



महाराष्ट्र राज्य तंत्रशिक्षण मंडळ, मुंबई
(स्वायत्व) (ISO 9001:2015) (ISO/IEC 27001:2013)

अभियांत्रिकी आणि तंत्रज्ञान पदविका

शिक्षण पुस्तिका
(Learning Material)

**PROBABILITY AND
STATISTICS**

(314316)

K Scheme

प्राॅबेबिलिटी अॅण्ड
स्टॅटिस्टिक्स

संगणक अभियांत्रिकी गट
(के स्कीम)

मराठी-इंग्रजी (द्विभाषिक) माध्यम
(अभियांत्रिकी व तंत्रज्ञानातील चौथे सत्र पदविका)

शिक्षण पुस्तिका
(Learning Material)
प्रोबॅबिलिटी अँड स्टॅटिस्टिक्स
Probability and Statistics
(314316)
संगणक अभियांत्रिकी गट
(के-स्किम)
K-Scheme
मराठी-इंग्रजी (द्विभाषिक) माध्यम
(अभियांत्रिकी व तंत्रज्ञानातील चौथे सत्र पदविका)



महाराष्ट्र राज्य तंत्रशिक्षण मंडळ, मुंबई
(स्वायत्त)(ISO 9001:2015) (ISO/IEC 27001:2013)

प्रोबॅबिलिटी अँड स्टॅटिस्टिक्स
Probability and Statistics
(314316)

मार्गदर्शक

प्रा. विजय नामदेवराव कुकरे
विभाग प्रमुख, संगणक विभाग

समन्वयक

प्रा. मंगला माळकर
विभाग प्रमुख, संगणक विभाग

संकलक

प्रा. चारुलता पाटील - व्याख्याता
प्रा. शिल्पा केशरवाणी- व्याख्याता
प्रा. श्रीरक्षा अपरंजी - व्याख्याता
प्रा. माधुरी अभंग- व्याख्याता
प्रा. सोनाली मराठे- व्याख्याता



महाराष्ट्र राज्य तंत्र शिक्षण मंडळ

(स्वायत्त) (ISO: ९००१:२०१५) (ISO/IES: २७००१-२०१३)

शासकीय तंत्रनिकेतन इमारत, चौथा मजला, ४९, खेरवाडी, बांद्रा (पूर्व), मुंबई - ४०० ०५१.

दूरध्वनी क्र.: ०२२-६२५४२१७० / १६१

Email : director@msbte.com

Web : www.msbte.org.in



प्रास्ताविक

महाराष्ट्र राज्यातील पदविका स्तरावरील तंत्रशिक्षणामध्ये विद्यार्थ्यांचे रोजगार कौशल्य विकसित करून विद्यार्थ्यांचा सर्वांगीण विकास घडवून आणण्याकरिता महाराष्ट्र राज्य तंत्रशिक्षण मंडळ कटिबद्ध आहे. उद्योगधंद्यातील बदलत्या तंत्रज्ञानाशी संबंधित गरजा लक्षात घेऊन महाराष्ट्र राज्य तंत्र शिक्षण मंडळाकडून पदविका अभ्यासक्रम वेळोवेळी अद्यावत करण्यात येतो. अभियांत्रिकी पदविका अभ्यासक्रम शिकत असतांना संकल्पनात्मक ज्ञान, सुसंगत संदर्भ, प्रश्न विचारणे, विश्वसनीय पुरावे, कारणमीमांसा आणि सुस्पष्ट निकष यांचा वापर करून अर्थाची उकल करण्याची, विश्लेषण व मूल्यमापन करण्याची तसेच तर्काने अनुमान काढण्याची क्षमता म्हणजेच चिकित्सक विचार विद्यार्थ्यांमध्ये अधिक दृढ होतील असा मला विश्वास आहे. जेव्हा विद्यार्थी ज्ञान मिळवण्याच्या माध्यमाशी पूर्णपणे परिचित आणि सोयीस्कर असतात, तेव्हा त्यांच्यासाठी वर्गातील चर्चेत भाग घेणे सोपे होते, संकल्पनात्मक व सैद्धांतिक बाबींचे आकलन परिपूर्ण होते, संज्ञानात्मक क्षमता सुधारते आणि त्यांचा आत्मविश्वास देखील वाढतो या सर्व गोष्टींचा विचार करून मंडळाकडून शैक्षणिक सामुग्रीची निर्मिती करण्यात आलेली आहे. भारत देश हा खेड्यापाड्यातून विकसित झालेला देश असून ग्रामीण भागातील विद्यार्थ्यांना तांत्रिक शिक्षण घेतांना भाषेचा अडसर न येता तांत्रिक बाबींचा आशय समजून घेणे शक्य होईल या दृष्टिकोनातून महाराष्ट्र राज्य तंत्र शिक्षण मंडळाने पदविका स्तरावरील तांत्रिक शिक्षणाकरिता विद्यार्थ्यांना मराठी-इंग्रजी द्विभाषिक माध्यमाचा पर्याय शैक्षणिक वर्ष २०२१-२२ पासून उपलब्ध करून दिलेला आहे.

