find the expected number of tails when this coin is tossed twice.

Sol^n: $P(T) = K$ $\qquad P(H) + P(T) = 1$

$P(H) = 3K$ $\qquad \Rightarrow K + 3K = 1$

$\Rightarrow 4K = 1$

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

$\therefore P(T) = \frac{1}{4}$

$P(H) = \frac{3}{4}$

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

X : Number of tails

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

x	0	1	2
$f(x)$	$\frac{9}{16}$	$\frac{6}{16}$	$\frac{1}{16}$

$$E(X) = 0 \times \frac{9}{16} + 1 \times \frac{6}{16} + 2 \times \frac{1}{16}$$

$$= 0 + \frac{6}{16} + \frac{2}{16}$$

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

4.12 If a dealer's profit, in units of \$5000, on a new automobile can be looked upon as a random variable X having the density function.

$$f(x) = \begin{cases} 2(1-x), & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

find the avg. profit per automobile.

Sol^n

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

$$= \int_{-\infty}^{0} x \cancel{f(x)}^{0} dx + \int_{0}^{1} x \cdot f(x)\, dx + \int_{1}^{\infty} \cancel{x f(x)\, dx}^{0}$$

$$= \int_0^1 2x(1-x)\, dx$$

$$= \int_0^1 2x - 2x^2\, dx$$

$$= \left[x^2 - \frac{2x^3}{3}\right]_0^1$$

$$= 1 - \frac{2}{3} = \frac{1}{3}$$

The total profit = $\$5000 \times \frac{1}{3}$

$\approx \$1666.7$

4.10 Two tire-quality experts examine stacks of tires and assign a quality rating to each tire on a 3-point scale. Let X denote the rating given by expert A and Y denote the rating given by B. The following table gives the joint distribution for X and Y.

$f(x,y)$		$y=1$	$y=2$	$y=3$
x	1	0.10	0.05	0.02
	2	0.10	0.35	0.05
	3	0.03	0.10	0.20

find μ_x and μ_y

Sol^n

$$g(x) = \sum_y f(x,y)$$

$$g(1) = \sum_y f(1,y)$$

$$= f(1,1) + f(1,2) + f(1,3)$$

$$= 0.10 + 0.05 + 0.02$$

$$= 0.17$$

$$g(2) = 0.10 + 0.35 + 0.05 = 0.50$$

$$g(3) = 0.03 + 0.10 + 0.20 = 0.33$$

x	1	2	3
$g(x)$	0.17	0.50	0.33

$\mu_x = 1\times 0.17 + 2\times 0.50 + 3\times 0.33$

$= 0.17 + 1 + 0.99$

$= 2.16$

Y	1	2	3
$g(y)$	0.23	0.50	0.27

$\mu_y = 1\times 0.23 + 2\times 0.5 + 3\times 0.27$

$= 0.23 + 1 + 0.81$

$= 2.04$

4.20 A continuous random variable X has the density function

$$f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

Find the expected value of $g(X) = e^{2X/3}$

Soln $E(g(x)) = \int_{-\infty}^{\infty} g(x) f(x)\, dx$

$= \int_{-\infty}^{0} \cancel{g(x) f(x)\, dx} + \int_{0}^{\infty} g(x) f(x)\, dx$

$= \int_{0}^{\infty} e^{2x/3} \cdot e^{-x}\, dx$

$= \int_{0}^{\infty} e^{-x/3}\, dx$

$= \left. \frac{e^{-\frac{x}{3}}}{-1/3} \right]_0^{\infty}$

$= -3 e^{-x/3} \Big]_0^{\infty}$

$= -3 (e^{-\infty} - e^{0})$

$= 3$

4.23 Suppose that X and Y have the following joint probability function:

$f(x,y)$		$x=2$	$x=4$
y	1	0.10	0.15
	3	0.20	0.30
	5	0.10	0.15

(a) Find the expected value of $g(X,Y) = XY^2$

(b) Find μ_x and μ_y

X, Y are a discr

Sol^n X, Y are discrete random variable with probability distribution function is $f(x)$

$g(x,y) = xy^2$

$E(g(X,Y)) = \sum_x \sum_y g(x,y) f(x,y)$

$= \sum_x \sum_y x y^2 f(x,y)$

$= 2f(2,1) + 18f(2,3) + 50f(2,5) + 4f(4,1) + 36f(4,3) + 100f(4,5)$

$= 2 \times 0.10 + 18 \times 0.20 + 50 \times 0.10 + 4 \times 0.15 + 36 \times 0.30 + 100 \times 0.15$

$= \cancel{0.2 + 0.36 + 5 + 0.6 + 1.08} + 15$

$= 35.2$

(b) $\mu_X = \sum_x x g(x)$

X	2	4
$g(x)$	0.40	0.60

$g(x) = \sum_y f(x,y)$

$g(2) = \sum_y f(2,y)$

$= f(2,1) + f(2,3) + f(2,5) =$

$= 0.10 + 0.20 + 0.10$

$= 0.40$

$\therefore \mu_x = 2 \times 0.40 + 4 \times 0.60$

$= 0.80 + 2.4$

$= 3.20$

$\mu_Y = \cancel{\sum_y y g(y)} \sum_y y h(y)$

$h(y) = \sum_y f(x,y)$

Y	1	3	5
$h(y)$	0.25	0.50	0.25

$h(y) = \cancel{0.25 + 0.50 + 0.25} = \cancel{1}$

$\mu_y = 1 \times 0.25 + 3 \times 0.50 + 5 \times 0.25$

$= 0.25 + 1.5 + 1.25$

$= 3$

4.26 Let X and Y be random variables with joint density function

$$f(x,y) = \begin{cases} 4xy, & 0 < x, y < 1, \\ 0 & \text{elsewhere} \end{cases}$$

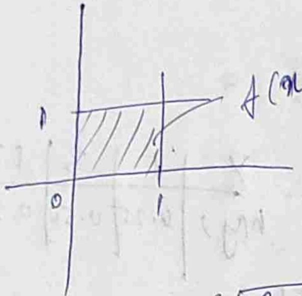
Find the expected value of $Z = \sqrt{X^2 + Y^2}$

Sol^n: $g(x)\ z = \sqrt{x^2+y^2}$

$$E[g(X,Y)] = \sum_x \sum_y g(x,y) f(x,y)$$

$$= \sum_x \sum_y \sqrt{x^2+y^2} f(x,y)$$

$$= \cancel{\sum_x \sum_y}$$



$f(x,y) = 4xy$

$$E(Z) = E(\sqrt{x^2+y^2})$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sqrt{x^2+y^2}\, f(x,y)\, dx\, dy$$

$$= \int_0^1\int_0^1 \sqrt{x^2+y^2}\ 4xy\ dx\ dy$$

$$= 4\int_0^1 x\sqrt{x^2+y^2}\,dx \int_0^1 y\, dy$$

$$= \int_0^1 \left[\int_0^1 4xy\sqrt{x^2+y^2}\, dx\right] dy$$

$$= y\int_{y^2}^{y^2+1} 2\sqrt{t}\, dt$$

$x^2+y^2 = t$

$2x\, dx = dt$

$$= 2y \times \frac{2}{3} t^{3/2}\Big]_{y^2}^{y^2+1}$$

$$= \frac{4y}{3}\left[(y^2+1)^{3/2} - (y^2)^{3/2}\right]$$

$$= \frac{4y}{3}\left[(y^2+1)^{3/2} - y^3\right]$$

$$= \int_0^1 \frac{4y}{3}\left[(y^2+1)^{3/2} - y^3\right] dy$$

$$= \int_0^1 \frac{4y}{3}(y^2+1)^{3/2}\, dy - \int_0^1 \frac{4y^4}{3}\, dy$$

$$= \cancel{\frac{4}{3}\int_0 \frac{du}{2.dy} \cdot 2(u)^{3/2} \frac{du}{2y}}$$

$y^2+1 = u$

$2y.dy = du$

$dy = \frac{du}{2y}$

$$= \int_1^2 \frac{2}{3}(u)^{3/2} du - \left[\frac{4}{3}\frac{y^5}{5}\right]_0^1$$

$$= \frac{2}{3} \times \frac{2}{5} u^{5/2}\Big]_1^2 - \left[\frac{4}{15} y^5\right]_0^1$$

$$= \frac{4}{15} \times 4\sqrt{2} - \frac{8}{15}$$

$$= \cancel{\frac{16\sqrt{2}}{15} - \frac{4}{15}}$$

$$= \frac{8}{15}(2\sqrt{2} - 1)$$

Variance and covariance of random variables

Defⁿ (variance) Let X is a random variable ~~which~~ with probability function $f(x)$ and mean μ. The variance of X is σ^2

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

For discrete:

$$\sigma^2 = \sum_x (x-\mu)^2 f(x) \qquad \mu = E(x) = \sum_x x f(x)$$

For continuous

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$$

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

Note: Let X be a random variable with probability function $f(x)$ then $\sigma^2 = E(x^2) - \mu^2$

Q Prove that $\sigma^2 = E(x^2) - \mu^2$ where X is a discrete random variable and μ is mean.

Solⁿ Proof: L.H.S $= \sigma^2 = \sum_x (x-\mu)^2 f(x)$

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

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

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

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

$$= E(x^2) - \mu^2 = \text{R.H.S}$$

Standard deviation

The +ve sqrt. of variance is known as standard deviation

$$\sigma = +\sqrt{\sigma^2}$$

Theorem

Let X be a random variable with probability function $f(x)$ and mean μ_X then the variance of ~~and~~ $g(x)$ is $\sigma^2_{g(x)}$

For discrete

$$\sigma^2_{g(x)} = \sum (g(x) - \mu_X)^2 f(x)$$

For continuous

$$\sigma^2 = \int_{-\infty}^{\infty} (g(x) - \mu_X)^2 f(x)\, dx$$

Ex 4.9 Let the random variable X represent the number of defective parts for a machine when 3 parts are sampled from a production line and tested. The following is the probability distribution of X

x	0	1	2	3
$f(x)$	0.51	0.38	0.10	0.01

Calculate σ^2 using defn & $\sigma^2 = E(X^2) - \mu^2$

Soln (a) $\mu = \sum_x x f(x)$

$= (0)\times 0.51 + 1\times 0.38 + 2\times 0.10 + 3\times 0.01$

$= 0.61$

$\sigma^2 = \sum_x ((x-\mu)^2 f(x))$

$= (0-0.61)^2 \times 0.51 + (1-0.61)^2 \times 0.38 + (2-0.61)^2 \times 0.10 + (3-0.61)^2 \times 0.1$

$= 0.4979$

(b) $X, f(x)$ $\quad \sigma^2 = E(X^2) - \mu^2$

$E(X^2) = \sum_x x^2 f(x)$

$= 0^2 \times 0.51 + 1^2 \times 0.38 + 2^2 \times 0.10 + 3^2 \times 0.01$

$= 0.38 + 0.40 + 0.09$

$= 0.87$

$\sigma^2 = 0.87 - (0.61)^2$

$= 0.4979$

Ex 4.10 The weekly demand for a drinking-water product, in thousands of litres, from a local chain of efficiency stores is a continuous random variable X having the probability density

$$f(x) = \begin{cases} 2(x-1), & 1 < x < 2 \\ 0, & \text{elsewhere} \end{cases}$$

Find mean and variance of X.

Soln Mean of X is μ

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

$$= \int_{-\infty}^{1} x \cancel{f(x)}^{0} dx + \int_{1}^{2} x f(x)\,dx + \int_{2}^{\infty} x \cancel{f(x)}^{0} dx$$

$$= \int_{1}^{2} 2x(x-1)\,dx$$

$$= \left[\frac{2x^3}{3} - x^2\right]_1^2$$

$$= \frac{16}{3} - 4 - \frac{2}{3} + 1$$

$$= \frac{5}{3}$$

variance $\sigma^2 = E(X^2) - \mu^2$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_1^2 2x^2(x-1)\,dx$$

$$= 2\int_1^2 (x^3 - x^2)\,dx$$

$$= 2\left[\frac{x^4}{4} - \frac{x^3}{3}\right]_1^2$$

$$= 2\left[4 - \frac{8}{3} - \frac{1}{4} + \frac{1}{3}\right]$$

$$= \frac{17}{6}$$

$$\sigma^2 = \frac{17}{6} - \left(\frac{5}{3}\right)^2$$

$$= \frac{1}{18}$$

Theorem : Let X be a random variable with probability function $f(x)$. The variance of the random variable $g(X)$ is

$$\sigma^2_{g(X)} = E\left((g(X) - \mu_{g(X)})^2\right)$$

$$\sigma^2_{g(X)} = \sum_x \left(g(x) - \mu_{g(X)}\right)^2 f(x)$$

If x is discrete

$$\sigma^2_{g(X)} = \int_{-\infty}^{\infty} \left(g(x) - \mu_{g(X)}\right)^2 f(x)\,dx$$

If x is continuous

Covariance - Let X and Y are two random variables with probability distribution $f(x,y)$. The covariance of X and Y

$$\sigma_{XY} = E\left((X - \mu_X)(Y - \mu_Y)\right)$$

$$\sigma_{XY} = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x,y)$$

If X, Y are discrete random variables.

$$\sigma_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x,y)\,dx\,dy$$

If X, Y are continuous random variables

Note : $\sigma_{XY} = E(XY) - \mu_X \mu_Y$

correlation coefficient - Let X and Y are two random variables with probability distribution function $f(x,y)$. The correlation coefficient of X and Y is ρ_{XY}

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \qquad -1 \le \rho_{XY} \le 1$$

Note ① If $\rho_{xy} = 0$; X, Y are ~~id~~ independent

② If ~~ρ_{xy} > +ve~~ $\rho_{xy} > 0$; X↑ Y↑, X↓ Y↓

③ If $\rho_{xy} < 0$; X↑ Y↓ , X↓ Y↑

Ex 4.11 Calculate the variance of $g(X) = 2X+3$, where X is a random variable with probability distribution

x	0	1	2	3
$f(x)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$

Sol^n $\sigma^2_{g(X)} = E((g(X) - \mu_{g(x)})^2)$

$= \sum_x (g(x) - \mu_{g(x)})^2 f(x)$

$\mu_{g(x)} = \sum_x g(x) f(x) = \sum_x (2x+3) f(x)$

$= \sum_x (2x+3) f(x)$

$= 3 \times \frac{1}{4} + 5 \times \frac{1}{8} + 7 \times \frac{1}{2} + 9 \times \frac{1}{8}$

$= 6$

$\sigma^2_{g(x)} = \sum_x (2x+3-6)^2 f(x)$

$= \sum_x (2x-3)^2 f(x)$

$= 9 \times \frac{1}{4} + 1 \times \frac{1}{8} + 1 \times \frac{1}{2} + 9 \times \frac{1}{8}$

