MATHEMATICS-IIA

Basic Probability Theory	2
Probability Distributions & Expectation	27
Chebychev's Inequality	66
Bivariate Distributions	72
Basic Statistics	90
Applied Statistics	104
Small Samples	117

WE ARE MAKAUTIANS

NOTE:

MAKAUT course structure and syllabus of 2nd Sem has been changed from 2019. Present syllabus of **MATHEMATICS IIA** [for CSE & IT] has been designed and structured with selected topics from other MATHEMATICS of different years and branches. Few new topics have been introduced also. Taking special care of this matter we are providing the relevant university solutions along with some model questions & answers for newly introduced topics, so that students can get an idea about university questions patterns.

BASIC PROBABILITY THEORY

☞ Chapter at a Glance

Random Experiment: Random Experiment is an in which
All possible outcomes of the experiments are known in advance but it is impossible to predict which outcome will occur in a particular performance and the experiment can be performed under identical conditions for infinite number of times, at least conceptually.
The experiments of "tossing an unbiased coin", "rolling an unbiased die" are very well known examples of random experiments.
Event: The outcomes of an experiment are known as events.
Simple and Composite/compound Events: An event A associated with random experiment E is called a *simple event* if it contains exactly one event point.
An event A associated with random experiment E is called a *composite/compound event* if it contains more than one event point.
Certain Event: An event which is sure to happen in every performance of an experiment is called *Certain Event* and denoted by S.
Impossible Event: An event which cannot occur at any performance of an experiment is called *Impossible Event* and denoted by ϕ.
Mutually Exclusive Events: Two or more events are said to be *mutually exclusive events* if they cannot happen simultaneously in any performance of a random experiment.
If events A, B are mutually exclusive, then $A \cap B = \phi$. (null set)
Exhaustive Set of Events: Set of all possible events of a random experiment constitute an *exhaustive set of events*.
Independent Events: A set of events is said to be independent if occurrence of any event does not depend on the occurrence of the other.
Event Space: The set of all possible outcomes of a random experiment is called event space of the experiment and denoted by S. The elements of S are called event points.
For example, in the experiment of 'tossing two unbiased coins', the event space is given by $S = \{HH, HT, TH, TT\}$, where H stands for 'Head' and T stands for 'Tail'; and HH, HT, TH, TT are the event points of the experiment 'tossing two unbiased coins'.
Complementary Event: The complementary event of an event A is denoted by A^c or A' or $\overline{A} = S - A$.

Classical Definition of Probability:

Let E be a random experiment and S be the event space with finite number of event points, say n; all of which are equally likely and exhaustive. Let A be an event associated with random

We ARE MAKAUTIANS

experiment E having m event points, then the probability of the event A is denoted as $P(A)$ and defined as $P(A)=\frac{m}{n}$.

Axioms of probability:

Let E be a random experiment described by the event space S and A be any event connected with E i.e. $A \subseteq S$ then the probability of the event A is denoted by $P(A)$ such that the following axioms are satisfied:

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

(ii) $P(S)=0$

(iii) If $A_1, A_2, A_3, \ldots$ be a finite or infinite sequence of pairwise mutually exclusive events then $P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1)+P(A_2)+P(A_3)+\ldots$

Conditional probability:

The conditional probability of an event A on the hypothesis that another event B has happened, denoted by $P(A|B)$ and defined by $P(A|B)=\frac{P(A \cap B)}{P(B)}$, provided $P(B) \neq 0$.

Similarly, the conditional probability $P(B|A)=\frac{P(B \cap A)}{P(A)}$, provided $P(A) \neq 0$.

Some Important Results:

1. The two events A and B are
   - i) Equally likely if $P(A)=P(B)$
   - ii) Mutually exclusive if $P(A \cap B)=0$
   - iii) Exhaustive set of events if $P(A+B)=1$
   - iv) Independent if $P(A \cap B)=P(A).P(B)$
2. For any two events A and B, $P(A \cup B)=P(A)+P(B)-P(A \cap B)$.
3. For any two mutually exclusive events A and B, $P(A \cup B)=P(A)+P(B)$.
4. If A and B are independent events, then $P(A \cap B)=P(A).P(B)$.
5. If A and B are independent events, then $P(A|B)=P(A)$.
6. If A and B are independent events, then $P(B|A)=P(B)$.

We ARE MAKAUTIANS

Bayes' theorem:

If $A_1, A_2, \ldots\ldots, A_n$ be a set of pairwise mutually exclusive and exhaustive set of events, then for any arbitrary event X, such that $P(X) \neq 0$.

Then $P(A_i \mid X) = \dfrac{p(A_i).p(X \mid A_i)}{\sum_{k=1}^{n} p(A_k)\, p(X \mid A_k)}$; $i = 1, 2, \ldots, n$

Bernoulli's Trials:

Let E be a random experiment in which the event space consists of only two event points "success" and "failure". Let E be repeated $n$times under identical conditions, then we get a sequence of n independent trials of E. These trials are known as *Bernoulli's sequence* of trials.

Binomial Law:

If A_i denotes the event 'exactly i successes' $(i \leq n)$ and consequently '$(n-i)$ failures' connected to the compound experiment resulting from Bernoulli trials with repetitions of the experiment for n times, then

$$P(A_i) = \binom{n}{i} p^i q^{n-i} \text{ for } i = 0, 1, \ldots, n$$

where $p(0 < p < 1)$ is the probability of "success" in each trial and $q = 1 - p$ is the probability of failure.

Poisson Approximation to Binomial Law:

Taking $p = \dfrac{\mu}{n}$, where $\mu(> 0)$ and considering $n \to \infty$ then $p \to 0$.

Hence, we get $P(A_i) \cong e^{-\mu} \dfrac{\mu^i}{i!}$, $(i = 0, 1, \ldots \infty)$

Multiple Choice Type Questions

1. The chance that a leap year selected at random will contain 53 Wednesdays is **[WBUT 2007]**

a) $\frac{2}{7}$ b) 0 c) 1 d) $\frac{6}{7}$

Answer: (a)

We ARE MAKAUTIANS

2. If $P(A)=\frac{1}{2}, P(B)=\frac{1}{3}, P(AB)=\frac{1}{4}$, then the value of $P(A\cup B)$ is [WBUT 2007]

a) $\frac{6}{7}$ b) $\frac{3}{7}$ c) 1 d) $\frac{7}{12}$

Answer: (d)

3. If A and B are two events with $P(A)=0.4$, $P(B)=0.3$ and $P(A\cap B)=0.2$, then $P(A\cap B^c)$ is [WBUT 2015, 2017]

a) 0.1 b) 0.2 c) 0.3 d) 0.4

Answer: (a)

4. A fair die is thrown. The probability that either an odd number or a number greater than 4 will turn up is [WBUT 2015, 2017]

a) $\frac{2}{5}$ b) $\frac{3}{7}$ c) $\frac{2}{7}$ d) $\frac{2}{3}$

Answer: (d)

5. A coin is tossed. Event {H}, {T} are [WBUT 2007]

a) mutually exclusive b) independent events

c) dependent events d) both (a) and (c)

Answer: (a)

6. For any two events A and B, $P(AB)\le P(A)\le P(A+B)\le P(A)+P(B)$

a) True b) False [WBUT 2008]

Answer: (a)

7. Three unbiased coins are tossed simultaneously. This is repeated four times. The probability of getting at least one head each time is [WBUT 2008]

a) $\left(\frac{1}{8}\right)^4$ b) $\left(\frac{2}{8}\right)^4$ c) $\left(\frac{3}{8}\right)^4$ d) $\left(\frac{7}{8}\right)^4$

Answer: (d)

8. Let *A* and *B* the two events, the $P(A\cap B)\ge$ [WBUT 2008]

a) $P(A)+P(B)$ b) $P(A)$ c) $P(B)$ d) $P(A)+P(B)-1$

Answer: (d)

We ARE MAKAUTIANS

9. If $A \subset B$ then $P(B-A)$ is equal to [WBUT 2008]

a) $P(A)-P(B)$ b) $P(B)-P(A)$ c) $P(A)-P(AB)$ d) none of these

Answer: (b)

10. Given $P(A)=\frac{1}{2}$, $P(B)=\frac{1}{3}$, $P(AB)=\frac{1}{4}$. The value of $P(\bar{A}\bar{B})$ is [WBUT 2008]

a) $\frac{5}{12}$ b) $\frac{1}{12}$ c) $\frac{3}{4}$ d) $\frac{7}{12}$

Answer: (a)

11. A purse contains 4 copper coins, 3 silver coins and an another purse contains 6 copper coins and 2 silver coins. A purse is chosen at random and a coin is taken out of it. The probability that it is a copper coin is [WBUT 2008]

a) $\frac{4}{7}$ b) $\frac{3}{4}$ c) $\frac{3}{7}$ d) $\frac{37}{56}$

Answer: (d)

12. The probability that a leap-year selected at random will contain 53 Sunday is [WBUT 2009]

a) $\frac{3}{7}$ b) $\frac{2}{7}$ c) $\frac{5}{7}$ d) $\frac{4}{9}$

Answer: (b)

13. If a coin is tossed 6 times in succession, the probability of getting at least one head is [WBUT 2009]

a) $\frac{63}{64}$ b) $\frac{3}{64}$ c) $\frac{7}{63}$ d) None of these

Answer: (a)

14. The probability that the 4 children of a family have different birthdays is [WBUT 2009]

a) 0.9836 b) 0.4735 c) 0.9 d) 0.757

Answer: (a)

15. If A and B are any two events such that $P(A \cap B)=\frac{1}{2}$, $P(A^c \cap B^c)=\frac{1}{3}$ and $P(A)=P(B)=p$, then the value of p is [WBUT 2009, 2018]

a) $\frac{7}{12}$ b) $\frac{5}{6}$ c) $\frac{1}{3}$ d) $\frac{1}{2}$

Answer: (a)

WE ARE MAKAUTIANS

16. If A and B are any two events, then [WBUT 2009]

a) $P(A+B)=P(A)+P(B)$
b) $P(A+B)=P(A)+P(B)-P(AB)$
c) $P(A+B)=P(A)+1$
d) $P(A+B)=1-P(B)$

Answer: (b)

17. If A and B are any two events such that A is the subset of B, then [WBUT 2009]

a) $P(A)=P(B)$
b) $P(A)\le P(B)$
c) $P(A)\ge P(B)$
d) $P(A)=P(B)=1$

Answer: (b)

18. Three unbiased coins are tossed simultaneously. This is repeated four times. Then the probability of getting at least one head each time is [WBUT 2010]

a) $(1/8)^4$
b) $(2/8)^4$
c) $(7/8)^4$
d) $(3/8)^4$

Answer: (a)

19. If $P(A)=\frac{1}{3}$, $P(B)=\frac{1}{4}$, $P(A\cup B)=\frac{1}{2}$ then $P(B|A)$ is [WBUT 2010, 2012]

a) $\frac{3}{4}$
b) $\frac{4}{3}$
c) $\frac{1}{4}$
d) $\frac{1}{3}$

Answer: (c)

20. A box contains 6 white and 4 black balls. One ball is drawn. What is the probability that it is white? [WBUT 2010]

a) $\frac{2}{5}$
b) $\frac{3}{5}$
c) $\frac{1}{5}$
d) none of these

Answer: (b)

21. If $\overline{A}$ is the complement of the event A, then [WBUT 2010]

a) $P(\overline{A})=1-P(A)$
b) $P(\overline{A})=P(A)$
c) $P(\overline{A})=1+P(A)$
d) none of these

Answer: (a)

22. If two events A and B are independent, then [WBUT 2015]

a) $P(A\cap B)=P(A)P(B)$
b) $P(A+B)=P(A)+P(B)$
c) $P(A-B)=P(A)P(B)$
d) $P(A\cap B)=P(A)P(B/A)$

Answer: (a)

We ARE MAKAUTIANS

23. The probability that A passes the exam is 0.9 and B passes the exam is 0.8. The probability that at least one of them passes is [WBUT 2016]

a) 0.98 b) 0.97 c) 0.9 d) 0.72

Answer: (a)

24. A problem in Mathematics is given to three students A, B and C. The chances of solving the problem by A, B and C are $\frac{1}{3}, \frac{1}{4}$ and $\frac{1}{5}$ respectively. The probability that the problem will be solved is [WBUT 2014]

a) $\frac{2}{5}$ b) $\frac{3}{5}$ c) $\frac{1}{60}$ d) $\frac{47}{60}$

Answer: (b)

25. The probability that a leap year selected at random will contain 53 Sundays is [WBUT 2014]

a) 2/53 b) 52/53 c) 1/7 d) 2/7

Answer: (d)

26. Let A and B be two events with $P(A)=\frac{1}{2}, P(B)=\frac{1}{3}, P(A\cap B)=\frac{1}{4}$. Then $P(A/B)$ is [WBUT 2014]

a) 3/4 b) 5/4 c) 7/4 d) 2

Answer: (a)

27. The probability of an event A is $\frac{1}{3}$, that of $A+B$ is $\frac{1}{2}$ and that of AB is $\frac{1}{4}$. Then the probability of B is [WBUT 2013]

a) $\frac{1}{12}$ b) $\frac{5}{12}$ c) $\frac{1}{6}$ d) none of these

Answer: (b)

28. Let A and B be two events such that $P(A\cup B)=\frac{7}{8}$; $P(A\cap B)=\frac{1}{4}$, and $P(\overline{A})=\frac{5}{8}$; then $P(A|\overline{B})$ is [WBUT 2018]

a) $\frac{1}{9}$ b) $\frac{1}{8}$ c) $\frac{1}{2}$ d) None of these

Answer: (d)

We ARE MAKAUTIANS

29. A fair die is thrown. The probability that either an odd number or a number greater than 4 will turn up is **[WBUT 2018]**

a) $\frac{2}{5}$ b) $\frac{2}{3}$ c) $\frac{2}{7}$ d) None of these

Answer: (c)

Short Answer Type Questions

1. What is axiomatic definition of probability? **[WBUT 2014]**

Answer:

Let E be a random experiment described by the event space S and A be any event connected with E i.e. $A \subseteq S$ then the probability of the event A is denoted by $P(A)$ such that the following axioms are satisfied:

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

(ii) $P(S) = 0$

(iii) If $A_1, A_2, A_3, \ldots$ be a finite or infinite sequence of pairwise mutually exclusive events then $P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$

2. If A and B are not mutually exclusive events, then show that **[WBUT 2007]**

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

Answer:

The events $A - A \cap B$, $A \cap B$ and $B - A \cap B$ are pairwise mutually exclusive.

We have $A = (A - A \cap B) + A \cap B$, $B = (B - A \cap B) + A \cap B$

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

Then $P(A) = P(A - A \cap B) + P(A \cap B) \Rightarrow P(A - A \cap B) = P(A) - P(A \cap B)$,

Similarly, $P(B) = P(B - A \cap B) + P(A \cap B) \Rightarrow P(B - A \cap B) = P(B) - P(A \cap B)$

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

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

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

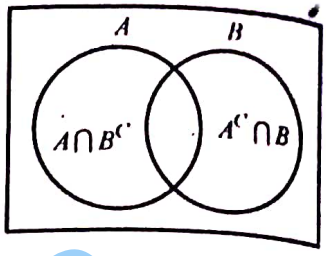
3. Show that the probability that exactly one of the events A and B occur is **[WBUT 2006, 2015]**

$P(A) + P(B) - 2P(AB)$.

WE ARE MAKAUTIANS

Answer:

The probability of occurrence of exactly one of the events A and B occur is



$$P(A\bar{B}+\bar{A}B)=P(A\bar{B})+P(B\bar{A})$$
$$=P(A)-P(AB)+P(B)-P(AB)$$
$$=P(A)+P(B)-2P(AB)$$
$$[\because A\bar{B}=A(S-B)=A-AB,\ \therefore P(A\bar{B})=P(A)-P(AB)]$$

4. If A and B are mutually independent events, prove that A^c and B^c are also mutually independent events. **[WBUT 2007, 2009]**

OR,

If $P(A\cap B)=P(A)P(B)$, then prove that $P(A^c\cap B^c)=P(A^c)P(B^c)$. **[WBUT 2012]**

Answer:

$$P(A^c\cap B^c)=P([A\cup B]^c)$$
$$=1-P(A\cup B)=1-P(A)-P(B)+P(A\cap B)$$
$$=1-P(A)-P(B)+P(A)P(B) \qquad \text{[By hypothesis]}$$
$$=1-P(A)-P(B)\{1-P(A)\}=\{1-P(A)\}\{1-P(B)\}$$
$$=P(A^c)P(B^c).$$

Hence, A^c and B^c are independent.

5. If A and B are two events such that $P(A^C\cup B^C)=5/6$, $P(A)=1/2$ and $P(B^C)=2/3$, show that A and B are independent. **[WBUT 2009, 2018]**

Answer:

We see

$$\frac{5}{6}=P(A^C\cup B^C)=1-P(A\cap B)\ \therefore P(A\cap B)=\frac{1}{6}.$$

$$P(A)=\frac{1}{2} \text{ and } P(B)=1-P(B^C)=1-\frac{2}{3}=\frac{1}{3}.$$

Clearly, $P(A\cap B)=P(A)P(B)$

Hence, A and B are independent.

We ARE MAKAUTIANS

6. If A and B are any two events such that $P(A\cap B)=\frac{1}{2}$, $P(A^c\cap B^c)=\frac{1}{3}$, $P(A)=P(B)=p$, then find the value of p. **[WBUT 2013]**

Answer:

We have $P(A\cap B)=\frac{1}{2}$, $P((A\cap B)^c)=\frac{1}{3}$ or, $1-P(A\cup B)=\frac{1}{3}$

or, $P(A\cup B)=\frac{2}{3}$, $P(A)=P(B)=p$

Now $\frac{2}{3}=P(A\cup B)=P(A)+P(B)-P(A\cap B)=p-\frac{1}{2}$

or, $2p=\frac{7}{6}$ or, $p=\frac{7}{12}$

7. If A and B are two events with $P(A)=0.4$, $P(B)=0.3$ and $P(A\cap B)=0.2$, find $P(A^c\cap B)$. **[WBUT 2004, 2006]**

Answer:

Since, $P(A\cup B)=P(A)+P(B)-P(A\cap B)=0.4+0.3-0.2=0.5$

Now, $P(A^c\cap B)=P(A\cap B)-P(A)=0.5+0.4=0.1$

8. Find the probability that a leap year selected at random contain 53 Wednesday. **[WBUT 2002, 2015]**

OR,

Find the probability that a leap year selected at random contain 53 Tuesdays. **[WBUT 2012]**

Answer:

We know that number of days in a leap year is 366, which contains 52 weeks and two extra days.

So, a leap year contains 52 Wednesday and 2 extra day.

Now, the leap year will contain 53 Wednesdays only if one of the two extra days is Wednesday.

Now, the two extra days will be either of the following:

(i) Sunday, Monday; (ii) Monday, Tuesday; (iii) Tuesday, Wednesday; (iv) Wednesday, Thursday; (v) Thursday, Friday; (vi) Friday, Saturday; (vii) Saturday, Sunday.

Hence, out of the above 7 cases (possible cases), 2 cases are in favour of Wednesday.

We ARE MAKAUTIANS

Therefore, the required probability $= \dfrac{\text{No. of favorable cases}}{\text{No. of possible cases}} = \dfrac{2}{7}$.

9. A and B play a game in which their chances of winning are in the ratio 3:2. A's chances of winning three games out of the five games played. **[WBUT 2017]**

Answer:

Chance of winning a game of player $A = \dfrac{3}{5}$;

Hence, probability of success of player A, $p = \dfrac{3}{5}$ and that of failure $q = \dfrac{2}{5}$

Hence, the probability of winning 3 games out of 5 games is $= \binom{n}{i} p^i q^{n-i}$

$$= \binom{5}{3}\left(\frac{3}{5}\right)^3\left(\frac{2}{5}\right)^{5-3} = \frac{5\times 4}{2}\left(\frac{3}{5}\right)^3\left(\frac{2}{5}\right)^2 = \frac{216}{625} = 0.3456$$

10. Let S be a sample space and two mutually exclusive events A and B such that $A \cup B = S$. If $P(\ )$ denotes the probability of the event, then find the maximum value of $P(A)\times P(B)$. **[WBUT 2017]**

Answer:

$P(A\cup B) = P(S) \Rightarrow P(A)+P(B)=1 \Rightarrow P(B) = 1-P(A)$

Let $P(A) = p$, then $P(B) = 1-p$

Let $P(A)\times P(B) = p(1-p) = f(p)$, (Say)

Now, $\dfrac{df}{dp} = 1-2p$ and $\dfrac{df}{dp} = 1-2p = 0 \Rightarrow p = \dfrac{1}{2}$

$\dfrac{d^2 f}{dp^2} = -2(<0)$

Therefore, $f(p)$ is maximum for $p = \dfrac{1}{2}$ *i.e.*, for $P(A) = \dfrac{1}{2}$ and hence, $P(B) = \dfrac{1}{2}$.

Therefore, the maximum value of $f(p) = \dfrac{1}{2}\times\dfrac{1}{2} = \dfrac{1}{4}$

That is the maximum value of $P(A)\times P(B) = \dfrac{1}{4}$

We ARE MAKAUTIANS

11. A pair of dice is thrown. Find the probability of getting a sum of 7, when it is known that the digit in the first die is greater than that of the second. [WBUT 2009]

Answer:

Let S denote the sample space and let A denote the event of getting a sum of 7 points when it is known that the digit in the first die is greater than that of the second.

Clearly, $S=\{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),....,(6,6)\}$

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

Hence, $P(A)=\dfrac{n(A)}{n(S)}=\dfrac{3}{36}=\dfrac{1}{12}$

12. The probabilities of solving a problem by three students A, B and C are $\frac{2}{7}$, $\frac{3}{8}$ and $\frac{1}{2}$ respectively. If all of them try independently, find the probability that the problem is solved. [WBUT 2009]

Answer:

Let X, Y, Z are the events that the problem solved by students A, B, C respectively.

Hence $P(X)=\dfrac{2}{7}, P(Y)=\dfrac{3}{8}, P(Z)=\dfrac{1}{2}$

Now, P (The problem is solved)

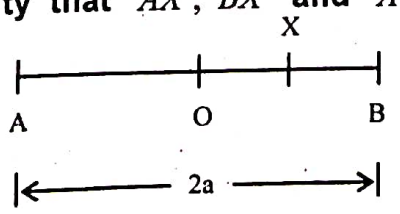
$=P(X\cup Y\cup Z)=1-P\left((X\cup Y\cup Z)^c\right)$

$=1-P\left(X^C\cap Y^C\cap Z^C\right)$ [$\because X, Y, Z$ are independent]

$=1-P\left(X^C\right)P\left(Y^C\right)P\left(Z^C\right)=1-\left(1-\dfrac{2}{7}\right)\left(1-\dfrac{3}{8}\right)\left(1-\dfrac{1}{2}\right)$

$=1-\dfrac{5}{7}\times\dfrac{5}{8}\times\dfrac{1}{2}=1-\dfrac{25}{112}=\dfrac{87}{112}$

13. A point X is chosen at random on a line segment AB of length $2a$ with mid-point O. Find the probability that AX, BX and AO can form the sides of a triangle. [WBUT 2015]



Answer:

We have to find

$P(AX, BX$ and AO form a triangle)

$=P(AX+BX>AO, BX+AO>AX, AO+AX>BX)$

We ARE MAKAUTIANS

$$= P(x+2a-x>a; 2a-x+a>x; a+x>2a-x)$$

$$= P\left(x<\frac{3}{2}a; x>a\right) = P\left(a<x<\frac{3}{2}a\right) = \frac{\frac{1}{2}a}{2a} = \frac{1}{4}$$

14. A group consists of equal number of men and women. Of this group 20% of the men and 50% of the women are unemployed. If a person is selected at random from this group, find the probability of the selected being employed. [WBUT 2016]

Answer:

Let n denote the no. of men and so n is also the no. of women.

So no. of unemployed men is $\frac{20n}{100}$ and no. of unemployed women is $\frac{50n}{100}$.

$\therefore P$ (the selected is unemployed) $= \frac{(0.2+0.5)n}{2n} = \frac{0.7}{2} = 0.35$

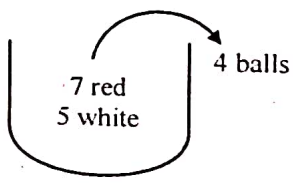
15. A bag contains 7 red and 5 white balls. 4 balls are drawn at random. What is the probability that (i) all of them are red (ii) two of them be white and two red? [WBUT 2009]

Answer:

Clearly 4 balls can be drawn from the bag in ${}^{12}C_4$ ways which are equally likely, mutually exclusive and exhaustive.

Now, P (all are red) $= \frac{{}^{7}C_4}{{}^{12}C_4}$

P (Two white & 2 red) $= \frac{{}^{7}C_2 \times {}^{5}C_2}{{}^{12}C_4}$.



16. An urn contains 3 white and 5 black balls. One ball is drawn and its colour is unnoted, kept aside and then another ball is drawn. What is the probability that it is (i) black (ii) white? [WBUT 2009]

Answer:

P(The second ball is black)

= P(1st ball is white & The 2nd ball is black OR 1st ball is black & 2nd ball is black)

= P(1st ball is white & 2nd ball is black) + P(1st ball is black & 2nd ball is black)

Since the respective events are mutually exclusive.

$$= \frac{3}{8}\times\frac{5}{7}+\frac{5}{8}\times\frac{4}{7} = \frac{35}{56} = \frac{5}{8}$$

Similarly, P(The second ball is white) $= \frac{3}{8}\times\frac{2}{7}+\frac{5}{8}\times\frac{3}{7} = \frac{3}{8}$

WE ARE MAKAUTIANS

17. State and prove Bayes' theorem. **[WBUT 2008, 2012, 2014, 2018]**

Answer:

Bayes' theorem:

If $A_1, A_2, \ldots, A_n$ be a set of pairwise mutually exclusive and exhaustive set of events, then for any arbitrary event X, such that $P(X) \neq 0$.

Then $P(A_i \mid X) = \dfrac{P(A_i).P(X \mid A_i)}{\sum_{k=1}^{n} P(A_k)P(X \mid A_k)}$; $i = 1, 2, \ldots, n$

Proof: By the theorem of conditional probability $(A_i \mid X) = \dfrac{P(A_i \cap X)}{P(X)}$; $P(X) \neq 0$

Since $A_1, A_2, \ldots, A_n$ are mutually exclusive and exhaustive set of events, $S = \bigcup_{i=1}^{n} A_i$, where S is the sample space.

Therefore, $X = X \cap S = X \cap \left(\bigcup_{i=1}^{n} A_i\right) = \bigcup_{i=1}^{n} (X \cap A_i)$

And $(X \cap A_i) \cap (X \cap A_j) = \phi$ for $i \neq j$

$$P(X) = P\left(\bigcup_{i=1}^{n} (A_i \cap X)\right) = \sum_{i=1}^{n} P(A_i \cap X)$$

$$P(A_i \mid X) = \frac{P(A_i \cap X)}{P(X)} = \frac{P(A_i \cap X)}{\sum_{i=1}^{n} P(A_i \cap X)} = \frac{P(A_i).P(X \mid A_i)}{\sum_{i=1}^{n} P(A_i)P(X \mid A_i)}$$

This completes the proof.

18. A box contains 6 white and 4 black balls. One ball is drawn at random. What is the probability that the ball drawn is white? **[WBUT 2013]**

Answer:

The required probability $= \dfrac{{}^{6}C_1}{{}^{10}C_1} = \dfrac{3}{5}$

19. The chance of a student passing an exam. is 20%. The chance of a student passing the exam and getting above 90% marks in it is 5%. Given that a student passes the examination, find the probability of his/her getting above 90% marks. **[WBUT 2016]**

Answer:

Let A denote the event of a student passing an exam and B denote the event of a student getting above 90%. Then $P(A) = 0.2$ and $P(A \cap B) = 0.05$

The required probability $= P(B|A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{0.05}{0.2} = 0.25$

WE ARE MAKAUTIANS

20. The probability that a doctor diagnoses a patient correctly is 80%. The probability that a patient dies after correct diagnosis is 40% and the probability that he dies after wrong diagnosis is 70%. The patient died. What is the probability that the disease was correctly diagnosed? [WBUT 2015, 2018]

Answer:

Let A denotes the event that doctor diagnoses a patient correctly and B denotes the event that a patient dies.

$\therefore P(A)=0.8, \quad P(B|A)=0.4, \quad P(B|\bar{A})=0.7$

Now, $P(A|B)=\dfrac{P(A).P(B|A)}{P(A).P(B|A)+P(\bar{A}).P(B|\bar{A})}=\dfrac{0.8\times 0.4}{0.8\times 0.4+0.2\times 0.7}$

$$=\frac{0.32}{0.32+0.14}=\frac{0.32}{0.46}=\frac{32}{46}=\frac{16}{23}$$

21. 3 identical boxes, I, II and III contain respectively 4 white and 3 red balls, 3 white and 7 red balls, 2 white and 3 red balls. A box is chosen at random and a ball is drawn out of it. If the ball is found to be white, what is the probability that Box II is selected. [WBUT 2007]

Answer:

We have, Box I: 4 white and 3 red balls
Box II: 3 white and 7 red balls
Box III: 3 white and 3 red balls

Let A_i denote the event of choosing the ith box, $i=1,2,3$ and B denote the event of the ball being white.

Then the required probability

$$=P(A_2|B)=\frac{P(B|A_2)P(A_2)}{\sum_{i=1}^{3}P(B|A_i)P(A_i)}=\frac{\frac{1}{3}\times\frac{3}{10}}{\frac{1}{3}\times\frac{4}{7}+\frac{1}{3}\times\frac{3}{10}+\frac{1}{3}\times\frac{2}{5}}=\frac{21}{89}$$

22. In answering a question on a multiple choice test, a student either knows the answer or he guesses. Let P be the probability that he knows the answer and $1-p$ be the probability that he guesses. Assume that a student who guesses the answer will be correct with probability $\frac{1}{5}$. What is the conditional probability that a student knew the answer to a question given that he answered it correctly? [WBUT 2008]

We ARE MAKAUTIANS

Answer:

Let A denote the event that the student knows the answer and B denote the event that he will answer correctly.

Then by hypothesis,

$$P(A) = p\ ,\ P(A^c) = 1 - p = P(\text{he guesses})$$

$$P(B|A^c) = \frac{1}{5}$$

The required probability $= P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

23. Consider a box, which contains 5 defective and 10 non-defective lamps. If 8 lamps are drawn at random in succession without replacement, then find the probability that the 8th lamp drawn is the 5th defective one. [WBUT 2005]

OR,

A box contains 5 defective and 10 non-defective lamps, 8 lamps are drawn at random in succession without replacement. What is the probability that the 8th lamp drawn is the 5th defective one? [WBUT 2007, 2008, 2015]

Answer:

Here, Since 8th lamp is the 5th defective, four defectives happened in the earlier seven lamp drawn.