राष्ट्रीय शैक्षणिक धोरण-२०२० प्रादेशिक भाषेतील शिक्षणास प्रोत्साहन देते, ज्यामुळे विद्यार्थ्यांना तांत्रिक अभ्यासक्रमांसाठी प्रादेशिक भाषांतुन शिक्षणाचे माध्यम निवडता येते. सदर धोरणामुळे प्रादेशिक भाषांमध्ये तांत्रिक सामग्री आणि अभ्यास सामग्रीचा विकास आणि भाषांतर निर्माण करण्याची आवश्यकता आहे. त्यास अनुसरून मंडळाने मराठी-इंग्रजी द्विभाषिक माध्यमाचा पर्याय द्वितीय व तृतीय वर्षाकरिताही उपलब्ध करून देण्यात आला आहे. तसेच त्याकरिताची शैक्षणिक सामग्रीही संबंधीत भागधारकरांना उपलब्ध करून देण्यात येत आहे.

पदविका स्तरावरील तंत्रशिक्षण अधिक दर्जेदार करण्यासाठी महाराष्ट्रातील अनुभवी व तज्ञ अध्यापकांनी व्यवहारिक मराठी भाषा व इंग्रजी भाषेतील तांत्रिक शब्दावली यांचा वापर करून मराठी - इंग्रजी भाषेचा सुवर्णमध्य साधण्याचा प्रयत्न केलेला आहे. मंडळाच्या स्तरावर गठीत सुकाणू समितीमार्फत सदर शैक्षणिक सामुग्रीचा दर्जा, तसेच इतर बाबींची तपासणी करण्यात आलेली आहे. त्यामुळे सदर शैक्षणिक सामुग्री अधिक सम्पन्न झालेली असून विद्यार्थी त्यांच्या व्यक्तिमत्त्वाचा सुसंवादी आणि सर्वांगीण विकास साधतील. परिणामतः विश्वस्तरीय मनुष्यबळाच्या गरजा पूर्ण करण्यात महाराष्ट्र राज्य अग्रेसर राहील व पर्यायाने राष्ट्रनिर्मिती करिता निश्चितच हातभार लागेल, असा मला विश्वास आहे.

अभियांत्रिकी पदविका अभ्यासक्रमातील प्रमुख विषयांची मराठी-इंग्रजी द्विभाषिक शैक्षणिक सामुग्री बनविण्यासाठी अध्यापक व सुकाणू समितीचे सदस्य यांनी दर्शविलेले समर्पण व वचनबद्धता कौतुकास पात्र आहे, या सर्वांचे मी मनः पूर्वक अभिनंदन करतो !


(प्रमोद नाईक)

संचालक

म. रा. तंत्र शिक्षण मंडळ, मुंबई.

अनुक्रमणिका

अ.नु.	युनिटचे नाव	पृष्ठ क्रमांक
1	प्रोबॅबिलिटी (Probability)	1
2	रँडम व्हेरिएबल अँड प्रोबॅबिलिटी डिस्ट्रिब्युशन (Random Variable And Probability Distribution)	28
3	नेटवर्क एनालिसिस (Network Analysis)	51
4	बायव्हेरिएट फ्रिक्वेन्सी डिस्ट्रीब्युशन अँड कोरिलेशन (Bivariate Frequency Distribution And Correlation)	79
5	रिग्रेशन एनालिसिस (Regression Analysis)	91

युनिट-1 प्रोबॅबिलिटी (Probability)

विषय निष्पत्ती (Course Outcome):

CO1: अभियांत्रिकी समस्यांचे समाधान करण्यासाठी संभाव्यतेच्या मूलभूत तत्वांचा वापर करा.

घटक निष्पत्ती (Theory Learning Outcome):

1. यादृच्छिक प्रयोगाचे परिणाम आणि नमुना अवकाश (Sample Space) स्पष्ट करा.
2. संभाव्यतेच्या विविध प्रकारच्या घटनांमध्ये फरक स्पष्ट करा.
3. संभाव्यतेच्या गुणधर्मांवर आधारित उदाहरणे सोडवा.
4. व्हेन आकृती (Venn Diagram) वापरून उदाहरणे सोडवा.
5. सशर्त संभाव्यतेवर (Conditional Probability) आधारित उदाहरणे सोडवा.
6. बेयसच्या सिद्धांतावर (Bayes' Theorem) आधारित उदाहरणे (प्रुफ) सोडवा.