$= 4$

Ex 4.15 Find the correlation coefficient between X and Y in example 4.13

$f(x,y)$	$x=0$	$x=1$	$x=2$	$h(y)$
$y=0$	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
$y=1$	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
$y=2$	$\frac{1}{28}$	0	0	$\frac{1}{28}$
$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	

Sol^n $\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$

x	0	1	2
$g(x)$	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

y	0	1	2
$g(y)$	15/28	6/14	1/28

$\sigma_{xy} = E(XY) - \mu_X \mu_Y$

$\mu_X = \sum_x x\, g(x) = \frac{15}{28} + 2 \times \frac{3}{28} = \frac{21}{28} = \frac{3}{4}$

$\mu_Y = \sum_y y\, h(y) = \frac{6}{14} + 2 \times \frac{1}{28} = \frac{14}{28} = \frac{1}{2}$

$E(XY) = \sum_x \sum_y xy\, f(x,y)$

$= 1 \times \frac{3}{14} = \frac{3}{14}$

$\therefore \sigma_{xy} = \frac{3}{14} - \left(\frac{3}{4}\right)\left(\frac{1}{2}\right) = \frac{3}{14} - \frac{3}{8} = \frac{9}{56}$

$\sigma_x^2 = E(x^2) - \mu_x^2$

$E(x^2) = \sum_x x^2 g(x)$

$= \frac{15}{28} + 4 \times \frac{3}{28} = \frac{27}{28}$

$\sigma_x^2 = \frac{27}{28} - \frac{9}{16} = \frac{45}{112}$

$\sigma_x = \sqrt{\frac{45}{112}}$

$\sigma_y^2 = E(y^2) - \mu_y^2$

$E(y^2) = \sum_y y^2 h(y)$

$= \frac{6}{14} + \frac{4}{28} = \frac{16}{28} = \frac{4}{7}$

$\sigma_y^2 = \frac{4}{7} - \frac{1}{4} = \frac{9}{28}$

$\sigma_y = \cancel{\frac{9}{28}} \frac{3}{\sqrt{28}}$

$\rho_{xy} = \frac{-\frac{9}{56}}{\sqrt{\frac{45}{112}} \times \sqrt{\frac{9}{28}}} = -\frac{1}{\sqrt{5}} = -0.44$

34, 35, 53

Mean and Variance of Linear combinations of random variables

Theorem 1: If a and b are two constants then

$E(ax+b) = aE(x) + b = \mu_{ax+b}$

Q - Determine the mean of the random variable $f(x) = 2x-1$, where x is a random variable with probability function $f(x)$

x	4	5	6	7	8	9
f(x)	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Sol^n From Theorem:

$\mu_{2x-1} = E(2x-1) = 2E(x) - 1$

$E(x) = \sum_x x f(x)$

$= 4 \times \frac{1}{12} + \frac{5}{12} + 6 \times \frac{1}{4} + \frac{7}{4} + 8 \times \frac{1}{6} + 9 \times \frac{1}{6}$

$= \frac{1}{3} + \frac{5}{12} + \frac{3}{2} + \frac{7}{4} + \frac{4}{3} + \frac{3}{2}$

$= \frac{5}{3} + \frac{5}{12} + \frac{7}{4} + 3$

$= \frac{20+5+21}{12} + 3$

$= \frac{46}{12} + 3 = \frac{23}{6} + 3$

$= \frac{23+18}{6} = \frac{41}{6}$

$\mu_{2x-1} = 2 \times \frac{43}{63} - 1 = 12.67$

Theorem 2: If X is a random variable and $g(x)$ and $h(x)$ is are two functions, then

$$E(g(x) \pm h(x)) = E(g(x)) \pm E(h(x))$$

Q - Find mean of the random variable $g(x) = 4x+3$ if X is a random variable with probability density function.

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

Sol^n $E(g(x)) = E(4x+3)$

$= E(4x) + E(3)$ (From thm. 2)

$= 4E(x) + 3$ ($\because E(\text{constant}) = \text{constant}$)

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

$$= \int_{-1}^{2} x \cdot \frac{x^2}{3}\, dx$$

$$= \left. \frac{x^4}{12} \right]_{-1}^{2}$$

$$= \frac{16}{12} - \frac{1}{12} = \frac{15}{12} = \frac{5}{4}$$

$\therefore E(g(x)) = 4E(X) + 3$

$= 4 \times \frac{5}{4} + 3 = 8$

Theorem 3: If X, Y are two random variables with probability functions, $g(X,Y)$, $h(X,Y)$ then

$$E(g(X,Y) \pm h(X,Y)) = E(g(X,Y)) \pm E(h(X,Y))$$

Note:

① If X and Y are two independent random variables then $E(XY) = E(X)\, E(Y)$

or $\sigma_{XY} = 0$

or $\rho_{XY} = 0$

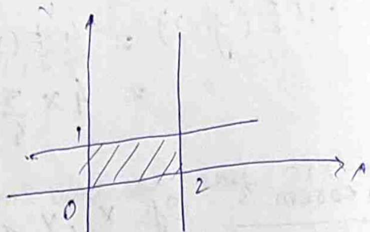
Ex 4.21 It is known that the ratio of gallium to arsenide does not affect the functioning of gallium-arsenide wafers which are the main components of microchips. Let X denote the ratio of gallium to arsenide and Y denote the ratio of function wafers retrieved during a 1 hr period. X, Y are independent random variables with the joint density function

$$f(x,y) = \begin{cases} \frac{x(1+3y^2)}{4}, & 0 < x < 2,\ 0 < y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

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

Sol^n $E(XY) = E(X)\cdot E(Y)$

$\mu_{xy} = \mu_x \mu_y$



$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\, dx\, dy$$

$$= \int_0^1\int_0^2 \frac{x^2 y(3y^2+1)}{4}\, dx\, dy$$

$$= \frac{5}{6}$$

$$E(X) = \int_{-\infty}^{\infty} x\, g(x)\cdot dx \quad \text{where } g(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

$$= \int_0^1 \frac{x(1+3y^2)}{4}\cdot dy$$

$$g(X) = \begin{cases} \frac{x}{2}, & 0<x<2 \\ 0, & \text{elsewhere} \end{cases}$$

$$= \frac{x}{4}\left[y+y^3\right]_0^1$$

$$= \frac{x(2)}{4} = \frac{x}{2}$$

$$E(X) = \int_0^2 \frac{x^2}{2}\, dx$$

$$= \left.\frac{x^3}{6}\right]_0^2$$

$$= \frac{8}{6}$$

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

$$E(Y) = \int_{-\infty}^{\infty} y\, h(y)\, .dy \quad \text{where } h(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$$

$$h(Y) = \begin{cases} \frac{3y^2+1}{2}, & 0<y<1 \\ 0, & \text{elsewhere} \end{cases}$$

$$= \int_0^2 \frac{x(1+3y^2)}{4}\, dx$$

$$= \left(\frac{3y^2+1}{4}\right)\frac{x^2}{2}\Big]_0^2$$

$$= \frac{3y^2+1}{2}$$

$$E(Y) = \int_0^1 y\cdot\left(\frac{3y^2+1}{2}\right) dy$$

$$= \int \frac{3y^3+y}{2}\, dy$$

$$= \frac{1}{2}\left[\frac{3y^4}{4} + \frac{y^2}{2}\right]_0^1$$

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

$$= \frac{5}{8}$$

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

$$= \left(\frac{4}{3}\right)\left(\frac{5}{8}\right)$$

$$= \frac{5}{6}$$

Theorem 4 If X and Y are two random variables with joint probability distribution $f(x, y)$ and a, b, c are constants. Then

$$\sigma^2_{aX+bY+c} = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\,\sigma_{XY}$$

If $b = 0$.

$$\sigma^2_{aX+b} = a^2\sigma_X^2 = a^2\sigma^2$$

If $a = 1, b = 0$ $\quad \sigma^2_{X+c} = \sigma_X^2 = \sigma^2$

If $b = 0, c = 0$, $\quad \sigma^2_{aX} = a^2\sigma_X^2 = a^2\sigma^2$

Note: (1) If X and Y are independent random variables then $\sigma^2_{aX+bY+C} = a^2\sigma_X^2 + b^2\sigma_Y^2$

(2) $\sigma^2_{aX-bY} = a^2\sigma_X^2 + b^2\sigma_Y^2$

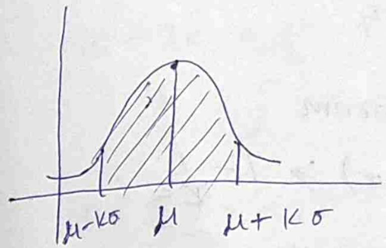
(3) If $X_1, X_2, X_3, \ldots, X_n$ are independent then

$$\sigma_{a_1X_1 + a_2X_2 + \ldots + a_nX_n}$$
$$= a_1^2\sigma_{X_1}^2 + a_2^2\sigma_{X_2}^2 + \ldots + a_n^2\sigma_{X_n}^2$$

Theorem 5: (Chebyshev's theorem)

The probability that any random variable X will assume a value K standard deviation of the mean is atleast $1 - \frac{1}{K^2}$

$$P(\mu - K\sigma < X < \mu + K\sigma) \geq 1 - \frac{1}{K^2}$$



Ex 4.22 If X and Y are random variables with variances $\sigma_X^2 = 2$ and $\sigma_Y^2 = 4$ and covariance $\sigma_{XY} = -2$, find variance of random variable $Z = 3X - 4Y + 8$

Sol^n $\sigma_X^2 = 2, \quad \sigma_Y^2 = 4, \quad \sigma_{XY} = -2$

$Z = 3X - 4Y + 8$

$$\sigma_Z^2 = \sigma^2_{3X-4Y+8}$$
$$= (3)^2 \times \sigma_X^2 + (-4)^2\sigma_Y^2 + 2(3)(-4)\,\sigma_{XY}$$
$$= 9 \times 2 + 16 \times 4 - 24 \times (-2) \quad \text{(Thm 4)}$$
$$= 18 + 64 + 48$$
$$= 130$$

4.77 A random variable X has a mean $\mu = 10$ and a variance $\sigma^2 = 4$. using Chebyshev's theorem

(a) $P(|X-10| \geq 3)$;

(b) $P(|X-10| < 3)$;

(c) $P(5 < X < 15)$;

(d) the value of constant c such that

$$P(|X-10| \geq c) \leq 0.04$$

Soln From Chebyshev's theorem

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

Here $\mu = 10$, $\sigma^2 = 4 \Rightarrow \sigma = 2$

(a) $P(|X-10| \geq 3)$

$= 1 - P(|X-10| < 3)$

$= 1 - P(7 < X < 13)$

$|x-10| < 3$

$-3 < x-10 < 3$

$7 < x < 13$

$10 - 2k = 7$, $10 + 2k = 13$

$k = \frac{3}{2}$, $k = \frac{3}{2}$

$\therefore P(7 < X < 13) \geq 1 - \frac{1}{k^2}$

$= 1 - \frac{1}{\frac{9}{4}}$

$= \frac{5}{9}$

$\therefore P(|X-10| \geq 3) \leq 1 - \frac{5}{9}$

$= \frac{4}{9}$

b) $P(|X-10| < 3) \geq \frac{5}{9}$

c) $P(5 < X < 15)$

$10 - 2k = 5$, $10 + 2k = 15$

$k = \frac{5}{2}$, $k = \frac{5}{2}$

$P(5 < X < 15) \geq 1 - \frac{1}{(5/2)^2}$

$= 1 - \frac{3}{25}$

$= \frac{22}{25}$

d) $P(|X-10| \geq c) \leq 0.04$

$1 - P(|X-10| < c) \leq 0.04$

$\Rightarrow P(|X-10| < c) \geq 0.96$

$\Rightarrow P(10 - c < X < 10 + c) \geq 0.96$

$1 - \frac{1}{k^2} = 0.96$

$1 - 0.96 = \frac{1}{k^2}$

$0.04 = \frac{1}{k^2}$

$k = \frac{1}{\sqrt{0.04}} = 5$

$\mu - 5\sigma = 10 - c$

$\Rightarrow 10 - 10 = 10 - c$

$\Rightarrow c = 10$

Ch-1 - 3, 4, 7, 10, 11, 12, 14, 21, 29, 30, 35

Ch-2 - 38, 42, 44, 49, 50, 56, 60, 62, 66, 76

Ch-3 - 4, 7, 10, 12, 15, 20, 23, 26

Ch-4
- 34, 35, 53
- 57, 58, 60, 64, 67, 74, 75, 77

Bernoulli's experiment (CHAPTER - 5)

① If the expt. contains n no. of trials

② all trials of are independent

③ Outcomes are of two types - success and failure

④ The probability of success p is same or equal in all the trials.

(Binomial Distribution)

'x' no. of success in n ~~no.~~ no. of trials having probability

$$f(x) = {}^nC_x\, p^x q^{n-x}, \quad x = 0, 1, 2, 3, \ldots n$$

Q- A coin is tossed twenty times then what is the ~~what~~ probability of getting 15 heads in tossing

Sol^n $n = 20$, $p = \frac{1}{2}$, $q = 1 - p = \frac{1}{2}$

$x = 15$

$n - x = 5$

$$f(x) = {}^{20}C_{15}\left(\frac{1}{2}\right)^{15}\left(\frac{1}{2}\right)^{5}$$

$$f(x) = {}^{20}C_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{20-x}$$

$$= {}^{20}C_x\left(\frac{1}{2}\right)^{20}, \quad x = 0, 1, 2, 3, \ldots\ldots, 20$$

what is the probability of getting atmost 3 heads in tossing 20 coins

$$P(X \le 3) = \sum_{x \le 3} f(x) = f(0) + f(1) + f(2) + f(3)$$

$$= \left(\frac{1}{2}\right)^{20} + 20\left(\frac{1}{2}\right)^{20} + 190\left(\frac{1}{2}\right)^{20} + 1140\left(\frac{1}{2}\right)^{20}$$

Q- what is the probability of five 6's in rolling 20 dice.