P(8th lamps is the 5th defective)

$= P$(Out of 7 drawn 4 are defective and the 8th lamp is defective

$= P$(Out of 7 drawn 4 are defective) P(8th lamp is defective)

$$= \frac{{}^5C_4 \times {}^{10}C_3}{{}^{15}C_7} \times \frac{1}{8} = \frac{5 \times \dfrac{10 \times 9 \times 8}{3 \times 2}}{\dfrac{15 \times 14 \times 13 \times 12 \times 11 \times 10 \times 9}{7 \times 6 \times 5 \times 4 \times 3 \times 2}} \times \frac{1}{8} = \frac{5}{429}$$

24. It is seen that a cricketer becomes out within 10 runs in 3 out of 10 innings. If he plays 4 innings, what is the probability that he will becomes (i) out twice (ii) out at least once within 10 runs. [WBUT 2015]

Answer:

Let p denote the prob. That a cricketer becomes out within 10 runs

Then $p = 3/10$

The cricketer plays 4 innings.

Now P (He will be out 2 times) $= {}^4C_2 \left(3/10\right)^2 \left(7/10\right)^2$

P (He will be out at least once) $= 1 - P$ (He will not be out in 4 innings)

$$= 1 - {}^4C_0 \left(3/10\right)^0 \left(7/10\right)^4 = 1 - \left(7/10\right)^4$$

WE ARE MAKAUTIANS

25. Three groups of children consists of 3 girls and 1 boy, 2 girls and 2 boys and 1 girl and 3 boys respectively. One child is selected from each group. Show that the probability that the 3 selected children consisting of 1 girl and 2 boys is 13/32.
[WBUT 2005, 2007]

Answer:

Here Group I: 3 girls, 1 boy
Group II: 2 girls, 2 boys
Group III: 1 girl, 3 boys

$\therefore P(1 \text{ girl and } 2 \text{ boys}) = P\left((G_1 \cap B_2 \cap B_3) \cup (B_1 \cap G_2 \cap B_2) \cup (B_1 \cap B_2 \cap G_3)\right)$

Where G_i denotes a girl from the ith group and B_i denotes a boy from the ith group.

$$= P(G_1)P(B_2)P(B_3) + P(B_1)P(G_2)P(B_2) + P(B_1)P(B_2)P(G_3)$$

$$= \frac{3}{4}\times\frac{2}{4}\times\frac{3}{4} + \frac{1}{4}\times\frac{2}{4}\times\frac{3}{4} + \frac{1}{4}\times\frac{2}{4}\times\frac{1}{4} = \frac{9+3+1}{32} = \frac{13}{32}$$

26. There are 500 misprints in a book of 500 pages. What is the probability that a given page will contain at most 3 misprints?
[WBUT 2007]

Answer:

Let X denote the no. of misprints, then the req. probability is

$$= P(X \le 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$$

$$= \frac{e^{-1}1^0}{0!} + \frac{e^{-1}1^1}{1!} + \frac{e^{-1}1^2}{2!} + \frac{e^{-1}1^3}{3!} = e^{-1}\left(1+1+\frac{1}{2}+\frac{1}{6}\right) = \frac{8}{3e}$$

27. In a shooting competition the probability of a man hitting a target is $\frac{1}{5}$. If he fires 5 times, what is the probability of hitting the target at least twice?
[WBUT 2012]

Answer:

Here $p = \frac{1}{5}$, $n = 5$, $\therefore q = \frac{4}{5}$

$$P(X \ge 2) = 1 - P(X=1) - P(X=0)$$

$$= 1 - {}^5C_1\left(\frac{1}{5}\right)^1\left(\frac{4}{5}\right)^4 - {}^5C_0\left(\frac{4}{5}\right)^5$$

$$= 1 - 0.4096 - 0.3277 = 0.2627$$

28. The overall percentage of failures in a certain examination is 40. What is the probability that out of a group of 6 candidates at least 4 passed the examination?
[WBUT 2008]

We ARE MAKAUTIANS

Answer:

Let p denote the probability of passing.

Then $p = 1 - 0.40 = 0.60$.

$\therefore P$ (at least 4 passed) $= P(4 \text{ passed}) + P(5 \text{ passed}) + P(6 \text{ passed})$

$$= {}^6C_4 (0.60)^4 (0.40)^2 + {}^6C_5 (0.60)^5 (0.40)^4 + {}^6C_6 (0.60)^6 (0.40)^0$$

29. It is observed that a cricket player becomes out with in 10 runs in 3 out of 10 innings. If he played 4 innings, what is the probability that he will become (i) out twice (ii) out at least once within 10 runs? [WBUT 2007]

Answer:

Let A denote the event of the cricketer's being out within 10 runs

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

P(he will be out within 10 runs twice in 4 innings) $= {}^4C_2 . \left(\frac{3}{10}\right)^2 . \left(\frac{7}{10}\right)^2$

P(he will be out within 10 runs at least once)

$$= 1 - P(\text{he will not be out within 10 run over}) = 1 - \left(\frac{7}{10}\right)$$

30. A bag contains 5 white and 4 black balls. If 3 balls are drawn at random, what are the probabilities of the following? [WBUT 2007]

i) 2 of them are white

ii) Atmost one of them is white

iii) at least two are white.

Answer:

i) Now, P(2 of them are white) $= \frac{{}^5C_2 \times {}^4C_1}{{}^9C_3}$

P (at most one of them is white) $= P$(none is white exactly one is white)

$= P$(none is white) + P(exactly one is white)

$$= \frac{{}^4C_3}{{}^9C_3} + \frac{{}^5C_1 \times {}^4C_2}{{}^9C_3} = \frac{1}{21} + \frac{5}{14} = \frac{17}{42}$$

ii) The total no. of equally likely, mutually exclusive and exhaustive cases is

$${}^9C_3 = \frac{9 \times 8 \times 7}{6} = 84$$

We ARE MAKAUTIANS

iii) P(at least two are white)

$= P$(Two are white or all three are white only)

$= P$(Two are white) + P(all three are white)

$$= \frac{{}^5C_2 \times {}^4C_1}{{}^9C_3} + \frac{{}^5C_3}{{}^9C_3} = \frac{40}{84} + \frac{10}{84} = \frac{50}{84} = \frac{25}{42}$$

31. One ticket is selected at random from 100 tickets numbered 00, 01, 02, ...98, 99. If X and Y denote the sum and product of the digit on the tickets respectively. Find $P(X = 9|Y = 0)$. [WBUT 2007]

Answer:

We know that $P(X=9|Y=0) = \dfrac{P[(X=9)\cap(Y=0)]}{P(Y=0)}$

Now, $P(X=9) = P\{09,18,27,36,45,54,63,72,81,90\} = \dfrac{10}{100} = \dfrac{1}{10}$

$P(Y=0) = P\{00,01,02,03,04,05,06,07,08,09,10,20,30,40,50,60,70,80,90\} = \dfrac{19}{100}$

$P[(X=9)\cap(Y=0)] = P(\{09,90\}) = \dfrac{2}{100} = \dfrac{1}{50}$

$$\therefore P(X=9|Y=0) = \frac{\frac{2}{100}}{\frac{19}{100}} = \frac{2}{19}$$

32. The manufacturing process of an article consists of two parts x and y. The probabilities of defect in parts x and y are 10% and 15% respectively. What is the probability that the assembled product will not have any defect? [WBUT 2009]

Answer:

Let A denote the event of getting an article defective in the first process and B denote the event of getting an article defective in the second process by the given condition,

$P(A) = \dfrac{10}{100} = \dfrac{1}{10}$ & $P(B) = \dfrac{15}{100} = \dfrac{3}{20}$

Now, P(no defect) = P(no defect in 1st process and no defect in 2nd process)

= P(no defect in 1st process) × P(no defect in 2nd process)

$$= \frac{9}{10} \times \frac{17}{20} = \frac{153}{200}$$

We ARE MAKAUTIANS

Long Answer Type Questions

1. *A* and *B* throw alternatively with a pair of dice. *A* wins if he throws 8 before *B* throws 5 and *B* wins if he throws 5 before *A* throws 8. Find the probability that *A* wins.

[WBUT 2006]

Answer:

Let X, Y be the event "A throws 8" i.e., "A wins" and "B throws 5" i.e., "B wins" in a single throw (trial) of a pair of dice.

So, $X = \{(2,6),(3,5),(4,4),(6,2),(5,3)\}$ and $Y = \{(1,4),(2,3),(3,2),(4,1)\}$.

Hence, $(X) = \frac{5}{36}$; $P(\bar{X}) = 1 - \frac{5}{36} = \frac{31}{36}$

$P(Y) = \frac{4}{36} = \frac{1}{9}$; $P(\bar{Y}) = 1 - \frac{1}{9} = \frac{8}{9}$

Since, the experiment performed infinite times, so it is an infinite sequence of Bernoulli's trial and hence the event "A wins" is

$$(X) \cup (\bar{X}, \bar{Y}, X) \cup (\bar{X}, \bar{Y}, \bar{X}, \bar{Y}, X) \cup \ldots\ldots$$

Hence, the probability that "A wins" is

$$= P\left[(X) \cup (\bar{X}, \bar{Y}, X) \cup (\bar{X}, \bar{Y}, \bar{X}, \bar{Y}, X) \cup \ldots\ldots\right]$$

$$= P(X) + P(\bar{X}\bar{Y}X) + P(\bar{X}\bar{Y}\bar{X}\bar{Y}X) + \ldots\ldots$$

$$= P(X) + P(\bar{X})P(\bar{Y})P(X) + P(\bar{X})P(\bar{Y})P(\bar{X})P(\bar{Y})P(X) \ldots\ldots$$

[all successive trials are independent]

$$= \frac{5}{36} + \frac{31}{36}.\frac{8}{9}.\frac{5}{36} + \frac{31}{36}.\frac{8}{9}.\frac{31}{36}.\frac{8}{9}.\frac{5}{36} + \ldots\ldots$$

$$= \frac{5}{36} + \left(\frac{31}{36}.\frac{8}{9}\right).\frac{5}{36} + \left(\frac{31}{36}.\frac{8}{9}\right)^2.\frac{5}{36} + \ldots\ldots$$

$$= \frac{\frac{5}{36}}{1 - \frac{31}{36}.\frac{8}{9}} = \frac{45}{76}$$

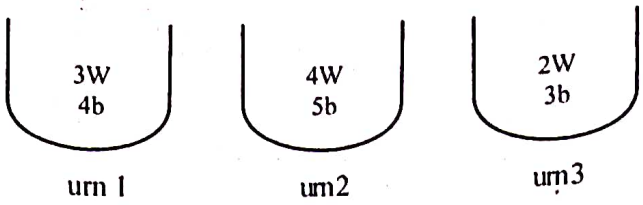
[As the series is an infinite G. P series]

2. There are three identical urns containing white and black balls. The first urn contains 3 white and 4 black balls, the 2nd urn contains 4 white and 5 black balls and the 3rd urn contains 2 white and 3 black balls. An urn is chosen at random and a ball is drawn from it. If the drawn ball is white, what is the probability that the 2nd urn chosen?

[WBUT 2009]

WE ARE MAKAUTIANS

Answer:



Let A_1, A_2, A_3 denote the events that the 1st, 2nd, 3rd urn is chosen respectively and B denote the event that the ball drawn is white by Bayes' theorem we have

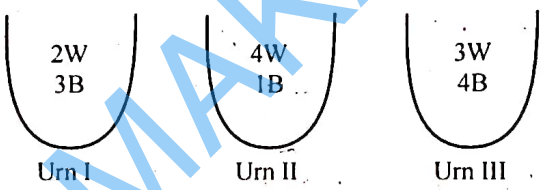
$$P(A_2/B) = \frac{P(A_2)P(B/A_2)}{P(A_1)P(B/A_1)+P(A_2)P(B/A_2)+P(A_3)P(B/A_3)}$$

$$= \frac{\frac{1}{3}\times\frac{4}{9}}{\frac{1}{3}\times\frac{3}{7}\times\frac{1}{3}\times\frac{4}{9}+\frac{1}{3}\times\frac{2}{5}} = 0.35 \text{ (approx..)}$$

3. Urn I has 2 white and 3 black balls. Urn II has 4 white and 1 black and Urn III 3 has white and 4 black balls. An Urn is selected at random and ball drawn at random is found to be white. Find the probability that Urn I was selected.

[WBUT 2010, 2015]

Answer:



Let A_i denote the event "selecting the ith urn" $(i = 1, 2, 3)$. Therefore the events A_1, A_2, A_3 are pairwise mutually exclusive and one of these necessarily occur.

Now, $P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$.

Again, let X denote the event "white ball".

Thus, $P(X|A_1) = \frac{2}{5}$, $P(X|A_2)\frac{4}{5}$, $P(X|A_3) = \frac{3}{7}$.

Therefore, $P(X) = P(A_1)P(X|A_1) + P(A_2)P(X|A_2) + P(A_3)P(X|A_3)$

$$= \frac{1}{3}\left[\frac{2}{5}+\frac{4}{5}+\frac{3}{7}\right] = \frac{19}{35}.$$

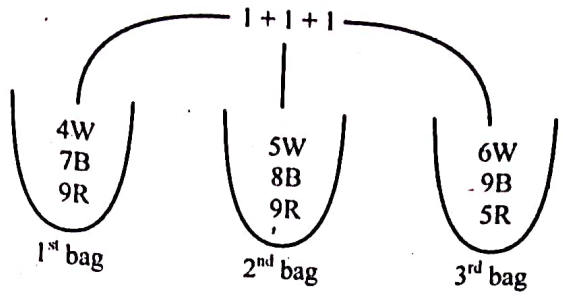
Now by Bayes' theorem we get,

$$P(A_1|X) = \frac{P(A_1)P(X|A_1)}{P(X)} = \frac{\frac{1}{3}\times\frac{2}{5}}{\frac{19}{35}} = \frac{2}{\cancel{15}_{3}}\times\frac{\cancel{35}^{7}}{19} = \frac{14}{57}.$$

We ARE MAKAUTIANS

4. Three bags contain respectively 4W, 7B, 9R; 5W, 8B, 7R and 6W, 9B, 5R balls. One ball is chosen at random from each bag. What is the probability that the balls are a) of same colour (b) 2W and IR?

Answer: **[WBUT 2011]**



a) $P(\text{all of same colour}) = P(\text{all are white or all are black or all are red})$

$$= P(\text{all are white}) + P(\text{all are black}) + P(\text{all are red})$$

$$= P(W_1 W_2 W_3) + P(B_1 B_2 B_3) + P(R_1 R_2 R_3)$$

$$= P(W_1)P(W_2)P(W_3) + P(B_1)P(B_2)P(B_3) + P(R_1)P(R_2)P(R_3)$$

$$= \frac{4}{20}\times\frac{5}{20}\times\frac{6}{20} + \frac{7}{20}\times\frac{8}{20}\times\frac{9}{20} + \frac{9}{20}\times\frac{7}{20}\times\frac{5}{20} = 0.117375$$

b) $P(2W \text{ and } 1R) = P(W_1 W_2 R_3) + P(W_1 R_2 W_3) + P(R_1 W_2 W_3)$

$$= \frac{4}{20}\times\frac{5}{20}\times\frac{5}{20} + \frac{4}{20}\times\frac{7}{20}\times\frac{6}{20} + \frac{9}{20}\times\frac{5}{20}\times\frac{6}{20} = 0.06725$$

5. A box contains 5 red balls and 10 white balls. Two balls are drawn at random from the box without replacement.
What is the probability that
i) the second ball is white,
ii) the first ball drawn is red, given the second ball drawn is white? **[WBUT 2012]**

Answer:

Let R_1 denote the event of a red ball in the 1st drawing and W_1 and W_2 denote the event of a white ball in the 1st and 2nd drawings respectively.

Then the required probability

$$= P(W_2) = P(R_1 \,\&\, W_2) + P(W_1 \,\&\, W_2)$$

$$= P(R_1)P(W_2|R_1) + P(W_1)P(W_2|W_1)$$

$$= \frac{5}{15}\times\frac{10}{14} + \frac{10}{15}\times\frac{9}{14} = \frac{2}{3}.$$

Similarly, $P(R_1|W_2) = \dfrac{P(R_1 \,\&\, W_2)}{P(W_2)} = \dfrac{50/210}{2/3} = \dfrac{5}{14}.$

We ARE MAKAUTIANS

6. In a bolt factory, the machines M_1, M_2 and M_3 manufacture respectively 25%, 35% and 40% of the total product. Of their outputs 5%, 4% and 2% respectively are defective bolts. One bolt is drawn at random from the product and is found to be defective. What is the probability that it was manufactured by machine M_3?

[WBUT 2003, 2012, 2013]

Answer:

Let A_i denote the event that the bolt is produced by machine $M_i\ (i=1,2,3)$.

Then, $P(A_1)=\frac{25}{100}$, $P(A_2)=\frac{35}{100}$, $P(A_3)=\frac{40}{100}$

Further, if A denote the event of being defective, then

$P(A|A_1)=\frac{5}{100}$, $P(A|A_2)=\frac{4}{100}$, $P(A|A_3)=\frac{2}{100}$.

Now the required probability $=P(A_3|A)=\frac{p(A_3)\times p(A|A_3)}{\sum_{i=1}^{3}p(A_i).p(A|A_i)}=\frac{80}{80+125+140}=\frac{16}{69}$

7. In a bolt factory, machines A,B,C manufacture respectively 25%, 35% and 40%. Of the total of their output 5%, 4%, 2% are defective bolts. A bolt is drawn at random from the product and is found to be defective. What are the probability that it was manufactured by machines A,B and C? **[WBUT 2012]**

Answer:

Let D denote the event of a bolt being defective.

By Bayes' theorem,

Then $$P(A|D)=\frac{P(D|A)P(A)}{P(D|A)P(A)+P(D|B)P(B)+P(D|C)P(C)}$$

$$=\frac{0.05\times0.25}{0.05\times0.25-0.04\times0.35+0.02\times0.40}=\frac{0.0125}{0.0125+0..0140+0.0080}$$

$$=\frac{0.0125}{0.0345}=0.3623$$

Similarly, $P(B|D)$ & $P(C|D)$ can be calculated.

8. In a bolt factory, machines A,B,C manufacture respectively 20%, 30% and 50% of the total output. Of the output 3%, 4% and 5% are defective. One bolt is drawn at random from the product and found defective. What is the probability that it is manufactured by A?

[WBUT 2017]

We ARE MAKAUTIANS

Answer:

Here $P(A)=0.2, P(B)=0.3, P(C)=0.5$

$P(D|A)=0.03, P(D|B)=0.04, P(D|C)=0.05$

$\therefore$ The req. prob. $=P(A|D)$

$$=\frac{P(A)P(D|A)}{P(A)P(D|A)+P(B)P(D|B)+P(C)P(D|C)}$$

$$=\frac{0.2\times 0.03}{0.2\times 0.03+0.3\times 0.04+0.5\times 0.05}=0.139$$

9. Three identical boxes I, II and III contain respectively 4 white and 3 red balls, 3 white and 7 red balls, 2 white and 3 red balls. A box is chosen at random and a ball is drawn out of it. If the ball is found to be white, what is the probability that box II is selected? **[WBUT 2014, 2018]**

Answer:

Let A_i denote the event 'the ball is drawn from the ith box'(i = I, II, III). The events A_I, A_{II}, A_{III} are pairwise mutually exclusive and of these necessarily occurs.

Now, $P(A_I)=P(A_{II})=P(A_{III})=\frac{1}{3}$

Let, X denote the event 'White ball'.

Then, $P(X|A_I)=\frac{4}{7}$, $P(X|A_{II})=\frac{3}{10}$, $P(X|A_{III})=\frac{2}{5}$

$$\therefore P(X)=P(A_I)P(X|A_I)+P(A_{II})P(X|A_{II})+P(A_{III})P(X|A_{III})$$

$$=\frac{1}{3}\left(\frac{4}{7}+\frac{3}{10}+\frac{2}{5}\right)=\frac{89}{210}$$

$$\therefore P(A_{II}|X)=\frac{P(A_{II})P(X|A_{II})}{P(X)}=\frac{\frac{1}{3}\times\frac{3}{10}}{\frac{89}{210}}=\frac{21}{89}$$

Therefore the required probability is $\frac{21}{89}$.

10. At the beginning of each day, the three motors required for the operation of a machine are checked and tested. These motors then have probabilities 7/8, 5/6 and 2/3 of continuously operating throughout a working day. If the production can be maintained as long as two of the three motors are operating, calculate the probability that the machine will not fail in a particular day. Also find the probability of the machine failing on a particular day. **[WBUT 2014]**

We ARE MAKAUTIANS

Answer:

Let A_i denote the event that i^{th} motor is continuously operating through a working day. $(i = 1, 2, 3)$.

$\therefore P(A_1) = 7/8,\ P(A_2) = \frac{5}{6},\ P(A_3) = 2/3.$

Then the probability that at least two motors are continuously performing is

$$\frac{7}{8}\times\frac{5}{6}\times\frac{1}{3}+\frac{7}{8}\times\frac{1}{6}\times\frac{2}{3}+\frac{1}{8}\times\frac{5}{6}\times\frac{2}{3}+\frac{7}{8}\times\frac{5}{6}\times\frac{2}{3}=\frac{43}{48}$$

Hence the probability that the machine will not fail is a particular day is $\frac{43}{48}$.

$\therefore$ The probability that the machine will fail is a particular day is $1-\frac{43}{48}=\frac{5}{48}$.

We ARE MAKAUTIANS

PROBABILITY DISTRIBUTIONS & EXPECTATION

☞ Chapter at a Glance

Random Variable:

Let S be the event space of a random experiment E. A random variable X is a real-valued function of the events on S given by $X : S \to R$. The set of all values which X assumes, is called the spectrum of X.

Probability Distribution Function:

The distribution function of a random variable X (a function of real variable x) is denoted by $F(x)$ and defined in $(-\infty, \infty)$ by $F(x) = P(-\infty < X \le x)$.

Important Properties of Distribution Function:

(i) The distribution function is monotonic non-decreasing everywhere

i.e., $F(x_2) \ge F(x_1)$ for $x_2 \ge x_1$

(ii) $F(-\infty) = 0$; $F(\infty) = 1$

(iii) $P(a < x \le b) = F(b) - F(a)$

(iv) The distribution function $F(x)$ is continuous on the right at all points and it has a jump discontinuity on the left at $x = a$, provided $P(X = a) > 0$; the height of jump is equal to $P(X = a)$.

Probability mass function (*pmf*):

Let us consider a discrete random variable X having spectrum $X = x_i$, $(i = 0, \pm 1, \pm 2, \ldots)$ with probability $P(X = x_i) = f_i$.

Then f_i is known as *probability mass function* (*p.m.f*).

The probability mass function f_i satisfy the following conditions:

(i) $f_i \ge 0$, for all i

(ii) $\sum_{i=-\infty}^{\infty} f_i = 1$

Probability density function (*pdf*):

For a continuous random variable X,

WE ARE MAKAUTIANS

$$P(a < x \le b) = F(b) - F(a) = \int_a^b F'(x)\,dx$$

$$\Rightarrow P(a < x \le b) = \int_a^b f(x)\,dx \text{, where } F'(x) = f(x)$$

The function $f(x)$ is called *probability density function (p.d.f)* of random variable X.

The probability density function $f(x)$ satisfy the following conditions:

(i) $f(x) \ge 0$, for all x

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

Note: The distribution function $F(x) = P(-\infty < X \le x) = \int_{-\infty}^{x} f(x)\,dx$

Expectation or Mean of Random Variable:

The expectation of a r.v. X is denoted by $m = E(X)$ or m_x and defined by

$$m = E(X) = \begin{cases} \sum_{i=-\infty}^{\infty} x_i f_i & \text{for discrete case} \\ \int_{-\infty}^{\infty} x f(x)\,dx & \text{for continuous case} \end{cases}$$

Variance and Standard Deviation:

Variance of a r.v. X is denoted by $\text{var}(X)$ and defined by $\text{var}(X) = E\{(X-m)^2\}$

And positive square root of $\text{var}(X)$ is called *Standard Deviation* (S.D) of r.v. X is denoted by σ_x or σ.

Hence, $d = +\sqrt{\text{var}(X)}$ by $\sigma = +\sqrt{\text{var}(X)}$.

Moments:

1. The kth order moment of a r.v. X about a fixed point 'a' is $E\{(X-a)^k\}$.
2. The kth order moment of a r.v. X about origin is $\alpha_k = E(X^k)$. Hence, $\alpha_0 = 1$, $\alpha_1 = E(X)$ is the mean.
3. The kth order central moment of a r.v. X is k^{th} order moment about it's mean and given by $\mu_k = E\{(X-m)^k\}$. Hence, for any distribution $\mu_0 = 1$, $\mu_1 = 0$.

We ARE MAKAUTIANS

Again, $\mu_2 = E\left\{(X-m)^2\right\} = \text{var}(X)$; that is, 2nd order central moment is known as variance of r.v. X.

Some Important Properties of Expectation & Variance:

1. $E(a) = a$; a is constant
2. $E(aX) = aE(X)$
3. $E(aX+b) = aE(X)+b$
4. $\text{var}(a) = 0$; a is constant
5. $\text{var}(aX) = a^2\,\text{var}(X)$
6. $\text{var}(aX+b) = a^2\,\text{var}(X)$; a, b are constants
7. $\text{var}(X) = E(X^2) - \{E(X)\}^2$
8. $\text{var}(X) = E\{X(X-1)\} - m(m-1)$ where m is the mean.

Some distribution with p.d.f/p.m.f and their Mean, Variance:

Sl. No.	Distribution and Parameters	p.d.f /p.m.f	Mean and Variance
1	X is Binomial distribution with parameter n & p; that is symbolically, $X \sim B(n,p)$	$f_i = \binom{n}{i} p^i (1-p)^{n-i}$; $i = 0,1,\ldots,n$	Mean $= np$ Variance $= npq$
2	X is Poisson distribution with parameter λ, that is symbolically, $X \sim P(\lambda)$	$f_i = \dfrac{e^{-\lambda}\lambda^i}{i!}$, $i = 0, 1, 2, \cdots\infty$	Mean $= \lambda$ Variance $= \lambda$
3	X is distributed normally with parameters m and σ; that is symbolically, $X \sim N(m,\sigma)$	$f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$; $-\infty < x < \infty$	Mean $= m$ Variance $= \sigma^2$
4	X is Gamma distribution with parameter $l, (l>0)$ that is symbolically, $X \sim \Gamma(l)$	$f(x) = \begin{cases} \dfrac{e^{-x}x^{l-1}}{\Gamma(l)} & ; \quad 0 < x < \infty \\ 0 & ; \quad \text{elsewhere} \end{cases}$	Mean $= l$ Variance $= l^2$
5	X is Exponential distribution with parameter μ, $(\mu > 0)$	$f(x) = \begin{cases} \mu e^{-\mu x} & \text{if} \quad x \geq 0 \\ 0 & \text{if} \quad x < 0 \end{cases}$	Mean $= \dfrac{1}{\mu}$ Variance $= \dfrac{1}{\mu^2}$

We ARE MAKAUTIANS

Important Results:

1. Normal distribution with mean $m=0$ and standard deviation $\sigma=1$ is known as **Standard Normal distribution.**
2. **Standardised Variable:** If X be any random variable with mean m and standard deviation σ, then $Y=\dfrac{X-m}{\sigma}$ has mean $m=0$ and standard deviation $\sigma=1$ is called **Standardised Random Variable.**

Theorem: Let X be a continuous random variable with *p.d.f* $f_X(x)$. Let $y=g(x)$ be a continuously differentiable and strictly monotonic function of x everywhere, so that inverse function $x=g^{-1}(y)$ exist uniquely. Then the p.d.f of the random variable $Y=g(X)$ is given by

$$f_Y(y)=f_X(x)\left|\frac{dx}{dy}\right|$$

Multiple Choice Type Questions

1. A random variable X has the following probability density function:

$$f(x)=\begin{cases} kx, & 1\le x\le 2 \\ 0, & \text{otherwise} \end{cases}$$

[WBUT 2017]

The value of k is

a) 1 **b) 2** **c) 2/3** **d) none of these**

Answer: (c)

2. A random variable X has the following probability density function:

$$f(x)=\begin{cases} kx & 0\le x\le 1 \\ 0 & \text{otherwise} \end{cases}$$

[WBUT 2014]

The value of k is

a) 1 **b) 2** **c) 4** **d) none of these**

Answer: (b)

3. A random variable X has the following p.d.f **[WBUT 2012, 2015]**

$$f(x)=k \qquad -2<x<2$$
$$=0 \qquad \text{otherwise}$$

Then the value of the constant k is

a) $\frac{1}{8}$ **b)** $\frac{1}{2}$ **c)** $\frac{1}{4}$ **d)** $\frac{1}{12}$

Answer: (c)

We ARE MAKAUTIANS

4. The distribution function $F(X)$ of a random variable X is given by where $(-\infty < X \le \infty)$ [WBUT 2010]

a) $P(-\infty < X < \infty)$
b) $P(-\infty < X \le x)$
c) $P(-\infty \le X < \infty)$
d) none of these

Answer: (b)

5. The random variable X has distribution function F. If $F(0)=\frac{1}{6}$, $F(1)=\frac{1}{2}$, $F(3)=\frac{3}{4}$ then $P(0 < X \le 3)$ is [WBUT 2010]

a) $\frac{3}{4}$
b) $\frac{1}{3}$
c) $\frac{7}{12}$
d) none of these

Answer: (a)

6. The mean and standard deviation of standard normal variable are [WBUT 2017]

a) 0, –1
b) 1, 1
c) 0, 1
d) 1, 0

Answer: (c)

7. For a random variable X with mean 0, the value of $P(-4\sigma < X < 4\sigma)$ will be at least [WBUT 2016]

a) $\frac{1}{16}$
b) $\frac{1}{4}$
c) $\frac{15}{16}$
d) $\frac{13}{16}$

Answer: (c)

8. If X is normally distributed with zero mean and unit variable, then the expectation of X^2, is [WBUT 2010]

a) 1
b) 0
c) 8
d) 2

Answer: (a)

9. If the random variable X has Binomial distribution with parameters n and p, then the mean and variance are respectively [WBUT 2010]

a) np & $np(1-p)$
b) n/p & $2n(1-p)$
c) $np(1-p)$ & np
d) none of these

Answer: (a)

We ARE MAKAUTIANS

10. The distribution for which the mean and variance are equal is [WBUT 2013]
a) Poisson b) normal c) binomial d) exponential
Answer: (a)

11. The mean and variance of a distribution is given to be 10 and 6 respectively. Then the distribution is [WBUT 2014]
a) Standard Normal Distribution b) Binomial Distribution
c) Poisson Distribution d) None of these
Answer: (b)

12. For two random variables X and Y, $E(XY)=E(X)\cdot E(Y)$ hold if X and Y are [WBUT 2010]
a) independent b) uncorrelated
c) continuous variate d) none of these
Answer: (a)

13. Let X be a random variable and λ be a real number. Then $\frac{\text{var}(\lambda X)}{\text{var}(X)}$ is [WBUT 2016]
a) λ b) λ^2 c) $\frac{1}{\lambda}$ d) $\frac{1}{\lambda^2}$
Answer: (b)

14. The variance of a random variable X is [WBUT 2012, 2015]
a) $[E(X)]^2$ b) $E(X^2)$
c) $E(X^2)-[E(X)]^2$ d) $E(X^2)-[E(X)]$
Answer: (c)

15. If $\text{var}(X)=\frac{2}{3}$, then $\text{var}(3X+5)$ is [WBUT 2018]
a) 8 b) 2 c) 6 d) 11
Answer: (c)

16. In a Binomial (n, p) distribution, if its mean and variance are 2 and 4/3 respectively, then the values of n and p are [WBUT 2013]
a) $8,\frac{1}{4}$ b) $6,\frac{1}{3}$ c) $4,\frac{1}{2}$ d) none of these
Answer: (b)

We ARE MAKAUTIANS

17. If X be binomially distributed with mean 2 and variance $\frac{4}{3}$, then the value of n and p are respectively [WBUT 2018]

a) $5, \frac{1}{2}$ b) $7, \frac{1}{5}$ c) $6, \frac{1}{3}$ d) none of these

Answer: (c)

18. The mean of Binomial variate is [WBUT 2013]

a) np b) $np(1-p)$ c) $\sqrt{np}$ d) none of these

Answer: (a)

19. Let X be a Poisson Random Variateand $E(X)=\lambda$. Then $E\left[(X+1)^2\right]$ will be [WBUT 2015]

a) λ b) $\lambda^2+2\lambda$ c) $\lambda^2+2\lambda+1$ d) $\lambda^2+3\lambda+1$

Answer: (d)

20. For a Poisson distribution $P(X)$ is $P(1)=P(2)$, then $P(0)$ is [WBUT 2010]

a) $1/e$ b) $1/e^2$ c) $1/e^3$ d) none of these

Answer: (b)

21. In a Poisson distribution if $2P(x=1)=P(x=2)$, then the variance is [WBUT 2014]

a) 0 b) –1 c) 4 d) 2

Answer: (c)

22. If the exponential distribution is given by the probability density function $f(x)=e^{-x}$, $0<x<\infty$, then the mean of the distribution is [WBUT 2013]

a) 1 b) 3 c) $\frac{1}{3}$ d) none of these

Answer: (a)

23. A fair die is thrown 180 times. The expected number of sixes is [WBUT 2016]

a) 10 b) 20 c) 30 d) 40

Answer: (c)

We ARE MAKAUTIANS

Short Answer Type Questions

1. If R.V X denotes number on the face of a dice, write down the spectrum of X and corresponding probability distribution. [WBUT 2014]

Answer:

Spectrum of X is $\{1, 2, 3, 4, 5, 6\}$.