1.1 डेफिनेशन आणि टाईप्स ऑफ इव्हेंट (Definition and types of events)

(घटनांची व्याख्या आणि प्रकार):

प्रयोग आणि चाचणी (Experiment and Trial): दोन किंवा अधिक परिणाम देणारी कोणतीही क्रिया म्हणजे प्रयोग आणि प्रयोग करणे म्हणजे चाचणी.

रँडम एक्सपेरिमेंट (Random Experiment): रँडम एक्सपेरिमेंट म्हणजे असा प्रयोग ज्यामध्ये सर्व संभाव्य परिणाम माहित असतात आणि ते समान परिस्थितीत पुनरावृत्ती करता येतात परंतु कोणत्याही विशिष्ट चाचणीचे निकाल आधीच सांगणे शक्य नसते.

उदाहरण: नाणे फेकणे (Tossing a coin- टॉसिंग अ कॉईन)

फासा फेकणे (Throwing a die- थ्रोइंग अ डाय)

सॅम्पल स्पेस (sample space): रँडम एक्सपेरिमेंटच्या संभाव्य परिणामांना मूलभूत परिणाम म्हणतात आणि सर्व मूलभूत परिणामांच्या संचाला सॅम्पल स्पेस म्हणतात. सॅम्पल स्पेस सामान्यतः S या अक्षराने दर्शविली जाते.

उदाहरण: दोन निष्पक्ष (fair- फेअर) नाणी फेकताना, सॅम्पल स्पेस $S = \{HH, HT, TH, TT\}$ असते

डाय टाकताना सॅम्पल स्पेस $S = \{1, 2, 3, 4, 5, 6\}$ असते.

1.1.1 इव्हेंट (Event): सॅम्पल स्पेस S च्या कोणत्याही उपसमूहाला (subset- **सबसेट**) इव्हेंट म्हणतात.

समजा, एक निष्पक्ष (fair) सहा बाजू असलेला फास टाकला जातो. संपूर्ण संभाव्य निकालांचे एकूण संकलन (Sample Space) म्हणजेच

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

समजा, E ही इव्हेंट "फासावर सम संख्या येणे" अशी परिभाषित (defined) केली आहे.

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

म्हणजेच, E ही इव्हेंट नमुना जागेचा (Sample Space) उपसंच (Subset) आहे आणि हा फासा टाकण्याच्या निकालांपैकी एक आहे.

ऑकरन्स ऑफ इव्हेंट (Occurrence of an Event):

इव्हेंट E घटित झाली असे म्हणतात **जेव्हा निकाल** ω (ओमेगा) **हा** E **च्या अंतर्गत असेल**

$$\omega \in E$$

म्हणजेच, दुसऱ्या शब्दांत, जर प्रयोगाचा निकाल निरीक्षित केल्यानंतर तो निकाल E मध्ये असला, तर E इव्हेंट घडल्याचे मानले जाते.

उदाहरण: एका निष्पक्ष सहा बाजूच्या फासावर (डायवर) 4 येण्याची इव्हेंट (A) गृहित धरा. जर फास टाकल्यावर 4 आला, तर इव्हेंट A घडली. 52 पत्त्यांच्या संचामधून (Deck) एक पत्ता निवडला जातो. गृहित धरा, इव्हेंट B = "राजा (King) काढणे". जर निवडलेला पत्ता राजा असेल, तर इव्हेंट B घडली.

नॉन ऑक्करन्स ऑफ इव्हेंट (Non-occurrence of an Event):

जर प्रयोग करून मिळालेला निकाल दिलेल्या इव्हेंट संच (E) मध्ये नसेल, तर ती इव्हेंट घडली नाही असे म्हणतात.

उदाहरण: इव्हेंट A = "फास टाकल्यावर 4 येणे". जर फास टाकल्यावर 2, 3, 5 किंवा 6 आला, तर इव्हेंट A घडली नाही.

इव्हेंट B = "राजा (King) काढणे".

जर काढलेला पत्ता क्वीन (Queen), एस (Ace) किंवा इतर कुठलाही राजा नसलेला पत्ता असेल, तर इव्हेंट B घडली नाही.

कॉम्प्लिमेंट ऑफ इव्हेंट (Complement of an Event)(घटनेचा पूरक)

समजा, A ही इव्हेंट आहे जी S या संपूर्ण नमुना जागेचा एक उपसंच आहे. नमुना जागेतील (S) असे सर्व संभाव्य निकाल जे A मध्ये नाहीत, ते A च्या पूरक घटनेत येतात. A ची पूरक इव्हेंट A^c किंवा A^c ने दर्शविली जाते.