Sol^n $n = 20$

$p = \frac{1}{6}$

'x' no. of six in 20 no. of trials

$$f(x) = {}^{20}C_x\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{20-x}, \quad x = 0, 1, 2, 3, \ldots, 20$$

$$f(x=5) = {}^{20}C_5\left(\frac{1}{6}\right)^5\left(\frac{5}{6}\right)^{15}$$

Q- The ~~patient~~ probability that a patient recovers from a blood disease is 0.4. If 15 people are known to have ~~co~~ contracted this disease then what is the probability that atleast 10 survive

Sol^n

$n = 15$

$q = 0.4$, $p = 0.6$

'x' no. of success in 15 no. of trials

$f(x) = {}^{15}C_x (0.4)^x (0.6)^{15-x}$

$f(10) = {}^{15}C_{10} (0.4)^{10} (0.6)^5$

Binomial Distribution

A ~~binomial's~~ Bernouli's trials can result in a success with probability p and a failure with probability ~~p and a~~ $q = 1-p$, Then the probability distribution of the binomial random variable X, the number of success in n independent trials is

$$b(x; n, p) = {}^nC_x\, p^x q^{n-x} \quad , x = 0, 1, 2, \ldots, n$$

Note: The mean and variance of the binomial distribution are $\mu = np$ and $\sigma^2 = npq$

Ex 5.1 The probability that a certain kind of component will survive a shock test is 3/4. Find the probability that exactly 2 of the next 4 components tested survive.

Sol^n X : No. of components survive

$n = 4$, $p = 3/4$, $q = \frac{1}{4}$

$P(X=2) = b(2; n, p)$

$= b(2; 4, \frac{3}{4})$

$= {}^4C_2 \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^2$

$= \frac{4!}{2!\,2!} \cdot \frac{9}{16} \cdot \frac{1}{16}$

$= \frac{24}{4} \times \frac{9}{16} \times \frac{1}{16}$

$= \frac{27}{128}$

Ex 5.2 The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that (a) atleast 10 survive (b) from 3 to 8 survive, (c) exactly 5 survive?

Sol^n Let X be a random variable.

X : The no. of people who survive

$n = 15$, $p = 0.4$, $q = 0.6$

7.41 × 10

(a) $P(X \geq 10)$

$$= \sum_{x=10}^{15} b(x; n, p)$$

$$= b(10; 15, 0.4) + b(11; 15, 0.4) + b(12; 15, 0.4) + b(13; 15, 0.4) + b(14; 15, 0.4) + b(15; 15, 0.4)$$

$$= {}^{15}C_{10}(0.4)^{10}(0.6)^5 + {}^{15}C_{11}(0.4)^{11}(0.6)^4 + {}^{15}C_{12}(0.4)^{12}(0.6)^3 + {}^{15}C_{13}(0.4)^{13}(0.6)^2 + {}^{15}C_{13}(0.4)^{14}(0.6) + {}^{15}C_{15}(0.4)^{15}$$

$$= \frac{15!}{5!\,10!}\, 0.0244 + 0.0074 + \ldots$$

$$= 0.0338$$

(b) $P(3 \leq X \leq 8)$

$$= \sum_{x=0}^{8} b(x; n, P) - \sum_{x=0}^{2} b(x; n, P)$$

$$= 0.9050 - 0.0271$$

$$= 0.8779$$

(c) $P(X = 5)$

$$= \sum_{x=0}^{5} b(x; n, P) - \sum_{x=0}^{4} b(x; n, p)$$

$$= 0.4032 - 0.2173$$

$$= 0.1859$$

Multinomial distribution

If a given trial can result in the K outcomes $E_1, E_2, \ldots, E_K$ with probabilities $P_1, P_2, \ldots, P_K$ then the probability distribution of the random variables $X_1, X_2, \ldots X_K$ denotes the number of occurance of $E_1, E_2, \ldots E_K$ in n independent trials is

$$f(x_1, x_2, \ldots, x_K, P_1, P_2, \ldots P_K, n) = \binom{n}{x_1, x_2, \ldots x_n} P_1^{x_1} P_2^{x_2} \cdots P_K^{x_K}$$

where $\sum_{i=1}^{K} x_i = n$, $\sum_{i=1}^{K} P_i = 1$ and

$$\binom{n}{x_1, x_2, \ldots x_K} = \frac{n!}{x_1!\,x_2!\ldots x_K!}$$

Ex 5.4 The complexity of arrivals and departures of planes at an airport is such that computer simulation is often used to model the "ideal" conditions. For a certain airport with three runways, it is known that in the ideal setting the following are the probabilities that the individual runways are accessed by a randomly arriving commercial jet:

Runway 1 : $P_1 = 2/9$

Runway 2 : $P_2 = 1/6$

Runway 3 : $P_3 = 11/18$

What is the probability that 6 randomly arriving airplanes are distributed in the following ~~function~~ fashion?

Runway 1 : 2 airplanes
Runway 2 : 1 airplanes
Runway 3 : 3 airplanes

RW1 RW2 RW3

$p_1 = \frac{2}{9}$ $p_2 = \frac{1}{6}$ $p_3 = \frac{11}{18}$

Soln $n = 6$

$x_1 = 2, \ x_2 = 1, \ x_3 = 3$

$P(X_1 = 2, X_2 = 1, X_3 = 3)$

$= {}^{n}C_{x_1, x_2, x_3} \, p_1^{x_1} p_2^{x_2} p_3^{x_k}$

$= {}^{6}C_{2,1,3} \left(\frac{2}{9}\right)^2 \left(\frac{1}{6}\right)^1 \left(\frac{11}{18}\right)^3$

$= \frac{6!}{2!\,3!} \left(\frac{4}{81}\right)\left(\frac{1}{6}\right)\left(\frac{11^3}{18^3}\right)$

$= 0.1127$

5.21 According to a genetic theory, a certain cross of guinea pigs will results red, black, and white offspring in the ratio 8:4:4. Find the probability that among 8 offspring, 5 will be red, 2 black and 1 white.

Soln R B W

X_1 X_2 X_3

8 : 4 : 4

$p_1 = \frac{1}{2}$, $p_2 = \frac{1}{4}$, $p_3 = \frac{1}{4}$

$n = 8, \ x_1 = 5, \ x_2 = 2, \ x_3 = 1$

$P(X_1 = 5, X_2 = 2, X_3 = 1)$

$= {}^{n}C_{x_1, x_2, x_3} \, p_1^{x_1} p_2^{x_2} p_3^{x_3}$

$= {}^{8}C_{5,2,1} \left(\frac{1}{2}\right)^5 \left(\frac{1}{4}\right)^2 \left(\frac{1}{4}\right)^1$

$= \frac{8!}{5!\,2!} \left(\frac{1}{32}\right)\left(\frac{1}{16}\right)\left(\frac{1}{4}\right)$

$=$

5.15 It is known that 60% of mice inoculated with a serum are protected from a certain disease. If 5 mice are inoculated, find the probability that

(a) none contracts the disease;

(b) fewer than 2 contract the disease;

(c) more than 3 contract the disease

Soln 60% protected from diseases

$n = 5, \ p = 40\% = 0.40, \quad q = 60\% = 0.60$

X : Number of mice contracted the disease

a) $P(X = 0) = {}^{n}C_x \, p^x q^{n-x}$

$= {}^{5}C_0 (0.40)^0 (0.60)^{5-0}$

$= 0.07776$

b) $P(X<2) = P(X=0) + P(X=1)$

$= 0.07776 + {}^5C_1 (0.40)(0.60)^4$

$= 0.\cancel{2592} + 0.07776 + 0.2592$

$= 0.33696$

c) $P(X>3) = 1 - P(X \le 3) \;/\; P(X=4) + P(X=5)$

$= {}^5C_4 (0.40)^4 (0.60) + {}^5C_5 (0.40)^5 (0.60)^0$

$= 0.08\cancel{2}9$

9, 11, 15, 16, 17, 22

Hypergeometric distribution

The probability distribution of the hypergeometric distribution

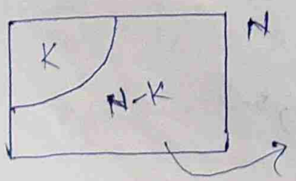
The probability distribution in a random sample size of n

X the number of success in a random sample size of n selected from N items of which K are success and N-K are failure is

$$P(X=x) = h(x; N, n, K) = \frac{{}^KC_x \; {}^{N-K}C_{n-x}}{{}^NC_n}$$

~~what~~ Mean $\mu = n\frac{K}{N}$

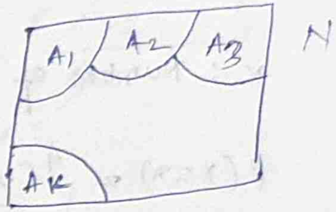
variance $\sigma^2 = \left(\frac{N-n}{N-1}\right)(n)\left(\frac{K}{N}\right)\left(1-\frac{K}{N}\right)$



n selected

Note: The hypergeometric distribution is a process of without replacement

Multivariate hypergeometric distribution



If N items can be partitioned into $A_1, A_2, \dots A_k$ with $a_1, a_2, \dots, a_k$ elements then the probability distribution of the random variable $X_1, X_2, \dots X_k$ representing the number of elements selected from $A_1, A_2, \dots A_k$ in a random sample size n is

$$P(X_1 = x_1, X_2 = x_2, \dots X_k = x_k)$$

$$= f(x_1, x_2, \dots x_k; a_1, a_2, \dots, a_k; N, n)$$

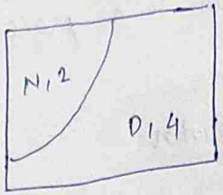
$$= \frac{{}^{a_1}C_{x_1} \; {}^{a_2}C_{x_2} \cdots {}^{a_k}C_{x_k}}{{}^NC_n}$$

where $\sum_{i=1}^{k} a_i = N$

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

5.31 A random committee of size 3 is selected from 4 doctors and 2 nurses. write a formula for the probability distribution of the random variable X representing the number of doctors on the committee. Find $P(2 \le X \le 3)$

Sol^n



$N = 6$
$n = 3$

X : Number of doctors in the committee $\quad x = 1, 2, 3$

$$P(X=x) = \frac{{}^4C_x \times {}^2C_{3-x}}{{}^6C_3}$$

$$P(2 \le X \le 3) = P(X=2) + P(X=3)$$

$$= \frac{{}^4C_2 \times {}^2C_1}{{}^6C_3} + \frac{{}^4C_3 \times {}^2C_0}{{}^6C_3}$$

$$= \frac{24+4}{20} = \frac{4}{5}$$

5.43 A foreign student club lists as its members 2 ~~Canadian~~ Canadians, 3 Japanese, 5 Italians and 2 Germans. If a committee of 4 is selected at random, find the probability that

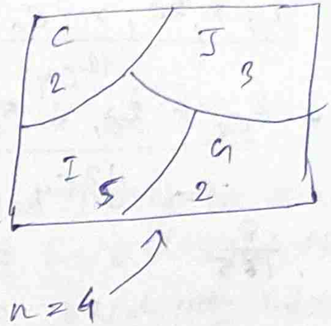
(a) all nationalities are represented

(b) all nationalities except Italian are represented

Sol^n

C	J	I	G
2	3	5	2

$N = 12$
$n = 4$



$A_1 = C \quad a_1 = 2 \quad N = 12 \quad x_1 = 1$

$A_2 = J \quad a_2 = 3 \quad n = 4 \quad x_2 = 1$

$A_3 = I \quad a_3 = 5 \quad x_3 = 1$

$A_4 = G \quad a_4 = 2 \quad x_4 = 1$

(a) $P(X_1 = 1, X_2 = 1, X_3 = 1, X_4 = 1)$

$$= \frac{{}^2C_1 \times {}^3C_1 \times {}^5C_1 \times {}^2C_1}{{}^{12}C_4}$$

$$= \frac{120}{495}$$

(b)

C	J	I	G
1	1	0	2
1	2	0	1
2	1	0	1

$P(X_1 = 1, X_2 = 1, X_3 = 0, X_4 = 2)$

$P(X_1 = 1, X_2 = 2, X_3 = 0, X_4 = 1)$

$P(X_1 = 2, X_2 = 1, X_3 = 0, X_4 = 1)$

c) $\dfrac{{}^2C_1 \times {}^3C_1 \times {}^5C_0 \times {}^2C_2}{{}^{12}C_4} + \dfrac{{}^2C_1 \times {}^3C_2 \times {}^5C_0 + {}^2C_1}{{}^{12}C_4}$

$+ \dfrac{{}^2C_2 \times {}^3C_1 \times {}^5C_0 \times {}^2C_1}{{}^{12}C_4}$

$= \dfrac{8}{165}$

31, 32, 43, 44, 47

Negative Binomial and geometric distribution

Negative Binomial distribution

If ~~accepted~~ repeated ~~exponential~~ independent trials can result in a success probability p and failure probability is $q = 1-p$ then the probability distribution of the random variables x, which is the number of trial on which kth success is

$$b^*(x; k, p) = {}^{x-1}C_{k-1}\, p^k q^{x-k}$$

where $x = k, k+1, k+2, \ldots$

In particular $k = 1$, the negative binomial is called geometric distribution and its probability function is

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

Note: The mean and variance of geometric distribution is $\mu = \dfrac{1}{p}$

$\sigma^2 = \dfrac{1-p}{p^2}$

Binomial

$$b(x; n, p) = {}^nC_x\, p^x q^{n-x}, \quad x = 1, 2, 3, \ldots$$

5.49 The probability that a person living in a certain city owns a dog is estimated to be 0.3. Find the probability that the tenth person randomly interviewed in that city is the fifth one to own a dog.

Sol^n: Let X be a random variable.

X: a person interviewed and owns a dog

$p = 0.3 \Rightarrow q = 1 - 0.3 = 0.7$

$b^*(x; k, p) = {}^{x-1}C_{k-1}\, p^k q^{x-k}$

$= {}^9C_4 (0.3)^5 (0.7)^5$

$= \dfrac{9!}{4!\,5!} (0.3)^5 (0.7)^5$

$= \dfrac{9 \times 8 \times 7 \times 6}{4 \times 3 \times 2} (0.3)^5 (0.7)^5$

$= 0.0514$

5.50 Find the probability that a person flipping a coin

(a) the third head on the seventh flip;

(b) the first head on the fourth flip;

Sol^n: X: random variable of getting head

(a) $p = 0.5$, $q = 0.5$

$x = 7$, $k = 3$

$b^*(x;k,p) = {}^6C_2 (0.5)^3 (0.5)$

$= \frac{6!}{2!\,4!} (0.5)^7$

$= \frac{3 \times 5}{\ldots} (0.5)^7$

$= 0.117$

(b) $\cancel{b^*(x;k,p)}$ $k = 1,\ x = 4$

$b^*(x;k,p) = {}^3C_0 (0.5)(0.5)^3$

$= (0.5)^4$

$= 0.0625$

$g(4, 0.5) = (0.5)(0.5)^3$

$= (0.5)^4$

$= 0.0625$

5.51 Three people toss a fair coin and the odd one pays for coffee. If the coins all turn up the same, they are tossed again. Find the probability that fewer than 4 tosses are needed.