Corresponding Probability distribution is

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

2. Evaluate k, if $f(x) = \begin{cases} k; x = 1, 2, 3, 4, 5, 6 \\ 0; \quad \text{otherwise} \end{cases}$ is a p.m.f. Write down the corresponding probability distribution. [WBUT 2014]

Answer:

Since, $f(x)$ is a $p.m.f.$,

therefore, $\sum_{x_i=1}^{6} f(x_i) = 1$

$$\Rightarrow \quad k + k + k + k + k + k = 1$$

$$\Rightarrow \quad 6k = 1$$

$$\Rightarrow \quad k = 1/6$$

Corresponding Probability distribution is as follows:

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

3. From the discrete distribution find k: [WBUT 2011]

X:	0	1	2	3	4	5	6	7
$P(X)$:	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

Answer:

We have,

$$0 + k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1$$

or, $\quad 10k^2 + 9k - 1 = 0$

or, $\quad (10k - 1)(k + 1) = 0$

$\therefore \quad k = \frac{1}{10}, -1$

As $k = -1$ is inadmissible (Since, Probability is ≥ 0), the required value of $k = \frac{1}{10}$.

We ARE MAKAUTIANS

4. A random variable X has the following probability function: **[WBUT 2011]**

X:	0	1	2	3	4
$P(X=x)=f(x)$:	0	$5k$	$3k$	k	k

Determine the value of k.

Answer:

We have,

$$5k+3k+k+k=1$$

or, $10k=1$

$\therefore$ $k=\dfrac{1}{10}$.

5. A random variable X has the following probability mass function: **[WBUT 2011]**

x	0	1	2	3	4	5	6
$P(X=x)$	k	$3k$	$5k$	$7k$	$9k$	$11k$	$13k$

a) Find the value of k

b) Find $P(X<4), P(X\geq 5), P(3<X\leq 5)$

c) Obtain the distribution function $F(x)$

d) What is the smallest value of x for which $P(X\leq x)>0.5$?

Answer:

a) We have

$$k+3k+5k+7k+9k+11k+13k=1$$

or, $49k=1$

or, $k=\dfrac{1}{49}$

b) $P(X<4)=P(X=0)+P(X=1)+P(X=2)+P(X=3)$

$$=\frac{1}{49}+\frac{3}{49}+\frac{5}{49}+\frac{7}{49}=\frac{16}{49}$$

$$P(X\geq 5)=P(X=5)+P(X=6)=\frac{11}{49}+\frac{13}{49}=\frac{24}{49}$$

$$P(3<X\leq 5)=P(X=4)+P(X=5)=\frac{9}{49}+\frac{11}{49}=\frac{20}{49}$$

c) We present $F(x)$ by the following table

X:	0	1	2	3	4	5	6
$F(X)$:	1/49	4/49	9/49	16/49	25/49	36/49	1

We ARE MAKAUTIANS

d) Clearly $P(X \le 4) = \frac{25}{49} > 0.5$

Hence, $x = 4$.

6. A random variable X has the following probability function: **[WBUT 2012, 2013]**

X	0	1	2	3	4	5	6	7
$P(x)$	0	K	$2K$	$2K$	$3K$	K^2	$2K^2$	$7K^2+K$

Obtain the value of K and estimate $P(X<6)$ and $P(0<X<5)$.

Answer:

Here, the Random Variable is discrete and the $p.m.f$ satisfy that

$$\sum P(x) = 1$$

$$0 + K + 2K + 2K + 3K + K^2 + 2K^2 + 7K^2 + K = 1$$

$$\Rightarrow \quad 10K^2 + 9K - 1 = 0$$

$$\Rightarrow \quad (10K-1)(K+1) = 0$$

$$\therefore \quad K = \frac{1}{10}, -1$$

Since $P(x) \ge 0$

We get, $K = \frac{1}{10}$.

$$P(X<6) = 0 + K + 2K + 2K + 3K + K^2$$

$$= 8K + K^2 = 8 \cdot \frac{1}{10} + \frac{1}{100} = \frac{8}{10} + \frac{1}{100} = \frac{80+1}{100} = \frac{81}{100}$$

$$P(0<X<5) = 0 + K + 2K + 2K + 3K = 0 + 8K = 8 \cdot \frac{1}{10} = \frac{4}{5}$$

$$P(X<6) = \frac{81}{100}, \; P(0<X<5) = \frac{4}{5}$$

7. Examine whether the function $|x|$ in (–1, 1) and zero elsewhere is a density function. **[WBUT 2013]**

Answer:

The given function is

$$f(x) = \begin{cases} |x| & \text{for } x \in (-1,1) \\ 0 & \text{elsewhere} \end{cases}$$

Clearly $f(x) \ge 0$ on $\forall x$ and $\int_{-\infty}^{\infty} f(x)\,dx = \int_{-1}^{0} (-x)\,dx + \int_{0}^{1} x\,dx = 1$

Hence $f(x)$ is a density function as $f(x) \ge 0$ and $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

We ARE MAKAUTIANS

8. Show that a function $f(x)=\begin{cases}|x| & \text{for } x\in(-1,1)\\ 0 & \text{elsewhere}\end{cases}$ **is a probability density function. Find the corresponding distribution function.** **[WBUT 2018]**

Answer:

We see $f(x)\geq 0$

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

$$= \int_{-\infty}^{-1} 0dx + \int_{-1}^{1}|x|dx + \int_{1}^{\infty} 0dx = 2\int_{0}^{1} xdx = 2\left[\frac{x^2}{2}\right]_0^1 = 1$$

Hence, $f(x)$ is a *p.d.f.*

We know the distribution function $F(x)$ is defined by

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

$$=\begin{cases}0 & \text{if } x\leq -1\\ \int_{-1}^{x}|t|dt & \text{if } -1<x<1\\ 1 & \text{if } x\geq 1\end{cases} = \begin{cases}0 & \text{if } x\leq -1\\ \frac{x^2-1}{2} & \text{if } -1<x<0\\ \frac{x^2+1}{2} & \text{if } 0<x<1\\ 1 & \text{if } x\geq 1\end{cases}$$

9. A random variable X has the *p.d.f.* **[WBUT 2012]**

$f(x)=3x, 0<x<1$

$=0$, otherwise

Then find $P(2X+3>2)$

Answer:

$$P(2X+3>2)=P(2X>-1)=P\left(X>-\frac{1}{2}\right)=\int_{-\frac{1}{2}}^{\infty} f(x)dx = \int_0^1 3xdx = \left[\frac{3x^2}{2}\right]_0^1 = \frac{3}{2}$$

10. A random variable X has the following: **[WBUT 2011]**

p.d.f. $f(x)=cx^2, 0\leq x\leq 1$

= 0, otherwise

Find **(i)** c

(ii) $P\left(0<x<\frac{1}{2}\right)$

WE ARE MAKAUTIANS

Answer:

(i) Since $f(x)$ is a p.d.f., therefore,

$$\int_{-\infty}^{\infty} f(x)dx = 1 \Rightarrow \int_0^1 cx^2 dx = 1 \Rightarrow c\left[\frac{x^3}{3}\right]_0^1 = 1 \Rightarrow \frac{c}{3} = 1 \Rightarrow c = 3$$

(ii) $P\left(0 < x < \frac{1}{2}\right) = \int_0^{\frac{1}{2}} f(x)dx = 3\int_0^{\frac{1}{2}} x^2 dx = 3\left[\frac{x^3}{3}\right]_0^{\frac{1}{2}} = \frac{1}{8}$

11. Show that $f(x)$ given by **[WBUT 2010]**

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

is a probability density function for a suitable value of k.

Calculate the probability that the random variable lies between $\frac{1}{2}$ and $\frac{3}{2}$.

Answer:

Clearly, $f(x) \geq 0$ for all $x \in R$ if $k \geq 2$

Now, $f(x)$ to be a $p.d.f$, $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\Rightarrow \int_0^1 f(x)dx + \int_1^2 f(x)dx = 1$$

$$\Rightarrow \int_0^1 xdx + \int_1^2 (k - x)dx = 1$$

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

$$\Rightarrow \frac{1}{2} + \left(2k - 2 - k + \frac{1}{2}\right) = 1$$

$$\Rightarrow k = 2$$

Now, $P\left(\frac{1}{2} < X < \frac{3}{2}\right) = \int_{1/2}^1 xdx + \int_1^{3/2} (2 - x)dx$

$$= \left[\frac{x^2}{2}\right]_{1/2}^1 + \left[2x - \frac{x^2}{2}\right]_1^{3/2}$$

$$= \left(\frac{1}{2} - \frac{1}{8}\right) + \left(3 - \frac{9}{8} - 2 + \frac{1}{2}\right) = 2 - \frac{10}{8} = \frac{3}{4}$$

We ARE MAKAUTIANS

2. Show that $f(x) = x \qquad 0 < x < 1$ **[WBUT 2010]**

$= k - x, \quad 1 < x < 2$

$= 0$ **otherwise**

is a density function for a suitable value of k **. Calculate** $P(1/3 < x < 1/2)$**.**

Answer:

Clearly, $f(x) \geq 0$ for all $x \in R$ if $k \geq 2$

Further,

$\int_{-\infty}^{\infty} f(x)dx = 1$ as $f(x)$ is a p.d.f.

or, $\int_0^1 xdx + \int_1^2 (k-x)dx = 1$

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

or, $\frac{1}{2} + 2k - 2 - k + \frac{1}{2} = 1$

or, $k = 2$

$$P\left(\frac{1}{3} < x < \frac{1}{2}\right) = \int_{\frac{1}{3}}^{\frac{1}{2}} f(x)dx = \int_{\frac{1}{3}}^{\frac{1}{2}} xdx = \left[\frac{x^2}{2}\right]_{\frac{1}{3}}^{\frac{1}{2}} = \frac{1}{8} - \frac{1}{18} = \frac{9-4}{72} = \frac{5}{72}$$

13. If $f(x) = kx$ **,** $x = 1, 2, \ldots, n$ **be the p.m.f. of a random variable** x**, then find the value of** k **.** **[WBUT 2017]**

Answer:

Since $f(x)$ is a p.m.f, we get $\sum_{x=1}^{n} f(x) = 1$

i.e., $k(1 + 2 + \cdots + n) = 1 \qquad$ or, $k = \frac{2}{n(n+1)}$

14. Lifetime of an electric bulb is a random variable.with p.d.f $f(x) = kx^2$**, where** x **is measured in years. If the minimum and maximum lifetimes of bulb are 1 year and 2 years respectively, then find the value of** k **.** **[WBUT 2016]**

Answer:

Let X denote the lifetime of an electric bulb. The p.d.f. of X is $f(x) = kx^2$

We ARE MAKAUTIANS

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

$$\Rightarrow \int_{1}^{2} kx^2\,dx = 1$$

$$\Rightarrow k\left[\frac{x^3}{3}\right]_1^2 = 1$$

$$\Rightarrow k\left[\frac{8}{3} - \frac{1}{3}\right] = 1$$

$$\Rightarrow k = \frac{3}{7}$$

5. The probability density function (p.d.f.) of a random variable (r.v.) X is $f(x) = k(x-2)(3-x)$ for $2 \le x \le 3$, then find the value of the constant *k*.

[WBUT 2018]

Answer:

$$f(x) = k(x-2)(3-x)$$

Since the given function is p.d.f.

$$\therefore \quad \int_{-a}^{a} f(x)\,dx = 1$$

or $$\int_{2}^{3} k(x-2)(3-x)\,dx = 1$$

or $$k\int_{2}^{3} (x-2)(3-x)\,dx = 1$$

or $$k\left[-\frac{x^3}{3} + \frac{5x^2}{2} - 6x\right]_2^3 = 1$$

or, $$k\left[\frac{8}{3} - \frac{27}{3} + \frac{45}{2} - \frac{20}{2} - 18 + 12\right] = 1$$

or, $$k\left[\frac{-38+75-36}{6}\right] = 1$$

or, $$k\left[\frac{1}{6}\right] = 1 \quad \Rightarrow \quad k = 6 \quad \textbf{(Ans.)}$$

We ARE MAKAUTIANS

16. The length of the life of tyre manufactured by a company follows a continuous distribution given by the density function

$$f(x) = k/x^3,\ 1000 \le x < 1500$$
$$= 0, \text{ elsewhere}$$

Find k and find the probability that a randomly selected tyre would function for at least 1200 hours. **[WBUT 2013]**

Answer:

Since $f(x)$ is a density function, $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\therefore \int_{1000}^{1500} x^3 dx = 1 \quad \text{or} \quad k\left[\frac{x^{-2}}{-2}\right]_{100}^{1500} = 1$$

or, $k = 36 \times 10^5$

$$\text{Now } P(X \ge 1200) = \int_{1200}^{1500} \frac{k}{x^3} dx = k\left(\frac{1}{1200^2} - \frac{1}{1500^2}\right) = \frac{9}{20}$$

17. The probability density of a random variable Z is given by **[WBUT 2014]**

$$f(z) = \begin{cases} kze^{-z^2}, & \text{for } z > 0 \\ 0 & \text{for } z \le 0 \end{cases}$$

Find the value of k and find out the corresponding distribution function of Z.

Answer:

As $f(z)$ is a $p.d.f.$, we have $f(z) \ge 0$ and $\int_{-\infty}^{\infty} f(z)dz = 1$

Thus $\int_0^{\infty} kze^{-z^2} dz = 1$

Putting $t = z^2$, $2z\,dz = dt$, when $z = 0, t = 0$ and when $z \to \infty, t \to \infty$.

$$\therefore \int_0^{\infty} e^{-t} dt = \frac{2}{k}$$

$$\Rightarrow \left[-e^{-t}\right]_0^{\infty} = \frac{2}{k}$$

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

$$\Rightarrow k = 2$$

We ARE MAKAUTIANS

The distribution function $F(z)$ is given by

$$F(z)=\int_{-\infty}^{z} f(x)dx$$

$$=\begin{cases}\int_{0}^{\infty} 2xe^{-x^2} & \text{if} \quad z>0\\ 0 & \text{if} \quad z\le 0\end{cases}$$

$$=\begin{cases}\left[-e^{-t}\right]_0^z; & z>0\\ 0 & z\le 0\end{cases}$$

$$=\begin{cases}1-e^{-z} & \text{if} \quad z>0\\ 0 & \text{if} \quad z\le 0\end{cases}$$

18. In a certain city, the daily consumption of electric power (in millions of kilowatt hours) is a random variable having the probability density

$$f(x)=\frac{1}{9}xe^{-x/3}, x>0$$
$$=0, \qquad x\le 0$$

If the city's power plant has a daily capacity of 12 million kilowatt-hours, what is the probability that this power supply will be inadequate on any given day?

[WBUT 2014]

Answer:

The required probability is (Since daily capacity of a city is 12 million kilowatt-hour)

$$=P(X>12)=\int_{12}^{\infty} f(x)dx=\frac{1}{9}\int_{12}^{\infty} xe^{-\frac{x}{3}}dx$$

19. Find the mathematical expectation of the number of points obtained in a single throw of an unbiased die.

[WBUT 2013]

Answer:

Let X denote the number of points obtained in a single throw of an unbiased die and the associated probability distribution is given as follows:

$X(=x_i)$	1	2	3	4	5	6
$P(X=x_i)=f_i$	1/6	1/6	1/6	1/6	1/6	1/6

Hence, the expectation is $E(X)=1\times\frac{1}{6}+2\times\frac{1}{6}+3\times\frac{1}{6}+4\times\frac{1}{6}+5\times\frac{1}{6}+6\times\frac{1}{6}=3\frac{1}{2}$

WE ARE MAKAUTIANS

20. If a random variable X has Binomial distribution with parameter n and p, find the mean and variance. **[WBUT 2005]**

Answer:

The $p.m.f$ is $f_i = P(X=i) = \binom{n}{i} p^i (1-p)^{n-i}$; $i = 0, 1, \ldots, n$

Mean is

$$= E(X) = \sum_{i=0}^{n} i f_i = \sum_{i=0}^{n} i \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=0}^{n} i \cdot \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i}$$

$$= np \cdot \sum_{i=1}^{n} \frac{(n-1)!}{(i-1)!\{(n-1)-(i-1)\}!} p^{i-1} (1-p)^{(n-1)-(i-1)}$$

$$= np \cdot \sum_{i=1}^{n} \binom{n-1}{i-1} p^{i-1} (1-p)^{(n-1)-(i-1)}$$

$$= np.\left[p + (1-p)\right]^{n-1} = np$$

Again,

$$E(X^2) = \sum_{i=0}^{n} i^2 f_i = \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=0}^{n} \{i(i-1)+i\} \cdot \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i}$$

$$= n(n-1)p^2 \cdot \sum_{i=2}^{n} \frac{(n-2)!}{(i-2)!\{(n-2)-(i-2)\}!} p^{i-2} (1-p)^{(n-2)-(i-2)} + np$$

$$= n(n-1)p^2 \cdot \left[p + (1-p)\right]^{n-2} + np$$

$$= n(n-1)p^2 + np$$

Therefore,

$$\text{var}(X) = E(X^2) - \{E(X)\}^2 = n(n-1)p^2 + np - (np)^2 = np(1-p) = npq$$

21. Define Poisson distribution and find its mean and variance. **[WBUT 2013]**

OR,

Find the mean and variance of Poisson distribution with parameter λ. **[WBUT 2012]**

Answer:

The Poisson distribution with parameter λ is defined by the probability mass function $f_i = P(X = x_i)$ is given by

$$f_i = \frac{e^{-\lambda}\lambda^i}{i!}, \quad i = 0, 1, 2, \ldots\ldots, n, n+1, \ldots\ldots \text{to } \infty.$$

The constant λ is called the parameter of the distribution.

We ARE MAKAUTIANS

Let X be a Poisson variate, i.e., $X \sim P(\lambda)$

Then Mean $= E(X) = \sum_{i=0}^{\infty} i.\frac{e^{-\lambda}\lambda^i}{i!}$

$$= e^{-\lambda}\left\{\frac{1\cdot\lambda}{1!}+\frac{2\cdot\lambda^2}{2!}+\frac{3\cdot\lambda^3}{3!}+\cdots\cdots\right\} = \lambda e^{-\lambda}\left\{1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\cdots\cdots\right\}$$

$$= \lambda e^{-\lambda}\cdot e^{\lambda} = \lambda.$$

Again, $E(X^2) = \sum_{i=0}^{\infty} i^2 \frac{e^{-\lambda}\lambda^i}{i!}$

$$= e^{-\lambda}\sum_{i=0}^{\infty}\{i(i-1)+i\}.\frac{\lambda^i}{i!}$$

$$= e^{-\lambda}\sum_{i=2}^{\infty}\frac{\lambda^i}{(i-2)!} + e^{-\lambda}\sum_{i=1}^{\infty}\frac{\lambda^i}{(i-1)!} = e^{-\lambda}\cdot\lambda^2\cdot e^{\lambda} + e^{-\lambda}\cdot\lambda\cdot e^{\lambda}$$

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

Hence, Variance is $\text{var}(X) = E(X^2) - \{E(X)\}^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

22. Define Normal distribution and find its mean, variance and standard deviation.

[WBUT 2015]

Answer:

The probability density function of Normal distribution is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}; \; -\infty < x < \infty$$

where, m and $\sigma(>0)$ are two parameters and denoted as $X \sim N(m, \sigma)$.

Mean $E(X) = \int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}}dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-m+m)e^{-\frac{(x-m)^2}{2\sigma^2}}dx$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-m)e^{-\frac{(x-m)^2}{2\sigma^2}}dx + \frac{m}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}}dx$$

[Let, $u = \frac{(x-m)^2}{2\sigma^2} \Rightarrow du = \frac{x-m}{\sigma^2}dx \Rightarrow \sigma^2 du = (x-m)dx$]

$$= \frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-u}du + \frac{m}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}}dx$$

We ARE MAKAUTIANS

$$= 0 + m \times \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx \qquad \left[\text{Since, } \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = 1\right]$$

$$= 0 + m \times 1 = m$$

Thus, $\boxed{\text{Mean } E(X) = m}$

Now, Variance

$$\text{var}(X) = E(x-m)^2 = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-m)^2 e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} v^2 e^{-\frac{v^2}{2}} dv \qquad \left[\text{Taking } v = \frac{x-m}{\sigma}\right]$$

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} v^2 e^{-\frac{v^2}{2}} dv = \frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} v^2 e^{-\frac{v^2}{2}} dv \quad \left[\text{Taking } t = \frac{v^2}{2}\right]$$

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} 2te^{-t} \frac{dt}{\sqrt{2t}}$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \int_{0}^{\infty} \sqrt{t}e^{-t} dt = \frac{2\sigma^2}{\sqrt{\pi}} . \Gamma\left(\frac{3}{2}\right) = \frac{2\sigma^2}{\sqrt{\pi}} . \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{2\sigma^2}{\sqrt{\pi}} . \frac{1}{2}\sqrt{\pi} = \sigma^2$$

Thus, Variance $= \sigma^2$

And hence Standard Deviation $= +\sqrt{\text{var}(X)} = \sigma$

23. Determine the Mean, Variance and Standard Deviation of Gamma distribution.
[WBUT 2017]

Answer:

The *p.d.f.* of Gamma distribution with parameter $l(>0)$ is

$$f(x) = \begin{cases} \dfrac{e^{-x}x^{l-1}}{\Gamma(l)} & ; \quad 0 < x < \infty \\ 0 & ; \quad \text{elsewhere} \end{cases}, \qquad l > 0$$

$$\text{Now, Mean} = E(X) = \int_0^{\infty} x . \frac{e^{-x}x^{l-1}}{\Gamma(l)} dx = \frac{1}{\Gamma(l)} \int_0^{\infty} e^{-x} x^{(l+1)-1} dx = \frac{\Gamma(l+1)}{\Gamma(l)} = l$$

$$\text{Again, } E(X^2) = \int_0^{\infty} x^2 . \frac{e^{-x}x^{l-1}}{\Gamma(l)} dx = \frac{1}{\Gamma(l)} \int_0^{\infty} e^{-x} x^{(l+2)-1} dx = \frac{\Gamma(l+2)}{\Gamma(l)} = l(l+1)$$

Hence, $\text{var}(X) = E(X^2) - \{E(X)\}^2 = l(l+1) - l = l^2$.

And hence Standard Deviation $= +\sqrt{\text{var}(X)} = l$

We ARE MAKAUTIANS

24. Determine the mean and variance of exponential distribution. **[WBUT 2013, 2017]**

Answer:

The $p.d.f$ of exponential distribution is

$$f(x)=\begin{cases}\mu e^{-\mu x} & \text{if} \quad x\geq 0\\ 0 & \text{if} \quad x<0\end{cases}, \qquad \mu>0$$

Now $\quad \text{Mean}=E(X)=\int_0^\infty x\mu e^{-\mu x}dx$

Putting $\mu x=y,\ \mu dx=dy$

$$\therefore \qquad E(X)=\int_0^\infty ye^{-y}\frac{dy}{\mu}=\frac{1}{\mu}\int_0^\infty yd\left(-e^{-y}\right)=\frac{1}{\mu}\left[-ye^{-y}\right]_0^\infty=\frac{1}{\mu}$$

Similarly, $E\left(X^2\right)=\int_0^\infty x^2\mu e^{-\mu x}dx=\frac{2}{\mu^2}$

Hence, $\text{var}(X)=E\left(X^2\right)-\{E(X)\}^2=\frac{2}{\mu^2}-\frac{1}{\mu^2}=\frac{1}{\mu^2}$.

25. Show that for the exponential distribution **[WBUT 2017]**

$$f(x)=\frac{1}{\sigma}e^{\left(-\frac{x}{\sigma}\right)},\ 0<x<\infty$$

mean and standard deviation both equal to σ.

Answer: ***Refer to Question No. 24 of Short Answer Type Questions.***

26. The mean and standard deviation of a binomial distribution are respectively 4 and $\sqrt{\frac{8}{3}}$. Find the value of n and p. Hence, evaluate $P(X=0)$. **[WBUT 2006]**

Answer:

The mean and standard deviation of a binomial distribution with parameter n and p are np and $\sqrt{npq}$.

Therefore, mean $np=4$ .... (1)

And s.d. $\sqrt{npq}=\sqrt{\frac{8}{3}}$ .... (2)

From (1) and (2), we have $npq=\frac{8}{3}\Rightarrow np(1-p)=\frac{8}{3}$.

We ARE MAKAUTIANS

$$\Rightarrow 4(1-p)=\frac{8}{3} \text{ [using (1)]}$$

$$\Rightarrow p=\frac{1}{3}$$

Hence, from (1), $np=4 \Rightarrow n=4\times 3=12$

2nd Part: $P(X=0)=f_0=n_{C_0}p^0(1-p)^{n-0}=12_{C_0}\left(\frac{1}{3}\right)^0\left(1-\frac{1}{3}\right)^{12-0}=\left(\frac{2}{3}\right)^{12}$

27. If the mean of a binomial distribution is 3 and the variance is $\frac{3}{2}$, find the probability of obtaining at most 3 success. **[WBUT 2007]**

Answer:

The mean and variance of a binomial distribution with parameter n and p are np and npq respectively.

Therefore, mean $np=3$ .... (1)

and variance $npq=\frac{3}{2}$ .... (2)

From (1) and (2), we have $\frac{npq}{np}=\frac{3}{2}\times\frac{1}{3}$

$$\Rightarrow 1-p=\frac{1}{2} \qquad \Rightarrow p=\frac{1}{2}$$

Hence, from (1) $np=3 \Rightarrow n=3\times 2=6$

Now, the probability of obtaining at most 3 success is $=P(X\le 3)$

$$=P(X=0)+P(X=1)+P(X=2)+P(X=3)$$

$$=f_0+f_1+f_2+f_3$$

$$=n_{C_0}p^0(1-p)^{n-0}+n_{C_1}p^1(1-p)^{n-1}+n_{C_2}p^2(1-p)^{n-2}+n_{C_3}p^3(1-p)^{n-3}$$

$$=6_{C_0}\left(\frac{1}{2}\right)^0\left(1-\frac{1}{2}\right)^{6-0}+6_{C_1}\left(\frac{1}{2}\right)^1\left(1-\frac{1}{2}\right)^{6-1}+6_{C_2}\left(\frac{1}{2}\right)^2\left(1-\frac{1}{2}\right)^{6-2}+6_{C_3}\left(\frac{1}{2}\right)^3\left(1-\frac{1}{2}\right)^{6-3}$$

$$=\left(\frac{1}{2}\right)^6+6\left(\frac{1}{2}\right)^6+15\left(\frac{1}{2}\right)^6+20\left(\frac{1}{2}\right)^6=\frac{42}{64}=\frac{21}{32}$$

We ARE MAKAUTIANS

28. A point X is chosen at random on a line segment AB of length $2a$ with mid-point O. Find the probability that AX, BX and AO can form the sides of a triangle. **[WBUT 2015]**

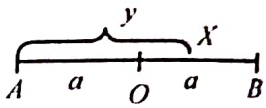
OR,

P is a point on the straight line AB of length $2a$. If O be the mid-point of AB, find the probability that AO, AP, PB form the sides of a triangle. **[WBUT 2017]**

Answer:

Let $AB = 2a$ and X be any point taken at random on the line segment AB. Let $AX = y$, and y is uniformly distributed and hence the $p.d.f$ is given by $f(y) = \begin{cases} \frac{1}{2a}, & 0 < y < 2a \\ 0, & \text{elsewhere} \end{cases}$

Now, AX, BX and AO will form the sides of a triangle if



$$\left.\begin{array}{l} AX + BX > AO \Rightarrow y + (2a - y) > a \Rightarrow 2a > a \\ BX + AO > AX \Rightarrow 2a - y + a > y \Rightarrow 2y < 3a \Rightarrow y < \frac{3a}{2} \\ AO + AX > BX \Rightarrow a + y > 2a - y \Rightarrow 2y > a \Rightarrow y > \frac{a}{2} \end{array}\right\} \Rightarrow \frac{a}{2} < y < \frac{3a}{2}$$

Therefore, the required probability $= \int_{\frac{a}{2}}^{\frac{3a}{2}} f(y)dy = \int_{\frac{a}{2}}^{\frac{3a}{2}} \frac{1}{2a} dy = \frac{1}{2a}[y]_{\frac{a}{2}}^{\frac{3a}{2}} = \frac{1}{2a}\left[\frac{3a}{2} - \frac{a}{2}\right] = \frac{1}{2}$

29. A random variable X has the density function $f(x) = x$, $0 \le x \le 1$, $f(x) = \frac{1}{2}$, $1 < x \le 2$. Find the mean of X. **[WBUT 2011, 2015]**

Answer:

Mean of $X = \bar{X} = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x^2 dx + \frac{1}{2}\int_1^2 x dx = \frac{1}{3} + \frac{3}{4} = \frac{13}{12}$.

30. If X is normally distributed with zero mean and unit variance, Find $E(X^2)$. **[WBUT 2012]**

Answer:

Since, X is normally distributed with zero mean and unit variance; So, $E(X) = 0$; $\text{var}(X) = 1$.

$\text{var}(X) = 1 \Rightarrow E(X^2) - \{E(X)\}^2 = 1$

We ARE MAKAUTIANS

$$\Rightarrow E(X^2)-0=1$$
$$\Rightarrow E(X^2)=1$$

1. Prove that $\operatorname{var}(aX+b)=a^2\operatorname{var}(X)$, a, b are constants. **[WBUT 2016]**

nswer:

et, $Y=aX+b$, a, b are constants and let m be the mean of Y.