1.1.2 संभाव्यतेची व्याख्या (Definition of Probability)

एखाद्या घटनेची (A) संभाव्यता ही त्या घटनेच्या घडण्याच्या एकूण शक्य मार्गांची संख्या, एकूण संभाव्य निकालांच्या संख्येने भाग दिल्यावर मिळते.

$$P(A) = (\text{इव्हेंट A घडण्याच्या शक्य मार्गांची संख्या}) / (\text{A घटनेच्या घडण्याचे एकूण मार्ग})$$

आता आम्हाला माहित आहे की कोणतीही इव्हेंट (Event) नमुना जागेचा (sample space) उपसंच आहे. त्यामुळे रँडम एक्सपेरिमेंट करताना विविध घटनांच्या प्रकारांमधला फरक आपल्याला माहित पाहिजे हे आपल्याला सोप्या आणि गुंतागुंतीच्या रँडम एक्सपेरिमेंटांसाठी संभाव्यता (Probability) मोजणे सोपे होते. आम्हाला माहित आहे की इव्हेंट (इव्हेंट्स)मुळात संच असतात, जेणेकरून ते त्यांच्याकडे असलेल्या घटकांवर (Elements) आधारित त्यांचे वर्गीकरण (Classification) करता येते..

1.1.3 घटनांचे विविध प्रकार खालील प्रमाणे आहे:

i. एकघटकी इव्हेंट (Simple Event): घटनेमध्ये एकल बिंदू किंवा नमुना जागेतून एकच परिणाम असतो; त्याला एकघटकी इव्हेंट म्हणतात.

उदाहरण: फास फेकल्यानंतर 2 पेक्षा कमी आकडा हि इव्हेंट, ज्याला $E = \{1\}$ म्हणून दर्शविले जाते, हे एका एकघटकी इव्हेंटचे उदाहरण आहे.

ii. संयुक्त इव्हेंट (Compound Event- कंपाउंड इव्हेंट): संयुक्त इव्हेंट साध्या घटनेच्या अगदी उलट आहे म्हणजेच. कोणत्याही घटनेमध्ये नमुना जागेच्या एकापेक्षा जास्त नमुना बिंदू असतो अशा घटनेला संयुक्त इव्हेंट म्हणतात.

समजा $S = \{44, 55, 65, 75, 85, 95\}$

$$E_1 = \{55, 65, 75\} \text{ and } E_2 = \{44, 95\}$$

$\therefore E_1$ आणि E_2 हे कंपाउंड इव्हेंट आहे.

iii. शुअर इव्हेंट (Sure Event) (निश्चित इव्हेंट): निश्चित इव्हेंट ही अशी इव्हेंट असते जी नेहमीच घडते किंवा जिचे होणे निश्चित असते. जर एखाद्या प्रयोगातील सर्व संभाव्य निष्कर्ष (outcomes) एका घटनेत समाविष्ट असतील, तर ती इव्हेंट निश्चित इव्हेंट मानली जाते. जर S हे कोणत्याही प्रयोगाचे नमुना अवकाश (sample space) असेल, तर S ही स्वतःच निश्चित इव्हेंट असते, कारण त्यामध्ये सर्व संभाव्य निष्कर्ष समाविष्ट असतात.

त्यानुसार, $P(S) = 1$, म्हणजेच निश्चित घटनेची संभाव्यता नेहमी 1 असते.

उदाहरण:

a) एक फासा (डाय) टाकताना, "7 पेक्षा कमी संख्या येणे" हा शुअर इव्हेंट आहे.

एका सहा बाजू असलेल्या साध्या पास्यावर 1, 2, 3, 4, 5, 6 या संख्यांचे येऊ शकतात.

7 हा अंकच नाही, त्यामुळे नेहमीच 1 ते 6 पैकी एक अंक येईल.

हा शुअर इव्हेंट आहे (Sure Event) आहे आणि याची संभाव्यता $P(S) = 1$ आहे.

b) नाणे टाकल्यावर "छापा (head) किंवा काटा (tail) मिळणे" ही निश्चित इव्हेंट आहे.

नाणे टाकल्यावर फक्त दोनच शक्यता असतात: छापा (Head) किंवा काटा (Tail).