Sol^n: Let X be a random variable

X: no. of tosses one pays for coffee.

$p = \frac{6}{8} = \frac{3}{4}$, $q = 1 - \frac{3}{4} = \frac{1}{4}$

$k = 1$

$\sum_{x=1}^{3} b^*(x;k,p) = \sum_{x=1}^{3} g(x;p)$

$= \cancel{b^*(1;1,\frac{3}{4}) + b^*(2;1,\frac{3}{4}) +}$

$= g(1;0.75) + g(2;0.75) + g(3;0.75)$

$= (\frac{3}{4})(\frac{1}{4})^0 + (\frac{3}{4})(\frac{1}{4}) + (\frac{3}{4})(\frac{1}{4})^2$

$= \frac{3}{4} + \frac{3}{16} + \frac{3}{64}$

$= \frac{48 + 12 + 3}{64}$

$= \frac{63}{64} = 0.9843$

5.70 A company purchases large lots of a certain kind of electronic device. A method is used that rejects a lot if 2 or more defective units are found in a random sample of 100 units

(a) What is the mean number of defective units found in a sample of 100 units if the lot is 1% defective?

(b) What is the variance?

Sol^n: Mean $\mu = np$

variance $\sigma^2 = npq$

$p = 1\% = 0.01$ $\Rightarrow q = 1 - 0.01 = 0.99$

$n = 100$

Mean $\mu = (100 \times 0.01)$

$\mu = 1$

variance $= 100 \times 0.01 \times 0.99$

$= 0.99$

Poisson's Distribution

The probability distribution of the poisson's random variable X representing the number of outcomes occuring in a given interval of time or specified region denoted by t is

$$P(x;\mu) = \frac{e^{-\mu}\mu^x}{x!}, \quad x = 0,1,2,\ldots \qquad \mu = \lambda t$$

where λ is the avg. number of outcomes per unit time, distance, area, volume etc.

The mean and variance of the poisson's distribution are λt and np

Note:

Let X be a binomial random variable with probability distribution $b(x;n,p)$

$$b(x;n,p) \xrightarrow{n\to\infty} p(x;\mu)$$

5.58 A certain area of the eastern united states is on average, hit by 6 hurricanes a year. Find the probability that in a given year that area will be hit by

(a) fewer than 4 hurricanes

(b) anywhere from 6 to 8 hurricanes

Soln Let X be a random variable

X : no. of hurricanes hitting in a particular area

$\lambda = 6, \ t = 1$

(a) $P(X<4) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$

$\therefore P(x,\mu) = \frac{e^{-\mu}\mu^x}{x!}, \ x = 0,1,2,3$

$\mu = \lambda t = 6\times 1 = 6$

$\Rightarrow P(0,6) = \frac{e^{-6}6^0}{0!}, \ P(1,6) = \frac{e^{-6}6^1}{1!}$

$P(2,6) = \frac{e^{-6}6^2}{2!} \quad P(3,6) = \frac{e^{-6}6^3}{3!}$

$\Rightarrow \frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!} + \frac{e^{-6}6^2}{2!} + \frac{e^{-6}6^3}{3!}$

$\Rightarrow e^{-6}(1+6+18+36)$

$\Rightarrow e^{-6}(61)$

$= 0.1512$

(b) $P(6\le X\le 8) = P(X\le 8) - P(X\le 5)$

$= 0.8472 - 0.4457$

$= 0.4015$

$\Rightarrow \frac{e^{-6}6^6}{6!} + \frac{e^{-6}6^7}{7!} + \frac{e^{-6}6^8}{8!}$

$\Rightarrow e^{-6}\left(\frac{6^6}{6!} + \frac{6^7}{7!} + \frac{6^8}{8!}\right)$

$= e^{-6} 6^6 \left(\frac{1}{6!} + \frac{6}{7!} + \frac{36}{8!} \right)$

$= 0.4014$

5.69 The probability that a person will die when he or she contracts a virus infection is 0.001. Of the next 4000 people infected, what is the mean number who will die?

Sol^n $\mu = np = 4000 \times (0.001) = 4$

Discrete Distribution

1. Binomial
2. Multinomial
3. Hypergeometric
4. Multivariance hypergeometric
5. Negative binomial
6. Poisson

$b(x; n, p) = {}^nC_x p^x q^{n-x}, \quad x = 0, 1, 2, \ldots n, \quad \mu = np, \quad \sigma^2 = npq$

$f(x_1, x_2, \ldots, p_1, p_2, \ldots, p_k) = {}^nC_{x_1, x_2, \ldots x_n} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$

$h(x; N, n, K) = \frac{{}^KC_x \; {}^{N-K}C_{n-x}}{{}^NC_n}, \quad \mu = \frac{nK}{N}$

$\sigma^2 = \left(\frac{N-n}{N-1}\right) n \frac{K}{N} \left(1 - \frac{K}{N}\right)$

$f(x_1, x_2, \ldots x_k; a_1, a_2 \ldots a_k, N, n) = \frac{{}^{a_1}C_{x_1} \; {}^{a_2}C_{x_2} \ldots {}^{a_k}C_{x_k}}{{}^NC_n}$

$\sum_{i=1}^{k} x_i = n, \quad \sum_{i=1}^{k} a_i = N$

$b^*(x; k, p) = {}^{x-1}C_{k-1} p^k q^{x-k}, \quad x = k, k+1, k+2, \ldots$

$k = 1$

$g(x, p) = pq^{x-1}, \quad x = 1, 2, 3, \ldots$

$\mu = \frac{1}{p}, \quad \sigma^2 = \frac{1-p}{p^2}$

$p(x; \mu) = \frac{e^{-\mu} \mu^x}{x!}, \quad x = 0, 1, 2, \ldots$

$\mu = \lambda t, \quad n \to \infty$

Some continuous probability distribution
Chapter - 6

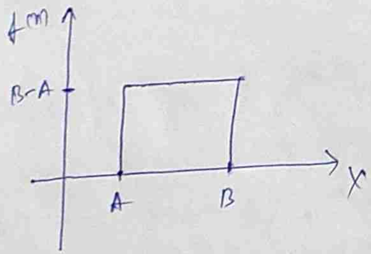
① uniform Distribution
② Normal Distribution
③ Exponential Distribution
④ Gamma Distribution

① uniform Distribution

Let X be a cts. random variable. The density function of the random variable X on the interval $[A, B]$ is

$$f(x) = \begin{cases} \frac{1}{B-A}, & A \le x \le B \\ 0, & \text{otherwise} \end{cases}$$

is called ~~uniform~~ continuous uniform distribution



The mean and variance of the uniform distribution is

$$\mu = \frac{A+B}{2}$$

$$\sigma^2 = \frac{(B-A)^2}{12}$$

Q- Find the mean and variance of uniform distribution on the interval $[A, B]$

Sol^n X, $f(x)$, $\mu = \int_{-\infty}^{\infty} x f(x)\, dx$

$$f(x) = \begin{cases} \frac{1}{B-A}, & A \le x \le B \\ 0, & \text{otherwise} \end{cases}$$

$$\mu = \int_{-\infty}^{A} x \cancel{f(x)}^{0} dx + \int_{A}^{B} x f(x)\, dx + \int_{B}^{\infty} x \cancel{f(x)}^{0} dx$$

$$= \int_{A}^{B} x \cdot \frac{1}{B-A} \cdot dx$$

$$= \left[\frac{x^2}{2}\right]_A^B \cdot \frac{1}{B-A}$$

$$= \left(\frac{B^2 - A^2}{2}\right) \times \frac{1}{B-A}$$

$$= \frac{A+B}{2}$$

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

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx$$

$$= \int_{A}^{B} x^2 \frac{1}{B-A} \cdot dx$$

$$= \frac{1}{B-A}\left[\frac{x^3}{3}\right]_A^B$$

$$= \frac{B^3-A^3}{3} \times \frac{1}{B-A}$$

$$= \frac{\cancel{(B-A)}(A^2+AB+B^2)}{3} \times \frac{1}{\cancel{B-A}}$$

$$\cancel{(A+B)}^2 = \frac{A^2+AB+B^2}{3}$$

$$\sigma^2 = \frac{A^2+AB+B^2}{3} - \frac{(A+B)^2}{4}$$

$$= \frac{(B-A)^2}{12}$$

The standard deviation of this uniform distribution is

$$\sigma = \frac{B-A}{2\sqrt{3}}$$

6.2 suppose X follows a cts uniform distribution from 1 to 5. Determine the conditional probability $P(X>2.5 \mid X \le 4)$

Sol^n^ $P(X>2.5 \mid X \le 4)$

X, $f(x)$

$$f(x) = \begin{cases} \frac{1}{4}, & 1 \le x \le 5 \\ 0, & \text{otherwise} \end{cases}$$

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

$$P(X > 2.5 \mid X \le 4)$$

$$= \frac{P((X>2.5) \cap (X \le 4))}{P(X \le 4)}$$

$$= \frac{P(2.5 < X < 4)}{P(X \le 4)}$$

$$P(2.5 < X < 4) = \int_{2.5}^{4} \frac{1}{4} \cdot dx$$

$$= \frac{1}{4}[x]_{2.5}^{4}$$

$$= 1.5 \times \frac{1}{4}$$

$$= \frac{\cancel{15}\,3}{\cancel{40}\,8} = \frac{3}{8}$$

$$P(X \le 4) = \cancel{1 - P(X>4)}$$

$$= \int_{-\infty}^{4} f(x)\,dx$$

$$= \int_{-\infty}^{1} \cancel{f(x)}^{\,0}\,dx + \int_{1}^{4} f(x)\,dx$$

$$= \frac{1}{4}[x]_1^4$$

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

$\therefore P(X > 2.5 \mid X \le 4) = \frac{\cancel{3}}{8} \times \frac{4}{\cancel{3}} = \frac{1}{2}$

Q.4 A bus arrives every 10 minutes at a bus stop. It is assumed that the waiting time for a particular individual is a random variable with a cts. uniform distribution.

(a) what is the probability that the individual waits more than 7 minutes?

(b) what is the probability that the individual waits between 2 and 7 minutes?

Sol^n: X : continuous random variable that waiting time (in mins) of a passenger at a bus stop.

$$f(x) = \begin{cases} \frac{1}{10}, & 0 \le x \le 10 \\ 0, & \text{otherwise} \end{cases}$$

$$P(X > 7) = \int_7^\infty f(x)\,dx = \int_7^{10} f(x)\,dx + \cancel{\int_{10}^\infty f(x)\,dx}^{\,0}$$

$$= \int_7^{10} \frac{1}{10}\,dx$$

$$= \frac{1}{10}[x]_7^{10}$$

$$= \frac{3}{10}$$

$$P(2 < X < 7) = \int_2^7 f(x)\,dx$$

$$= \frac{1}{10}[x]_2^7$$

$$= \frac{\cancel{5}^{1}}{\cancel{10}_{2}}$$

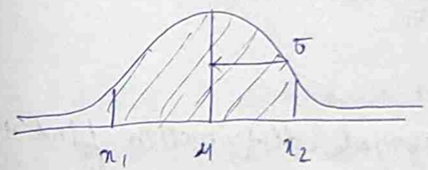
$$= \frac{1}{2}$$

Normal Distribution

A cts. random variable X having bell shape distribution is known as Normal distribution.

The density function of the normal distribution is

$$n(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty$$



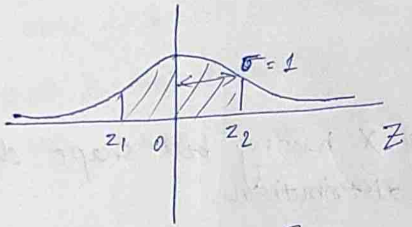
$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} n(x; \mu, \sigma)\,dx$$

Standard Normal Distribution

The standard Normal distribution is denoted by Z

$Z = \frac{X - \mu}{\sigma}$, $\mu \to$ mean
$\sigma \to$ standard deviation

$\mu = 0$, $\sigma = 1$



$$P(z_1 < Z < z_2) = \int_{z_1}^{z_2} n(z;0,1)\,dz$$

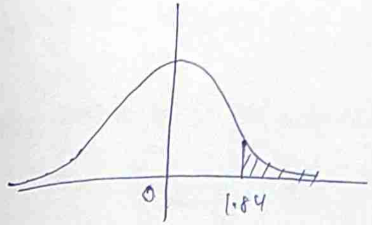
$$= \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} . dz$$

Ex 6.2 Given a standard normal distribution find the area under the curve that lies

(a) to the right of $z = 1.84$ and

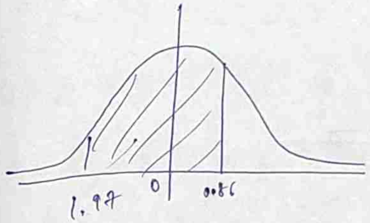
(b) between $z = -1.97$ and $z = 0.86$

Sol^n (a) $P(Z > 1.84)$



$P(Z > 1.84) = 1 - P(Z \leq 1.84)$
$= 1 - 0.9671$
$= 0.0329$

(b)



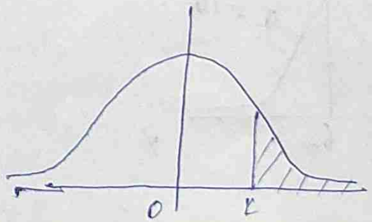
$P(-1.97 < X < 0.86) = P(Z < 0.86) - P(Z < -1.97)$
$= 0.8051 - \cancel{0.2440}\ 0.02440$
$= \cancel{0.5611}\ 0.7807$

Ex 6.3 Given a standard normal distribution, find the value of k such that

(a) $P(Z > k) = 0.3015$ and

(b) $P(k < Z < -0.18) = 0.4197$

Sol^n $P(Z > k) = 0.3015$



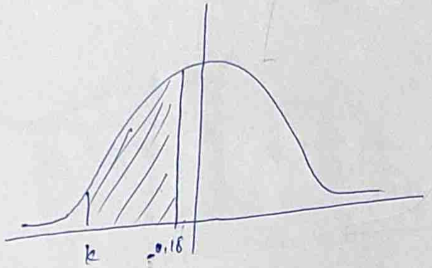
$P(Z > K) = 1 - P(Z \leq K)$

$= 1 - 0.3015$

$= 0.6985$

$\therefore K = 0.52$

(b) $P(K < Z < -0.18) = 0.4197$



$P(K<Z<-0.18) = P(Z<-0.18) - P(Z<K)$

$\Rightarrow 0.4197 = 0.4286 - P(Z<K)$

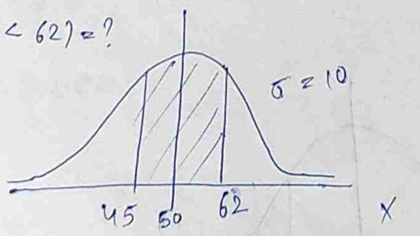
$\Rightarrow P(Z<K) = 0.4286 - 0.4197$

$= 0.0089$

$\therefore K = -2.37$

Ex 6.4 Given a random variable X having a normal distribution with $\mu = 50$ and $\sigma = 10$, Find the probability that X assumes a value between 45 and 62.