Therefore, $m=E(Y)=aE(X)+b$

$$\operatorname{var}(Y)=E\{(Y-m)^2\}=E\{(aX+b-aE(X)-b)^2\}=E\{(aX-aE(X))^2\}$$
$$=a^2E\{(X-E(X))^2\}=a^2\operatorname{var}(X)$$

32. The mean and standard deviation of a Binomial distribution are 4 and $\sqrt{3}/3$ respectively. Find n and p. **[WBUT 2012]**

Answer:

The mean and standard deviation of a binomial distribution with parameter n and p are np and $\sqrt{npq}$.

Therefore, mean $np=4$ .... (1)

and s.d. $\sqrt{npq}=\frac{1}{\sqrt{3}}$ .... (2)

From (1) and (2), we have $npq=\frac{1}{3}$

$$\Rightarrow np(1-p)=\frac{1}{3}$$
$$\Rightarrow 4(1-p)=\frac{1}{3} \quad \text{[using (1)]}$$
$$\Rightarrow p=\frac{11}{12}$$

Hence, from (1) $np=4 \Rightarrow n=4\times\frac{12}{11}=12$

33. If a Poisson variate is such that $P(X=1)=P(X=2)$, then find $P(X=4)$. **[WBUT 2013]**

We ARE MAKAUTIANS

Answer:

We see $P(X=1)=\dfrac{e^{-\lambda}\lambda^1}{1!}=e^{-\lambda}\lambda$

$P(X=2)=\dfrac{e^{-\lambda}\lambda^2}{2!}=\dfrac{e^{-\lambda}\lambda^2}{2}$

As $P(X=1)=P(X=2)$, we get $e^{-\lambda}.\lambda=\dfrac{e^{-\lambda}.\lambda^2}{2}$ $\therefore \lambda=2$

Now $P(X=4)=\dfrac{e^{-2}2^4}{4!}=\dfrac{2}{3}e^{-2}$

34. A car hire firm has two cars, which it hires out day by day. The No. of demands for a car on each day is distributed as a Poisson distribution with mean 1.5. Calculate the proportion of days on which (i) neither car is used, (ii) some demands are refused. [WBUT 2013]

Answer:

Here $\lambda=1.5$,

Now $P(X=0)=\dfrac{e^{-1.5}.(1.5)^0}{0!}=e^{-1.5}=0.2231$

$P(X>2)=1-P(X\le 2)=1-P(X=0)-P(X=1)-P(X=2)$

$$=1-e^{-1.5}\left(1+\frac{3}{2}+\frac{9}{8}\right)=0.19$$

35. If X has the standard normal variate, then find the probability density function of $Z=X^2$. [WBUT 2016

Answer:

The p.d.f. of random variable X is $f_X(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$; $-\infty<x<\infty$

Now, $z=x^2 \Rightarrow \dfrac{dz}{dx}=2x \Rightarrow \dfrac{dx}{dz}=\dfrac{1}{2x}=\dfrac{1}{2\sqrt{z}}$

Again, as $-\infty<x<\infty \Rightarrow 0<z<\infty$

We know that $f_Z(z)=f_X(x)\left|\dfrac{dx}{dz}\right|$

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

We ARE MAKAUTIANS

Thus, the probability density function of $Z = X^2$ is given by

$$f_Z(z) = \frac{1}{2\sqrt{2\pi z}} e^{-\frac{z}{2}};\ 0 < z < \infty$$

36. Show that the variance of Binomial distribution cannot exceed $\frac{n}{4}$ for n independent trials.

[WBUT 2016]

Answer:

We know, if the random variable X having Binomial distribution with parameters n and p then

$$\text{var}(X) = npq = np(1-p) = -n(p^2 - p)$$

$$= -n\left\{\left(p - \frac{1}{2}\right)^2 - \frac{1}{4}\right\} = n\left\{\frac{1}{4} - \left(p - \frac{1}{2}\right)^2\right\} \le \frac{n}{4}$$

37. A continuous random variable X has the pdf $f(x) = \frac{1}{2} - ax$, $0 \le x \le 4$, a is a constant. Determine a and the probability that X lies between 2 and 3.

[WBUT 2017]

Answer:

Since $f(x)$ is a p.d.f., we get

$$\int_{-\infty}^{\infty} f(x)dx = 1 \quad \Rightarrow \int_0^4 \left(\frac{1}{2} - ax\right)dx = 1 \Rightarrow \left[\frac{x}{2} - a\frac{x^2}{2}\right]_0^4 = 1 \quad \Rightarrow 2 - 8a = 1 \quad \Rightarrow a = \frac{1}{8}$$

And the required probability is

$$P(2 < X < 3) = \int_2^3 \left(\frac{1}{2} - \frac{x}{8}\right)dx = \left[\frac{x}{2} - \frac{x^2}{16}\right]_2^3 = \frac{3}{2} - \frac{9}{16} - 1 + \frac{1}{4} = \frac{3}{16}$$

38. The incidence of occupational disease in an industry is such that workers have a 10% chance of suffering from it. What is the chance that out of 6 workers, at least 3 will suffer from the disease?

[WBUT 2017]

Answer:

Let X be the random variable denotes "the number of workers suffering from occupational disease". It follows binomial distribution with parameters n, p.

Here, $p = 0.1$, $n = 6$

Therefore, $P(X \ge 3) = 1 - P(X < 3) = 1 - [P(X=0) + P(X=1) + P(X=2)]$

We ARE MAKAUTIANS

$$=1-\left[\binom{6}{0}(0.1)^0(0.9)^{6-0}+\binom{6}{1}(0.1)^1(0.9)^{6-1}+\binom{6}{2}(0.1)^2(0.9)^{6-2}\right]$$

$$=1-\left[(0.9)^6+6\times0.1\times(0.9)^5+15\times(0.1)^2\times(0.9)^4\right]$$

$$=1-[0.5314+0.3543+0.0984]=1-0.9841=0.0159$$

39. If 3% of the bolts manufactured by a company are defective, find the probability that in a sample of 200 bolts (a) 5 will be defective, (b) none is defective. $\left[e^{-6}=0.00248\right]$ [WBUT 2017]

Answer:

Let X denotes "defective bolts" and p denote the probability of defective bolt.

Here, $p=0.03, n=200$

Therefore, $\mu=np=200\times0.03=6$

(a) $P(X=5)=\dfrac{e^{-6}6^5}{5!}=\dfrac{0.00248\times7776}{120}=0.1607$

(b) $P(X=0)=\dfrac{e^{-6}6^0}{0!}=\dfrac{0.00248\times1}{1}=0.00248$

40. If X is normally distributed with mean 3 and standard deviation 2, find c such that $P(X>c)=2P(X\le c)$. Given that $\int_{-\infty}^{0.43}\phi(t)dt=0.6666$. [WBUT 2018]

Answer:

Given that, $\mu=3$, $\sigma=2$.

So, $X\sim N(\mu,\sigma^2)=N(3,4)$.

Therefore, $Z=\dfrac{X-3}{4}\sim N(0,1)$

Also, $P(X\le c)=2P(X\le c)$ gives

$\Rightarrow 1-P(X\le c)=2P(X\le c)$

$\Rightarrow P(X\le c)=\dfrac{1}{3}$

$\Rightarrow P\left(\dfrac{X-3}{2}\le\dfrac{c-3}{2}\right)=\dfrac{1}{3}$

$\Rightarrow P\left(Z\le\dfrac{c-3}{2}\right)=0.3333$

We ARE MAKAUTIANS

Again from the given relation, $\int_{-\infty}^{-0.43} \phi(t)\,dt = 0.6666 \Rightarrow \int_{-\infty}^{-0.43} \phi(t)\,dt = 0.3333$

Hence, we get $P\left(Z \le \frac{c-3}{2}\right) = \int_{-\infty}^{-0.43} \phi(t)\,dt$

$\Rightarrow \frac{c-3}{2} = -0.43$

$\Rightarrow c = 3 - 0.86 = 2.14$

Long Answer Type Questions

1. The probability density function of a random variable X is **[WBUT 2010]**

$f(x) = K(x-1)(2-x)$, **for** $1 \le x \le 2$.

= 0, otherwise.

Determine (i) the value of the constant k and (ii) $P\left(\frac{5}{4} \le X \le \frac{3}{2}\right)$

Answer:

(i)Since $f(x)$ is a Probability density function.

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

or, $-K\int_{1}^{2}\left[x^2 - 3x + 2\right]dx = 1$ or, $-K\left[\frac{x^3}{3} - 3\frac{x^2}{2} + 2x\right]_1^2 = 1$

or, $-K\left[\frac{8}{3} - \frac{1}{3} - 6 + \frac{3}{2} + 4 - 2\right] = 1$ or, $-K\left[\frac{7}{3} - 4 + \frac{3}{2}\right] = 1$

or, $-K\left[\frac{14 - 26 + 9}{6}\right] = 1$ or, $\frac{K}{6} = 1$ or, $K = 6$

(ii) $\therefore P\left(\frac{5}{4} \le x \le \frac{3}{2}\right) = 6\int_{5/4}^{3/2}(x-1)(2-x)dx = -6\int_{5/4}^{3/2}(x^2 - 3x + 2)dx$

$= -6\left[\frac{x^3}{3} - 3\frac{x^2}{2} + 2x\right]_{5/4}^{3/2} = -6\left[\frac{9}{8} - \frac{27}{8} + 3 - \frac{125}{192} + \frac{75}{32} - \frac{5}{2}\right]$

$= -6\left[\frac{3}{32} - \frac{125}{192} + \frac{1}{2}\right] = -6\left[\frac{18 - 125 + 96}{192}\right] = \frac{11}{32}$

WeARE MAKAUTIANS

2. Show that **[WBUT 2010]**

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

is a *p.d.f* of a random variable X for a suitable value k. Find the value of k. Then find the distribution function of the random variable X. Calculate the probability that the random variable lie between $\frac{1}{2}$ and $\frac{3}{2}$.

Answer:

Since, $f(x)$ is a density function, we get

$$\int_{-\infty}^{\infty} f(x)dx=1 \Rightarrow \int_0^1 xdx+\int_0^1 (k-x)dx=1$$

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

$$\Rightarrow \frac{1}{2}+2k-2-k+\frac{1}{2}=1$$

$$\Rightarrow k=2$$

Thus, $f(x)\ge 0$, for all $x\in\Re$ and is a *p.d.f* if $k\ge 2$.

Distribution Function:

For, $-\infty<x<0$, $F(x)=\int_{-\infty}^{x} f(x)dx=0$

For, $0\le x<1$, $F(x)=\int_{-\infty}^{0} f(x)dx+\int_0^x f(x)dx=0+\int_0^x xdx=\frac{x^2}{2}$

For, $1\le x<2$, $F(x)=\int_{-\infty}^{0} f(x)dx+\int_0^1 f(x)dx+\int_1^x f(x)dx$

$$=0+\int_0^1 x\,dx+\int_1^x (2-x)\,dx$$

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

For, $2\le x<\infty$, $F(x)=\int_{-\infty}^{0} f(x)dx+\int_0^1 f(x)dx+\int_1^2 f(x)dx+\int_2^x f(x)dx$

$$=0+\int_0^1 x\,dx+\int_1^2 (2-x)\,dx+\int_2^x f(x)\,dx$$

We ARE MAKAUTIANS

$$=\frac{1}{2}+\left(4-2-2+\frac{1}{2}\right)+0=1$$

Thus, the distribution function is

$$F(x)=\begin{cases} 0 & ; -\infty<x<0 \\ \frac{x^2}{2} & ; 0\le x<1 \\ -1+2x-\frac{x^2}{2} & ; 1\le x<2 \\ 1 & ; 2\le x<\infty \end{cases}$$

Now, $P\left(\frac{1}{2}<X<\frac{3}{2}\right)=\int_{\frac{1}{2}}^{\frac{3}{2}} f(x)\,dx=\int_{\frac{1}{2}}^{1} x\,dx+\int_{1}^{\frac{3}{2}}(2-x)\,dx=\left[\frac{x^2}{2}\right]_{\frac{1}{2}}^{1}+\left[2x-\frac{x^2}{2}\right]_{1}^{\frac{3}{2}}$

$$=\frac{1}{2}-\frac{1}{8}+3-\frac{9}{8}-2+\frac{1}{2}=\frac{3}{4}$$

3. The p.d.f. of a random variable X is $f(x)=k(x-1)(2-x)$, for $1\le x\le 2$. Determine (i) the value of constant k (ii) the distribution function $F(x)$ and (iii) $P\left(\frac{5}{4}\le X\le\frac{3}{2}\right)$.

[WBUT 2014, 2015, 2018]

Answer:

i) Since the given function is a $p.d.f$

$\therefore \int_{-\alpha}^{\alpha} f(x)dx=1$

or, $\int_1^2 k(x-1)(2-x)\,dx=1$

or, $k\int_1^2(-x^2+3x-2)\,dx=1$

or, $k\left[-\frac{x^3}{3}+3\frac{x^2}{2}-2x\right]_1^2=1$

or, $k\left[-\frac{1}{3}(8-1)+\frac{3}{2}(4-1)-2(2-1)\right]=1$

or, $k\left[\frac{7}{3}+\frac{9}{2}-2\right]=1$

or, $k\left[\frac{-14+27-12}{6}\right]=1$

or, $k=6$

WE ARE MAKAUTIANS

ii) Distribution Function:

$$F(x)=\int_{-\alpha}^{x} f(x)\,dx$$

$$\text{or, } F(x)=\begin{cases} =0, & -\infty<x<1 \\ =\int_1^x f(x)\,dx=6\left[-\frac{x^3}{3}+\frac{3x^2}{2}-2x\right]_1^x \\ \quad =6\left[\frac{1}{3}-\frac{x^3}{3}+\frac{3x^2}{2}-\frac{3}{2}-2x+2\right] \\ \quad =2-2x^3+9x^2-9-12x+12 \\ \quad =5-12x+9x^2-2x^3; & 1\le x\le 2 \\ =\int_{-\alpha}^{x} f(x)dx=\int_1^2 f(x)dx+\int_2^{\alpha} f(x)dx \\ \quad =1 & x>2 \end{cases}$$

iii) $P\left(\frac{5}{4}\le x\le\frac{3}{2}\right)=F\left(\frac{3}{2}\right)-F\left(\frac{5}{4}\right)$

$$=-\frac{12\times 2}{8}+9\times\frac{11}{16}-2\times\left(\frac{27}{8}-\frac{125}{64}\right)$$

$$=-3+\frac{99}{16}-2\times\left(\frac{216-125}{64}\right)$$

$$=-3+\frac{99}{16}-\cancel{2}\times\frac{91}{\cancel{64}\,32}$$

$$=-3+\frac{107}{32}=\frac{11}{32}$$

4. If $F(x)$ be the distribution function of a random variable X, then prove that $F(-\infty)=0$ and $F(\infty)=1$. **[WBUT 2016]**

Answer:

1st Part:

Let A_n be the event $(-\infty<X\le -n)$; $n=1,2,\ldots\ldots\infty$

Thus, $\{A_n\}$ be a decreasing sequence of events such that $\lim_{n\to\infty} A_n=\Phi$ (impossible event).

Now, $P(A_n)=P(-\infty<X\le -n)=F(-n)$ and $P\left(\lim_{n\to\infty} A_n\right)=P(\Phi)=0$

We ARE MAKAUTIANS

We also know that for a decreasing sequence of events $\{A_n\}$, $\lim_{n\to\infty} P(A_n) = P\left(\lim_{n\to\infty} A_n\right)$

$$\Rightarrow \lim_{n\to\infty} P(A_n) = 0$$

$$\Rightarrow \lim_{n\to\infty} F(-n) = 0$$

$$\text{i.e., } F(-\infty) = 0$$

2nd Part:

Let us now consider, $B_n = (-\infty < X \le n)$; $n = 1, 2, \cdots\cdots \infty$

Thus, $\{B_n\}$ is a monotonic increasing sequence of events such that

$$\lim_{n\to\infty} B_n = (-\infty < X < \infty) = S \text{ (Sample Space).}$$

Therefore, $P\left(\lim_{n\to\infty} B_n\right) = P(S) = 1$

and $P(B_n) = P(-\infty < X \le n) = F(n)$

So that $\lim_{n\to\infty} P(B_n) = \lim_{n\to\infty} F(n) = F(\infty)$

Now, for a monotonic increasing sequence of events $\{B_n\}$, $\lim_{n\to\infty} P(B_n) = P\left(\lim_{n\to\infty} B_n\right)$

$$\Rightarrow F(\infty) = 1$$

5. Let $F(x)$ be a distribution function of a random variable *X*, then show that

i) $F(x)$ is monotonically increasing on the real line.

ii) $F(x)$ is continuous from the right at each point.

iii) $\lim_{x\to\infty} F(x) = 1$ i.e., $F(\infty) = 1$ and $\lim_{x\to\infty} F(x) = 0$ i.e., $F(-\infty) = 0$. **[WBUT 2010]**

Answer:

i) If $-\infty < x_1 < x_2 < \infty$, then $P(x_1 < x \le x_2) = F(x_2) - F(x_1) \ge 0$

$$\Rightarrow F(x_2) \ge F(x_1) \quad \text{if } x_2 \ge x_1$$

So, $F(x)$ is monotonically increasing on the real line.

ii) Let us now consider an event $A_n = \left(a < X \le a + \frac{1}{n}\right)$; $n = 1, 2, \cdots\cdots \infty$

Thus, $\{A_n\}$ is a decreasing sequence and $\lim_{n\to\infty} A_n = \Phi$ (Impossible event).

Therefore, $P\left(\lim_{n\to\infty} A_n\right) = P(\Phi) = 0$

Again, $P(A_n) = P\left(a < X \le a + \frac{1}{n}\right) = F\left(a + \frac{1}{n}\right) - F(a)$

We ARE MAKAUTIANS

$$\therefore \quad \lim_{N\to\infty} P(A_n) = \lim_{n\to\infty} F\left(a+\frac{1}{n}\right) - F(a)$$
$$= F(a+0) - F(a)$$

Therefore, $F(a+0) - F(a) = 0 \quad \Rightarrow F(a+0) = F(a)$

So, $F(x)$ is continuous at $x = a$ on the right. Since, $x = a$ is an arbitrary point in $-\infty < x < \infty$, $F(x)$ is continuous on the right everywhere.

iii) Consider

$$H_n = \{a < X \le a + \frac{1}{n}\}. \text{ Clearly } H_n \downarrow \phi \text{ as } n \to \infty.$$
$$\therefore \lim_{n\to\infty} P(H_n) = 0$$
$$or, \lim_{n\to\infty}\left\{F\left(a+\frac{1}{n}\right) - F(a)\right\} = 0$$
$$i.e., F(a+0) = F(a).$$

6. The radius of a circle has distribution given by the p.d.f. **[WBUT 2011, 2015]**

$$f(x) = 1, 1 < x < 2$$
$$= 0, \text{ otherwise}$$

Find the mean and variance of the area of the circle.

Answer:

1st Method:

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

Let Y denote the area of the circle of radius X.

Then $\quad Y = \pi X^2$

$\therefore \quad dy = 2\pi x\, dx$

$$f(y) = f_X(x)\left|\frac{dx}{dy}\right| = \begin{cases} 1\cdot\dfrac{\sqrt{\pi}}{2\pi\sqrt{y}} & ; \pi < y < 4\pi \\ 0 & ; \text{elsewhere} \end{cases}$$

$$= \begin{cases} \dfrac{1}{2\sqrt{\pi}} y^{-1/2} & \pi < y < 4\pi \\ 0 & \text{elsewhere} \end{cases}$$

Hence, $E(Y) = \int_{-\infty}^{\infty} y f(y)\, dy = \dfrac{1}{2\sqrt{\pi}} \int_{\pi}^{4\pi} y^{1/2} dy = \dfrac{1}{2\sqrt{\pi}} \times \dfrac{2}{3}\left[y^{3/2}\right]_{\pi}^{4\pi} = \dfrac{7\pi^{\frac{3}{2}}}{3\sqrt{\pi}} = \dfrac{7\pi}{3}$

We ARE MAKAUTIANS

$$E\left(Y^2\right)=\int_{-\infty}^{\infty} y^2 f(y)\,dy=\frac{1}{2\sqrt{\pi}}\int_{\pi}^{4\pi} y^{3/2}dy=\frac{1}{2\sqrt{\pi}}\times\frac{2}{5}\left[y^{5/2}\right]_{\pi}^{4\pi}=\frac{31\pi^{\frac{5}{2}}}{5\sqrt{\pi}}=\frac{31\pi^2}{5}$$

$$\operatorname{var}(Y)=E\left(Y^2\right)-\{E(Y)\}^2=\frac{31\pi^2}{5}-\frac{49\pi^2}{9}=\frac{34\pi^2}{45}.$$

2nd Method:

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

Let Y denote the area of the circle of radius X.

Then $Y=\pi X^2$.

Thus, Mean is $E(Y)=E(\pi X^2)=\pi E(X^2)=\pi\int_1^2 x^2dx=\pi\left[\frac{x^3}{3}\right]_1^2=\frac{7\pi}{3}$

And variance is $Var(Y)=E(Y^2)-\{E(Y)\}^2=E(\pi^2X^4)-\{E(Y)\}^2$

$$=\pi^2\int_1^2 x^4dx-\left(\frac{7\pi}{3}\right)^2=\pi^2\left[\frac{x^5}{5}\right]_1^2-\left(\frac{7\pi}{3}\right)^2$$

$$=\frac{31\pi^2}{5}-\frac{49\pi^2}{9}=\frac{34\pi^2}{45}$$

7. A fair coin is tossed 400 times, using normal approximation to binomial distribution. Find the probability of obtaining (i) exactly 200 heads, (ii) between 190 and 210 heads, both inclusive, given that the area under standard normal curve between $z=0$ and $z=0.05$ is $=0.0199$ and between $z=0$ and $z=1.05$ is 0.3531.

[WBUT 2010]

Answer:

Here, p (probability of success) $=\frac{1}{2}$,

N (number of trial) $=400$.

Now, we consider $z=\frac{X-np}{\sqrt{npq}}\sim N(0,1)$, for large n.

Here $np=200$, $npq=400\times\frac{1}{2}\times\left(1-\frac{1}{2}\right)=100$.

i) When $X=200$, we consider $199.5\le X\le 200.5$

i.e., $P(199.5\le X\le 200.5)=P\left(\frac{199.5-200}{\sqrt{100}}\le z\le\frac{200.5-200}{\sqrt{100}}\right)$

We ARE MAKAUTIANS

$$= P(-0.05 \le z \le 0.05) = 2P(0 \le z \le 0.05)$$
$$= 2 \times 0.4801 = 0.9602$$

ii) When $189.5 \le X \le 210.5$

$$P(189.5 \le X \le 210.5) = P(-1.05 \le z \le 1.05) = 2P(0 \le z \le 1.05) = 2 \times 0.3531 = 0.7062$$

8. If the weekly wage of 10,000 workers in a family follows normal distribution with mean and standard derivation Rs. 70 and Rs. 5 respectively, find the expected number of workers whose weekly wages are:

a) between Rs. 66 and Rs. 72 **[WBUT 2011, 2016, 2017]**

b) less than Rs. 66

c) more than Rs. 72

[Given $\frac{1}{\sqrt{2\pi}}\int_0^z e^{-\frac{t^2}{2}}dt = 0.1554$ **or** 0.2881 **according as** $z = 0.4$ **or** 0.8**]**

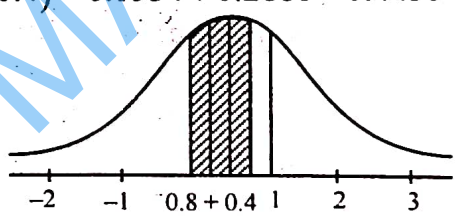
Answer:

Let X denote weekly wave of a worker

Then $X \sim N(70, 25)$

a) Now, $P(66 \le X \le 72) = P\left(\frac{66-70}{5} \le \frac{X-70}{5} \le \frac{72-70}{5}\right)$

$$= P(-0.8 \le Z \le 0.4) = 0.1554 + 0.2881 = 0.4435$$



Expected no. of workers with weekly wage between Rs.66 and Rs.72

$$= 10,000 \times 0.4435 = 4435$$

b) $P(X \le 66) = P\left(\frac{X-70}{5} < \frac{66-70}{5}\right) = P(Z < 0.8) = 0.2881$

$\therefore$ Expected no. of workers with wage less than Rs.66 $= 10,000 \times 0.2881 = 2881$

c) $P(X > 72) = P\left(\frac{X-70}{5} > \frac{72-70}{5}\right)$

$$= P(Z > 0.4) = 0.5 - P(0 < Z \le 0.4) = 0.5 - 0.1554 = 0.3446$$

Therefore, expected no. of workers with weekly wage more than Rs.72

$$= 10,000 \times 0.3446 = 3446.$$

We ARE MAKAUTIANS

9. The mean of a normal distribution is 50 and 5% of the values are greater than 60. Find the standard deviation of the distribution (Area under standard normal curve between $z=0$ and $z=1.64$ is 0.45) **[WBUT 2017]**

Answer:

Let X denote the normal variate whose mean is 50.

Then $Z=\dfrac{X-50}{\sigma}$ is the standard normal variate.

By the given condition

$$P(X>60)=0.05$$

or, $$P\left(\frac{X-50}{\sigma}>\frac{60-50}{\sigma}\right)=0.05$$

or, $$P\left(Z>\frac{10}{\sigma}\right)=0.05$$

But it is given $P(Z>1.64)=0.05$

So, $$\frac{10}{\sigma}=1.64$$

or, $$\sigma=\frac{10}{1.64}=6.098$$

10. Assuming that the lifespan of a type of transistor is normal, find mean and standard deviation if 84% of the transistors have lifespan less than equal to 65.2 months and 68% have lifespan lying between 65.2 and 62.8 months.

Given $\left[P(z\le 0.9)=0.84;\ P(z\le -0.9)=0.16\right]$. **[WBUT 2012, 2014]**

Answer:

Let X be a random variable denotes the lifetime of the transistor.

$\therefore P(X<65.2)=0.84$ .... (1)

and $P(62.8<X<65.2)=0.68$ .... (2)

Let us consider, $X\approx N(m,\ \sigma^2)$

Now, $$P\left(\frac{X-m}{\sigma}<\frac{65.2-m}{\sigma}\right)=0.84$$

or, $$P\left(Z<\frac{65.2-m}{\sigma}\right)=0.84$$

Given that, $P(Z\le 0.9)\le 0.84$

WeARE MAKAUTIANS

$\therefore \frac{65.2-m}{\sigma}=0.9$

or, $65.2=m+0.9\sigma$ .... (3)

Again, $P(62.8<X<65.2)=0.68$

or, $P\left(\frac{62.8-m}{\sigma}<Z<\frac{65.2-m}{\sigma}\right)=0.68$

or, $P\left(Z<\frac{62.8-m}{\sigma}\right)=0.84-0.68$

or, $P\left(Z<\frac{62.8-m}{\sigma}\right)=0.16$

Given that, $P(Z\le -0.9)=0.16$

$\therefore \frac{62.8-m}{\sigma}=-0.9$

or, $62.8=m-0.9\sigma$ .... (4)

Adding equations (3) and (4), we get

$2m=128$

or, $m=64$

Again from equation (3): $0.9\sigma=\frac{1.2}{1}$

or, $\sigma=\frac{1.2}{.9}=\frac{12}{9}=\frac{4}{3}$

11. 50% of a product have weight 60 kg or more and 10% have weight less than 55kg. Assuming the distribution of weight to be normal (μ, σ^2), evaluate $(\mu-3\sigma, \mu+3\sigma)$.

For a standard normal variate Z, $P(Z\le 1.28)=0.90$ approximately.

[WBUT 2015, 2017]

Answer:

Let X be the random variable denotes the weight of the product.

$\therefore P(X\ge 60)=0.5$ .... (1)

$P(0<X<55)=0.1$ .... (2)

$\therefore P\left(\frac{X-m}{\sigma}\ge\frac{60-m}{\sigma}\right)=0.5$

We ARE MAKAUTIANS

$$P\left(Z \geq \frac{60-m}{\sigma}\right) = 0.5$$

$$\frac{0-m}{\sigma} = 0 \quad \Rightarrow 60 - m = 0$$

$$m = 60$$

gain, $P\left(0 < \frac{X-m}{\sigma} < \frac{55-m}{\sigma}\right) = 0.1$

or, $P\left(0 < Z < \frac{55-m}{\sigma}\right) = 0.1 \Rightarrow P\left(Z \geq \frac{55-m}{\sigma}\right) = 0.9$

$$\Rightarrow P\left(\frac{55-m}{\sigma} < Z < 0\right) = 0.1$$

Since $P(Z \leq 1.28) = 0.90 \Rightarrow P(0 < Z < 1.28) = 0.1$

$$\therefore \frac{55-m}{\sigma} = -1.18 \Rightarrow 55 - m = 1.28\sigma \Rightarrow \sigma = \frac{5}{1.28} = 3.906$$

$$\therefore (\mu - 3\sigma, \mu + 3\sigma) = (m - 3\sigma, m + 3\sigma)$$

$$= (60 - 11.72, 60 + 11.72)$$

$$= (48.28, 71.72)$$

2. A car hire farm has two cars which it hires out day by day. The number of emands for a car on each day is distributed as a Poisson distribution with verage number of demands per day is 1.5. Calculate the number of days pproximately) in a year on which some demand is refused and number of days in year (approximately) on which neither car is used $\left[e^{-1.5} = 0.2231\right]$. [WBUT 2016]

nswer:

t X be the random variable corresponding to the demand of car on any day. Then X llows poisson distribution with parameter $\mu = 1.5$ and the p.m.f

$$P(X = i) = \frac{e^{-\mu}\mu^i}{i!}; \; i = 0, 1, \ldots$$

(i) Some demand is refused if the number of demands is more than two *i.e.*, $X > 2$

$$\therefore P(X > 2) = P(X = 3) + P(X = 4) + \ldots\ldots$$

$$= 1 - \left[P(0) + P(1) + P(2)\right]$$

$$= 1 - \left[\frac{e^{-1.5}(1.5)^0}{0!} + \frac{e^{-1.5}(1.5)^1}{1!} + \frac{e^{-1.5}(1.5)^2}{2!}\right]$$

We ARE MAKAUTIANS

$$= 1 - \left[e^{-1.5} + e^{-1.5} \times 1.5 + e^{-1.5} \times 1.125\right]$$

$$= 1 - e^{-1.5}\left[1 + 1.5 + 1.125\right] = 1 - e^{-1.5} \times 3.625$$

$$= 1 - 0.2231 \times 3.625 = 1 - 0.8087375 = 0.1912625$$

Therefore, number of days in a year when some demand of car is refused

$$= 365 \times 0.1912625 = 69.81 = 70 \text{ days.}$$

(ii) If the car is not used, then demand $X = 0$

Therefore, $P(X = 0) = \dfrac{e^{-1.5}(1.5)^0}{0!} = e^{-1.5} = 0.2231$

So, number of days in a year when the demand is zero $= 365 \times 0.2231 = 81.4315 \cong 81$ days.

13. The mean yield for one acre plot is 662 kilos with a standard deviation 32 kilos. Assuming normal distribution how many one acre plots in a batch of 1000 plots would you expect to have yield over 700 kilos? (Given that $\phi(1.19) = 0.3830$)

[WBUT 2016]

Answer:

Here, $\mu = 662$ kilos, $\sigma = 32$ kilos.