इतर कोणताही निष्कर्ष (outcome) येऊ शकत नाही. त्यामुळे, "छापा (हेड) किंवा काटा (टेल) मिळणे" ही निश्चित इव्हेंट आहे आणि याची संभाव्यता $P(S) = 1$ आहे.

c) पृथ्वी सूर्यभोवती फिरते हे निश्चित सत्य आहे.

iv. इम्पॉसिबल इव्हेंट (Impossible Event) (अशक्य इव्हेंट)

अशक्य इव्हेंट ही ती इव्हेंट असते जी कधीच घडू शकत नाही. गणितीयदृष्ट्या, अशक्य इव्हेंट रिक्त संचाने (empty set – एम्पटी सेट – $\emptyset$) दर्शवली जाते, आणि तिची संभाव्यता $P(A) = 0$ असते.

उदाहरण: सहा बाजूच्या फास्यावर (डायवर) "7" संख्या येणे ही अशक्य इव्हेंट आहे.

नमुना अवकाश (Sample Space) = $S = \{1, 2, 3, 4, 5, 6\}$

येथे 7 हा अंकच नाही, त्यामुळे "फास्यावर 7 येणे" ही अशक्य इव्हेंट आहे.

म्हणून, $A = \{\emptyset\}$, आणि $P(A) = 0$.

अशक्य घटनांमध्ये कोणतेही संभाव्य निष्कर्ष (outcomes) नसतात. म्हणून त्यांची संभाव्यता $P(A) = 0$ असते.

v. डिपेन्डन्ट इव्हेंट (Dependent Event) अवलंबित इव्हेंट: जर इव्हेंट A ची इव्हेंट, इव्हेंट B ची संभाव्यता बदलते, तर इव्हेंट A आणि B अवलंबून असतात.

उदाहरण: समजा, एका पिशवीत 3 काळे आणि 4 लाल चेंडू आहेत. समजा 3 काळे आणि 4 लाल चेंडू असलेल्या पिशवीतून एक चेंडू काढला गेला आहे. समजा A ही अशी इव्हेंट आहे जो चेंडू काढला गेला आहे तो लाल चेंडू आहे. जेव्हा दुसऱ्यांदा चेंडू काढला जातो तेव्हा B ही पुन्हा लाल चेंडू येण्याची इव्हेंट असते, दुसऱ्यांदा लाल चेंडू मिळण्याची संभाव्यता बदलते कारण आता पिशवीत 3 काळे आणि 3 लाल चेंडू आहेत. अशा प्रकारे बदलीशिवाय पिशवीमधून एकामागून एक दोन लाल चेंडू येण्याची संभाव्यता अवलंबून असते. त्यामुळे ही इव्हेंट अवलंबित इव्हेंट (डिपेन्डन्ट इव्हेंट) आहे.

vi. इन डिपेन्डन्ट इव्हेंट (Independent Event) (स्वतंत्र इव्हेंट): जर एका घटनेच्या घडण्यामुळे दुसऱ्या घटनेची संभाव्यता बदलत नसेल, तर त्या घटनांना स्वतंत्र इव्हेंट (Independent Events) म्हणतात.

जर इव्हेंट A ची इव्हेंट B ची संभाव्यता बदलत नसेल, तर इव्हेंट A आणि B स्वतंत्र असतात.

उदाहरण:

a) नाणे टाकणे (Tossing a Coin- टॉसिंग अ कॉईन):

जर आपण नाणे दोन वेळा टाकले, तर पहिल्यांदा छापा (हेड) किंवा काटा (टेल) येण्याचा दुसऱ्या वेळच्या निकालावर काहीही परिणाम होत नाही.

पहिल्या नाण्याच्या निकालाने दुसऱ्या नाण्याच्या निकालाची संभाव्यता बदलत नाही.

त्यामुळे नाणे फेकणे ही एक स्वतंत्र इव्हेंट आहे..

b) पत्त्यांचा खेळ (Playing Cards) – बदलीसह काढणे (With Replacement):

समजा, 52 पत्त्यांच्या गठ्यातून एक पत्ता काढला आणि तो परत ठेवला. आता दुसऱ्यांदा काढलेला पत्ता पूर्वी काढलेल्या पत्त्यावर अवलंबून नाही, कारण संपूर्ण 52 पत्त्यांचा संच पुन्हा तसाच आहे. त्यामुळे ही स्वतंत्र इव्हेंट आहे.

c) दोन वेगवेगळे फासे टाकणे (Rolling Two Dice- रोलिंग अ डार्स):

जर एकाच वेळी दोन फासे टाकले, तर पहिल्या फास्यावर काय आले याचा दुसऱ्या फास्यावर काहीही परिणाम होत नाही. प्रत्येक फास्याची संभाव्यता स्वतंत्र असते.