Sol^n $P(45<X<62) = ?$



we know that $Z = \frac{X-\mu}{\sigma}$

$P(45 < X < 62) = P(\quad < Z < \quad)$

$\Rightarrow P\left(\frac{45-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{62-\mu}{\sigma}\right)$

$\Rightarrow P(-0.5 < Z < 1.2)$

$= P(Z<1.2) - P(Z<-0.5)$

$= 0.8849 - 0.3085$

$= 0.5764$

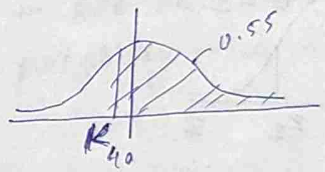
Q − Find the value of K

$P(X > K) = 0.55$

$\mu = 40, \sigma = 6$

Sol^n $P(X > K) = 0.55$

$Z = \frac{X-\mu}{\sigma}$



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

$= P\left(Z > \frac{K-40}{6}\right)$

$= 1 - P\left(Z \leq \frac{K-40}{6}\right)$

$1 - P\left(Z \leq \frac{K-40}{6}\right) = 0.55$

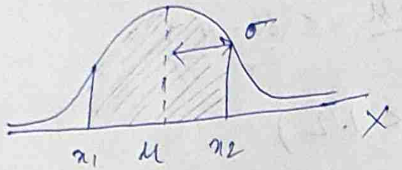
$\Rightarrow P\left(Z \leq \frac{K-40}{6}\right) = 1 - 0.55$

$= 0.45$

$\frac{K-40}{6} = -0.12 \Rightarrow K = 39.28$

Application of ~~Binomial~~ Normal distribution

$$n(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty$$

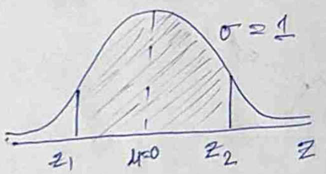


$$P(x_1 \le X \le x_2) = \int_{x_1}^{x_2} n(x,\mu,\sigma)\,dx$$

= Area under the curve from x_1 to x_2

Standard Normal distribution

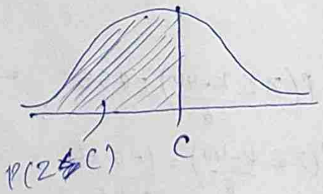
$$Z = \frac{X-\mu}{\sigma}$$



$$P(Z_1 \le Z \le Z_2) = \int_{Z_1}^{Z_2} n(z,0,1)\,dz$$

$$n(z,0,1) = \frac{1}{\sqrt{2\pi}} e^{\frac{-(x)^2}{2}}$$

$$P(z > C) = 1 - P(Z \le C)$$



Ex 6.7 A certain type of storage battery last, on average, 3.0 yrs with a standard deviation of 0.5 year. Assuming that battery life is normally distributed, find the probability that a given battery will last less than 2.3 years.

Soln. $\mu = 3, \quad \sigma = 0.5$

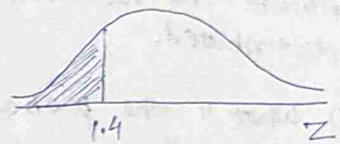
X: Life span of a battery in year

$$P(X < 2.3) = P\left(\frac{X-\mu}{\sigma} < \frac{2.3-\mu}{\sigma}\right)$$

$$= P\left(Z < \frac{2.3-3}{0.5}\right)$$

$$= P(Z < -1.4)$$

$$= 0.0808$$



Ex 6.8 An electrical firm manufactures light bulb that have a life, before burn-out, that is normally distributed with mean equal to 800 hours and a standard deviation of 40 hours. Find the probability that a bulb burns between 778 and 834 hrs.

Soln. X: Life span of a bulb before burnout

$\mu = 800, \quad \sigma = 40$

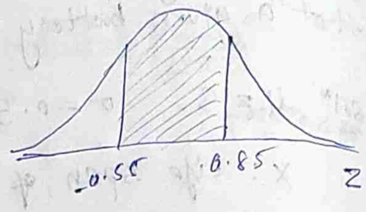
$$P(778 < X < 834)$$

$$= P\left(\frac{778-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{834-\mu}{\sigma}\right)$$

$$= P\left(\frac{778-800}{40} < Z < \frac{834-800}{40}\right)$$

$= P(-0.55 < Z < 0.85)$

$= P(Z < 0.85) - P(Z < -0.55)$



$= 0.8023 - 0.2912$

$= 0.5112$

Ex 6.15 A lawyer commutes daily from his suburban home to his midtown office. The avg. time for a one-way trip is 24 minutes, with a standard deviation of 3.8 minutes. Assume the distribution of trip times to be normally distributed.

(a) What is the probability that a trip will take atleast 1/2 hour?

(b) If the office opens at 9:00 A.M. and the lawyer leaves his house at 8:45 AM daily, what percentage of the time is he late for work?

(c) If he leaves the house at 8:35 AM and coffee is served at the office from 8:50 A.M. until 9:00 AM, what is the probability that he misses coffee?

(d) Find the length of time above which we find the slowest 15% of the trips.

(e) Find the probability that 2 of the next 3 trips will take at 1/2 hour

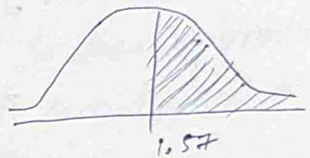
Solⁿ X : Time for oneway trip from home to office (mins)

$\mu = 24, \quad \sigma = 3.8$

(a) $P(X \geq 30)$

$= P\left(\frac{X-\mu}{\sigma} \geq \frac{30-\mu}{\sigma}\right)$

$= P\left(Z \geq \frac{30-24}{3.8}\right)$



$= P(Z \geq 1.57)$

$= \cancel{P(Z \geq} \; 1 - P(Z < 1.57)$

$= 1 - 0.9418$

$= 0.0582$

(b) Home 8:45 AM — Office 9:00 AM

$P(X > 15)$

$= P\left(\frac{X-\mu}{\sigma} > \frac{15-\mu}{\sigma}\right)$

$= P\left(Z > \frac{15-24}{3.8}\right)$

$= P(Z > -2.37)$

$= 1 - P(Z \leq -2.37)$

$= \cancel{1 - 0.9911} \; 1 - 0.0089$

$= \cancel{0.0089} \; 0.9911$

13, 15, 22 0, 10,

(c) Home 8:35 AM — Coffee time 8:50 - 9:00 AM — Office 9 AM

$$P(X > 25) = P\left(\frac{X-\mu}{\sigma} > \frac{25-\mu}{\sigma}\right)$$

$$= P\left(Z > \frac{1}{3.8}\right)$$

$$= P(Z > 0.26)$$

$$= 1 - P(Z \le 0.26)$$

$$= 1 - 0.6026$$

$$= 0.3974$$

(d) $P(X \geq C) = 15\%$

$$\Rightarrow P\left(\frac{X-\mu}{\sigma} \geq \frac{C-24}{\sigma}\right) = 0.15$$

$$\Rightarrow P\left(Z \geq \frac{C-24}{3.8}\right) = 0.15$$

$$\Rightarrow \frac{C-24}{3.8} = 1.04$$

$$\Rightarrow C = 1.04 \times 3.8 + 24$$

$$\Rightarrow C = 27.952$$

(e) $p = 0.0582$

$q = 1 - 0.0582 = 0.9418$

${}^{n}C_{y}\, p^{y} q^{n-y}$, $n = 3$, $y = 2$

${}^{3}C_{2}\,(0.0582)^2(0.9418)^1$

$= 3 \times 0.0033 \times 0.9418 = 0.0095$

Normal approximation to Binomial

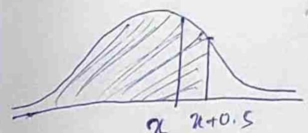
A coin is tossed 400 times. Find the probability obtaining exactly 205 heads?

Let X be a binomial random variable with mean μ and variance σ^2. For large n X has approximately a normal distribution with mean $\mu = np$ and $\sigma^2 = npq = np(1-q)$ and

$$P(X \le x) = \sum_{k=0}^{x} b(k; n, p)$$

$\approx$ Area under the curve to the left of $x + 0.5$

$$= P\left(Z \le \frac{x+0.5-\mu}{\sigma}\right)$$



$n > 30$

Note: ① $P(X \le x) = P\left(Z \le \frac{x+0.5-\mu}{\sigma}\right)$

② $P(X < x) = P\left(Z \le \frac{x-0.5-\mu}{\sigma}\right)$

③ $P(x_1 \le X \le x_2) = P\left(\frac{x_1-0.5-\mu}{\sigma} \le Z \le \frac{x_2+0.5-\mu}{\sigma}\right)$

~~④ $P\left(\frac{x_1+0.5-\mu}{\sigma} \le Z \le \frac{x_2-0.5-\mu}{\sigma}\right)$~~

④ $P(x_1 < X < x_2) = P\left(\frac{x_1+0.5-\mu}{\sigma} \le Z \le \frac{x_2-0.5-\mu}{\sigma}\right)$

⑤ $P(X > x) = 1 - P(X \le x)$

⑥ $P(X \geq x) = 1 - P(X < x)$

Ex 6.15 The probability that a patient recovers from a rare blood disease. If 100 people are known to have contracted this disease, what is the probability that fewer than 30 survive?

Soln: X : No. of persons recover from blood disease

$p = 0.4$, $q = 1 - 0.4 = 0.6$

$n = 100$, $\mu = np = 40$

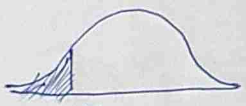
$\sigma = \sqrt{npq} = \sqrt{24}$

$p(X < 30) = P\left(Z < \frac{29.5 - \mu}{\sigma}\right)$

$= P\left(Z < \frac{29.5 - 40}{\sqrt{24}}\right)$

$= P(Z < -2.14)$

$= 0.0162$



6.24 A coin is tossed 400 times. Use the normal curve approximation to find the probability of obtaining
(a) between 185 and 210 heads inclusive;
(b) exactly 205 heads;
(c) fewer than 176 or more than 227 heads

Soln: $n = 400$, X : No. of heads

$p = \frac{1}{2}$, $q = \frac{1}{2}$, $\mu = np = 200$

$\sigma = \sqrt{npq} = 10$

(a) $P(185 \le X \le 210) = P\left(\frac{185 - 0.5 - 200}{10} \le Z \le \frac{210 + 0.5 - 200}{10}\right)$

$= P(-1.55 \le Z \le 1.05)$

$= P(Z \le 1.05) - P(Z \le -1.55)$

$= 0.8531 - 0.0606$

$= 0.7927$

(b) $P(X = 205) = P(204 < X < 206)$

$= P\left(\frac{204.5 - \mu}{\sigma} < Z < \frac{205.5 - \mu}{\sigma}\right)$

$= P\left(\frac{204.5 - 200}{10} < Z < \frac{205.5 - 200}{10}\right)$

$= P(0.45 < Z < 0.55)$

$= P(Z < 0.55) - P(Z < 0.45)$

$= 0.7088 - 0.6736$

$= 0.0352$

(c) $P(X < 176) + P(X > 227)$

$= P\left(Z \le \frac{175.5 - 200}{10}\right) + 1 - P(X \le 227)$

$= P(Z \le -2.45) + 1 - P\left(Z \le \frac{227.5 - 200}{10}\right)$

$= P(Z \le -2.45) + 1 - P(Z \le 2.75)$

$= 0.0071 + 1 - 0.9970$

$= 0.0101$

24, 26, 29, 34,

Gamma and Exponential function

Def^n (Gamma function)

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx \quad , \text{for } \alpha > 0$$

Properties

① $\Gamma(n) = (n-1)!$, n is +ve Integer

② $\Gamma(n) = (n-1)\,\Gamma(n-1)$

③ $\Gamma(1) = 1$

④ $\Gamma(1/2) = \sqrt{\pi}$

Gamma distribution

The continuous random variable X has a gamma distribution with parameter α and β if the density function is given by

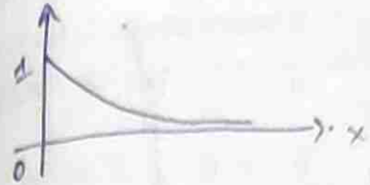
$$f(x, \alpha, \beta) = \begin{cases} \dfrac{1}{\beta^{\alpha}\,\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} & x > 0 \\ 0 & , \text{otherwise} \end{cases}$$

where $\alpha > 0$, $\beta > 0$

Note: If $\alpha = 1$ with the parameter β the gamma distribution is known as Exponential distribution

Exponential distribution

$$f(x;\beta) = \begin{cases} \dfrac{1}{\beta} e^{-x/\beta} & , \; x > 0 \\ 0 & , \text{otherwise} \end{cases} \qquad \beta > 0$$



Mean and Variance

$\mu = \alpha\beta$, $\sigma^2 = \alpha\beta^2$ (for gamma distribution)

$\mu = \beta$, $\sigma^2 = \beta^2$ (for exponential distribution)

Ex 6.19 In a biomedical study with rats, a dose-response investigation is used to determine the effect of the dose of a toxicant on their survival time. The toxicant one that is frequently discharged into the atmosphere from jet fuel. For a certain dose of the toxicant, the study determines that the survival time, in weeks, has a gamma distribution with $\alpha = 5$ and $\beta = 10$. What is the probability that a rat survives no longer than 60 weeks

Sol^n X : Survival time (in weeks)

$\alpha = 5$, $\beta = 10$

$P(X < 60)$

$$f(x) = \begin{cases} \dfrac{1}{10^5\,\Gamma(5)} x^4 e^{-\frac{x}{10}} & , \; x > 0 \\ & \text{otherwise} \end{cases}$$

$$P(X < 60) = \int_{0}^{60} f(x)\,dx$$

$$= \int_{0}^{60} \frac{1}{10^5 \Gamma(5)} x^4 e^{-\frac{x}{10}}\,dx$$

Incomplete Gamma function

$$\boxed{F(x;\alpha) = \int_0^x \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)}\,dy}$$