Let X denote yield of an one acre plot.

Then $P(X > 700 \text{ kilos}) = P\left(\dfrac{X-\mu}{\sigma} > \dfrac{700-662}{32}\right)$

$= P\left(Z > \dfrac{38}{32}\right)$, denoting $Z = \dfrac{X-\mu}{\sigma}$

$= P(Z > 1.19) = 0.5 - 0.3850 = 0.115$

Hence, number of plots yielding more than 700 kilos $= 1000 \times 0.115 = 115$

14. Suppose that an airplane engine will fail, when in flight, with probability $(1-p)$ independently from engine to engine; suppose that the airplane will make a successful flight if at least 50% of its engines remain operative. For what values of p is a four-engine plane preferable to a two-engine plane?

[WBUT 2014]

Answer:

A four-engine plane is preferable to a two-engine plane if 4-engine plane flies more successfully than a 2-engine plane i.e., if two or more engines of a 4-engine plane remain operative

P (1 or more engines of a 2-engine plane remain operative)

But, P (2 or more engines remains operative)

$$= {}^4C_2 p^2 (1-p)^2 + {}^4C_3 p^3 (1-p) + {}^4C_4 p^4$$

We ARE MAKAUTIANS

P (1 or more engines remains operative) $= {}^2C_1 p(1-p) + {}^2C_2 p^2$

Hence the required condition is

$$ {}^4C_2 p^2 (1-p)^2 + {}^4C_3 p^3 (1-p) + {}^4C_4 p^4 > {}^2C_1 p(1-p) + {}^2C_2 p^2 $$

or, $$ 6p^2(1-p)^2 + 4p^3(1-p) + p^4 - 2p(1-p) - p^2 > 0 $$

or, $$ 6p^2 - 12p^3 + 6p^4 + 4p^3 - 4p^4 + p^4 - 2p + 2p^2 - p^2 > 0 $$

or, $$ 3p^4 - 8p^3 + 7p^2 - 2p > 0 $$

or, $$ 3p^3 - 8p^2 + 7p - 2 > 0 \qquad \text{[As } p > 0 \text{ by definition]} $$

$$ 3p^2 - 5p + 2 > 0 \qquad \text{as } p \neq 1 $$

$$ p^2 - \frac{5}{3}p + \frac{2}{3} > 0 $$

$$ \left(p - \frac{5}{6}\right)^2 > \frac{25}{36} - \frac{2}{3} = \frac{1}{36} $$

or, $$ p > \frac{5}{6} + \frac{1}{6} $$

or, $$ p < \frac{5}{6} - \frac{1}{6} \qquad \text{or, } < \frac{2}{3} $$

We ARE MAKAUTIANS

CHEBYCHEV'S INEQUALITY

☞ Chapter at a Glance

Statement:

If X be a Random Variable with mean $E(X)=\mu$ and variance $(X)=\sigma^2$,

then $(|X-\mu|\geq\varepsilon)\leq\frac{\sigma^2}{\in^2}$; where $\varepsilon>0$. or, $(|X-\mu|\geq\varepsilon)\leq 1-\frac{\sigma^2}{\in^2}$.

Taking $\varepsilon=k\sigma(k>0)$, we get $P(|X-\mu|\geq k\sigma)\leq\frac{1}{\sigma^2}$ or, $(|X-\mu|\geq k\sigma)\leq 1-\frac{1}{\sigma^2}$.

Multiple Choice Type Questions

1. For a random variable X with mean 0, the value of $P(-4\sigma<X<4\sigma)$ will be at least **[WBUT 2016]**

a) $\frac{1}{16}$ b) $\frac{1}{4}$ c) $\frac{15}{16}$ d) $\frac{13}{16}$

Answer: (c)

Short Answer Type Questions

1. State Chebychev's inequality. **[MODEL QUESTION]**

Answer:

If X be a Random Variable with mean $E(X)=\mu$ and variance $(X)=\sigma^2$,

Then $P(|X-\mu|\geq\varepsilon)\leq\frac{\sigma^2}{\varepsilon^2}$; where $\varepsilon>0$. or, $(|X-\mu|\geq\varepsilon)\geq 1-\frac{\sigma^2}{\varepsilon^2}$

Taking $\varepsilon=k\sigma(k>0)$, we get $P(|X-\mu|\geq k\sigma)\leq\frac{1}{\sigma^2}$ or, $(|X-\mu|\geq k\sigma)\leq 1-\frac{1}{\sigma^2}$

2. Show by Tchebyshev's inequality that if a die is thrown 3,600 times, the probability of number of sixes that lies between 550 and 650 is at least $\frac{4}{5}$. **[WBUT 2017]**

WE ARE MAKAUTIANS

Answer:

Answer:

The Tchebyshev's inequality is

$$P\left(\left|X-E(X)\right|\le\varepsilon\right)\ge 1-\frac{\text{var}(X)}{\varepsilon^2}$$

Let X denote the no. of sizes in 3600 throws of a die

$$\therefore \quad E(X)=np=3600\times\frac{1}{6}=600$$

$$\text{var}(X)=npq=3600\times\frac{1}{6}\times\frac{5}{6}=500$$

The inequality gives

$$P\left(\left|X-600\right|\le 50\right)\ge 1-\frac{500}{50\times 50}, \text{ taking } \varepsilon=50$$

or, $$P\left(550\le X\le 650\right)\ge 1-\frac{1}{5}=\frac{4}{5}.$$

3. A random sample of size $n=100$ is taken from an infinite population with the mean $\mu=75$ and the variance $\sigma^2=256$. Based on Chebyshev's theorem with what probability can we assert that the value we obtain for $\overline{X}$ will fall between 67 and 83? **[WBUT 2016]**

Answer:

Here, $\mu=75,\ \sigma^2=256,\ n=100$

By Chebyshev's theorem, we get

$$P\left(\left|\overline{X}-E(\overline{X})\right|\ge\varepsilon\right)\le\frac{Var(\overline{X})}{\varepsilon^2}$$

or, $$P\left(\left|\overline{X}-75\right|\ge\varepsilon\right)\le\frac{256/100}{\varepsilon^2}$$

or, $$P\left(75-\varepsilon<\overline{X}<75+\varepsilon\right)\ge 1-\frac{2.56}{\varepsilon^2}$$

Taking $\varepsilon=8$, we get $P(67<\overline{X}<83)\ge 1-0.04=0.96$

So the probability that $\overline{X}$ lies between 67 and 83 is at least 0.96.

4. If X is the number scored in a throw of a fair dice, show that the Tchebycheff's inequality gives $P\left(\left|X-\mu\right|\ge 2.5\right)\le 0.47$, where μ is the mean of X, while the actual probability is zero. **[WBUT 2017]**

We ARE MAKAUTIANS

Answer:

The spectrum of r.v. X is $\{1,2,3,4,5,6\}$ and $(X=i)=\frac{1}{6}$.

$\therefore \mu = E(X) = 1.\frac{1}{6}+2.\frac{1}{6}+3.\frac{1}{6}+4.\frac{1}{6}+5.\frac{1}{6}+6.\frac{1}{6}=\frac{7}{2}$

and $E(X^2)=1^2.\frac{1}{6}+2^2.\frac{1}{6}+3^2.\frac{1}{6}+4^2.\frac{1}{6}+5^2.\frac{1}{6}+6^2.\frac{1}{6}=\frac{91}{6}$

Therefore, $\text{var}(X)=E(X^2)-\{E(X)\}^2=\frac{91}{6}-\left\{\frac{7}{2}\right\}^2=2.9167$

Now, by Tchebychev's inequality, we get, $P(|X-\mu|\geq\varepsilon)\leq\frac{\sigma^2}{\epsilon^2}$

Taking $=2.5$, we get $P(|X-\mu|\geq 2.5)\leq\frac{2.9167}{2.5^2}=0.47$

Hence, this is the required inequality.

The actual probability is $P(|X-\mu|\geq 2.5)=P(|X-3.5|\geq 2.5)=1-P(|X-3.5|<2.5)$

$$=1-P(-2.5<X-3.5<2.5)=1-P(1<X<6)=1-1=0$$

Hence, the result.

5. Use Tchebychev's inequality to show that for $n\geq 36$, the probability that in n throws of a fair die the number of sixes lies between $\frac{1}{6}n-\sqrt{n}$ and $\frac{1}{6}n+\sqrt{n}$ is at least $\frac{31}{36}$. **[WBUT 2008]**

Answer:

Let X denote the no. of sixes in n throws of a fair die.

Here $p=\frac{1}{6}$ $\quad\therefore E(X)=\frac{n}{6}$, $\text{var}(X)=n\cdot\frac{1}{6}\cdot\frac{5}{6}=\frac{5n}{36}$

By Chebyshev's inequality,

$$P\left(\frac{1}{6}n-\sqrt{n}<X<\frac{1}{6}n+\sqrt{n}\right)=P\left(\left|X-\frac{n}{6}\right|<\sqrt{n}\right)$$

$$=1-P\left(\left|X-\frac{n}{6}\right|\geq\sqrt{n}\right)\geq 1-\frac{5}{36}=\frac{31}{36},$$

Since the no. of sixes cannot be negative, $\frac{1}{6}n-\sqrt{n}\geq 0$

or, $n^2\geq 36n$ or, $n\geq 36$.

[Recall: $P(|X-E(X)|\geq\varepsilon)\leq\frac{\text{var}(X)}{\varepsilon^2}$ or, $P(|X-E(X)|<\varepsilon)\geq 1-\frac{\text{var}(X)}{\varepsilon^2}$]

We ARE MAKAUTIANS

Long Answer Type Questions

1. Show by Tchebycheff's inequality that in 2000 throws with a coin, the probability that the number of heads lies between 900 and 1100 is at least $\frac{19}{20}$. [WBUT 2017]

Answer:

From Tchebychev's inequality we have

$$P\left(\left|X-E(X)\right|\le\varepsilon\right)\ge 1-\frac{\operatorname{var}(X)}{\varepsilon^2}$$

Let X denote the no. of heads in 2000 throws of a coin. Then, X is Binomial $\left(2000,\frac{1}{2}\right)$ variate.

$$E(X)=np=2000\times\frac{1}{2}=1000$$

$$\operatorname{var}(X)=npq=2000\times\frac{1}{2}\times\frac{1}{2}=500$$

ow, $P(900<X<1100)=P(-100<X-1000<100)$

$$=P\left(\left|X-1000\right|<100\right)=1-P\left(\left|X-1000\right|\ge 100\right)$$

Then by Chebychev's inequality taking $\varepsilon=100$, we get

$$P\left(\left|X-1000\right|\ge 100\right)\le\frac{\operatorname{var}(X)}{100^2}=\frac{500}{100^2}=\frac{1}{20}$$

Therefore, $1-P\left(\left|X-1000\right|\ge 100\right)\ge 1-\frac{1}{20}=\frac{19}{20}$

Thus, $P(900<X<1100)\ge\frac{19}{20}$

Hence, the proof.

2. Show by Tchebycheff's inequality that in 1600 throws with a coin, the probability that the number of heads lies between 700 and 900 is at least $\frac{24}{25}$.

[MODEL QUESTION]

Answer:

From Tchebychev's inequality, we have

$$P\left(\left|X-E(X)\right|\le\varepsilon\right)\ge 1-\frac{\operatorname{var}(X)}{\varepsilon^2}$$

We ARE MAKAUTIANS

Let X denote the no. of heads in 1600 throws of a coin. Then, X is Binomial $\left(2000, \frac{1}{2}\right)$ variate.

$$E(X) = np = 1600 \times \frac{1}{2} = 800$$

$$\text{var}(X) = npq = 1600 \times \frac{1}{2} \times \frac{1}{2} = 400$$

Now, $P(700 < X < 900) = P(-100 < X - 800 < 100)$

$$= P(|X - 800| < 100) = 1 - P(|X - 800| \geq 100)$$

Then by Chebychev's inequality taking $\varepsilon = 100$, we get

$$P(|X - 800| \geq 100) \leq \frac{\text{var}(X)}{100^2} = \frac{400}{100^2} = \frac{1}{25}$$

Therefore, $1 - P(|X - 1000| \geq 100) \geq 1 - \frac{1}{25} = \frac{24}{25}$

Thus, $P(700 < X < 900) \geq \frac{24}{25}$

Hence, the proof.

3. A random variable *x* has the function e^{-x}, $x \geq 0$. **[MODEL QUESTION]**

Show that Tchebycheff's inequality gives $P[|X - 1| > 2] < \frac{1}{4}$ and show that actual probability is e^{-3}.

Answer:

Let X be the r.v. with probability function $f(x) = \begin{cases} e^{-x} & \text{when } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$

Then $E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^{\infty} x e^{-x}\,dx = \left[-xe^{-x}\right]_0^{\infty} + \int_0^{\infty} e^{-x}\,dx = 1$

$$E(X^2) = \int_0^{\infty} x^2 e^{-x}\,dx = \left[-x^2 e^{-x}\right]_0^{\infty} + 2\int_0^{\infty} x e^{-x}\,dx = 2$$

$\therefore$ Var $(X) = E(X^2) - \{E(X)\}^2 = 2 - 1 = 1.$

Now the Tchebycheff's inequality is

$$P(|X - E(X)| > \varepsilon) < \frac{\text{Var}(X)}{\varepsilon^2} \quad \text{or } P(|X - 1| > 2) < \frac{1}{4}$$

We ARE MAKAUTIANS

Again, $P(|X-1|>2) = 1 - P(|X-1| \le 2) = 1 - P(-1 \le X \le 3) = 1 - \int_{-1}^{3} f(x)\,dx$

$$= 1 - \int_{0}^{3} e^{-x}\,dx = 1 - \left[-e^{-x}\right]_{0}^{3} = 1 - \left(1 - e^{-3}\right) = e^{-3}$$

We ARE MAKAUTIANS

BIVARIATE DISTRIBUTIONS

☞ Chapter at a Glance

Let X and Y be two random variables, the joint distribution function of X and Y is defined by $F(x,y)=P(-\infty<X\le x,\ -\infty<Y\le y)$.

If X and Y are Discrete R.V., $P(X=x_i, Y=y_j)=f_{ij}$, is known as joint probability mass function (*p.m.f*) if it satisfy the following conditions:

1. $f_{ij}\ge 0,\ \forall\, i,j$
2. $\sum_{j=-\infty}^{\infty}\sum_{i=-\infty}^{\infty} f_{ij}=1$

and distribution function is given by

$$F(x,y)=\sum_{\substack{j\\ y_j\le y}}\sum_{\substack{i\\ x_i\le x}} f_{ij}$$

If X and Y are Continuous R.V.,

$P\left(x-\frac{dx}{2}\le X\le x+\frac{dx}{2},\ y-\frac{dy}{2}\le Y\le y+\frac{dy}{2}\right)=f(x,y)$, is known as probability density function (*p.d.f*) if it satisfy the following conditions:

1. $f(x,y)\ge 0,\ \forall\, x,y$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy=1$

and distribution function is given by

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

Properties:

- $F(-\infty,y)=0,\ F(x,-\infty)=0,\ F(\infty,\infty)=1$
- $P(a<X<b, c<Y<d)=F(b,d)-F(a,d)-F(b,c)+F(a,c)$
- $\frac{\partial^2 F}{\partial x\partial y}=f(x,y)$, at any points of continuity of $f(x,y)$.

We ARE MAKAUTIANS

Marginal Probability Function:

For Discrete Case:

- $P(X=x_i)=P[X=x_i\cap Y=y_1]+P[X=x_i\cap Y=y_1]+\ldots\ldots$

$=\sum_j P[X=x_i\cap Y=y_j]=\sum_j f_{ij}=f_{i1}+f_{i2}+\ldots\ldots=f_{i\bullet}$

Similarly,

- $P(Y=y_j)=P[X=x_1\cap Y=y_j]+P[X=x_2\cap Y=y_j]+\ldots\ldots$

$P[X=x_i\cap Y=y_j]=\sum_i f_{ij}=f_{1j}+f_{2j}+\ldots\ldots=f_{\bullet j}$

ie $P(X=x_i)$ and $P(Y=y_j)$ or, in other notation $f_{i\bullet}$ and $f_{\bullet j}$ are known as **Marginal** *m.f* of X and Y respectively.

gain, $f_{i|j}=P(X=x_i\mid Y=y_j)=\dfrac{P(X=x_i\cap Y=y_j)}{P(Y=y_j)}=\dfrac{f_{ij}}{f_{\bullet j}}$

And $f_{j|i}=P(Y=y_j\mid X=x_i)=\dfrac{P(Y=y_j\cap X=x_i)}{P(X=x_i)}=\dfrac{f_{ij}}{f_{i\bullet}}$

are known as **conditional probability mass function.**

r Continuous Case:

$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy$ and $f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\,dx$ are known as **marginal probability density function** of X and Y respectively.

$$F_X(x)=F(x,\infty)=\int_{-\infty}^{x}\int_{-\infty}^{\infty} f(x,y)\,dy\,dx=\int_{-\infty}^{x}\left[\int_{-\infty}^{\infty} f(x,y)\,dy\right]dx=\int_{-\infty}^{x} f_X(x)\,dx$$

and $F_Y(y)=\int_{-\infty}^{y} f_Y(y)\,dy$ are known as **marginal probability distribution function** of X and Y respectively.

$f_X(x\mid y)=\dfrac{f(x,y)}{f_Y(y)}$ is **the conditional density function** of X, given Y and

$F_X(x\mid y)=\int_{-\infty}^{x} f_X(x\mid y)\,dx$ is the **conditional distribution function** of X, given Y.

WE ARE MAKAUTIANS

Similarly, $f_Y(y|x)=\dfrac{f(x,y)}{f_X(x)}$ is the conditional density function of Y, given X; and

$F_Y(y|x)=\int_{-\infty}^{y} f_Y(y|x)dy$ is the conditional distribution function of Y, given X.

Independent Random Variable:

- The random variables X and Y are said to be independent if
  $P(-\infty < X \le x, -\infty < Y \le y) = P(-\infty < X \le x).P(-\infty < Y \le y)$
  or, $F(x,y)=F_X(x).F_Y(y)$
- If the random variables X and Y are discrete and independent, then
  i) $f_{ij}=f_{i*}.f_{*j}$
  ii) $f_{i|j}=f_{i*}$
  iii) $f_{j|i}=f_{*j}$
- If the random variables X and Y are continuous and independent, then
  i) $f(x,y)=f_X(x).f_Y(y)$
  ii) $f_X(x|y)=f_X(x)$
  iii) $f_Y(y|x)=f_Y(y)$

Theorem: If the joint *p.d.f* of two-dimensional random variable (X,Y) be $f_{XY}(x,y)$ and if $U=g(X,Y)$, $V=h(X,Y)$ are other two random variables, then the joint *p.d.f* of (U,V) is given by $f_{UV}(u,v)=f_{XY}(x,y)\left|\dfrac{\partial(x,y)}{\partial(u,v)}\right|$.

Multiple Choice Type Questions

1. The random variables X and Y are said to be independent if **[MODEL QUESTION]**

a) $f(x,y)=f_X(x)+f_Y(y)$

b) $f_X(x|y)=f_X(x)$

c) $f_X(x|y)=\dfrac{f(x,y)}{f_Y(y)}$

d) None of these

Answer: (b)

We ARE MAKAUTIANS

2. If random variables X and Y are independent, then [MODEL QUESTION]

a) $F(x,y)=F_X(x)+F_Y(y)$

b) $F(x,y)=F_X(x)/F_Y(y)$

c) $F(x,y)=F_X(x).F_Y(y)$

d) None of these

Answer: (c)

3. Conditional distribution of Y given $X=x$, $P(Y=y|X=x)$ is [MODEL QUESTION]

a) $\dfrac{P(Y=y)}{P(X=x)}$

b) $\sum_i P(X=x_i,Y)$

c) $\dfrac{P(X=x,Y=y)}{P(X=x)}$

d) $\dfrac{P(X=x,Y=y)}{P(Y=y)}$

Answer: (c)

4. Two random variables X and Y are independently distributed if all of the following conditions hold, with the exception of [MODEL QUESTION]

a) $P(Y=y|X=x)=P(Y=y)$

b) $P(X=x|Y=y)=P(X=x)$

c) If the conditional distribution of Y given X equals the marginal distribution of Y

d) $E(Y)=E[E(Y|X)]$

Answer: (d)

5. $\text{var}(aX+bY)=$ [MODEL QUESTION]

a) $a^2\sigma_X^2+b^2\sigma_Y^2$

b) $a^2\sigma_X^2+2ab\sigma_X\sigma_Y+b^2\sigma_Y^2$

c) $a\sigma_X^2+b\sigma_Y^2$

d) none of these

Answer: (b)

Short Answer Type Questions

1. A card is drawn from a full pack of 52 cards. If X denotes the number on the card (assuming that the jack, queen and king respectively correspond to the numbers 11, 12 and 13) and Y takes values 1, 2, 3, 4 for spade, heart, diamond and club respectively, find the distribution of the two-dimensional variate (X,Y) and show that X and Y are independent. [MODEL QUESTION]

We ARE MAKAUTIANS

Answer:
Let X denote the number on the card and Y takes values 1, 2, 3, 4 for spade, heart, diamond and club respectively.
Let the individual spectra of X and Y are
$x_i = i$, for $i = 1, 2, \ldots, 13$
$y_j = j$, for $j = 1, 2, 3, 4$

Therefore, the spectrum of (X, Y) is $(x_i, y_j) = (i, j)$; $i = 1, 2, \ldots, 13$; $j = 1, 2, 3, 4$.
Therefore, the joint distribution is given by

$$f_{ij} = P(X = i, Y = j) = \frac{1}{52}, \text{ for all } i, j.$$

The marginal distributions are given by

$$f_{i*} = \frac{1}{13}, \text{ for all } i$$

$$f_{*j} = \frac{13}{52} = \frac{1}{4}, \text{ for all } j$$

Thus, $f_{ij} = f_{i*} \times f_{*j} = \frac{1}{52}$, for all i, j.
Therefore, the random variables X and Y are independent.

2. The joint p.m.f of (X, Y) is given by **[MODEL QUESTION]**

$$f(x, y) = k(2x + 3y), \; x = 0, 1, 2; \; y = 1, 2, 3$$

(a) Find the value of k. (b) Find the marginal distributions.
Answer:
Let the individual spectra of X and Y are

$$x_i = i \quad (i = 0, 1, 2)$$
$$y_j = j \quad (j = 1, 2, 3)$$

Therefore, the spectrum of (X, Y) is given by $(x_i, y_j) = (i, j)$; $i = 0, 1, 2$; $j = 1, 2, 3$ and the distribution is given in the following table.

$X \setminus Y$	1	2	3	Marginal Distribution of X $f_{i*} = \sum_j f_{ij}$
0	$3k$	$6k$	$9k$	$18k$
1	$5k$	$8k$	$11k$	$24k$
2	$7k$	$10k$	$13k$	$30k$
Marginal Distribution of Y $f_{*j} = \sum_i f_{ij}$	$15k$	$24k$	$33k$	$72k$

We ARE MAKAUTIANS

(a) Since, we know that

$$\sum_{j=-\infty}^{\infty}\sum_{i=-\infty}^{\infty} f_{ij}=1 \Rightarrow \sum_{j=1}^{3}\sum_{i=0}^{2} f_{ij}=1 \Rightarrow 72k=1 \Rightarrow k=\frac{1}{72}$$

(b) Marginal distribution of X is

$$P(X=0)=\frac{18}{72},\ P(X=1)=\frac{24}{72},\ P(X=2)=\frac{30}{72}$$

Marginal distribution of Y is

$$P(Y=1)=\frac{15}{72},\ P(Y=2)=\frac{24}{72},\ P(Y=3)=\frac{33}{72}$$

3. The joint probability density function of the random variables (X, Y) is

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

Find

(a) The marginal *p.d.f* of X and Y.

(b) The mean and variance of X and Y.

(c) Test whether X and Y are independent. **[MODEL QUESTION]**

Answer:

(a) The marginal *p.d.f* of X is $f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy$

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

The marginal *p.d.f* of Y is $f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\,dx$

$$=\int_{0}^{y} 2dx=2y,\ 0<y<1$$

(b) Mean of X is $E(X)=\int_0^1 x f_X(x)\,dx=2\int_0^1 x(1-x)\,dx=\frac{1}{3}$

Mean of Y is $E(Y)=\int_0^1 y f_Y(y)\,dy=2\int_0^1 y^2\,dy=\frac{2}{3}$

Again, $E(X^2)=\int_0^1 x^2 f_X(x)\,dx=2\int_0^1 x^2(1-x)\,dx=\frac{1}{6}$

We ARE MAKAUTIANS

and $\quad E(Y^2) = \int_0^1 y^2 f_Y(y)\,dy = 2\int_0^1 y^3\,dy = \frac{1}{2}$

Thus, variance of X and Y is

$$\text{var}(X) = E(X^2) - [E(X)]^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}$$

$$\text{var}(Y) = E(Y^2) - [E(Y)]^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

(c) Since, $f(x, y) = 2 \neq f_X(x) \cdot f_Y(y) = 4(1-x)Y$

Hence, X and Y are not independent.

4. The joint probability density function of the random variables X and Y is

$$f(x, y) = \begin{cases} k(1-x-y); & x \geq 0,\ y \geq 0,\ x+y \leq 1 \\ 0; & \text{elsewhere} \end{cases}$$

where k is constant.Find

(a) The value of k.

(b) The marginal *p.d.f* of X.

(c) The *p.d.f* of Y when $X = \frac{1}{2}$ and hence find the mean value of Y when $X = \frac{1}{2}$.

[MODEL QUESTION]

Answer:

(a) Since, $f(x, y)$ is a density function,

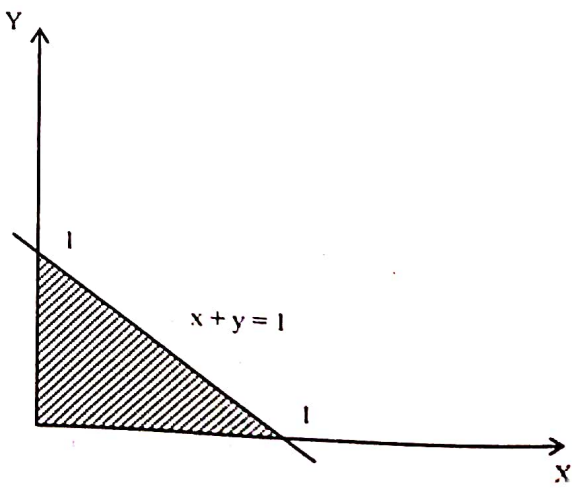
$$\therefore \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,dx\,dy = 1$$

$$\Rightarrow \quad \int_{x=0}^{1}\int_{y=0}^{1-x} k(1-x-y)\,dx\,dy = 1$$

$$\Rightarrow \quad k\int_{x=0}^{1}\left[(1-x)y - \frac{y^2}{2}\right]_0^{1-x} dx = 1$$

$$\Rightarrow \quad k\int_{x=0}^{1}\left[(1-x)^2 - \frac{(1-x)^2}{2}\right] dx = 1$$

$$\Rightarrow \quad \frac{k}{2}\int_{x=0}^{1}(1-x)^2\,dx = 1$$



WeARE MAKAUTIANS

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

$$\Rightarrow \quad \frac{k}{2}\cdot\frac{1}{3} = 1 \quad \Rightarrow \quad k = 6$$

(b) The marginal $p.d.f$ of X is $f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,dy$

$$= 6\int_0^{1-x}(1 - x - y)\,dy = 3(1-x)^2, \; 0 < x < 1$$

(c) The $p.d.f$ of Y when $X = \frac{1}{2}$ is

$$f_Y\left(y \mid x = \frac{1}{2}\right) = \frac{f\left(x = \frac{1}{2}, y\right)}{f_X\left(x = \frac{1}{2}\right)} = \frac{6\left(1 - \frac{1}{2} - y\right)}{3\left(1 - \frac{1}{2}\right)^2} = 4(1 - 2y), \; 0 < y < 1$$

Now, the mean value of Y when $X = \frac{1}{2}$ is

$$E\left(Y \mid X = \frac{1}{2}\right) = \int_{-\infty}^{\infty} y\, f_Y\left(y \mid x = \frac{1}{2}\right)dy = \int_0^1 y \cdot 4(1 - 2y)\,dy = 4\left[\frac{y^2}{2} - 2\cdot\frac{y^3}{3}\right]_0^1 = -\frac{2}{3}$$

5. The joint probability density function of the random variables X and Y is

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

Find a) $E(XY)$ and b) $E(X + Y)$. **[MODEL QUESTION]**

Answer:

a) $$E(XY) = \int_0^1\int_0^1 xy\, f(x, y)\,dx\,dy = \int_0^1\int_0^1 xy(x + y)\,dx\,dy = \int_0^1\left[\frac{x^2}{2} + \frac{x}{3}\right]dx = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

b) $$E(X + Y) = \int_0^1\int_0^1 (x + y)\, f(x, y)\,dx\,dy = \int_0^1\int_0^1 (x + y)^2\,dx\,dy$$

$$= \int_0^1\left[x^2 + x + \frac{1}{3}\right]dx = \frac{1}{3} + \frac{1}{2} + \frac{1}{3} = \frac{7}{6}$$

We ARE MAKAUTIANS

Long Answer Type Questions

1. A ball is drawn from an urn consisting 9 balls numbered 0, 1,..., 8 of which the first 4 are white, the next 3 are red and the last 2 are black. If the colours white, red and black are numbered as 0, 1 and 2 respectively, find the joint distribution of the random variables denoting "number of the ball and the colour number". Calculate the marginal distributions and the conditional distribution of the number of the ball given that the colour is white. **[MODEL QUESTION]**

Answer:

Let X and Y be the random variable denotes "number of the balls" and "colour number".

The individual spectra of X and Y are

$$x_i = i \quad (i = 0, 1, 2, \cdots\cdots, 8)$$

$$y_j = j \quad (j = 0, 1, 2)$$

Therefore, the spectrum of (X, y) is given by

$$(x_i, y_j) = (i, j);\ i = 0, 1, \cdots\cdots, 8;\ j = 0, 1, 2$$

and the distribution is given in the following table.

$Y \backslash X$	0	1	2	3	4	5	6	7	8	$f_{\cdot j}$
0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	0	0	0	0	0	$\frac{4}{9}$
1	0	0	0	0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	0	0	$\frac{3}{9}$
2	0	0	0	0	0	0	0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{2}{9}$
$f_{i\cdot}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	

(a) From the above table, the joint distribution of (X, Y) is given by

$$f_{ij} = P(X = x_i, Y = y_j) = \begin{cases} \frac{1}{9} & i = 0, 1, 2, 3;\ j = 0 \\ \frac{1}{9} & i = 4, 5, 6;\ j = 1 \\ \frac{1}{9} & i = 7, 8;\ j = 2 \\ 0 & \text{elsewhere} \end{cases}$$

We ARE MAKAUTIANS

(b) Marginal distribution of X is $P(X=i)=\frac{1}{9};\ i=0,1,\ldots\ldots,8$

Marginal distribution of Y is $P(Y=0)=\frac{4}{9},\ P(Y=1)=\frac{3}{9},\ P(Y=2)=\frac{2}{9}$.