vii. मुच्युअली एक्सक्लुजिव्ह ऑर डिस्जॉईन्ट इव्हेंट (Mutually Exclusive or Disjoint Events) (परस्पर अपवर्जित किंवा विसंवादी इव्हेंट)

परस्पर अपवर्जित इव्हेंट (Mutually Exclusive Events) म्हणजे एकाच वेळी घडू न शकणाऱ्या इव्हेंट. दुसऱ्या शब्दांत, जर A आणि B या दोन इव्हेंट एकाच वेळी घडू शकत नसतील, तर त्या परस्पर अपवर्जित (डिस्जॉईन्ट) असतात. जर A आणि B परस्पर अपवर्जित असतील, तर त्यांचा सामायिक भाग (intersection- इंटरसेक्शन) शून्य असेल तर

$$A \cap B = \emptyset.$$

याचाच अर्थ, A आणि B चा कोणताही सामान्य निकाल (outcome- आउटकम) नसतो. अशा प्रकारे, इव्हेंट A आणि त्याचे पूरक A^c हे परस्पर अपवर्जित इव्हेंट आहेत.

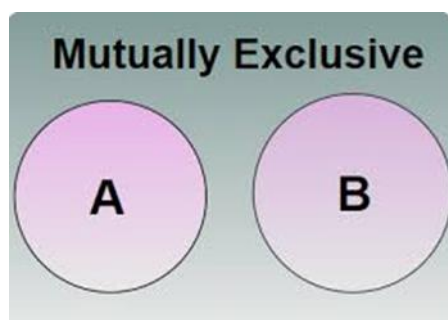


Figure 1.1: मुच्युअली एक्सक्लुजिव्ह (Mutually Exclusive)

नाणे फेकताना, जर निकाल छापा (Head) असेल तर त्याच चाचणीत काटा (टेल). मिळण्याची शक्यता नसते. नाणे फेकले असता छापा (हेड) किंवा काटा (टेल) यापैकी फक्त एकच येऊ शकते, दोन्ही एकाच वेळी एका सहा बाजूच्या फास्यावर 1, 2, 3, 4, 5, 6 हे निकाल मिळू शकतात.

समजा, $A =$ सम संख्या येणे $= \{2, 4, 6\}$

$B =$ विषम संख्या येणे $= \{1, 3, 5\}$

येथे A आणि B मध्ये कोणताही सामान्य निकाल नाही, त्यामुळे त्या परस्पर अपवर्जित इव्हेंट आहेत. कोणताही येऊ शकत नाहीत. त्यामुळे छापा (हेड) आणि काटा (टेल). या परस्पर अपवर्जित इव्हेंट आहेत.

viii. एक्सहौस्टिव्ह इव्हेंट (Exhaustive Events) (सर्वसमावेशी इव्हेंट):

कोणत्याही प्रयोगातील सर्व संभाव्य निष्कर्षांचा संपूर्ण संच (Total Possible Outcomes- टोटल पॉसिबल आउटकम) म्हणजे सर्वसमावेशी इव्हेंट (Exhaustive Events).

दुसऱ्या शब्दांत, जर A आणि B या दोन घटनांचा एकत्रित संच $(A \cup B)$ संपूर्ण नमुना अवकाश (Sample Space, S) असेल, तर त्या सर्वसमावेशी इव्हेंट असतात. जर A आणि B या दोन सर्वसमावेशी व्यापक असतील, तर

$$A \cup B = S$$

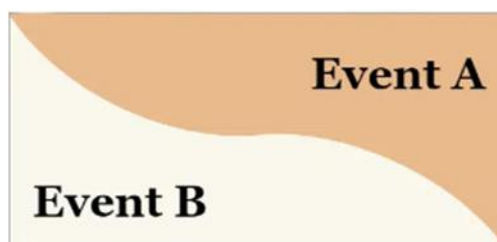


Figure 1.2: एक्सहौस्टिव्ह इव्हेंट (Exhaustive Events)

याचाच अर्थ, A आणि B मिळून संपूर्ण संभाव्यता व्यापतात, त्यामुळे कोणताही इतर निष्कर्ष उरत नाही.

उदाहरण: एखादे नाणे फेकले असता नमूना अवकाश (Sample Space) = $S = \{H, T\}$

समजा $A =$ "छापा (H) येणे" आणि $B =$ "काटा (T) येणे" या इव्हेंट आहेत

$$\therefore A = \{H\} \text{ आणि } B = \{T\}$$

येथे, $A \cup B = \{H, T\} = S$, त्यामुळे हे A आणि B सर्वसमावेशी इव्हेंट आहेत.