Let $y = \frac{x}{10} \Rightarrow x = 10y$

$dx = 10\,dy$

when $x = 0$, $y = 0$

$x = 60$, $y = 6$

$$= \int_0^6 \frac{1}{\cancel{10^5}\Gamma(5)} \times \cancel{10^4} y^4 e^{-y} \times \cancel{10}\,dy$$

$$= \int_0^6 \frac{1}{\Gamma(5)} y^4 e^{-y}\,dy$$

$= F(6,5)$

$= 0.7150$

Ex 6.20 It is known, from previous data, that the length of time in months between customer complaints about a certain product is a gamma distribution with $\alpha = 2$ and $\beta = 4$. Changes were made to tighten quality control requirements. Following these changes, 20 months passed. ~~Changes were made to tighten quality control requirements~~ before the 1st complaint. Does it appear as if the quality control tightening was effective

Sol^n $\alpha = 2$, $\beta = 4$

$$f(x) = \begin{cases} \frac{1}{4^2\Gamma(2)} x e^{-\frac{x}{4}}\,dx, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

X : time to the first complaint

$$P(X \geq 20) = 1 - P(X < 20)$$

$$= 1 - \int_0^{20} \frac{1}{4^2\Gamma(2)} x e^{-\frac{x}{4}}\,dx$$

Let $y = \frac{x}{4} \Rightarrow x = 4y$

$dx = 4\,dy$

when $x = 0$, $y = 0$

$x = 20$, $y = 5$

$$= 1 - \int_0^5 \frac{\cancel{4}y\, e^{-y}}{\cancel{4^2}\Gamma(2)} \cdot \cancel{4}\,dy$$

$$= 1 - \int_0^5 \frac{y\, e^{-y}}{\Gamma(2)}\,dy$$

$= 1 - F(5;2) = 1 - 0.9600 = 0.04$

6.41 If a random variable X has the gamma distribution with $\alpha = 2$ and $\beta = 1$, find $P(1.8 < X < 2.4)$.

Sol^n $f(x,\alpha,\beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, & x>0 \\ 0, & \text{otherwise} \end{cases}$

$= \begin{cases} \frac{1}{\Gamma(2)} x e^{-x}, & x>0 \\ 0, & \text{otherwise} \end{cases}$

$P(1.8 < X < 2.4) = \int_{1.8}^{2.4} f(x)\,dx$

$= \int_{1.8}^{2.4} x e^{-x} dx$

$= -xe^{-x} + \int e^{-x} dx$

$= -xe^{-x} - e^{-x} \Big]_{1.8}^{2.4}$

$= -2.4e^{-2.4} - e^{-2.4} + 1.8e^{-1.8} + e^{-1.8}$

$= -3.4e^{-2.4} + 2.8e^{-1.8}$

$= 0.1544$

6.54 The lifetime, in weeks, of a certain type of transistor is known to follow a gamma distribution with mean 10 weeks and standard deviation $\sqrt{50}$ weeks.

(a) What is the probability that a transistor of this type will last at most 50 weeks?

(b) What is the probability that a transistor of this type will not survive the first weeks?

Sol^n X : Lifetime of transistor (in weeks)

$\mu = 10, \ \sigma = \sqrt{50}$ (Gamma distribution)

$\mu = \alpha\beta \qquad \alpha\beta = 10$ — ①

$\sigma^2 = \alpha\beta^2 \qquad \alpha\beta^2 = 50$ — ②

$\Rightarrow \beta = 5, \ \alpha = 2$

$f(x) = \begin{cases} \frac{1}{25\,\Gamma(2)} x e^{-\frac{x}{5}}, & x>0 \\ 0, & \text{otherwise} \end{cases}$

(a) $P(X \le 50) = \int_0^{50} f(x)\,dx$

$= \int_0^{50} \frac{1}{25\,\Gamma(2)} x e^{-x/5} dx$

$= \int_0^{10} \frac{1}{25\,\Gamma(2)} 5y e^{-y} 5dy$

let $\frac{x}{5} = y$

$x = 5y$

$dx = 5dy$

when $x=0$, $y=0$

$x=50$, $y=10$

$$= \int_0^{10} \frac{y e^{-y}\, dy}{\Gamma(2)}$$

$$= F(10;2)$$

$$= 1 = 0.9995$$

6.46 The life, in years, of a certain type of electrical switch has an exponential distribution with an average life $\beta = 2$. If 100 of these switches are installed in different system, what is the probability that at most 30 fail during the first year?

Sol: X : Life of a battery (in years)

Exponential distribution

$\beta = 2$

$$f(x) = \begin{cases} \frac{1}{2} e^{-\frac{x}{2}}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

$n = 100$

$$P(X < 1) = \int_0^1 f(x)\, dx$$

$$= \int_0^1 \frac{1}{2} e^{-\frac{x}{2}} dx$$

$$= \frac{1}{2} \frac{e^{-x/2}}{(-1/2)} \Big]_0^1$$

$$= -e^{-x/2} \Big]_0^1$$

$$= -e^{-1/2} + 1$$

$p = 0.3934$, $q = 1 - 0.3934 = 0.6065$

$n = 100$, $\mu = np$, $\sigma = \sqrt{npq}$

$= 39.34$, $\sigma = 10$

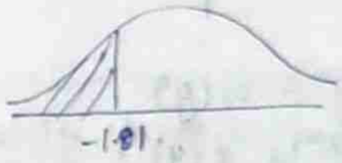
Y: No. of battery fail in a year

$$P(Y \le 30) = P\left(\frac{Y - \mu}{\sigma} \le \frac{30.5 - \mu}{\sigma}\right)$$

$$= P\left(Z \le \frac{30.5 - 39.34}{4.8}\right)$$

$$= P(Z \le -1.81)$$

$$= 0.0352$$



Functions of Random Variables
Chapter 7

Thm Suppose X is a discrete random variable and $f(x)$ be the probability distribution function and $Y = f(X)$ is a one-to-one function between X and Y and $y = f(x)$ can be written uniquely as $x = w(y)$, so,

$$g(y) = f(w(y))$$

Ex 7.1 Let X be a geometric random variable with probability distribution

$$f(x) = \frac{3}{4}\left(\frac{1}{4}\right)^{x-1}, \quad x = 1, 2, 3, \ldots$$

Find the probability function of $Y = X^2$

Soln: $y = x^2 \Rightarrow x = \sqrt{y} = w(y)$

The probability fn. of Y is $g(y)$

$$g(y) = f(w(y))$$
$$= f(\sqrt{y})$$
$$= \frac{3}{4}\left(\frac{1}{4}\right)^{\sqrt{y}-1}, \quad y = 1, 4, 9, \ldots$$

Ex Let X be a random variable with probability

$$f(x) = \begin{cases} \frac{1}{3}, & x = 1, 2, 3, \\ 0, & \text{elsewhere} \end{cases}$$

Find the probability distribution of the random variable $Y = 2X - 1$

Soln: ~~$x \cdot f(x) = \{$~~

$Y = 2X - 1$

$g(y)$?

$y = 2x - 1$
$\Rightarrow 2x = y + 1$
$\Rightarrow x = \frac{y+1}{2} = w(y)$

$$g(y) = f(w(y))$$
$$= f\left(\frac{y+1}{2}\right)$$

$$g(y) = \begin{cases} \frac{1}{3}, & y = 1, 3, 5 \\ 0, & \text{elsewhere} \end{cases}$$

Thm 2 Let X (continuous random variable) and $f(x)$ is the probability function and $Y = f(X)$ is a one-to-one fn. between X and Y and $y = f(x)$ can be written uniquely as $x = w(y)$. Then,

$$g(y) = f(w(y))\,|J|$$

where J is the Jacobian

$$J = \frac{dw}{dy} = w'(y)$$

Ex 7.3 Let X be a cts random variable with probability distribution

$$f(x) = \begin{cases} \frac{x}{12}, & 1 < x < 5 \\ 0, & \text{elsewhere} \end{cases}$$

Find the probability distribution of the random variable $Y = 2X - 3$

Sol^n $J = \frac{d\,w(y)}{dy}$

$y = 2x - 3$

$\Rightarrow 2x = y + 3$

$\Rightarrow x = \frac{y+3}{2} = w(y)$

$$J = \frac{d\,w(y)}{dy} = \frac{d\left(\frac{y+3}{2}\right)}{dy} = \frac{1}{2}$$

The probability function of Y is g(y)

$g(y) = f(w(y))\,|J|$

$= \frac{y+3}{24} \times \frac{1}{2}$

$= \frac{y+3}{48}$

$$g(y) = \begin{cases} \frac{y+3}{48}, & -1 < y < 7 \\ 0, & \text{elsewhere} \end{cases}$$

7.6 Given the random variable X with probability distribution

$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Find the probability distribution of $Y = 8X^3$

Sol^n $y = 8x^3$

$\Rightarrow x^3 = \frac{y}{8}$

$\Rightarrow x = \frac{\sqrt[3]{y}}{2} = w(y)$

$$J = \frac{d\,w(y)}{dy} = \frac{d\left(\frac{y^{\frac{1}{3}}}{2}\right)}{dy} = \frac{1}{2} \times \frac{1}{3} y^{-2/3}$$

$= \frac{1}{6} y^{-2/3}$

The probability function of Y is g(y)

$g(y) = f(w(y))\,|J|$

$= f\left(\frac{y^{1/3}}{2}\right) \times \frac{1}{6} y^{-2/3}$

$= \frac{\cancel{2} y^{\frac{1}{3}}}{\cancel{2}} \times \frac{1}{6} y^{-2/3}$

$= \frac{y^{-\frac{1}{3}}}{6}$

$$g(y) = \begin{cases} \frac{y^{-\frac{1}{3}}}{6}, & 0 < y < 8 \\ 0, & \text{elsewhere} \end{cases}$$

Thm 3 Suppose X_1 and X_2 are discrete random variables with joint probability $f(x_1, x_2)$. Let $Y_1 = u_1(X_1, X_2)$ and $Y_2 = u_2(X_1, X_2)$ and (Y_1, Y_2) so that the equation $y_1 = u_1(x_1, x_2)$ and $y_2 = u_2(x_1, x_2)$ uniquely written as $x_1 = w_1(y_1, y_2)$ and $x_2 = w_2(y_1, y_2)$. Then the probability function of Y_1 and Y_2 is
$g(y_1, y_2) = f(w(y_1, y_2), w_2(y_1, y_2))$.

* one-to-one corresponds between points (x_1, x_2) and (y_1, y_2).

7.3 Let X_1 and X_2 be discrete random variables with the joint multinomial distribution

$$f(x_1, x_2) = \binom{2}{x_1, x_2, 2-x_1-x_2}\left(\frac{1}{4}\right)^{x_1}\left(\frac{1}{3}\right)^{x_2}\left(\frac{5}{12}\right)^{2-x_1-x_2}$$

for $x_1 = 0, 1, 2$; $x_2 = 0, 1, 2$; $x_1 + x_2 \le 2$; and zero elsewhere. Find the joint probability of $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$

Soln: $y_1 = x_1 + x_2$ — ①

$y_2 = x_1 - x_2$ — ②

$y_1 + y_2 = 2x_1 \Rightarrow x_1 = \frac{y_1 + y_2}{2} = w_1$

$\therefore y_1 = \frac{y_1 + y_2}{2} + x_2$

$\Rightarrow x_2 = y_1 - \left(\frac{y_1 + y_2}{2}\right)$

$= \frac{y_1 - y_2}{2} = w_2$

$f(x)$: $x_1 = 0, 1, 2$

$x_2 = 0, 1, 2$

$x_1 + x_2 \le 2$

$$g(y_1, y_2) = {}^{2}C_{\frac{y_1+y_2}{2}, \frac{y_1-y_2}{2}, 2-y_1}\left(\frac{1}{4}\right)^{\frac{y_1+y_2}{2}}\left(\frac{1}{3}\right)^{\frac{y_1-y_2}{2}}\left(\frac{5}{12}\right)^{2-y_1}$$

$y_1 = x_1 + x_2$ $\qquad$ $y_2 = x_1 - x_2$

$x_1 = 0$, $x_2 = 0, 1, 2$

$x_1 = 1$, $x_2 = 0, 1$ $\qquad$ $y_1 = 0, 1, 2$, $\quad y_2 = 0, -1, -2, 1, 2$

$x_1 = 2$, $x_2 = 0$

$\therefore y_1 + y_2 \le 4$

Thm 4 Suppose X and Y are cts. random variable with joint probability $f(x_1, x_2)$. Let $Y_1 = u_1(X_1, X_2)$ and $Y_2 = u_2(X_1, X_2)$ and one-to-one corresponds between points (x_1, x_2) and (y_1, y_2) so that the eqn. $y_1 = u_1(x_1, x_2)$ and $y_2 = u_2(x_1, x_2)$ uniquely written as $x_1 = w_1(y_1, y_2)$ and $x_2 = w_2(y_1, y_2)$ then the probability function of Y_1 and Y_2 is $g(y_1, y_2) = f(w_1(y_1, y_2), w_2(y_1, y_2))\,|J|$

where $J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$

7.8. A dealer's profit, in units of \$5000, on a automobile is given by $Y = X^2$, where X is a random variable having the density function

$$f(x) = \begin{cases} 2(1-x), & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

(a) Find the probability density function of the random variable Y

(b) using the density function of Y, find the probability that the profit on the next new automobile sold by this dealership will be less than \$500

Sol: $Y = X^2$

$y = x^2 \Rightarrow x = \sqrt{y} = w(y)$

Y, $g(y) = f(w(y))|J|$

where $J = \frac{dw}{dy} = \frac{1}{2} y^{-\frac{1}{2}} = \frac{1}{2\sqrt{y}}$

$$g(y) = \begin{cases} \frac{1}{2} \times \frac{2(1-\sqrt{y})}{\sqrt{y}}, & 0 < y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

$$g(y) = \begin{cases} \left(\frac{1}{\sqrt{y}} - 1\right), & 0 < y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