Conditional probability

$$P(X=x_i \mid Y=0)=\frac{P(X=x_i,\ Y=0)}{P(Y=0)}=\begin{cases}\dfrac{1/9}{4/9}=\dfrac{1}{4} & ;\ i=0,1,2,3\\[2ex] \dfrac{0}{4/9}=0 & ;\ \text{Otherwise}\end{cases}$$

2. An urn contains 12 white, red and black balls, there being 4 balls of each colour numbered 0, 1, 2, 3. A ball is drawn at random from the urn. The colour of the balls is also assigned numbers 0 for white, 1 for red and 2 for black.
a) Find the joint distribution of the random variables "number of the ball and that of the colour". b) Calculate the marginal distributions of the individual random variables, Expectations and the conditional distribution of the number of the ball given that the colour is white.
c) Show that the number of the ball and that of the colour are independent random variables.
[MODEL QUESTION]

Answer:

Let X and Y be the random variable denotes "number of the balls" and "colour of the balls".

The individual spectra of X and Y are

$$x_i=i \quad (i=0,1,2,3);\quad y_j=j \quad (j=0,1,2)$$

Therefore, the spectrum of (X,Y) is given by $(x_i,y_j)=(i,j);\ i=0,1,2,3;\ j=0,1,2$ and the distribution is given in the following table.

$X \backslash Y$	0	1	2	Marginal Distribution of X $f_{i*}=\sum_j f_{ij}$
0	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
1	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
2	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
Marginal Distribution of Y $f_{*j}=\sum_i f_{ij}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	

We ARE MAKAUTIANS

(a) From the above table, we get $f_{ij} = P(X = x_i, Y = y_j) = \frac{1}{12}$ for all i, j.

(b) Marginal distribution of X is $f_{i*} = \sum_{j=0}^{2} f_{ij} = \frac{1}{4}$, $i = 0, 1, 2, 3$

$$\therefore \quad E(X) = 0 \times \frac{1}{4} + 1 \times \frac{1}{4} + 2 \times \frac{1}{4} + 3 \times \frac{1}{4} = \frac{6}{4} = \frac{3}{2}$$

Marginal distribution of Y is $f_{*j} = \sum_{i=0}^{3} f_{ij} = \frac{1}{3}$, $j = 0, 1, 2$

$$\therefore \quad E(Y) = 0 \times \frac{1}{3} + 1 \times \frac{1}{3} + 2 \times \frac{1}{3} = 1$$

Conditional probability $P(X = x_i \mid Y = 0) = \frac{P(X = x_i, Y = 0)}{P(Y = 0)} = \frac{1/12}{1/3} = \frac{1}{4}$, $(i = 0, 1, 2, 3)$

(c) From the above table, we get $f_{ij} = f_{i*} \times f_{*j}$, for all i, j.

Therefore, X and Y are independent.

3. The joint probability density function of the random variables (X, Y) is

$$f(x, y) = \begin{cases} Kxye^{-(x^2+y^2)}; & x \geq 0,\ y \geq 0 \\ 0; & \text{elsewhere} \end{cases}$$

(a) Find the value of K.
(b) Test whether X and Y are independent.
(c) Find the conditional density function of X given Y. **[MODEL QUESTION]**

Answer:

a) From the property *p.d.f*, we get $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,dx\,dy = 1$

$$\Rightarrow \quad \int_0^{\infty}\int_0^{\infty} Kxye^{-(x^2+y^2)}\,dx\,dy = 1$$

$$\Rightarrow \quad K\int_0^{\infty} xe^{-x^2}\,dx \times \int_0^{\infty} ye^{-y^2}\,dy = 1$$

[Taking $x^2 = u \Rightarrow du = 2x\,dx$; $y^2 = v \Rightarrow dv = 2y\,dy$]

$$\Rightarrow \quad K\int_0^{\infty} e^{-u}\frac{du}{2} \times \int_0^{\infty} e^{-v}\frac{dv}{2} = 1$$

WeARE MAKAUTIANS

$$\Rightarrow \quad K\left[-\frac{e^{-u}}{2}\right]_0^{\infty} \times \left[-\frac{e^{-v}}{2}\right]_0^{\infty} = 1$$

$$\Rightarrow \quad K \cdot \frac{1}{2} \cdot \frac{1}{2} = 1 \quad \Rightarrow \quad K = 4$$

b) The marginal $p.d.f$ of X is $f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,dy$

$$= \int_0^{\infty} 4xye^{-(x^2+y^2)}dy, \quad 0 \le x < \infty$$

$$= 4xe^{-x^2} \int_0^{\infty} ye^{-y^2}\,dy$$

$$= 4xe^{-x^2} \int_0^{\infty} e^{-t}\,\frac{dt}{2} \qquad [\text{Let } y^2 = t]$$

$$= 2xe^{-x^2}\left[-e^{-t}\right]_0^{\infty} = 2xe^{-x^2}$$

The marginal $p.d.f$ of Y is $f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\,dx$

$$= \int_0^{\infty} 4xye^{-(x^2+y^2)}dx, \quad 0 \le y < \infty$$

$$= 4ye^{-y^2} \int_0^{\infty} xe^{-x^2}\,dx$$

$$= 4ye^{-y^2} \int_0^{\infty} e^{-t}\,\frac{dt}{2} \qquad [\text{Let, } x^2 = t]$$

$$= 2ye^{-y^2}\left[-e^{-t}\right]_0^{\infty} = 2ye^{-y^2}$$

Now, since $f(x, y) = f_X(x) \cdot f_Y(y)$, so X and Y are independent.

c) The conditional density function of X given Y is

$$f_X(x \mid y) = \frac{f(x, y)}{f_Y(y)} = \frac{4xye^{-(x^2+y^2)}}{2ye^{-y^2}} = 2xe^{-x^2}$$

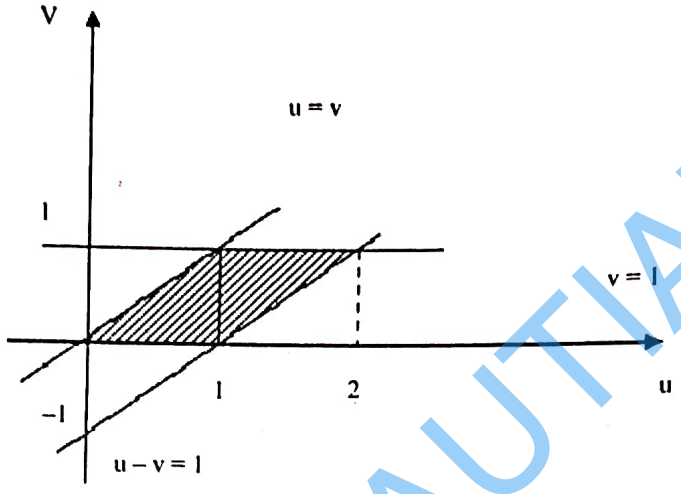
We ARE MAKAUTIANS

4. The joint probability density function of the random variables X and Y is

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

Find the distribution of $X+Y$. **[MODEL QUESTION]**

Answer:



Let $\quad U=X+Y,\ V=X$

In terms of real variable, $u=x+y,\ v=x \quad \Rightarrow \quad x=v,\ y=u-v$

Now, $$\frac{\partial(u,v)}{\partial(x,y)}=\begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}=\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix}=-1$$

Also, as x, y ranges from 0 to 1; u, v ranges from 0 to 2 and 0 to 1 respectively.

Now, the $p.d.f$ of (U,V) is given by

$$g(u,v)=f(x,y)\left|\frac{\partial(x,y)}{\partial(u,v)}\right|=f(x,y)\cdot 1=f(v,u-v)=v+(u-v)=u$$

For, $\quad 0<u<1;$

$$g_U(u)=\int_{-\infty}^{\infty} g(u,v)\,dv=u\int_0^u dv=u^2$$

For, $\quad 1<u<2;$

$$g_U(u)=\int_{-\infty}^{\infty} g(u,v)\,dv=u\int_{u-1}^{1} dv=u(2-u)$$

Hence, the marginal $p.d.f$ of $U=X+Y$ is given by

$$g_U(u)=\begin{cases} u^2 & ;\ 0<u<1 \\ u(2-u) & ;\ 1<u<2 \end{cases}$$

We ARE MAKAUTIANS

5. The joint probability density function of a two-dimensional random variable (X,Y) is

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

Find the density function of the random variable $\frac{X}{Y}$. **[MODEL QUESTION]**

Answer:

Let $U=\frac{X}{Y},\ V=X$

In terms of real variable, $u=\frac{x}{y},\ v=x \Rightarrow x=v,\ y=\frac{v}{u}$

Now, $$\frac{\partial(x,y)}{\partial(u,v)}=\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}=\begin{vmatrix} 0 & 1 \\ -\frac{v}{u^2} & \frac{1}{u} \end{vmatrix}=\frac{v}{u^2}$$

Again, as x, y ranges from 0 to ∞; u, v both ranges from 0 to ∞.

Now, the $p.d.f$ of (U,V) is given by

$$g(u,v)=f(x,y)\left|\frac{\partial(x,y)}{\partial(u,v)}\right|=f\left(v,\frac{v}{u}\right)\cdot\frac{v}{u^2}=\frac{v}{u^2}\cdot e^{-\left(v+\frac{v}{u}\right)};\ 0<u,\ v<\infty$$

Hence, the marginal $p.d.f$ of $U=\frac{X}{Y}$ is given by

$$g_U(u)=\int_{-\infty}^{\infty} g(u,v)\,dv=\int_0^{\infty}\frac{v}{u^2}\cdot e^{-\left(v+\frac{v}{u}\right)}dv=\frac{1}{u^2}\int_0^{\infty} ve^{-v\left(1+\frac{1}{u}\right)}dv$$

$$=\frac{1}{u^2}\left[-v\cdot\frac{u}{u+1}e^{-v\left(\frac{u+1}{u}\right)}-\left(\frac{u}{u+1}\right)^2 e^{-v\left(\frac{u+1}{u}\right)}\right]_{v=0}^{\infty}=\frac{1}{u^2}\cdot\left(\frac{u}{u+1}\right)^2=\frac{1}{(u+1)^2}$$

Hence, the marginal $p.d.f$ of $U=\frac{X}{Y}$ is

$$g_U(u)=\frac{1}{(u+1)^2};\ 0<u<\infty$$

We ARE MAKAUTIANS

6. Consider the following joint probability mass function: **[MODEL QUESTION]**

$$P_{XY}(x,y)=\frac{xy^2}{13}I_S(s,y) \text{ with } S=\{(1,1),\ (1,2),\ (2,2)\}$$

(a) Show that P_{XY} is a valid joint probability mass function

(b) What is $P(X+Y\le 3)$?

(c) Determine the marginal probability mass function of X and Y

(d) What are the expected values of X and Y?

(e) Are X and Y independent?

Answer:

a) The joint probability mass function of X and Y is given by the following table:

$X \backslash Y$	1	2
1	$\frac{1}{13}$	$\frac{4}{13}$
2	0	$\frac{8}{13}$

We first note that $P_{XY}(x,y)\ge 0$ for all $x,y=1,2$

Now, $\sum_{x=1}^{2}\sum_{y=1}^{2}P_{XY}(x,y)=\frac{1}{13}+\frac{4}{13}+0+\frac{8}{13}=1$

Hence, P_{XY} is indeed a valid joint probability mass function.

b) Let us denote B the set of values such that $X+Y\le 3$:

$$B=\{(1,1),\ (1,2),\ (2,1)\}\}$$

Therefore,

$$\begin{aligned}P(X+Y\le 3)&=\sum_{(x,y)\in B}P_{XY}(x,y)\\&=P_{XY}(1,1)+P_{XY}(1,2)+P_{XY}(2,1)\\&=\frac{1}{13}+\frac{4}{13}+0=\frac{5}{13}\end{aligned}$$

c) The marginal probability mass function of X is given by:

$$P_X(1)=\sum_{y=1}^{2}P_{XY}(1,y)=P_{XY}(1,1)+P_{XY}(1,2)=\frac{1}{13}+\frac{4}{13}=\frac{5}{13}$$

We ARE MAKAUTIANS

$$P_X(2)=\sum_{y=1}^{2}P_{XY}(2,y)=P_{XY}(2,1)+P_{XY}(2,2)=0+\frac{8}{13}=\frac{8}{13}$$

Similarly, the marginal probability mass function of Y is given by:

$$P_Y(1)=\sum_{x=1}^{2}P_{XY}(x,1)=P_{XY}(1,1)+P_{XY}(2,1)=\frac{1}{13}+0=\frac{1}{13}$$

$$P_Y(2)=\sum_{x=1}^{2}P_{XY}(x,2)=P_{XY}(1,2)+P_{XY}(2,2)=\frac{4}{13}+\frac{8}{13}=\frac{12}{13}$$

d) The expected values of X and Y is given by:

$$E(X)=\sum_{x=1}^{2}\sum_{y=1}^{2}x\ P_{XY}(x,y)$$

$$=\sum_{x=1}^{2}x\left\{\sum_{y=1}^{2}P_{XY}(x,y)\right\}$$

$$=\sum_{x=1}^{2}x\ P_X(x)=1\cdot\frac{5}{13}+2\frac{8}{13}=\frac{21}{13}$$

Similarly, expected value of Y is given by:

$$E(Y)=\sum_{x=1}^{2}\sum_{y=1}^{2}y\ P_{XY}(x,y)=\sum_{y=1}^{2}y\left\{\sum_{x=1}^{2}P_{XY}(x,y)\right\}$$

$$=\sum_{y=1}^{2}y\ P_Y(y)=1\cdot\frac{1}{13}+2\cdot\frac{12}{13}=\frac{25}{13}$$

e) It is easy to see that X and Y are not independent, since for example:

$$P_X(2)P_Y(1)=\frac{8}{13}\cdot\frac{1}{13}=\frac{8}{169}\neq 0=P_{XY}(2,1)$$

7. Let X and Y are two continued random variables with joint probability density function:

$$f_{XY}(x,y)=4xy\quad J_{[0,1]^2}(x,y)$$

(a) Verify that f_{XY} is a valid joint probability density function.

(b) What is $P(Y<X)$?

(c) What are the marginal probability density functions of X and Y?

(d) Are X and Y independent?

(e) What are the expected value of X and Y? **[MODEL QUESTION]**

We ARE MAKAUTIANS

Answer:

a) Note that f_{XY} can be rewritten as follows:

$$f_{XY}(x,y) = \begin{cases} 4xy & \text{if } 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

t is now clear that $f_{XY}(x,y) \ge 0$ for all $x, y \in R$

Now, $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dx\,dy = \int_0^1\int_0^1 4xy\,dx\,dy = \int_0^1 4x\left\{\int_0^1 y\,dy\right\}dx$

$$= \int_0^1 4x\left[\frac{y^2}{x}\right]_0^1 dx = \left[x^2\right]_0^1 = 1$$

Hence, f_{XY} is a valid joint probability density function.

b) The set B is the xy plane such that $(x,y) \in [0,1]^2$ and $y < x$ is:

$$B = \{\ (x,y):\ 0 \le y < x \le 1\}$$

Then, $\iint_B f_{XY}(x,y)dx\,dy = \int_{x=0}^{1}\int_{y=0}^{x} 4xy\,dx\,dy$

$$= \int_0^1 4x\left\{\int_{y=0}^{x} y\,dy\right\}dx = \int_0^1 4x\left[\frac{y^2}{x}\right]_0^x dx$$

$$= \int_0^1 2x^3 dx = \left[\frac{x^4}{2}\right]_0^1 = \frac{1}{2}$$

c) The marginal probability density function of X for $x \in [0,1]$, is given by:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_0^1 4xy\,dy = 4x\left[\frac{y^2}{2}\right]_0^1 = 2x$$

and $f_X(x) = 0$ otherwise. In a nutshell

$$f_X(x) = 2x\ I_{[0,1]}(x)$$

Similarly, we find that the marginal probability density function of Y is:

$$f_Y(y) = 2y\ I_{[0,1]}(y)$$

d) Yes, since for all $x, y \in R$:

$$f_X(x)\ f_Y(y) = 4xy\ I_{[0,1]}(x)\ I_{[0,1]}(y) = f_{XY}(x,y)$$

We ARE MAKAUTIANS

c) The expected value of X can be computed as follows:

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

$$= \int_{-\infty}^{\infty} x\, f_X(x)\,dx = \int_0^1 2x^2\,dx = \left[\frac{2x^3}{x}\right]_0^1 = \frac{2}{3}$$

Similarly, we find that the expected value of Y is $\frac{2}{3}$

We ARE MAKAUTIANS

BASIC STATISTICS

☞ Chapter at a Glance

Data for statistical studies are obtained by conducting either experiments or surveys. Experimental design is the branch of statistics that deals with the design and analysis of experiments. The methods of experimental design are widely used in the fields of agriculture, medicine, biology, marketing research, and industrial production.

In an experimental study, variables of interest are identified. One or more of these variables, referred to as the factors of the study are controlled so that data may be obtained about how the factors influence another variable referred to as the response variable, or simply the response. As a case in point, consider an experiment designed to determine the effect of three different exercise programs on the cholesterol level of patients with elevated cholesterol. Each patient is referred to as an experimental unit, the response variable is the cholesterol level of the patient at the completion of the program, and the exercise program is the factor whose effect on cholesterol level is being investigated. Each of the three exercise programs is referred to as a treatment.

Three of the more widely used experimental designs are the completely randomized design, the randomized block design, and the factorial design. In a completely randomized experimental design, the treatments are randomly assigned to the experimental units. For instance, applying this design method to the cholesterol-level study, the three types of exercise program (treatment) would be randomly assigned to the experimental units (patients).

The use of a completely randomized design will yield less precise results when factors not accounted for by the experimenter affect the response variable. Consider, for example, an experiment designed to study the effect of two different gasoline additives on the fuel efficiency, measured in miles per gallon (mpg), of full-size automobiles produced by three manufacturers. Suppose that 30 automobiles, 10 from each manufacturer, were available for the experiment. In a completely randomized design the two gasoline additives (treatments) would be randomly assigned to the 30 automobiles, with each additive being assigned to 15 different cars. Suppose that manufacturer 1 has developed an engine that gives its full-size cars a higher fuel efficiency than those produced by manufacturers 2 and 3. A completely randomized design could, by chance, assign gasoline additive 1 to a larger proportion of cars from manufacturer 1. In such a case, gasoline additive 1 might be judged to be more fuel efficient when in fact the difference observed is actually due to the better engine design of automobiles produced by manufacturer 1. To prevent this from occurring, a statistician could design an experiment in which both gasoline additives are tested using five cars produced by each manufacturer; in this way, any effects due to the manufacturer would not affect the test for significant differences due to gasoline additive. In this revised experiment, each of the manufacturers is referred to as a block, and the experiment is called a randomized block design. In general, blocking is used in order to enable comparisons among the treatments to be made within blocks of homogeneous experimental units.

We ARE MAKAUTIANS

ıctorial experiments are designed to draw conclusions about more than one factor, or variable. he term factorial is used to indicate that all possible combinations of the factors are considered. or instance, if there are two factors with a levels for factor 1 and b levels for factor 2, the periment will involve collecting data on ab treatment combinations. The factorial design can e extended to experiments involving more than two factors and experiments involving partial factorial designs.

Analysis of variance and significance testing

A computational procedure frequently used to analyze the data from an experimental study employs a statistical procedure known as the analysis of variance. For a single-factor experiment, this procedure uses a hypothesis test concerning equality of treatment means to determine if the factor has a statistically significant effect on the response variable. For experimental designs involving multiple factors, a test for the significance of each individual factor as well as interaction effects caused by one or more factors acting jointly can be made. Further discussion of the analysis of variance procedure is contained in the subsequent section.

Regression and correlation analysis

Regression analysis involves identifying the relationship between a dependent variable and one or more independent variables. A model of the relationship is hypothesized, and estimates of the parameter values are used to develop an estimated regression equation. Various tests are then employed to determine if the model is satisfactory. If the model is deemed satisfactory, the estimated regression equation can be used to predict the value of the dependent variable given values for the independent variables.

Regression model

In simple linear regression, the model used to describe the relationship between a single dependent variable y and a single independent variable x is $y = \beta_0 + \beta_1 x + \varepsilon$. β_0 and β_1 are referred to as the model parameters, and ε is a probabilistic error term that accounts for the variability in y that cannot be explained by the linear relationship with x. If the error term were not present, the model would be deterministic; in that case, knowledge of the value of x would be sufficient to determine the value of y.

In multiple regression analysis, the model for simple linear regression is extended to account for the relationship between the dependent variable y and p independent variables $x_1, x_2, \cdots\cdots\cdots, x_p$.

The general form of the multiple regression model is $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots\cdots\cdots \beta_p x_p + \varepsilon$.

The parameters of the model are the $\beta_0, \beta_1, \cdots\cdots\cdots, \beta_p$ and ε is the error term.

We ARE MAKAUTIANS

Short Answer Type Questions

1. The table below gives the frequency distribution of weights of 85 apples:

Weight in gms	**: 110-119**	**120-129**	**130-139**	**140-149**
Frequency	**: 5**	**8**	**12**	**18**
Weight in gms	**: 150-159**	**160-169**	**170-179**	**180-189**
Frequency	**22**	**9**	**7**	**4**

Determine the median weight of an apple. **[MODEL QUESTION]**

Answer:

Following table has been prepared for calculation of Median.

Class Interval Weight in gms. x	Class boundary	Frequency (f)	Cumulative Frequency
110 - 119	109.5 - 119.5	5	5
120 - 129	119.5 - 129.5	8	13
130 - 139	129.5 - 139.5	12	25←C ←N/2=42.5
140 - 149	←M	18→fm	43
150 - 159	$i=10\begin{cases}139.5-149.5\\149.5-159.5\end{cases}$	22	65
160 - 169	159.5 - 169.5	2	74
170 - 179	169.5 - 179.5	7	61
180 - 189	179.5 - 189.5	4	85
		$\sum f = N = 85$	

The formula of median of a distribution is

$$\text{Median} = L_1 + \frac{N/2 - c}{fm} \times i = 139.5 + \frac{42.5 - 2.5}{18} \times 10$$

$$= 139.5 + \frac{17.5 \times 10}{18} = 139.5 + 9.72 = 149.22$$

[**Ans.** Median weight of an apple is 149.22 gms.]

2. The mean marks of 100 students was found to be 40. Later on it was discovered that a mark 53 was misread as 83. Find the corrected mean mark.

[MODEL QUESTION]

Answer:

Mean marks of 100 student is 40

So, the total marks obtained by those 40 students are $100 \times 40 = 4000$

From this the wrong mark 83 is deducted and correct mark 53 has been added.

So, the correct total = 4,000 – 83 + 53 = 3,970.

WE ARE MAKAUTIANS

Now correct mean mark will be $\frac{3970}{100} = 39.7$

[Ans. Correct mean = 39.7]

3. The A. M. of $7, x-2, 10, x+3$ is 9 ; Find x. **[MODEL QUESTION]**

Answer:

$$\text{A.M.} = \frac{7+(x-2)+10+(x+3)}{4} = 9 \quad [\because n = 4]$$

$$\text{or, } \frac{7+x-2+10+x+3}{4} = 9 \qquad \text{or, } \frac{2x+18}{4} = 9$$

$$\text{or, } 2x+18 = 36 \quad \text{or, } 2x = 18 \qquad \text{or, } x = 9$$

[Ans. $x = 9$]

4. The table below gives the marks obtained in statistics by 60 students:

Class Interval:	**0-10**	**10-20**	**20-30**	**30-40**	**40-50**	**50-60**	**60-70**
Frequency :	**5**	**8**	**11**	**15**	**13**	**6**	**2**

With the help of the cumulative frequency diagram or otherwise determine the median work of a student. **[MODEL QUESTION]**

Answer:

To find out median the following table has been prepared.

Class interval x	Frequency f	Cumulative frequency less than more than	Cumulative frequency
0-10	5	5	60 → N
10-20	8	13	55
20-30	11	24	47
30-40	15	39	36
40-50	13	52	21
50-60	6	58	8
60-70	2	60 →N	2

Calculation of median by algebraic method i.e., by using formula

$$\text{Median} = l_1 + \frac{\frac{N}{2} - c}{fm} \times i = 30 + \frac{\frac{60}{2} - 24}{15} \times 10$$

$$= 30 + \frac{30-24}{15} \times 10 = 30 + \frac{6 \times 10}{15} = 30 + 4 = 34$$

[Ans. Median = 34 marks in statistics]

WeARE MAKAUTIANS

5. The heights of the students of a college are given below. Find the median height:

Height in cm.	160-163	164-167	168-171	172-175	176-179	180-183	184-187
Number of Students	22	80	98	148	14	43	5

[MODEL QUESTION]

Answer:

For calculation of Median height following table has been prepared.

Height in cm. (class limit)	Class boundary x	No. of Students f	Cumulative frequency (less than type)
160 - 163	159.5 - 163.5	22	22
164 - 167	163.5 - 167.5	80	102
168 - 171	167.5 – 171.5	98	$200 \rightarrow F \Rightarrow \frac{N}{2} = 250$
172 - 175	$1 \leftarrow$ 171.5 - 175.5	$148 \rightarrow fm$	348
176 - 179	175.5 - 179.5	104	452
180 – 183	179.5 – 183.5	43	495
184 - 187	83.5 – 187.5	5	$500 \rightarrow N$
		$N = 500$	

$$\text{Median} = l_1 + \frac{\frac{N}{2} - F}{fm} \times i = 171.5 + 1.35 = 172.85$$

[**Ans.** The median height is 172.85 cms.]

6. Construct the grouped frequency distribution from the following data and hence find its Arithmetic Mean.

[MODEL QUESTION]

Marks	Number of Students
Less than 10	175
Less than 20	360
Less than 30	680
Less than 40	790
Less than 50	900
Less than 60	1,000

We ARE MAKAUTIANS

Answer:

Statement showing grouped frequency distribution and calculations for Arithmetic Mean

Class limits (x)	Class frequency (f)	Mid x	$d = \frac{x-A}{f}$	fd
0.9	175	4.5	–2	–350
10 – 19	185	14.5	–1	–185
20 – 29	320	24.5 → A	0	0
30 – 39	110	34.5	1	110
40 – 49	110	44.5	2	220
50 –59	100	54.5	3	300
	$N = 1000$			$\sum fd = 95$

$i = 10$,

$$\text{A.M.} = A + \frac{\sum fd}{N} \times i = 24.5 + \frac{95}{1000} \times 10 = 24.5 + .95 = 25.45$$

[Ans. Mean marks is 25.45]

Long Answer Type Questions

1. a) Following are the daily wages (in rupees) of 30 workers in a factors:
60,45,41,32,47,45,50,37,53,17,26,39,59,68,44,12,30,25,36,18,40,62,46,29,32,54,41,14,32,30

Make a frequency distribution taking class-intervals of Rs.10 and find the mean of the distribution.

b) The A.M. of the following distribution is 28.8. Find the missing frequency.

Marks	**: 0-10**	**10-20**	**20-30**	**30-40**	**40-50**	**50-60**
Number of students	**: 4**	**6**	**20**	**?**	**7**	**3**

[MODEL QUESTION]

Answer:

a) Frequency distribution table of wages of 30 workers.

Class intervals	Tally Marks	Frequency	Mid x	d	fd
11 - 20	III	4	15.5	–2	–8
21 - 30	~~IIII~~	5	25.5	1	5
31 - 40	~~IIII~~ II	7	35.5 – A	0	0
41 - 50	~~IIII~~ III	8	45.5	1	8
51 - 60	IIII	4	55.5	2	8
61 - 70	II	2	65.5	3	6
		$N = 30$			$\sum fd = 19$

WE ARE MAKAUTIANS

Arithmetic Mean $= A + \frac{\sum fd}{N} \times i = 35.5 + \frac{9}{30} \times 10 = 35.5 + 3 = 38.5$

[**Ans.** Mean wages is Rs. 38.5]

b)

Marks	No. of Students	Mix x	$d = \frac{x-35}{10}$	fd
0 - 10	4	5	-3	-12
10 - 20	6	15	-2	-12
20 - 30	20	25	-1	-20
30 - 40	$f4$	$35 \rightarrow A$	0	0
40 - 50	7	45	1	7
50 - 60	3	55	2	6
	$N = 40 + f_4$			$\sum fd = -31$

A.M. $= A + \frac{\sum fd}{N} \times i$

or, $28.8 = 35 + \frac{-31}{40 + f_f} \times 10$

or, $28.8 - 35 = \frac{-310}{40 + f_4}$

or, $-6.2 = \frac{-310}{40 + f_4}$

or, $-6.2(40 + f_4) = -310$

or, $(40 + f_4) = \frac{-310}{-6.2} = 50$

or, $f_4 = 50 - 40 = 10$ [Ans. $f_4 = 10$]

2. a) Prove that G.M. ≥ H.M. for two positive numbers. **[MODEL QUESTION]**

b) The mean monthly salary paid to all employees in a certain company was Rs. 15,000. The mean monthly salaries paid to male and female employees are Rs. 15,200 and Rs. 14,200 respectively. Obtain the percentages of male and female employees in the company.

c) Find the missing values from the given information. The median and model daily incomes of 230 persons are Rs. 143.50 and Rs. 144 and the distribution is:

Daily Incomes (in Rs.)	110-120	120-130	130-140	140-150	150-160	160-170	170-180
Frequency	4	16	–	–	–	8	2

Answer:

a) Let the numbers are a and b.

Now G.M. $= \sqrt{ab}$ and H.M. $= \frac{2}{\frac{1}{a} + \frac{1}{b}} = \frac{2}{\frac{a+b}{ab}} = \frac{2ab}{a+b}$

We have to prove that G.M. ≥ H.M.

WE ARE MAKAUTIANS

$$w\ \sqrt{ab} - \frac{2ab}{a+b}$$

$$r, \frac{\sqrt{ab}}{a+b}\left[a+b-2\sqrt{ab}\right]$$

$$\text{or}, \frac{\sqrt{ab}}{a+b}\left[\sqrt{a}-\sqrt{b}\right]^2 \geq 0$$

Since a, b are + (ve).

Hence G.M $\geq$ H.M.

b) Suppose number of male employees = x and number of female employees = y

Therefore, $\dfrac{15200x+14200y}{x+y} = 15,000$

or, $15,200x + 14,200y = 15,000x + 15,000y$

or, $200x = 800y$ or, $\dfrac{x}{y} = \dfrac{200}{800} = \dfrac{1}{4} = 1:4$

c) Table showing calculation of missing frequencies (Suppose, a, bc)

Daily Incomes (in Rs.) (x)	Frequency (f)	Daily Incomes (in Rs.) (x)	Frequency (f)	Cumulative frequency
110 – 120	4	110 – 120	4	4
120 – 130	16	120 – 130	16	20
130 – 140	$a \rightarrow f_1$	130 – 140	a	$20 + aF$
$l_1 \leftarrow$ 140 – 150	$b \rightarrow f_0$	$l_1 \leftarrow$ 140 – 150	$b \rightarrow f_m$	$20 + a + bN/2 = 115$
150 – 160	$c \rightarrow f_2$	150 – 160	c	$20 + a + b + c$
160 – 170	8	160 – 170	8	$28 + a + b + c$
170 – 180	2	170 – 180	2	$30 + a + b + c$
	$N = 230$		$N = 230$	

Given, $4 + 16 + a + b + c + 8 + 2 = 230$

or, $a + b + c = 200$ .... (i)

$$\text{Mode} = l_1 + \left\{\left(\frac{f_0 - f_1}{2f_0 - f_1 - f_2}\right) \times i\right\}$$

$$\text{or}, 144 = 140 + \left[\left(\frac{b-a}{2b-a-c}\right) \times 10\right]$$

$$\text{or}, 4 = \frac{10b-10a}{2b-a-c} \quad \text{or}, 8b - 4a - 4c = 10b - 10a$$

We ARE MAKAUTIANS

or, $3a - b - 2c = 0$ .... (ii)

$$\text{edian} = l_1 + \left[\left(\frac{\frac{N}{2} - F}{f_m}\right) \times i\right]$$

$$\text{or, } 143.50 = 140 + \left[\frac{115 - (20 + a)}{b} \times 10\right]$$

$$\text{or, } 3.5 = \left(\frac{95 - a}{b}\right) \times 10$$

or, $3.56 = 950 - 10a$

or, $10a + 3.5b - 950 = 0$ .... (iii)

Solving the three equations, we get

$a = 600, b = 100, c = 40$ [Ans.]