1.2 अलजेब्रा ऑफ इव्हेंट (Algebra of Events) (घटनांचे बीजगणित):

1) घटनांचा संघ (Union of Events): इव्हेंट A आणि B च्या संघटनमध्ये A किंवा B मध्ये किंवा A आणि B दोन्हीमध्ये असलेल्या सर्व परिणामांचा समावेश असतो. ते $A \cup B$ ने दर्शविले जाते लक्षात घ्या की $\cup$ हे चिन्ह "किंवा (OR)" असे वाचले जाते.

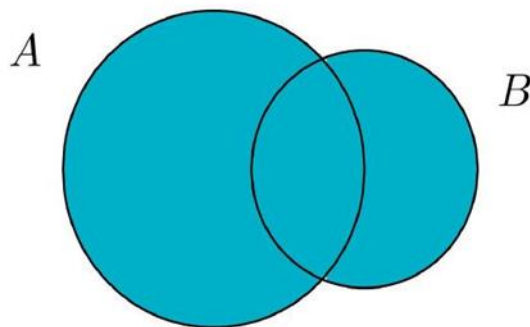


Figure. 1.3: घटनांचा संघ (Union of Events)

$$A \cup B = \{x: x \in A \text{ किंवा } x \in B \text{ किंवा दोन्हीमध्ये}\}$$

याचा अर्थ, $A \cup B$ मध्ये असलेले सर्व निष्कर्ष हे A , B किंवा दोन्हीमध्ये असतात. एका सहा बाजू असलेला फास (डाय) फेकला असता नमूना अवकाश (Sample Space),

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

समजा (A) सम संख्या येण्याची इव्हेंट आहे

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

आणि (B) 2 पेक्षा मोठी संख्या येण्याची इव्हेंट आहे.

$$\therefore B = \{3, 4, 5, 6\}$$

$$\therefore (A \cup B) \text{ घटनांचा संघ } A \cup B = \{2, 4, 6\} \cup \{3, 4, 5, 6\} = \{2, 3, 4, 5, 6\}$$

2) घटनांचा छेद (Intersection of Events):

इव्हेंट A आणि B च्या छेदनबिंदूमध्ये A आणि B दोन्हीमध्ये असलेल्या सर्व परिणामांचा समावेश असतो. ते $A \cap B$ ने दर्शविले जाते. लक्षात ठेवा की $\cap$ हे चिन्ह "आणि" असे वाचले जाते.

$$A \cap B = \{x: x \in A \text{ आणि } x \in B\}$$

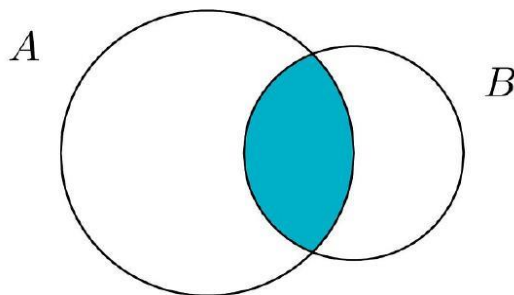


Figure 1.4: घटनांचा छेद (Intersection of Events)

उदाहरण: एका घटनेची अट अशी आहे की सहा बाजूच्या फास्याच्या पृष्ठभागावर सम आणि 2 पेक्षा मोठी संख्या मिळवणे A या घटनेची अट अशी आहे की पृष्ठभागावर "सम संख्या मिळणे"

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

B या घटनेची अट अशी आहे की पृष्ठभागावर "2 पेक्षा मोठी संख्या मिळणे"

$$\therefore B = \{3, 4, 5, 6\}$$

$$\therefore A \cap B = \{4, 6\} \text{ (कारण 4 आणि 6 हेच दोन्हीमध्ये आहेत)}$$

3) एलिमेंटरी प्रॉपर्टीज ऑफ प्रोबॅबिलिटी (Elementary Properties of Probability) संभाव्यतेचे मूलभूत गुणधर्म:

अ-नकारात्मकता (Non-Negativity- नॉन निगेटिव्हिटी) कोणत्याही घटनेची संभाव्यता ही कधीही ऋण (Negative) नसते. कोणत्याही इव्हेंट A साठी

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

सामान्यीकरण गुणधर्म (Normalization property- नॉर्मलायझेशन प्रॉपर्टी):

संपूर्ण नमुना जागेची संभाव्यता 1 असते.

$$P(S) = 1$$

उदाहरण नाणे फेकले असता नमुना अवकाश: $S = \{H, T\}$

$$P(H) = 1/2, P(T) = 1/2$$

$$\therefore P(S) = P(H) + P(T) = 1/2 + 1/2 = 1$$

कॉम्प्लिमेंट रूल (Complement Rule) (पूरक नियम): A जर एखादी इव्हेंट असेल तर A ची पूरक इव्हेंट (A न घडण्याची इव्हेंट) (A^c (A -complement)) आहे.