1, 2, 3, 5, 8, 12

$-\frac{1}{2}+1$

(b) $P\left(Y < \frac{500}{5000}\right)$

$= P(Y < 0.1)$

$= \int_0^{0.1} g(y)\,dy$

$= \int_0^{0.1} \left(\frac{1}{\sqrt{y}} - 1\right).dy$

$= \int_0^{0.1} \left(y^{-\frac{1}{2}} - 1\right).dy$

$= \int_0^{0.1} y^{-\frac{1}{2}}\,dy - \int_0^{0.1} 1.dy$

$= 2y^{\frac{1}{2}}\Big]_0^{0.1} - y\Big]_0^{0.1}$

$= \frac{2\sqrt{y}}{\sqrt{y}}\Big]_0^{0.1} - y\Big]_0^{0.1}$

$= 2\sqrt{0.1} - 0.1$

$= 0.532$

Moments and moment generating function

Let x be a random variable with probability function $f(x)$. The rth moment of X about origin is

$$E(x^r) = \mu_r'$$

where

$$E(x^r) = \sum_x x^r f(x) \quad \text{(Discrete)}$$

$$E(x^r) = \int_{-\infty}^{\infty} x^r f(x)\, dx \quad \text{(continuous)}$$

Note: If $r = 1$ then,

$$E(x^1) = \mu_1' = \mu$$

If $r = 2$, then

$$E(x^2) = \mu_2' = \sigma^2 + \mu^2 \quad [\sigma^2 = E(x^2) - \mu^2]$$

Moment generating function

The moment generating function of the random variable X is denoted by the function

$$M_x(t) = E(e^{tx})$$

$$= \sum_x e^{tx} f(x) \quad \text{(If } x \text{ is discrete)}$$

$$M_x(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx \quad \text{(If } x \text{ is continuous)}$$

Thm Suppose X is a random variable with moment generating function $M_x(t)$, then the derivative,

$$\frac{d^r}{dt^r}(M_x(t))\Big]_{t=0} = \mu_r'$$

Q- Find the moment generating function of the binomial random variable X and show that the mean is np and variance is npq

Soln Suppose X is a random variable and

$$f(x) = {}^nC_x\, p^x q^{n-x}, \quad x = 0, 1, 2, \ldots n$$

$$M_x(t) = E(e^{tx})$$

$$= \sum_{x=0}^{n} e^{tx} f(x)$$

$$= \sum_{x=0}^{n} e^{tx}\, {}^nC_x\, p^x q^{n-x}$$

$$= \sum_{x=0}^{n} {}^nC_x (pe^t)^x \cdot q^{n-x}$$

$$= (pe^t + q)^n \quad \left(\text{by binomial expansion } (a+b)^n = \sum_{x=0}^{n} {}^nC_x\, a^{n-x} b^x\right)$$

$$M_x(t) = (pe^t + q)^n$$

$$\frac{d}{dt} M_x(t)\Big]_{t=0} = n(pe^t + q)^{n-1}(pe^t)\Big]_{t=0}$$

$$\Rightarrow \mu_1' = n(p+q)^{n-1} p \quad [\because p + q = 1]$$

$$\Rightarrow \boxed{\mu = np}$$

$$\frac{d^2}{dt^2} M_x(t) = n(n-1)p^2 (e^t)^2 (pe^t+q)^{n-2} + npe^t (pe^t+q)^{n-1}$$

$$\left.\frac{d^2}{dt^2} M_x(t)\right]_{t=0} = n(n-1)p^2(p+q)^{n-2} + np(p+q)^{n-1}$$

$$= n(n-1)p^2 + np \qquad (\because p+q=1)$$

$$\mu_2' = n^2p^2 - np^2 + np$$

$$\sigma^2 + \mu^2 = n^2p^2 - np^2 + np$$

$$\sigma^2 + \cancel{n^2p^2} = \cancel{n^2p^2} - np^2 + np$$

$$\sigma^2 = np - np^2$$

$$\sigma^2 = np(1-p)$$

$$\boxed{\sigma^2 = npq}$$

Ex 7.17 A random variable X has the discrete uniform distribution

$$f(x;K) = \begin{cases} \frac{1}{K}, & x = 1, 2, \ldots, K \\ 0, & \text{elsewhere} \end{cases}$$

Show that the moment-generating function of X is

$$M_x(t) = \frac{e^t(1-e^{Kt})}{K(1-e^t)}$$

Sol^n $M_x(t) = \sum_{x=1}^{K} e^{tx} f(x)$

$$= \sum_{x=1}^{K} \frac{1}{K} e^{tx}$$

$$= \frac{1}{K}\left[e^t + e^{2t} + e^{3t} + \cdots e^{Kt}\right]$$

$$= \frac{x}{K} e^t [1 + e^{2t}$$

$$= \frac{1}{K}\left[e^t \frac{(1-e^{Kt})}{1-e^t}\right]$$

$$= \frac{e^t}{K} \frac{(1-e^{Kt})}{(1-e^t)}$$

Statistics

ch - 8

* Measure of locations

The sample mean and median

Sample Mean

$x_1, x_2, x_3, \ldots, x_n$ $\quad n \rightarrow$ samples

$\bar{x} \rightarrow$ sample mean

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Trimmed Mean

$x_1, x_2, \ldots, x_{n-1}, x_n$ $\quad n \rightarrow$ samples

10 % trimmed mean $= n \times 10\%$

$$= \frac{n \times 10}{100} = \frac{n}{10}$$

Arrange the data either in ascending or descending order

Delete n/10 elements from the left and right side

$\cancel{x_1^*}, \cancel{x_2^*}, x_3^* \ldots \ldots x_{n-2}, \cancel{x_{n-1}^*}, \cancel{x_n^*}$

$$\bar{x}_{10\%\ \text{trimmed}} = \frac{x_3^* + x_4^* + \ldots x_{n-2}^*}{n-4}$$

Q- Calculate the mean and 10 % trimmed mean of the following data

~~3.4,~~ 2.5, 3.4, 2.5, 4.8, 2.9, 3.6, 2.8, 3.3, 5.8, 3.7, 2.8, 4.4, 4.0, 5.2, 3.0, 4.8

Soln

$$\text{Mean}, \bar{x} = \frac{57}{15} = 3.8$$

$$10\% \text{ trimmed mean} = \frac{n \times 10}{100}$$

$$= \frac{15 \times 10}{100} = 1.5, \approx 1$$

2.5, 2.8, 2.8, 2.9, 3.0, 3.3, 3.4, 3.6, 3.7, 4.0, 4.4, 4.8, 4.8, 5.2, 5.8

$$\bar{x}_{10\%\ \text{trimmed}} = \frac{2.8 + 2.8 + \ldots + 4.8 + 5.2}{(15 - 2)}$$

$$= \frac{48.7}{13} = 3.74$$

Sample Median

Given that the observations of a sample are $x_1, x_2, \ldots x_n$, arranged in ascending order in their magnitude. Then the sample median is

$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}}, & \text{if } n \text{ is odd} \\ \frac{1}{2}\left\{x_{\frac{n}{2}} + x_{\frac{n+1}{2}}\right\}, & \text{if } n \text{ is even} \end{cases}$$

Q- Calculate the sample median

1.7, 2.2, 3.9, 3.11, 14.8

Soln: $n = 5$ (odd)

$\tilde{x} = x_{\frac{n+1}{2}} = x_{\frac{6}{2}} = x_3$

$\tilde{x} = 3.9$

Sample Variance

The sample variance of the sample $x_1, x_2, \ldots, x_n$ is denoted by s^2

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

Standard deviation is the +ve square root of s^2, i.e. $+\sqrt{s^2}$

Q- Calculate sample variance & standard deviation of 1.7, 2.2, 3.9, 3.11, 14.8

Soln: $\bar{x} = \frac{1.7 + 2.2 + 3.9 + 3.11 + 14.8}{5}$

$\Rightarrow \bar{x} = 5.142$

$$s^2 = \frac{(1.7 - 5.142)^2 + (2.2 - 5.142)^2 + (3.9 - 5.142)^2 + (3.11 - 5.142)^2 + (14.8 - 5.142)^2}{4}$$

$$\Rightarrow s^2 = \frac{11.847 + 8.655 + 1.542 + 4.129 + 93.276}{4}$$

$$\Rightarrow s^2 = \frac{119.449}{4} = 29.86$$

Let $x_1, x_2, \ldots, x_n$ are n random variables;

① Sample mean, $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$

② Sample variance, $s^2 = \frac{1}{n-1} \left\{ \sum_{i=1}^{n} (x_i - \bar{x})^2 \right\}$

③ Sample Mode (value of the sample occurs most)

④ Sample Median

$$x^2 = \begin{cases} x_{\frac{n+1}{2}}, & \text{if } n \text{ is odd} \\ \frac{1}{2}\left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}\right), & \text{if } n \text{ is even} \end{cases}$$

⑤ Sample standard deviation

$s = +\sqrt{s^2}$

⑥ Sample Range

$R = X_{Max} - X_{min}$

Alternate formula for variance

$$s^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} (x_i - \bar{x})^2 \right)$$

$$= \frac{1}{n-1} \left\{ \sum_{i=1}^{n} (x_i^2 - 2x_i\bar{x} + \bar{x}^2) \right\}$$

$$= \frac{1}{n-1} \left\{ \sum_{i=1}^{n} x_i^2 - 2\bar{x} \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \bar{x}^2 \right\}$$

$$= \frac{1}{n-1}\left\{\sum_{i=1}^{n} x_i^2 - 2\bar{x}.n\bar{x} + n\bar{x}^2\right\}$$

$$= \frac{1}{n-1}\left\{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right\}$$

$$= \frac{1}{n-1}\left\{\sum_{i=1}^{n} x_i^2 - n\left(\sum_{i=1}^{n} x_i\right)^2\right\}$$

$$S^2 = \frac{1}{n(n-1)}\left\{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right\}$$

Ex-8.3 Find the mean, median, mode, variance, standard deviation & range of the following sample data.

3, 6, 5, 4, 7, 6

Sol^n Mean $(\bar{x}) = \frac{1}{6}\sum_{i=1}^{n} x_i$

$= \frac{1}{6}[3+6+5+4+7+6]$

$= 5.16$

Median $(\tilde{x}) = \dfrac{x_{\frac{6}{2}} + x_{\frac{6}{2}+1}}{2}$

$= \dfrac{x_3 + x_4}{2}$

$= \dfrac{5+6}{2}$

$= 5.5$

Mode $= 6$

variance $= \frac{1}{\cancel{n-1}}\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]$

variance $(S^2) = \frac{1}{n(n-1)}\left\{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1} x_i\right)^2\right\}$

$\sum_{i=1}^{6} x_i = 31$, $\sum_{i=1}^{6} x_i^2 = 9+16+25+36+36+49$

$= 171$

$\therefore S^2 = \frac{1}{6\times5}[6\times171 - (31)^2]$

$S^2 = 2.16$

$S = \sqrt{2.16} = 1.46$

$R = x_{Max} - x_{Min}$

$= 7 - 3$

$= 4$

8.5 The numbers of incorrect answers on a true-false competency test for a random sample of 15 students were recorded as follows: 2, 1, 3, 0, 1, 3, 6, 0, 3, 3, 5, 2, 1, 4 and 2. Find (a) the mean, (b) the median (c) the mode

Sol^n 0, 0, 1, 1, 1, 2, 2, 2, 3, 3, 3, 3, 4, 5, 6

2, 3, 5, 7, 10, 12

Sampling distribution

The probability distribution of statistics is known as sampling distribution

Sampling distribution of Mean

Suppose that a random sample of n observations is taken from a normal population with mean μ and variance σ^2

If each observation X_i, $i = 1, 2, \ldots, n$ of the random sample have mean μ and variance σ^2

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \ldots + X_n)$$

The mean of $\bar{X}$ is $\dfrac{\mu + \mu + \ldots + \mu \text{ (n times)}}{n}$

$$\boxed{\text{Mean of } \bar{X} = \mu}$$

$$\boxed{\mu_{\bar{x}} = \mu}$$

The variance of $\bar{X}$ is

$$\sigma^2_{\bar{x}} = \frac{1}{n^2}(\sigma^2 + \sigma^2 + \ldots + \sigma^2) \text{ (n times)}$$

$$= \frac{n\sigma^2}{n^2}$$

$$\boxed{\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}}$$

Central limit theorem

If $\bar{X}$ is the mean of the random sample size of n taken from the normal population with mean μ and variance is σ^2 then

$$Z = \frac{\bar{X} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$

as $n \to \infty$, this is standard normal distribution $n(z; 0, 1)$

Ex 8.4 An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed, with mean equal to 800 hrs and a standard deviation of 40 hrs. Find the probability that a random space of 16 bulbs will have an avg. life of less than 775 hours.

Sol^n $n = 16$, $\mu_{\bar{x}} = 800$, $\sigma_{\bar{x}} = \dfrac{40}{\sqrt{16}}$

$$P(\bar{X} < 775)$$

$$= P\left(\frac{\bar{X} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)} < \frac{775 - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}\right)$$

$= P\left(Z < \frac{775 - 800}{10}\right)$

$= P(Z < -2.5)$

$= \cancel{0.0202}\ 0.0062$

8.17 If all possible samples of size 16 are drawn from a normal population with mean equal to 50 and standard deviation equal to 5, what is the probability that a sample mean $\bar{X}$ will fall in the interval from $\mu_{\bar{X}} - 1.9\sigma_{\bar{X}}$ to $\mu_{\bar{X}} - 0.4\sigma_{\bar{X}}$. Assume that the sample means can be measured to any degree of accuracy.

Sol^n $n = 16$, $\mu_{\bar{X}} = 50$, $\sigma_{\bar{X}} = \frac{5}{\sqrt{16}} = 5/4$

$P(\mu_{\bar{X}} - 1.9\sigma_{\bar{X}} < \bar{X} < \mu_{\bar{X}} - 0.4\sigma_{\bar{X}})$

$= P\left(50 - 1.9 \times \frac{5}{4} < \bar{X} < 50 - 0.4 \times \frac{5}{4}\right)$

$= P\left(-1.9 < \frac{\bar{X} - 50}{5/4} < -0.4\right)$

$= P(-1.9 < Z < -0.4)$

$= P(Z < -0.4) - P(Z < -1.9)$

$= 0.3446 - 0.0287$

$= 0.3159$

8.25 The random variable X, representing the number of cherries in a cherry puff, has the following probability distribution

x	4	5	6	7
$P(X=x)$	0.2	0.4	0.3	0.1

(a) Find the mean μ and the variance σ^2 of X

(b) Find the mean $\mu_{\bar{X}}$ and the variance $\sigma^2_{\bar{X}}$ of the mean $\bar{X}$ for random samples of 36 cherry puffs

(c) Find the probability that the avg. no. of cherries in 36 cherry puffs will be less than 5.5

Sol^n $\mu = \sum_x x f(x)$

$= 4 \times 0.2 + 5 \times 0.4 + 6 \times 0.3 + 7 \times 0.1$

$= 5.3$

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

$E(X^2) = \sum_x x^2 f(x) = 16 \times 0.2 + 25 \times 0.4 + 36 \times 0.3 + 49 \times 0.1$

$= 28.9$

$\sigma^2 = 28.9 - (5.3)^2 = 0.81$

(17, 19, 20, 23, 24, 20, 30)

$\mu_{\bar{x}} = \mu \qquad n = 36$

$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n} = \frac{0.81}{36} = 0.0225$

$P(\bar{x} < 5.5)$

$\Rightarrow P\left(\frac{\bar{x} - \mu_{\bar{x}}}{\left(\frac{\sigma}{\sqrt{n}}\right)} < \frac{5.5 - 5.3}{\sqrt{0.0225}}\right)$

$= P(Z < 1.33)$

$= 0.9082$

Sample estimation problems

Statistical inference

Statistical inference consists of those methods one makes inference about population

It may be decided into two major areas

1. Estimation
2. Tests of Hypothesis

Estimation < Point ~~evaluation~~ estitmation / Interval estimation

Point Estimation

Maximum likelyhood method

step 1. Consider $x_0, x_1, x_2, \ldots x_n$ are ~~no~~ n no. of independent samples of a population with probability function $f(x, \theta)$ where θ is the statistical parameter.