3. a) If the first moment of a distribution about the value 2 of the variable is 2, find the Mean.

b) Explain the terms Skewness and Kurtosis. For a moderately skewed distribution. Mean = 172, Median = 167 and S.D. = 60. Find the co-efficient of Skewness and Mode. **[MODEL QUESTION]**

Answer:

a) The first moment of a distribution about an arbitrary constant A is 2.

We have to find the Mean.

Mean = First moment about an arbitrary constant + the arbitrary constant.

In symbol, $\bar{x} = \mu_1' + A$

Here, $\bar{x} = 2 + 2 = 4$.

[**Ans.** Mean = 4]

(b) Skewness: Skewness refers to the extent of symmetry or lack of symmetry in the data, i.e., the values of the variable equidistance from mean does not have equal frequency. If we take the mean as middle point then frequencies of two ends have varied strength so far as their total frequencies are concerned.

Kurtosis: Kurtosis of a distribution refers to the degree of sharpness/peakedness of frequency curve of the distribution. In other words it is a measure of extent of concentration of values near mode showing peak (high or low) in the frequency curve.

Co-efficient of Skewness is measured as: $\frac{3(\text{Mean} - \text{Median})}{\text{Std. Deviation}}$

We ARE MAKAUTIANS

Here it is $\frac{3(172-167)}{60}=\frac{3\times 5}{60}=\frac{1}{4}=0.25.$

There is an empirical relation among means, mode and median for a moderately skewed tribution.

ean – Mode = 3 (Mean – Median)

, Mode = 3 Median – 2 Mean $=3\times 167-2\times 172=501-344=157$

[**Ans.** Co-efficient of skewness = 0.25 and Mode = 157]

4. a) If the first two Moments of a distribution about 1 are respectively 2 and 25, find the mean and S.D. of the distribution.

b) Find the first, second and third central moments from the following values of the first three raw moments obtained from a frequency distribution: $\mu_1'=-0.22$, $\mu_2'=1.34$ and $\mu_3'=0.58$. **[MODEL QUESTION]**

Answer:

a) The first moment about the value 1 is 2

i.e. $\frac{1}{n}\sum(x-1)=2$ or, $\frac{1}{n}\sum x-\frac{1}{n}\sum 1=2$

or, $\frac{1}{n}\sum x-1=2\left[\because \frac{1}{n}\sum 1=1\right]$ or, $\frac{1}{n}\sum x=2+1$

or, Mean $=\bar{x}=\frac{1}{n}\sum x=3$

Again, the second moment about the value 1 is 25.

i.e., $\frac{1}{n}\sum(x-1)^2=25$ or, $\frac{1}{n}\sum[x^2-2x+1]=25$

or, $\frac{1}{n}\sum x^2-\frac{1}{n}\sum 2x+\frac{1}{n}\sum 1=25$ or, $\frac{1}{n}\sum x^2-2\frac{1}{n}\sum x+1=25$

or, $\frac{1}{n}x^2-2.3+1=25\left[\because \frac{1}{n}\sum x=3\right]$

or, $\frac{1}{n}\sum x^2=25-1+6$ or, $\frac{1}{n}\sum x^2=30.$

Now, S.D. $=\sqrt{\frac{\sum x^2}{n}-\left(\frac{\sum x}{n}\right)^2}$

$=\sqrt{30-(3)^2}\left[\because \frac{1}{n}\sum x^2=30 \text{ and } \frac{1}{n}\sum x=3\right]$

$=\sqrt{21}-4.58$

We ARE MAKAUTIANS

Alternative Method:

$m_1 = \text{A.M.} = A + m_1$ i.e., $1 + 2 = 3$

and $m_2 = (\text{S.D.})^2 = m_2' - (m_1')^2$ i.e., $25 - (2)^2 = 21$

$\therefore \text{S.D.} = \sqrt{m_2} = \sqrt{21} = 4.58$

b) First Central moment $= m_1 = \mu_1' - \mu_1' = 0$

Second Central moment $= m_2 = \mu_2' - \mu_1'^2 = 1.34 - (-0.22)^2 = 1.34 - .0484 = 1.256$

Third Central moment $= m_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3 = -0.58 - 3(1.34)(-0.22) + 2(-0.22)^3$

$= -0.58 + 0.8844 - 0.021296 = 0.283104$

[**Ans.** $m_1 = 0, m_2 = 1.26, m_3 = 0.28$]

5. a) Calculate the correlation co-efficient of the following data:

X:	**63**	**60**	**67**	**61**	**69**	**70**
Y:	**61**	**65**	**64**	**63**	**68**	**63**

b) Find the equation of the line of regression of Y on X from the following data:

Husband's age (X)	:	**25**	**27**	**29**	**33**	**26**	**32**	**35**	**31**	**36**	**30**
Wife age (Y)	:	**20**	**24**	**24**	**29**	**21**	**26**	**18**	**30**	**29**	**28**

[MODEL QUESTION]

Answer:

a) For calculation of correlation co-efficient following table has been prepared.

x	y	$X = (x - \bar{x})$ $(x - 65)$	$Y = (y - \bar{y})$ $(y - 64)$	X^2	Y^2	XY
63	61	–2	–3	4	9	6
60	65	–5	1	25	1	–5
67	64	2	0	4	0	0
61	63	–4	–1	16	1	4
69	68	4	4	16	15	15
70	63	5	–1	25	1	–5
$\sum x = 390$	$\sum y = 384$	$\sum X = 0$	$\sum Y = 0$	$\sum X^2 = 90$	$\sum Y^2 = 28$	$\sum XY = 16$

$$\bar{x} = \frac{\sum x}{n} = \frac{390}{6} = 65, \quad \bar{y} = \frac{\sum y}{n} = \frac{384}{6} = 64, \; n = 6 \text{ pairs}$$

$$\text{Co-efficient of correlation} = r = \frac{\sum XY}{\sqrt{\sum X^2 \sum Y^2}} = \frac{16}{\sqrt{90 \times 28}} = \frac{16}{50.20} = 0.3187$$

[**Ans.** Co-efficient of correlation $r = +0.32$]

We ARE MAKAUTIANS

b) For calculation of equation of the line of regression of Y on X following table has been prepared.

Husband's Age (X)	Wife's age (Y)	$X = x - 30$	$Y = y - 26$	X^2	XY
25	20	–5	–6	25	30
27	24	–3	–2	9	6
29	24	–1	–2	1	2
33	29	3	3	9	9
26	21	–4	–5	16	20
32	26	2	0	4	0
35	28	5	2	25	10
31	30	1	4	1	4
36	29	6	3	36	18
30	28	0	2	0	0
$\sum x = 304$	$\sum y = 259$	$\sum x = 4$	$\sum y = -1$	$\sum x^2 = 126$	$\sum xy = 99$

10 pairs, $\bar{x} = \frac{\sum x}{n} = \frac{304}{10} = 30.4$ $\bar{y} = \frac{\sum y}{n} = \frac{259}{10} = 25.9$

Since $\sum x$ and $\sum y$ are not zero we apply the following formula to find equation of y on x. i.e., $y - \bar{y} = byx(x - \bar{x})$

$$\text{or,}(y - 25.9) = \frac{\frac{\sum xy}{n} - \frac{\sum x}{n} \cdot \frac{\sum y}{n}}{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2} \cdot (x - 30.4)$$

$$\text{or, } (y - 25.9) = \frac{\frac{99}{10} - \left(\frac{4}{10}\right)\left(\frac{-1}{10}\right)}{\frac{126}{10} - \left(\frac{4}{10}\right)^2} \cdot (x - 30.4)$$

$$\text{or,}(y - 25.9) = \frac{9.9 - \left(\frac{-4}{100}\right)}{12.6 - (.4)^2}(x - 30.4)$$

$$\text{or,}(y - 25.9) = \frac{9.9 + .04}{12.6 - .16} \cdot (X - 30.4)$$

$$\text{or,}y = .799(x - 30.4) + 25.9$$

$$\text{or,}y = .799x - 24.29 + 25.90 \qquad \text{or, } y = .799\,x + 1.61$$

[**Ans.** The equation of the line of regression of Y on X is Y = 0.799X + 1.61]

WE ARE MAKAUTIANS

6. a) If the two regression co-efficients are $b_{yx} = -0.4$ and $b_{xy} = -0.9$, find the value of the correlation co-efficient r.

b) Find the co-efficient of correlation from the following data.

x	:	10	12	13	16	17	20	25
y	:	19	22	24	27	29	33	37

c) From the following data, find the two regression equations:

x	:	1	2	3	4	5
y	:	2	3	5	4	6

Find also the most probable value of Y where $X = 2.5$.

Answer:

a) Given $b_{yx} = -0.4 b_{xy} = -0.9$ we have to find the value of correlation of co-efficient r.

We know $r^2 = b_{yx} \times b_{xy}$ or, $r = \pm\sqrt{b_{yx} \times b_{xy}}$ here $r = \pm\sqrt{-0.4 \times -0.9} = \pm\sqrt{.36} = \pm 0.6$

But since both the regression co-efficient are negative we take r negative -0.6. Because they always have same sign.

[**Ans.** $r = -0.6$]

b) For calculation of co-efficient of correlation following table has been prepared.

x	y	$x-16$ (X)	$y-27$ (Y)	X^2	Y^2	XY
10	19	−6	−8	36	64	48
12	22	−4	−5	16	25	20
13	25	−3	−3	9	9	9
16	27	0	0	0	0	0
17	29	1	2	1	4	2
20	33	4	6	16	36	24
25	37	9	10	81	100	90
$\sum x = 114$	$\sum y = 191$	$\sum X = 1$	$\sum Y = 2$	$\sum X^2 = 159$	$\sum Y^2 = 238$	$\sum XY = 193$

$$\text{Co-efficient of correlation} = r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2} \times \sqrt{n\sum Y^2 - (\sum Y)^2}}$$

$$= \frac{7 \times 193 - (1)(2)}{\sqrt{7 \times 59 - (1)^2} \times \sqrt{7 \times 238 - (2)^2}} = \frac{1351 - 2}{\sqrt{113 - 1} \times \sqrt{166 - 4}}$$

$$= \frac{1349}{\sqrt{112}\,\sqrt{162}} = \frac{1349}{33.35 \times 40.77} = \frac{1349}{1360} = +0.99$$

[**Ans.** Co-efficient of correlation, $r = +0.99$]

WeARE MAKAUTIANS

c) 1st Part:

To form the two regression equations following table has been prepared.

X	Y	$x=X-\bar{X}$ $X-3$	$y=Y-\bar{Y}$ $Y-4$	x^2	y^2	xy
1	2	−2	−2	4	4	4
2	3	−1	−1	1	1	1
3	5	1	1	0	1	0
4	4	0	0	1	0	0
5	6	2	2	4	4	4
$\sum x=15$	$\sum Y=20$	$\sum x=0$	$\sum y=0$	$\sum x^2=10$	$\sum y^2=10$	$\sum xy=9$

$n=5$

$$\bar{X}=\frac{\sum X}{n}=\frac{15}{5}=3 \qquad \bar{Y}=\frac{\sum X}{n}=\frac{20}{5}=4$$

The two regressions are as follows.

(i) Regression equation of x or y or, $(x-\bar{x})=b_{xy}(y-\bar{y})$

or, $x=\bar{x}=\frac{\sum xy}{\sum y^2}(y-\bar{y})$ or, $x-3=\frac{9}{10}(y-4)$

or, $x=\frac{9}{10}(y-4)+3$ or, $x=9(y-4+3)$

or, $x=.9y-3.6+3$ or, $x=.9y-.6$ or $x=-.6+.9y$

(ii) Regression equation of y on i.e., $(y-\bar{y})=b_{yx}(x-\bar{x})$

or, $y-\bar{y}=\frac{\sum xy}{\sum x^2}(x-\bar{x})$ or, $y-4=\frac{9}{10}(x-3)$

or, $y=\frac{9}{10}(x-3)+4$ or, $y=.9(x-3)+4$

or, $y=.9x-2.7+4$ or, $y=.9x+1.3$ or, $y=1.3+.9x$

2nd part:

We have to predict the value of Y when $X=2.5$. Therefore we use the equation Y on X.

i.e., $y=1.3+.9x$ or, $y=1.3+.9\times 2.5$ $[\because x=2.5]$

or, $y=1.3+2.25$ or, $y=3.55$

[Ans. Regressing equations are

(i) $x=-.6+.9y$ and (ii) $y=1.3+.9x$ and when y is 2.5, y will be 3.55]

We ARE MAKAUTIANS

APPLIED STATISTICS

☞ Chapter at a Glance

Estimation

It is often of interest to learn about the characteristics of a large group of elements such as individuals, households, buildings, products, parts, customers, and so on. All the elements of interest in a particular study form the population. Because of time, cost, and other considerations, data often cannot be collected from every element of the population. In such cases, a subset of the population, called a sample, is used to provide the data. Data from the sample are then used to develop estimates of the characteristics of the larger population. The process of using a sample to make inferences about a population is called statistical inference.

Characteristics such as the population mean, the population variance, and the population proportion are called parameters of the population. Characteristics of the sample such as the sample mean, the sample variance, and the sample proportion are called sample statistics. There are two types of estimates: point and interval. A point estimate is a value of a sample statistic that is used as a single estimate of a population parameter. No statements are made about the quality or precision of a point estimate. Statisticians prefer interval estimates because interval estimates are accompanied by a statement concerning the degree of confidence that the interval contains the population parameter being estimated. Interval estimates of population parameters are called confidence intervals.

Sampling and sampling distributions

Although sample survey methods will be discussed in more detail below in the section Sample survey methods, it should be noted here that the methods of statistical inference, and estimation in particular, are based on the notion that a probability sample has been taken. The key characteristic of a probability sample is that each element in the population has a known probability of being included in the sample. The most fundamental type is a simple random sample.

For a population of size N, a simple random sample is a sample selected such that each possible sample of size n has the same probability of being selected. Choosing the elements from the population one at a time so that each element has the same probability of being selected will provide a simple random sample. Tables of random numbers, or computer-generated random numbers, can be used to guarantee that each element has the same probability of being selected.

A sampling distribution is a probability distribution for a sample statistic. Knowledge of the sampling distribution is necessary for the construction of an interval estimate for a population parameter. This is why a probability sample is needed; without a probability sample, the sampling distribution cannot be determined and an interval estimate of a parameter cannot be constructed.

We ARE MAKAUTIANS

Multiple Choice Type Questions

1. The standard deviation of a sample mean for SRSWR is [WBUT 2013]

a) σ^2/n b) $\sigma/\sqrt{n}$ c) σ/n d) n

Answer: (b)

2. If t is a statistic such as $E(t^2)=5$ and $E(t)=2$, then the standard error of t is [WBUT 2013]

a) 0 b) 1 c) 2 d) none of these

Answer: (b)

3. A statistic t is said to be an unbiased estimator of a population parameter θ when [WBUT 2012, 2014, 2017]

a) $E(t)=\theta$ b) $E(t^2)=\theta$ c) $E(t^2)=\{E(\theta)\}^2$ d) $\{E(t)\}^2=E(\theta^2)$

Answer: (a)

4. The maximum likelihood estimate is a solution of the equation [WBUT 2012]

a) $\frac{d\angle(\theta)}{d\theta}=0$ b) $\frac{d\angle(\theta)}{d\theta}=$ constant c) $\frac{d\angle(\theta)}{d\theta}=\theta$ d) none of these

Answer: (a)

5. If $E(T)=5\theta+6$ then the unbiased estimator of θ is [WBUT 2016]

a) $\frac{1}{5}T-\frac{6}{5}$ b) $T-\frac{6}{5}$ c) $\frac{1}{5}T$ d) $T-2$

Answer: (a)

6. The area under the standard normal curve beyond the lines $z=\pm1.96$ is [MODEL QUESTION]

a) 95 per cent b) 90 per cent c) 5 per cent d) 10 per cent

Answer: (a)

7. If T_1, T_2 be statistic with expectation $E(T_1)=2\theta_1+3\theta_2$ and $E(T_2)=\theta_1+\theta_2$, then the unbiased estimators of the parameter θ_1 is [MODEL QUESTION]

a) $3T_2-T_1$ b) $3T_2+2T_1$ c) $2T_2-3T_1$ d) $3T_2\ T_1$

Answer: (a)

WE ARE MAKAUTIANS

8. A random sample of two individuals is to be drawn from a population of size 43. The possible numbers of distinct samples under SRSWR is [MODEL QUESTION]

a) 43×43 b) ${}^{43}C_2$ c) ${}^{43}P_2$ d) none of these

Answer: (b)

9. In random sampling with replacement from a population with s.d. σ, if the sample size is equal to the population size (= *N*), then the standard error of sample mean will be [MODEL QUESTION]

a) 0 b) σ c) $\sigma/\sqrt{n}$ d) none of these

Answer: (a)

10. The expected value of the sample variance of size *n* drawn from a population with mean *m* and standard deviation σ is [MODEL QUESTION]

a) σ^2 b) $n\sigma^2$ c) $\frac{n-1}{n}\sigma^2$ d) $\frac{\sigma^2}{n}$

Answer: (c)

Short Answer Type Questions

1. Let $x_1, x_2, \ldots\ldots, x_n$ be the values of a random sample from an exponential population that the $f(x) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$ for $x_i > 0$. Then find the maximum likelihood estimator of its parameter θ. [WBUT 2014]

Answer:

The m.l.f. $L(x_1, x_2, \ldots\ldots, x_n \theta) = \frac{1}{\theta}e^{-\frac{x_1}{\theta}} \cdot \frac{1}{\theta}e^{-\frac{x_2}{\theta}} \ldots\ldots \frac{1}{\theta}e^{-\frac{x_n}{\theta}} = \frac{1}{\theta^n}e^{-\frac{1}{\theta}(x_1+x_2+\ldots\ldots+x_n)}$

Taking log both sides we get

$$\log L = -\frac{1}{\theta}\left(\sum x_i\right) - n\log\theta$$

Now, $\frac{\partial \log L}{\partial \theta} = 0$ gives $\frac{\sum x_i}{\theta^2} = \frac{n}{\theta}$ or, $\theta = \frac{1}{n}\sum x_i = \bar{x}$.

2. From the random sample of size 40 drawn from a normal population of standard deviation 2, find the 99% confidence interval of the population mean. Find the interval if the mean of such a sample is 3. [MODEL QUESTION]

We ARE MAKAUTIANS

Answer:

Here $n = 40, \bar{x} = 3, \sigma = 2$

We know the confidence interval for population mean μ is

$$\left(\bar{x} - \frac{\sigma}{\sqrt{n}} 2.58, \bar{x} + \frac{\sigma}{\sqrt{n}} 2.58\right)$$

i.e. $$\left(3 - \frac{2}{\sqrt{40}} \times 2.58, 3 + \frac{2}{\sqrt{40}} \times 2.58\right)$$

i.e. $(2.18, 3.82)$.

3. From the random sample of size 49 drawn from a normal population of standard deviation2, find the 99% confidence interval of the population mean. Find the interval if the mean of such a sample is 3.

[Given $\int_0^{2.58} \phi(z)\,dz = 0.495$**]** **[MODEL QUESTION]**

Answer:

Here $\bar{x} = 3, \sigma = 2$ and the sample is a large one $(v \geq 30)$

$\therefore$ The confidence interval is

$$\left(\bar{x} - z_\alpha \frac{\sigma}{\sqrt{n}}, \bar{x} + z_\alpha \frac{\sigma}{\sqrt{n}}\right)$$

i.e. $$\left(3 - 2.58 \times \frac{2}{\sqrt{49}}, 3 + 2.58 \times \frac{2}{\sqrt{49}}\right)$$

i.e. $$\left(3 - \frac{5.16}{7}, 3 + \frac{5.16}{7}\right)$$

i.e. $$\left(\frac{21 - 5.16}{7}, \frac{21 + 5.16}{7}\right)$$

i.e. $$\left(\frac{15.84}{7}, \frac{26.6}{7}\right)$$

i.e. $(2.2628, 3.8)$

WE ARE MAKAUTIANS

4. What will be the likelihood function for estimating population proportion *P* of a binomially distributed population with parameters *P* and *N* when *N* is known?
[MODEL QUESTION]

Answer:

$$L(x_1, x_2, \cdots, x_n; p) = {}^N C_{x_1}\ {}^N C_{x_2} \cdots {}^N C_{x_r}\ p^{x_1+x_2+\cdots+x_N} (1-p)^{nN-\sum_{i=1}^{n} x_i}$$

5. Define point estimation and interval estimation. **[MODEL QUESTION]**

Answer:

When a parameter is estimated by a single value it is called a point estimation. But when an interval is taken in which the parameter is expected to lie, it is called an interval estimation.

6. A population has a normal distribution with parameters μ and σ^2. Find the MLE (maximum likelihood estimator) of σ^2 when μ is known. Also find the mean and variance of the MLE and show that the MLE is unbiased and consistent.
[MODEL QUESTION]

Answer:

The likelihood function is

$$L(x_1, x_2, \ldots, x_n; \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_1-\mu)^2}{2\sigma^2}} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_2-\mu)^2}{2\sigma^2}} \cdots\cdots \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_n-\mu)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$\therefore \log L = -\frac{n}{2}\log\sigma^2 - \frac{n}{2}\log(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\therefore \frac{\partial}{\partial\sigma^2}(\log L) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Equating to zero we therefore get

$$\sigma^2 = \frac{1}{n}\sum(x_i-\mu)^2 = S^2, \text{ say,}$$

Further $\frac{\partial}{\partial\sigma^2}\left(\frac{\partial}{\partial\sigma^2}(\log L)\right) = -\frac{n}{2\sigma^4} = -\frac{n}{2S^4}$ at $\sigma^2 = S^2$

which the negative.

Hence S^2 is the maximum likelihood estimator of σ^2.

We ARE MAKAUTIANS

Observe that $E(S^2)=\sigma^2$ which proves (i) the mean of the likelihood estimator is σ^2 and (ii) this estimator is unbiased.

7. A sample of 600 screws is taken from a large consignment and 75 are found to be defective. Set up a 99% confidence interval for the proportion of defectives for the population. **[MODEL QUESTION]**

Answer:

Here the sample is large and $P=\frac{75}{600}=\frac{1}{8}, n=600$

The 99% confidence interval is, $\left(p-2.58\sqrt{\frac{pq}{n}}, p+2.58\sqrt{\frac{pq}{n}}\right)$

i.e. $\left(\frac{1}{3}-2.58\sqrt{\frac{\frac{1}{3}\times\frac{2}{3}}{600}}, \frac{1}{3}+2.58\sqrt{\frac{\frac{1}{3}\times\frac{2}{3}}{600}}\right)$

i.e. $\left(\frac{1}{3}-2.58\sqrt{\frac{0.22}{600}}, \frac{1}{3}+2.58\sqrt{\frac{0.22}{600}}\right)$

i.e. $(0.284, 0.379)$

8. Find out the Maximum Likelihood Estimate for a population having Poisson distribution. **[MODEL QUESTION]**

Answer:

Let λ be the parameter of the Poisson distribution

$$L(x_1,x_2,x_n;\lambda)=\frac{e^{-\lambda}\lambda^{x_1}}{x_1!}.\frac{e^{-\lambda}.\lambda^{x_2}}{x_2!}.....\frac{e^{-\lambda}\lambda^{x_n}}{x_n!}=\frac{e^{-n\lambda}\lambda^{x_1+x_2+...x_n}}{x_1!x_2!.....x_n!}$$

$\frac{\partial L}{\partial\lambda}=0$ now gives $-n^{-n\lambda}\lambda^{\sum x_i}+e^{-nx}.\left(\sum x_i\right)\lambda^{\sum x_i-1}=0$

or, $-n+\frac{\sum x_i}{\lambda}=0$ or, $\lambda=\frac{\sum x_i}{n}=\bar{x}$

Clearly $\left(\frac{\partial^2 L}{\partial\lambda^2}\right)_{\lambda=\bar{x}}<0$. Hence the m. l. e. of λ is $\bar{x}$

9. A sample of 600 screws in taken from a large consignment and 75 are found to be defective. Set up a 99% confidence interval for the proportion of the defectives in the population. **[MODEL QUESTION]**

We ARE MAKAUTIANS

Answer:

Here, $n = 600$, $p = \frac{75}{600} = \frac{1}{8}$ $\therefore q = \frac{7}{8}$

So the 99% confidence interval for the proportion of defectives

$$= p \pm 2.58\sqrt{\frac{pq}{n}} = 0.125 \pm 2.58 \times 0.0135$$

i.e. $(0.0902,\ 0.15983)$.

10. Find the maximum likelihood estimate for the parameter λ of a Poisson distribution on the basis of a sample of size *n*. **[MODEL QUESTION]**

Answer:

The m.l.f. of a Poisson distribution

$$L(x_1, x_2, x_n, \lambda) = \frac{e^{-\lambda}.\lambda^{x_1}}{x_1!} \cdot \frac{e^{-\lambda}.\lambda^{x_2}}{x_2!} \cdots\cdots \frac{e^{-\lambda}.\lambda^{x_n}}{x_n!} = \frac{e^{-n\lambda}.\lambda^{\sum_{i=1}^{n} x_i}}{x_1!x_2!\ldots x_n!}$$

$$\therefore \frac{\partial L}{\partial \lambda} = \frac{1}{x_1!x_2!\ldots x_n!}\left\{e^{-n\lambda}(-n)\lambda^{\sum_{i=1}^{n} x_i} + \left(\sum_{i=1}^{n} x_i\right)\lambda^{\sum_{i=1}^{n} x_i - 1}.e^{-n\lambda}\right\}$$

Now $\frac{\partial L}{\partial \lambda} = 0$ gives $e^{-n\lambda}.\lambda^{\sum_{i=1}^{n} x_i}\left\{-n + \frac{\sum_{i=1}^{n} x_i}{\lambda}\right\} = 0$; or, $\lambda = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$

11. The marks obtained by 17 students is an examination have a mean 57 and variance 64. Find 99% confidence interval for the mean of the population of marks assuming it to be normal. [Given that $P(t > 3.250) = 0.005$ for 16 degree of freedom]. **[MODEL QUESTION]**

Answer:

Here $\bar{x} = 57$, $s^2 = 64$, $n = 17$

$\therefore$ The 99% confidence interval for the population mean is

$$\left(\bar{x} - t_{0.005}\frac{s}{\sqrt{n-1}},\ \bar{x} + t_{0.005}\frac{s}{\sqrt{n-1}}\right)$$

i.e., $\left(57 - 3.250 \times \frac{8}{4},\ 57 + 3.250 \times \frac{8}{4}\right)$ i.e., $(50.75,\ 63.25)$

We ARE MAKAUTIANS

12 Write the mean and variance of the sampling distribution of sample means for SRSWR and SRSWOR? **[MODEL QUESTION]**

Answer:

$E(\bar{X}) = \mu$ for SRSWR and also for SRSWOR.

$$Var(\bar{X}) = \frac{\sigma^2}{n} \text{ for SRWR } = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) \text{ for SRSWOR.}$$

Long Answer Type Questions

1. a) If T is an unbiased estimator of θ; show that $\sqrt{T}$ is biased estimate of $\sqrt{\theta}$. **[WBUT 2012]**

OR,

If T is an unbiased estimator of θ, then show that T^2 is a biased estimator of θ^2. **[WBUT 2014]**

Answer:

Since $E(X^2) \geq \{E(X)\}^2$, we get $\{E(\sqrt{T})\}^2 \leq E(T)$

or, $E(\sqrt{T}) \leq \sqrt{E(T)} = \sqrt{\theta}$.

Hence $\sqrt{T}$ is not an unbiased estimator of $\sqrt{\theta}$.

b) Prove that the sample mean $\bar{x}$ is an unbiased estimator of the population mean. **[WBUT 2012]**

Answer:

Since $X_1, X_2, \ldots\ldots X_n$ are r-v-s having the same mean μ, say, 'i' is $E(X_i) = \mu$ for every i,

we have $E(\bar{X}) = E\left(\frac{X_1 + X_2 + \ldots\ldots + X_n}{n}\right) = \frac{1}{n}\sum E(X_i) = \frac{1}{n} n\mu = \mu.$

Hence $\bar{X}$ is an unbiased estimated of the population mean μ.

c) If a population has normal distribution with parameter μ and σ, then prove that the statistic $\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$ is maximum likelihood estimate of σ^2 where μ is known. **[WBUT 2012, 2015]**

WE ARE MAKAUTIANS

Answer:

Let $X_1, X_2, \ldots\ldots, X_n$ be a sample of size n and $X_i \sim N(\mu, \sigma^2)$.
Then the likelihood function is

$$L(\sigma^2) = \frac{1}{\left(\sigma\sqrt{2\pi}\right)^n} e^{-\frac{\sum(x_i-\mu)^2}{2\sigma^2}} \text{ for } \mu \text{ known.}$$

$$\therefore \ln L = -\frac{n}{2}\ln\left(\sigma^2 2\pi\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\therefore \frac{\partial}{\partial\sigma^2}(\ln L) = -\frac{n}{2\sigma^2} + \frac{\sum(x_i-\mu)^2}{2\sigma^4}$$

Equating this to zero we get

$$-\frac{n}{2\sigma^2} + \frac{\sum(x_i-\mu)^2}{2\sigma^4} = 0$$

$$\text{or}, \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

Hence the maximum likelihood estimator of σ^2 is $\frac{1}{n}\sum_{i}^{n}(x_i-\mu)^2$

Further,

$$E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2\right\} = \frac{1}{n}\sum_{i=1}^{n}\left(x_i^2 - 2X_i\mu + \mu^2\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left\{E\left(X_i^2\right) - 2\mu E(X_i) + E\left(\mu^2\right)\right\}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(\mu_2' - 2\mu\mu + \mu^2\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(\mu_2' - \mu_2\right)$$

$$= \frac{1}{n}.n\mu_2' - \frac{1}{n}.n\mu_2 = \mu_2' - \mu_1' = \mu_2 = \sigma^2$$

Hence $\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$ is an unbiased estimator of σ^2.

2. Let $\bar{X}$ be the sample mean of samples of size n drawn at random from a population which is normally distributed with mean μ and variance σ^2. Find the standard error of the statistics $\bar{X}$. **[WBUT 2013]**

We ARE MAKAUTIANS

Answer:

Let the r.v. assume the values $X_1, X_2, \cdots\cdots, X_N$ and $x_1, x_2, \cdots\cdots, x_n$ constitute a sample (Note then $x_1, \cdots\cdots, x_n$ are n r.v.'s assuming value $X_1, X_2, \cdots\cdots, X_N$).

Then $\quad E(\bar{x}) = E\left(\dfrac{x_1 + x_2 + \cdots\cdots + x_n}{n}\right) = \dfrac{1}{n}\{E(x_1) + E(x_2) + \cdots\cdots + E(x_n)\}$

But $\quad E(x_i) = \dfrac{X_1}{N} + \dfrac{X_2}{N} + \cdots\cdots + \dfrac{X_N}{N},\ i = 1, 2, \ldots\ldots, n = \dfrac{1}{N}\sum_{j=1}^{N} X_j = \mu$

So, $\quad E(\bar{x}) = \dfrac{1}{n}\{\mu + \mu + \cdots\cdots + \mu\} = \dfrac{1}{n}\cdot n\mu = \mu$

$$\operatorname{var}(\bar{x}) = \operatorname{var}\left\{\frac{x_1 + x_2 + \cdots\cdots + x_n}{n}\right\}$$

$$= \frac{1}{n^2}\operatorname{var}(x_1) + \frac{1}{n^2}\operatorname{var}(x_2) + \cdots\cdots + \frac{1}{n^2}\operatorname{var}(x_n) + \frac{2}{n^2}\operatorname{cov}(x_1, x_2)$$

$$+ \frac{2}{n^2}\operatorname{cov}(x_1, x_3) + \cdots\cdots + \frac{2}{n^2}\operatorname{cov}(x_{n-1}, x_n)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{var}(x_i) \text{ since } x_1, x_2, \ldots\ldots, x_n \text{ are i.i.d.}$$

$$\operatorname{var}(x_i) = E(x_i^2) - \{E(x_i)\}^2 \text{ for } i = 1, 2, \ldots\ldots, n = E(x_i^2) - \mu^2$$

Now, $\quad E(x_i^2) = \dfrac{X_1^2}{N} + \dfrac{X_2^2}{N} + \cdots\cdots + \dfrac{X_N^2}{N} = \dfrac{1}{N}\sum_{j=1}^{N} X_j^2$.

Also, $\quad \sigma^2 = \dfrac{1}{N}\sum_{j=1}^{N} X_j^2 - \left\{\dfrac{1}{N}\sum_{j=1}^{N} X_j\right\}^2 = \dfrac{1}{N}\sum_{j=1}^{N} X_j^2 - \mu^2$

$\therefore \quad E(x_i^2) = \sigma^2 + \mu^2$

So, $\quad \operatorname{var}(x_i) = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$

and $\quad \operatorname{var}(\bar{x}) = \dfrac{1}{n^2}\{\sigma^2 + \sigma^2 + \cdots\cdots + \sigma^2\} = \dfrac{1}{n^2}\cdot n\sigma^2 = \dfrac{\sigma^2}{n}$

$\therefore \quad SE(\bar{x}) = \dfrac{\sigma}{\sqrt{n}}$.

We ARE MAKAUTIANS

3. Let $x_1, x_2, \ldots\ldots, x_n$ be the values of a random sample from an exponential population that is $f(x_i)=\frac{1}{\theta}e^{-\frac{x}{\theta}}$ for $x_i>0$. Then find the maximum likelihood estimator of its parameter θ. **[WBUT 2016]**

Answer:

The likelihood function is

$$L(x_1,x_2,\ldots\ldots,x_n;\theta)=\frac{1}{\theta}\cdot e^{-\frac{x_1}{\theta}}\cdot\frac{1}{\theta}e^{-\frac{x_2}{\theta}}\ldots\ldots\frac{1}{\theta}e^{-\frac{x_n}{\theta}}=\left(\frac{1}{\theta}\right)^n e^{-\frac{1}{\theta}(x_1+x_2+\ldots\ldots+x_n)}$$

$$\log L=-n\log\theta-\frac{1}{\theta}\left(\sum x_i\right)$$

$$\therefore\quad \frac{\partial}{\partial\theta}(\log L)=-\frac{n}{\theta}+\frac{1}{\theta^2}\left(\sum x_i\right)$$

Equating to zero we get, $-\frac{n}{\theta}+\frac{1}{\theta^2}\left(\sum x_i\right)=0$

or, $\theta=\frac{1}{n}\sum x_i=\bar{x}$

Clearly, $\frac{\partial^2}{\partial\theta}(\log L)<0$ at $\theta=\bar{x}$

Hence the maximum likelihood estimator of θ is $\bar{x}$

4. If a population has Poisson distribution with parameter λ, then show that the sample mean is the maximum likelihood estimate of λ. **[WBUT 2017]**

Answer:

The p.m.f. of the Poisson distribution $=f(x)=\frac{e^{-\lambda}\lambda^x}{x!},\ x=0,1,2,\ldots\ldots,\infty$

$\therefore$ The likelihood function is

$$L(x_1, x_2, \ldots\ldots, x_n, \lambda)=f(x_1, \lambda)\ldots\ldots f(x_n, \lambda)$$

$$=\frac{e^{-\lambda}\lambda^{x_1}}{x_1!}\cdot\frac{e^{-\lambda}\lambda^{x_2}}{x_2!}\cdot\ldots\ldots\cdot\frac{e^{-\lambda}\lambda^{x_n}}{x_n!}$$

$$=\frac{e^{-n\lambda}\lambda^{\Sigma x_i}}{x_1!x_2!\ldots\ldots x_n!}$$

$$\therefore\quad \log L=-n\lambda+n\bar{x}\log\lambda-\log(x_1!x_2!\ldots\ldots x_n!)$$

$$\therefore\quad \frac{\partial}{\partial\lambda}(\log L)=-n-\frac{n\bar{x}}{\lambda}-0$$

We ARE MAKAUTIANS

Equating to zero, we get, $-n+\frac{n\bar{x}}{\lambda}$

or, $\lambda=\bar{x}$

Again, $\frac{\partial^2}{\partial\lambda^2}(\log L)=\frac{-n\bar{x}}{\lambda^2}=-\frac{n}{\bar{x}}<0$ at $\lambda=\bar{x}$

Hence $\log L$ and therefore L is maximum at $\lambda=\bar{x}$.

Thus the maximum likelihood estimator of λ is $\bar{x}$.

5. If $x_1, x_2, x_3, x_4, x_5, x_6$ be an independent simple random sample from a normal population with unknown variance σ^2, find K so that $K\left[(X_1-X_2)^2+(X_3-X_4)^2+(X_5-X_6)^2\right]$ is an unbiased estimator of σ^2.

[WBUT 2017]

Answer:

Let μ denote the mean of the normal variate X

$$\text{var}(X)=\sigma^2$$

Since $x_1, x_2, \ldots\ldots x_n$ are independent, we have

$$E(X_iX_j)=E(X_i)E(X_j)=\mu^2 \text{ for } i\neq j$$

and $\quad E(X_i^2)-\{E(X_i)\}^2=\sigma^2$ for each i

or, $\quad E(X_i^2)=\sigma^2+\mu^2$

Since $K\left\{(X_1-X_2)^2+(X_3-X_4)^2+(X_5-X_6)^2\right\}$ is an unbiased estimator of σ^2, we have

$$E\left\{K\left[(X_1-X_2)^2+(X_3-X_4)^2+(X_5-X_6)^2\right]\right\}=\sigma^2$$

or, $\quad KE\left(X_1^2+X_2^2+X_3^2+X_4^2+X_5^2+X_6^2-2X_1X_2-2X_3X_4-2X_5X_6\right)=\sigma^2$

or, $\quad K\{E(X_1^2)+E(X_2^2)+E(X_3^2)+E(X_4^2)+E(X_5^2)+E(X_6^2)$

$$-2E(X_1X_2)-2E(X_3X_4)-2E(X_5X_6)\}=\sigma^2$$

or, $\quad K\left(6\sigma^2+6\mu^2-2\mu^2-2\mu^2-2\mu^2\right)=\sigma^2$

or, $\quad 6K=1$

or, $\quad K=\frac{1}{6}$

WE ARE MAKAUTIANS

6. The weights in pound of a sample of 12 packets of butter are 13, 7, 22, 15, 12, 18, 14, 21, 8, 17, 10, 23 taken at random from its population having standard deviation 5. Find 95% confidence interval for the mean of the population. [Given $Z_{0.05} = 1.96$]

[WBUT 2017]

Answer:

Given that, $\sigma = 5$, $Z_{0.05} = 1.96$, $n = 12$

$$\bar{x} = \frac{(13+7+22+15+12+18+14+21+8+17+10+23)}{12} = \frac{180}{12} = 15$$

$\therefore$ Confidence interval for the population mean μ is,

$$\left(\bar{x} - \frac{\sigma}{\sqrt{n}} \times 1.96, \bar{x} + \frac{\sigma}{\sqrt{n}} \times 1.96\right) = \left(15 - \frac{5}{\sqrt{12}} \times 1.96, 15 + \frac{5}{\sqrt{12}} \times 1.96\right) = (12.17, 17.82)$$

We ARE MAKAUTIANS

SMALL SAMPLES

☞ Chapter at a Glance

Hypothesis testing is a form of statistical inference that uses data from a sample to draw conclusions about a population parameter or a population probability distribution. First, a tentative assumption is made about the parameter or distribution. This assumption is called the null hypothesis and is denoted by H_0. An alternative hypothesis (denoted H_a), which is the opposite of what is stated in the null hypothesis, is then defined. The hypothesis-testing procedure involves using sample data to determine whether or not H_0 can be rejected. If H_0 is rejected, the statistical conclusion is that the alternative hypothesis H_a is true.

For example, assume that a radio station selects the music it plays based on the assumption that the average age of its listening audience is 30 years. To determine whether this assumption is valid, a hypothesis test could be conducted with the null hypothesis given as $H_0 : \mu = 30$ and the alternative hypothesis given as $H_a : \mu \neq 30$. Based on a sample of individuals from the listening audience, the sample mean age, $\bar{x}$, can be computed and used to determine whether there is sufficient statistical evidence to reject H_0. Conceptually, a value of the sample mean that is "close" to 30 is consistent with the null hypothesis, while a value of the sample mean that is "not close" to 30 provides support for the alternative hypothesis. What is considered "close" and "not close" is determined by using the sampling distribution of $\bar{x}$.

Ideally, the hypothesis-testing procedure leads to the acceptance of H_0 when H_0 is true and the rejection of H_0 when H_0 is false. Unfortunately, since hypothesis tests are based on sample information, the possibility of errors must be considered. A type I error corresponds to rejecting H_0 when H_0 is actually true, and a type II error corresponds to accepting H_0 when H_0 is false. The probability of making a type I error is denoted by α, and the probability of making a type II error is denoted by β.

In using the hypothesis-testing procedure to determine if the null hypothesis should be rejected, the person conducting the hypothesis test specifies the maximum allowable probability of making a type I error, called the level of significance for the test. Common choices for the level of significance are $\alpha = 0.05$ and $\alpha = 0.01$. Although most applications of hypothesis testing control the probability of making a type I error, they do not always control the probability of making a type II error. A graph known as an operating-characteristic curve can be constructed to show how changes in the sample size affect the probability of making a type II error.

A concept known as the p-value provides a convenient basis for drawing conclusions in hypothesis-testing applications. The p-value is a measure of how likely the sample results are, assuming the null hypothesis is true; the smaller the p-value, the less likely the sample results. If the p-value is less than α, the null hypothesis can be rejected; otherwise, the null hypothesis cannot be rejected. The p-value is often called the observed level of significance for the test.

We ARE MAKAUTIANS

A hypothesis test can be performed on parameters of one or more populations as well as in a variety of other situations. In each instance, the process begins with the formulation of null and alternative hypotheses about the population. In addition to the population mean, hypothesis-testing procedures are available for population parameters such as proportions, variances, standard deviations, and medians.

Hypothesis tests are also conducted in regression and correlation analysis to determine if the regression relationship and the correlation coefficient are statistically significant (see below Regression and correlation analysis). A goodness-of-fit test refers to a hypothesis test in which the null hypothesis is that the population has a specific probability distribution, such as a normal probability distribution. Nonparametric statistical methods also involve a variety of hypothesis-testing procedures.

Multiple Choice Type Questions

1. If $H_1(\mu > 60)$ is an alternative hypothesis, then the null hypothesis is **[WBUT 2012]**

a) $H_0(\mu < 60)$ b) $H_0(\mu \geq 60)$ c) $H_0(\mu \leq 60)$ d) none of these

Answer: (d)

3. In 'Goodness of fit' which of the following is used as test statistic **[WBUT 2014]**

a) normal variate b) *t* variate c) Poisson variate d) X^2 variate

Answer: (d)

3. A null hypothesis is a statistical hypothesis which is setup and whose validity is tested for possible **[WBUT 2015]**

a) acceptance b) rejection c) testing d) none of these

Answer: (b)

4. The power of the test in case of testing of hypothesis, is **[MODEL QUESTION]**

a) $1 - P$ (Type I Error) b) P (Type I Error) P (Type II Error)

c) $1 - P$ (Type II Error) d) $1 - P$ (Type I Error) P (Type II Error)

Answer: (c)

5. Which of the following is type II error? **[MODEL QUESTION]**

a) The error of accepting H_0 when H_0 is true

b) The error of rejecting H_0 when H_0 is false

c) The error of accepting H_0 when H_0 is false

d) The error of rejecting H_0 when H_0 is true

Answer: (c)

WE ARE MAKAUTIANS

Short Answer Type Questions

1. A random sample with observations 65, 71, 64, 71, 70, 69, 64, 63, 67, 68 is drawn from a normal population with variance 7.056. Test the hypothese's that the population mean is 69 at 1% level of significance. [WBUT 2012]

[Given that $P(0<z<2.58)=0.495$]

Answer:

Here the null hypothesis is $H_0(\mu=69)$

We have $\overline{X}=67.2$ $\sigma^2=7.056$

The test statistic, $z=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}=\dfrac{67.2-69}{2.07/\sqrt{10}}=\dfrac{-1.8}{2.07/3.162}=-\dfrac{1.8}{0.65}=-2.769$

2. Suppose that 100 tires made by a certain manufacturer lasted on the average 21819 miles with a standard deviation of 1295 miles. Test the null hypothesis $\mu=22000$ miles against the alternate hypothesis $\mu<22000$ miles at the 0.05 level of significance. [WBUT 2014]

Answer:

Set the null hypothesis H_0 as $H_0(\mu=22,000)$ against $H_1(\mu<22,000)$

The test statistics

$$Z=\frac{21819-22000}{1295/\sqrt{100}}\sim N(0,1) \text{ approx}$$

$$=\frac{-181}{129.5}=-1.398$$

As this value is greater than 1.645, we accept the null hypothesis.

3. A random sample of size 20 from a normal population gives the sample mean of 42 and the sample standard deviation of 6. Test the hypothesis that the population mean is 44. Value of t distribution with 19 degrees of freedom at 5% level is 2.09. [WBUT 2015]

Answer:

Null hypothesis is that the population mean is 44. $H_0(\mu=44)$. Alternative hypothesis is that the population mean is different from 44. $H_2(\mu\neq 44)$, i.e., both-sided alternatives.

WE ARE MAKAUTIANS

Since the population s.d. (σ) is not known, we use t test. Here, $n = 20, \bar{x} = 42$ and $S = 6$. Therefore, $t = \dfrac{42-44}{\frac{6}{\sqrt{19}}} = -1.45$

Degrees of freedom $= (20-1) = 19$

From the tables of t distribution, we find for 19 d.f. the percentage points of t are respectively $t_{.025} = 2.09$ and $t_{.005} = 2.86$. We have to use a two-tailed test, because the alternatives are both-sided. Since $|t| = 1.45$ is less than the 5% tabulated value corresponding to the two tails (viz. 2.09) there is no reason to reject the null hypothesis at 5% level of significance and we conclude that the population mean may be 44.

[99% confidence limits for μ are $42 \pm 2.86 \times \left(\dfrac{6}{\sqrt{19}}\right) = 42 \pm 3.94 = 38.06$ and 45.94]

4. What are Type I error and Type II error. **[MODEL QUESTION]**

Answer:

The error made in rejecting the null hypothesis H_O when it is true is called Type I error.

The error made in accepting the null hypothesis H_O when it is false is called Type II error.

5. If $x \geq 1$ is the critical region for testing the null hypothesis $H_0 : \theta = 2$ against the alternative hypothesis $H_1 : \theta = 1$, on the basis of a single observation from the population $f(x, \theta) = \theta \exp[-\theta_x]$, $0 \leq x < \infty = 0$, elsewhere, then find the type I error and type II error. **[MODEL QUESTION]**

Answer:

P (Type I error) = P (Rejecting H_0 when H_0 is true)

$$= P(x \geq 1 \text{ where } \theta = 2)$$

$$= \int_1^\infty 2e^{-2x} dx = \left[2\frac{e^{-2x}}{-2}\right]_1^\infty = e^{-2}$$

P (Type II error) = P (Accepting H_0 when H_0 is false)

$$= P(x < 1 \text{ when } \theta = 1) = \int_0^1 e^{-x} dx = \left[-e^{-x}\right]_0^1 = 1 - \frac{1}{e}$$

We ARE MAKAUTIANS

6. To test the unbiasedness of a die it is thrown six times and is accepted as unbiased if not more than one six is obtained. Find the probability of type I error.

[MODEL QUESTION]

Answer:

P (Type I error) = P (Rejecting H_0 when H_0 is true)

= P (More than one six is obtained in 6 throws with P=1/6)

= 1-P (Less than one six is obtained when P = 1/6)

$$= 1 - {}^6C_0\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^6 - {}^6C_1\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^5$$

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

7. The following table gives the number of road accidents that occurred during various days of the week in Kolkata. Test whether the accidents are uniformly distributed over the week.

[MODEL QUESTION]

Day	Sun	Mon	Tues	Wed	Thurs	Fri	Sat
No. of accidents	13	14	19	12	11	15	14

[Given $\chi^2_{0.05}(v=6)=12.59$ **.]**

Answer:

Here the null hypothesis H_0 is that the accidents are uniformly distributed over the week.

Thus we have

Day	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Total
f_0	13	14	19	12	11	15	14	98
f_e	14	14	14	14	14	14	14	98

$$X_6^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{1}{4}(1+0+25+4+9+1+0) = \frac{40}{14} = 2.8$$

Since the observed value of χ^2 is less than the tabulated value 12.59, we conclude that the accidents can be taken to be uniformly distributed.

8. The length of life X of certain computers is approximately normally distributed with mean 800 hours and st. dev. 40 hrs. If a random sample of 30 computers has an average life of 788 hrs, test the null hypothesis that $\mu = 800$ against the alternative hypothesis $\mu \neq 800$ hrs at 5% level of significance (Given $Z_{0.025} = 1.96$)

[MODEL QUESTION]

We ARE MAKAUTIANS

Answer:

Here the null hypothesis $H_0 : \mu = 800$ against $H_1 : \mu \neq 800$

So it is a case of two tailed test.

The test statistic Z is given by

$$Z = \frac{\bar{X} - E(\bar{X})}{S \cdot E \cdot (\bar{X})} = \frac{788 - 800}{40/\sqrt{30}}$$

Since $E(\bar{X}) = \mu$ and $S \cdot E(\bar{X}) = \frac{\sigma}{\sqrt{n}} = \frac{3\sqrt{30}}{10} = 1.643$

Since this observed value is numerically less than 1.96, the null hypothesis is accepted.

9. In order to test whether a coin is perfect the coin is tossed 5 times. The null hypothesis of perfectness is rejected if more than 4 heads are obtained. What is the probability of Type 1 error? Find the probability of Type II error when corresponding probability of head is 0.2. **[MODEL QUESTION]**

Answer:

P (Type I error) = P (Rejecting H_0 when H_0 is true)

= P (getting 5 heads in 5 tosses of the coin)

$$= {}^5C_5 \left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

P (Type II error) = P (Accepting H_0 when H_1 is true)

= P (getting at most 4 heads when p = 0.2)

= 1 - P (getting 5 heads when p = 0.2)

$$= 1 - {}^5C_5 (0.2)^5$$

$$= 1 - (0.2)^5$$

10. The following table gives the number of aircraft accidents that occurred during various days of by the week. Test whether the accidents are uniformly distributed over the week.

Days	Sun	Mon	Tues	Wed	Thurs	Fri	Sat
No. of accidents	13	14	19	12	11	15	14

[Given $\chi^2_{0.5,\,6} = 12.59$.] **[MODEL QUESTION]**

Answer:

Assume that the accidents are uniformly distributed. Thus the null hypothesis

H_0 : (The accidents are uniformly distributed) against the alternative hypothesis

H_1 (they are not).

WE ARE MAKAUTIANS

The test statistics is therefore

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

where $f_e = \dfrac{13+14+19+12+11+15+14}{7} = \dfrac{98}{7} = 14$

$$= \frac{(13-14)^2}{14} + \frac{(14-14)^2}{14} + \frac{(19-14)^2}{14} + \frac{(12-14)^2}{14} + \frac{(11-14)^2}{14} + \frac{(15-14)^2}{14} + \frac{(14-14)^2}{14}$$

$$= \frac{1}{14}(1+0+25+4+9+1+0) = \frac{40}{14} = 2.857 \text{ with 6 df.}$$

Since the observed values 2.857 is much less than the tabulated value 12.59 of χ^2 at 5% level of significance, we accept the null hypothesis and conclude that the accidents are uniformly distributed.

11. In a certain city 100 men in a sample of 400 were found to be smokes. In another city the number of smokes was 300 in a random sample of 800. Does this indicate that there is a greater proportion of smokes in the second city than in the first city? [MODEL QUESTION]

Answer:

Here $n_1 = 400,\ p_1 = 100/400 = 1/4;\ n_2 = 800,\ p_2 = 300/800 = 3/8$

We set up

$$\therefore \quad p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{400 \times \frac{1}{4} + 800 \times \frac{3}{8}}{1200} = \frac{1}{3} \quad \therefore\ q = \frac{2}{3}$$

$H_0 : P_1 = P_2$ against $H_1 : P_1 < P_2$ where P_1 and P_2 are the population proportions.

The test statistic

$$z = \frac{p_1 - p_2}{pq\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{\frac{1}{4} - \frac{3}{8}}{\frac{1}{3} \times \frac{2}{3}\sqrt{\frac{1}{400} + \frac{1}{800}}} = -9.186$$

Since this observed value is much less than the tabulated value of z at 1% level, we reject the null hypothesis and conclude that there is a greater proportion of smokers in the second city than in the first city.

We ARE MAKAUTIANS

12. The proportion of defective items in a large lot of items is *p*. To test the hypothesis $p = 0.2$, we take a random sample of 3 items and accept the hypothesis if the number of defectives in the sample is 6 or less. Find the probability of Type-I – error of the test. Find the probability of Type-II – error if $p = 0.3$?

[MODEL QUESTION]

Answer:

Here, $p = 0.2,\ n = 3$

$$P(\text{Type I error}) = P(\text{rejecting } H_0 \text{ when it is free})$$
$$= P(\text{rejecting when } p = 0.2)$$
$$= P(\text{getting 7 or 8 defectives when } p = 0.2)$$
$$= {}^8C_7 (0.2)^7 \cdot (0.8)^1 + {}^8C_8 (0.2)^8 \cdot (0.8)^0$$
$$= 8(0.2)^7 (0.8) + (0.2)^8.$$

$$P(\text{Type II error}) = P(\text{accepting } H_0 \text{ when it is false})$$
$$= P(\text{getting less than or equal to 6 heads when } H_1 \text{ is true})$$
$$= 1 - P(\text{getting 7 or 8 heads when } p = 0.3)$$
$$= 1 - {}^8C_7 (0.3)^7 (0.7)^1 - {}^8C_8 (0.3)^8 (0.7)^0.$$

13. A random sample with observations 65, 71, 64, 71, 70, 69, 64, 63, 67, 68 is drawn from a normal population with standard deviation $\sqrt{7.056}$. Test the hypothesis that the population mean is 69 at 1% level of significance.

[Given: $P(0 < Z < 2.58) = 0.495$]. **[MODEL QUESTION]**

Answer:

Here

$$\overline{X} = \frac{(65+71+64+71+70+69+64+63+67+68)}{10} = 67.2$$

We set the null hypothesis is $H_0(\mu = 69)$ against the alternative hypothesis $H_1(\mu \neq 69)$.

The test statistic

$$Z = \frac{\overline{X} - E(\overline{X})}{\frac{\sigma}{\sqrt{n}}} = \frac{67.2 - 69}{\frac{\sqrt{7.056}}{\sqrt{10}}} = \frac{-1.8}{0.84} = -2.143.$$

Since the observed value of $|Z| < 2.58$, we accept the null hypothesis and conclude therefore the population mean may be taken as 69 at 1% level.

We ARE MAKAUTIANS

14. Intelligence tests on two groups of boys and girls gave the following results:

	Mean	SD	N
Boys	70	20	250
Girls	75	15	150

Is there any significant difference in the mean scores obtained by boys and girls?

[MODEL QUESTION]

Answer:

We set the null hypothesis H_0 as H_0 (There is no difference in the mean scores by boys and girls) against the alternative hypothesis H_1 as Hi (There is difference in the said mean scores)

We now have the test statistic Z as

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{70-75}{\sqrt{\frac{400}{250} + \frac{225}{150}}} = \frac{-5}{\sqrt{1.6+1.5}} = \frac{-5}{\sqrt{3.1}} = \frac{-5}{1.76} = -2.84$$

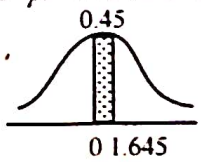
As the value of Z is numerically greater than 2.58, we reject the null hypothesis and conclude that there is significant difference in the mean scores.

Long Answer Type Questions

1. a) A machine part was designed to withstand an average pressure of 120 units. A random sample of size 100 from a large batch was tested and it was found that the average pressure which these parts can withstand is 105 units with a standard deviation of 20 units. Test at 5% level of significance whether the batch meet the specification. Suppose the population has normal distribution and given that $\phi(1.645) = 0.45$. **[WBUT 2015]**

Answer:

Here the null hypothesis $H_0 : (\mu = 120)$ against $H_1 (\mu < 120)$, $n = 100$ where μ denotes the average pressure that the machine part can withstand.



The test statistic $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{105-120}{20/\sqrt{100}} = \frac{-15}{2} = -7.5$

Since the deserved value -7.5 is less than the tabulated value of Z at 5% level, we reject the null hypothesis and conclude that the average pressure that the machine part can withstand is less than 120.

We ARE MAKAUTIANS

b) Survey of 320 families with 5 children each revealed the following distribution:

No. of boys:	5	4	3	2	1	0
No. of girls:	0	1	2	3	4	5
No. of family:	14	56	110	88	40	12

Is the result consistent with the hypothesis that male and female births are equally probable? The 5% value of X^2 with 5 d.o.f. is 11.07. **[WBUT 2015]**

Answer:

We set the null hypothesis as H_0 (male & female births are equally probable) of 5 children,

$$P(\text{all are boys}) = \left(\frac{1}{2}\right)^5$$

$$P(4 \text{ boys \& } 1 \text{ girl}) = {}^5C_4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^1 = 5\left(\frac{1}{2}\right)^5$$

$$P(3 \text{ boys \& } 2 \text{ girls}) = {}^5C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 = 10\left(\frac{1}{2}\right)^5$$

$$P(2 \text{ boys \& } 3 \text{ girls}) = {}^5C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3 = 10\left(\frac{1}{2}\right)^5$$

$$P(1 \text{ boy \& } 4 \text{ girls}) = {}^5C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^4 = 5\left(\frac{1}{2}\right)^5$$

$$P(\text{no boy \& } 5 \text{ girls}) = {}^5C_0 \left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^5 = \left(\frac{1}{2}\right)^5$$

Therefore of 320 families,

No. of families having all five boys $= \left(\frac{1}{2}\right)^5 \times 320 = 10$

No. of families having 4 boys & 1 girl $= 5\left(\frac{1}{2}\right)^5 \times 320 = 50$

No. of families having 3 boys & 2 girls $= 10\left(\frac{1}{2}\right)^5 \times 320 = 100$

No. of families having 2 boys & 3 girls $= 10\left(\frac{1}{2}\right)^5 \times 320 = 100$

No. of families having 1 boy & 4 girls $= 5\left(\frac{1}{2}\right)^5 \times 320 = 50$

We ARE MAKAUTIANS

No. of families having only girls $= \left(\frac{1}{2}\right)^5 \times 320 = 10$

So we have

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{(14-10)^2}{10} + \frac{(56-50)^2}{50} + \frac{(110-100)^2}{100} + \frac{(88-100)^2}{100} + \frac{(40-50)^2}{50} + \frac{(12-10)^2}{10} = 7.16$$

As the observed value 7.16 is less than the tabulated value 11.07 of χ^2 with 5 d.f. We accept the null hypothesis and conclude that male and female births are equally probable.

2. A random sample with observations 65, 71, 64, 71, 70, 69, 64, 63, 67, 68 is drawn from a normal population with a standard deviation $\sqrt{7.056}$. Test the hypothesis that the population mean is 69 at 1% level of significance.

[Given that $P(0 < z < 2.58) = 0.495$]. **[WBUT 2016]**

Answer:

Here $\overline{X} = (65+71+64+71+70+69+64+63+67+68)/10 = 67.2$

We set the null hypothesis H_o as $H_o(\mu = 69)$ against $H_1(\mu \neq 69)$

The test statistics is

$$Z = \frac{\overline{X} - E(\overline{X})}{SE(\overline{X})} = \frac{67.2 - 69}{\sqrt{7.056}/10} = 6.776$$

Since this value is greater than 2.58, it is significant at 1% level. So we reject the null hypothesis.

3. The following figures show the distribution of digits in numbers chosen at random from a telephone directory:

Digits	0	1	2	3	4	5	6	7	8	9	Total
Frequency	1026	1107	997	966	1075	933	1107	972	964	853	10,000

Test whether the digits may be taken to occur equally frequently in the directory.

[The tabulated $\chi^2_{0.05}$ for 9 d.f. $= 16.919$] **[WBUT 2017]**

Answer:

We set the null hypothesis H_0 as

H_0 (The digits occur equally frequently) against H_1 (They are not)

WE ARE MAKAUTIANS

Here $\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{(1026-1000)^2}{1000} + \frac{(1107-1000)^2}{1000}$

$$+\frac{(997-1000)^2}{1000} + \frac{(966-1000)^2}{1000} + \frac{(1075-1000)^2}{1000}$$

$$+\frac{(933-1000)^2}{1000} + \frac{(1107-1000)^2}{1000} + \frac{(972-1000)^2}{1000}$$

$$+\frac{(964-1000)^2}{1000} + \frac{(853-1000)^2}{1000}$$

$$= \frac{1}{1000}\{26\times26+107\times107+3\times3+34\times34+25\times25+67\times67$$

$$+107\times107+28\times28+56\times56+147\times147\}$$

$$= \frac{1}{1000}\{676+11,449+9+1,156+625+4,489+11,449$$

$$+784+3,136+21,609\}$$

$$= 55.382$$

Since $\chi^2_{0.05}$ for 9 d.f. = 16.919 and observed $\chi^2 = 55.382 > 16.919$, we reject the null hypothesis and conclude that the digits do not occur equally frequently in the directory.

We ARE MAKAUTIANS