2. Maximum weighted function

$L = f(x_1, \theta) \times f(x_2, \theta) \times \ldots \times f(x_n, \theta)$

$L = \prod_{i=1}^{n} f(x_i, \theta)$

$\ln(L) = \ln\left(\prod_{i=1}^{n} f(x_i, \theta)\right) \qquad \frac{\partial \ln(L)}{\partial \theta} = \frac{\partial}{\partial \theta}\left(\ln\left(\prod_{i=1}^{n} f(x_i, \theta)\right)\right)$

3. $\frac{\partial L}{\partial \theta} = \partial\left(\prod_{i=1}^{n} f(x_i, \theta)\right)$

let $\frac{\partial L}{\partial \theta} = 0$, $\hat{\theta}$

$\left[\frac{\partial^2 L}{\partial \theta^2}\right]_{\theta = \hat{\theta}} = -ve$, L has max at $\theta = \hat{\theta}$

Ex 9.20 Consider a Poisson distribution with probability mass function

$$f(x|\mu) = \frac{e^{-\mu} \mu^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Suppose that a random sample $x_1, x_2, \ldots x_n$ is taken from the distribution. What is the maximum likelihood estimate of μ?

Sol^n: consider $x_1, x_2, \dots, x_n$ are n no. of samples

$$f(x,\mu) = \frac{e^{-\mu}\mu^{x}}{x!}$$

The max^m likelyhood function:

$$L = \prod_{i=1}^{n} f(x_i, \theta)$$

$$L = \frac{e^{-\mu}\mu^{x_1}}{x_1!} \cdot \frac{e^{-\mu}\mu^{x_2}}{x_2!} \cdots \frac{e^{-\mu}\mu^{x_n}}{x_n!}$$

$$L = \frac{e^{-n\mu}\,\mu^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

$$\ln L = \ln\left(e^{-n\mu}\mu^{\sum_{i=1}^{n} x_i}\right) - \ln\left(\prod_{i=1}^{n} x_i!\right)$$

$$= \ln e^{-n\mu} + \ln \mu^{\sum_{i=1}^{n} x_i} - \ln\left(\prod_{i=1}^{n} x_i!\right)$$

$$\ln L = -n\mu + \left(\sum_{i=1}^{n} x_i\right)\ln\mu - \ln\left(\prod_{i=1}^{n} x_i!\right)$$

Differentiate w.r.t μ both sides

$$\frac{\partial(\ln L)}{\partial\mu} = -n + \frac{1}{\mu}\sum_{i=1}^{n} x_i - 0$$

$$\Rightarrow \frac{\partial(\ln L)}{\partial\mu} = 0 \Rightarrow -n + \frac{1}{\mu}\sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \frac{1}{\mu}\sum_{i=1}^{n} x_i = n$$

$$\Rightarrow \mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\Rightarrow \boxed{\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i} = \bar{x}$$

$$\frac{\partial^2 \ln L}{\partial\mu^2} = -\frac{1}{\mu^2}\sum_{i=1}^{n} x_i$$

$$\left.\frac{\partial^2(\ln L)}{\partial\mu^2}\right]_{\mu=\hat{\mu}} = \frac{-\sum_{i=1}^{n} x_i}{\frac{1}{n^2}\left(\sum_{i=1}^{n} x_i\right)^2}$$

$$= -n^2 \frac{1}{\sum_{i=1}^{n} x_i} \rightarrow -ve$$

Q.81 Suppose that there are n trials $x_1, x_2, \dots, x_n$ from a Bernoulli process with parameter p, the probability of a success. That is, the probability of r successes is given by ${}^{n}C_r\, p^r (1-p)^{n-r}$. Work out the maximum likelihood estimator for the parameter p.

81, 82, 85, 86, 87

Soln: $x_1, x_2, \ldots\ldots, x_n$

$$f(x,p) = {}^nC_x\, p^x (1-p)^{n-x}, \quad x = 0,1,2,\ldots$$

$$L = \prod_{i=1}^{n} f(x_i, p)$$

$$L = \prod_{i=1}^{n} {}^nC_x\, p^{\sum_{i=1}^{n} x_i} (1-p)^{\sum_{i=1}^{n}(n-x_i)}$$

$$\ln L = \ln\left(\prod_{i=1}^{n} {}^nC_{x_i}\right) + \left(\sum_{i=1}^{n} x_i\right)\ln p + \sum_{i=1}^{n}(n-x_i)\ln(1-p)$$

Diff. w.r.t p

$$\frac{\partial(\ln L)}{\partial p} = \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) + \frac{1}{p-1}\sum_{i=1}^{n}(n-x_i)$$

$$\frac{\partial(\ln L)}{\partial p} = 0$$

$$\Rightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) + \frac{1}{p-1}\sum_{i=1}^{n}(n-x_i) = 0$$

$$\text{or } \frac{(p-1)\sum_{i=1}^{n} x_i + p\left(\sum_{i=1}^{n}(n-x_i)\right)}{p(p-1)} = 0$$

$$\Rightarrow \frac{\cancel{p\sum_{i=1}^{n} x_i} - \sum_{i=1}^{n} x_i + pn^2 - \cancel{p\sum_{i=1}^{n} x_i}}{p(p-1)} = 0$$

$$\Rightarrow pn^2 - \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow p = \frac{\sum_{i=1}^{n} x_i}{n^2}$$

$$\Rightarrow p = \frac{\bar{x}}{n} = \hat{p}$$

$$\hat{p} = \frac{\bar{x}}{n}$$

Interval estimation

① confidence interval for mean μ with known variance.

② confidence interval for mean μ with unknown variance

confidence interval for mean μ with known variance

step 1 - Choose a confidence level $100(1-\alpha)\%$. where α is the level of significance

2 - Determine $Z_{\frac{\alpha}{2}}$ from Z distribution table where $Z_{\frac{\alpha}{2}}$ is the z-value leaving an area of $\alpha/2$ to the right

3 - Compute $\bar{x}$ from samples $x_1, x_2, \ldots, x_n$ where $\bar{x} = \frac{1}{n}\sum x_i$ n=sample size

4 - Compute the value of $K = \frac{\sigma}{\sqrt{n}} Z_{\frac{\alpha}{2}}$

5 - ~~conf~~ CONF $_{100(1-\alpha)\%}(\bar{x} - k \le \mu \le \bar{x} + k)$

Error

$$e = \frac{\sigma}{\sqrt{n}} Z_{\frac{\alpha}{2}}$$

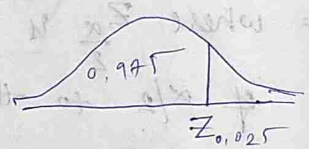
Ex 9.2 The avg. zinc concentration removed from a sample of measurements taken in 36 different locations in a river is found to be 2.6 grams per ~~million~~ milliliter. Find the 95% and 99% confidence intervals for the mean zinc concentration in the river. Assume that the ~~standard~~ population standard deviation is 0.3 gram per millimeter

soln $n = 36$, $\bar{x} = 2.6$

$100(1-\alpha)\% = 95\%$ $\sigma = 0.3$

$1 - \alpha = 0.95 \Rightarrow \alpha = 0.05$

$Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = Z_{0.025}$



$Z_{\frac{\alpha}{2}} = 1.96$

$\bar{x} = 2.6$

$k = \frac{\sigma}{\sqrt{n}} Z_{\frac{\alpha}{2}} = \frac{0.3}{\sqrt{36}} \times 1.96 = 0.098$

CONF $_{95\%}(\bar{x} - k \le \mu \le \bar{x} + k)$

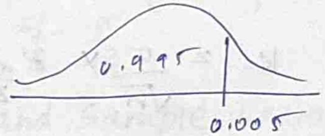
CONF $_{95\%}(2.502 \le \mu \le 2.698)$

$100(1-\alpha)\% = 99\%$

$1 - \alpha = 0.99$

$\alpha = 0.01$

$Z_{\frac{\alpha}{2}} = Z_{\frac{0.01}{2}} = Z_{0.005}$



$Z_{\frac{\alpha}{2}} = 2.57$

$\bar{x} = 2.6$

$k = \frac{\sigma}{\sqrt{n}} Z_{\frac{\alpha}{2}} = \frac{0.3}{\sqrt{36}} \times 2.57 = 0.1285$

CONF $_{99\%}(\bar{x} - k \le \mu \le \bar{x} + k)$

CONF $_{99\%}(2.4715 \le \mu \le 2.7285)$

9.2 An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed with a standard deviation of 40 hrs. If a sample 30 bulbs has an avg. life of 780 hrs, find a 96% confidence interval for the population mean of all bulbs produced by this firm

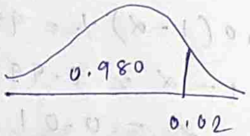
Sol^n $n = 30$, $\sigma = 40$

$\bar{x} = 780$

$100(1-\alpha)\% = 96\%$

$\alpha = 0.04$

$Z_{\frac{\alpha}{2}} = Z_{\frac{0.04}{2}} = Z_{0.02}$



$Z_{\frac{\alpha}{2}} = 2.05$

$K = \frac{\sigma}{\sqrt{n}} \times Z_{\frac{\alpha}{2}} = \frac{40}{\sqrt{30}} \times 2.05 \approx 14.97$

$\text{CONF}_{96\%} (\bar{x} - K \le \mu \le \bar{x} + K)$

$\text{CONF } 96\% \, (765.03 \le \mu \le 794.97)$

9.6 How large a sample is needed in excercise 9.2 if we wish to be 96% confident that our sample mean will be within 10 hrs of the true mean?

Sol^n $Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 10$

$\rightarrow 2.05 \times \frac{40}{\sqrt{n}} = 10$

$\rightarrow n \approx \cancel{2686}\ 68$

Confidence interval for mean μ with ~~uniform~~ unknown variance

Step 1 Choose a confidence level $100(1-\alpha)\%$, where α is the significance level

Step 2 Determine $t_{\frac{\alpha}{2}}$ from t-distribution table where $t_{\frac{\alpha}{2}}$ is the t-value with $\nu = n - 1$ degree of freedom, leaving the area to the right of $\frac{\alpha}{2}$

Step 3 Compute sample mean $\bar{x}$ and sample variance s^2: from sample

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} \mu \quad \text{and} \quad s^2 = \frac{1}{n(n-1)} \left[n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2 \right]$$

Step 4 Compute $K = \frac{s}{\sqrt{n}} t_{\frac{\alpha}{2}}$

Step 5 $\text{CONF}_{100(1-\alpha)\%} (\bar{x} - K \le \mu \le \bar{x} + K)$

Ex 9.5 The contents of seven similar containers of sulfuric acid are 9.8, 10.2, 10.4, 9.8, 10.0, 10.2, and 9.6 litres. Find a 96% confidence interval for the mean contents of all such containers, assuming an approximately normal distribution.

Sol^n^ $n = 7$

$100(1-\alpha) = 95$

$\Rightarrow \alpha = 0.05$

$t_{\frac{\alpha}{2}} = t_{\frac{0.05}{2}} = t_{0.025} \quad \nu = n-1 = 6$

$= 2.447$

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$= \frac{9.8 + 10.2 + 10.4 + 9.8 + 10 + 10.2 + 9.6}{7}$

$= \frac{70}{7} = 10$

$s^2 = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right]$

$s^2 = \frac{1}{7 \times 6}\left[7(700.48) - (70)^2\right]$

$s^2 = 0.08$

$s = \sqrt{0.08} = 0.283$

$K = \frac{0.283}{\sqrt{7}} \times 2.447 \approx 0.2617$

$CONF_{95\%} (10 - 0.2617 \le \mu \le 10 + 0.2617)$

$\Rightarrow CONF_{95\%} (9.7383 \le \mu \le 10.2617)$

Q.4 The heights of a random sample of 50 college students showed a mean of 174.5 cm and a standard deviation of 6.9 cm.

(a) Construct a 98% confidence interval for the mean height of all college students.

Sol^n^ $n = 50, \ \bar{x} = 174.5, \ s = 6.9$

$100(1-\alpha) = 98$

$\alpha = 0.02$

$z_{\frac{\alpha}{2}} = z_{\frac{0.02}{2}} = z_{0.01} \quad \nu = n-1 = 49$

$t_{\frac{\alpha}{2}} = t_{\frac{0.02}{2}} = t_{0.01}$

$= \frac{2.423 + 2.390}{2}$

$= 2.4065$

$K = \frac{6.9}{\sqrt{50}} \times 2.4065 \approx 2.348$

$CONF_{98\%} (174.5 - 2.348 \le \mu \le 174.5 + 2.348)$

$\Rightarrow CONF_{98\%} (172.152 \le \mu \le 176.848)$

Q.5 ~~A random samp~~

(b) What can be assert with confidence about the possible size of our error if we estimate the mean height of all college students to be 174.5 cm?

Sol^n Error at 98% is 2.348

9.5 A random sample of 100 automobile owners in the state of virginia shows that an automobile is driven on an avg. 23,500 km per year with a standard deviation of 3900 km. Assume the distribution of measurements to be approximately normal.

(a) construct a 99% confidence interval for the avg. no. of kms an automobile is driven annually in virginia

Sol^n $n = 100$, $\bar{x} = 23,500$, $s = 3900$

$100(1-\alpha) = 99$

$\alpha = 0.01$

$t_{\frac{\alpha}{2}} = t_{\frac{0.01}{2}} = t_{0.005}$ $\quad \nu = 100 - 1 = 99$

$= \frac{2.704 + 2.66}{2}$

$= 2.682$

$K = \frac{3900}{\sqrt{100}} \times 2.682 = 1045.98$

$CONF_{99\%} (23500 - 1045.98 \leq \mu \leq 23500 + 1045.98)$

$\Rightarrow CONF_{99\%} (22454.02 \leq \mu \leq 24,545.98)$