मोनोटोनिसिटी प्रॉपर्टी (Monotonicity Property) (एकसंधता गुणधर्म): जर A हा B चा उपसंच असेल म्हणजेच $A \subseteq B$ असेल तर इव्हेंट A ची संभाव्यता B पेक्षा कमी किंवा समान असेल.

$$P(A) \leq P(B)$$

जर A हा B च्या अंतर्गत असेल, तर B मध्ये असणारे सर्व संभाव्य निष्कर्ष (Outcomes) A मध्येही असतात. परंतु, B मध्ये A व्यतिरिक्त काही अतिरिक्त निष्कर्षही असू शकतात, त्यामुळे B ची संभाव्यता A च्या संभाव्यतेपेक्षा जास्त किंवा किमान समान असते.

उदाहरण: A या घटनेची अट अशी आहे की पृष्ठभागावर "सम (even – इव्हन) संख्या मिळणे"

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

$$\therefore P(A) = 3/6 = 1/2$$

B या घटनेची अट अशी आहे की पृष्ठभागावर "2 पेक्षा मोठी संख्या मिळणे"

$$\therefore B = \{3, 4, 5, 6\}$$

$$\therefore P(B) = \frac{4}{6} = \frac{2}{3}$$

येथे, $A \subseteq B$ असल्याने,

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

लॉ ऑफ टोटल प्रोबॅबिलिटी (Law of Total Probability) (एकूण संभाव्यतेचा नियम): समजा नमुना अवकाश (Sample Space) $S = \{A_1, A_2, A_3, \dots, A_n\}$ आहे तर सॅम्पल स्पेस S ची संभाव्यता ही नमुना अवकाश विभाजित करणाऱ्या परस्पर अपसंगत (Mutually Exclusive) घटनांच्या संभाव्यतेची बेरीज आहे.

जर $S = A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n$

तर, S ची संभाव्यता म्हणजे प्रत्येक उपघटनेच्या संभाव्यतेची बेरीज असते

$$P(S) = P(A_1) + P(A_2) + P(A_3) + \dots + P(A_n)$$

सब ऍडिटिव्ह प्रॉपर्टी (Sub-Additive Property) (उप-वजाबाकी गुणधर्म):

कोणत्याही दोन इव्हेंट A आणि B साठी, त्यांची संघट्ट इव्हेंट (Union) $A \cup B$ ची संभाव्यता A आणि B च्या संभाव्यतेच्या बेरीजेपेक्षा कमी किंवा समान असते

$$P(A \cup B) \leq P(A) + P(B)$$

1.3. दोन घटनांसाठी बेरीज प्रमेय

एडिशन थेरम ऑफ प्रोबॅबिलिटी (Addition Theorem of Probability)

संभाव्यतेचा बेरीज नियम: जर A आणि B या यादृच्छिक प्रयोगाशी संबंधित दोन इव्हेंट असतील तर त्यांची एकत्रित (Union) संभाव्यता $P(A \cup B)$ खालीलप्रमाणे दिली जाते:

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

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ आणि } B)$$

हा गुणधर्म दोनपेक्षा जास्त घटनांपर्यंत वाढवता येतो. जर A , B आणि C या यादृच्छिक प्रयोगाशी संबंधित तीन इव्हेंट असतील तर

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

विशेष बाब:

जर A आणि B परस्परविरोधी (Mutually Exclusive) इव्हेंट असतील, तर $A \cap B = \emptyset$, असेल तर $P(A \cap B) = 0$

म्हणून **संभाव्यतेचा बेरीज नियमानुसार**, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

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

जर A , B आणि C तीनही इव्हेंट परस्पर अपसंगत असतील

$$P(A \cap B) = P(B \cap C) = P(A \cap C) = P(A \cap B \cap C) = 0$$

म्हणून, $P(A \cup B \cup C) = P(A) + P(B) + P(C)$.

संभाव्यतेच्या गुणधर्मांवरील उदाहरणे:

1. निष्पक्ष (fair- फेअर) फासा (डाय) फेकला असता. 6 पेक्षा जास्त संख्या पृष्ठभागावर येण्याची संभाव्यता किती आहे?

उत्तर: एका निष्पक्ष (fair) सहा बाजूच्या फास्यावर सॅम्पल स्पेस (sample space)

$$S = \{1, 2, 3, 4, 5, 6\} \quad \therefore n(S) = 6$$

A या घटनेची अट अशी आहे की पृष्ठभागावर "६ पेक्षा जास्त संख्या मिळणे"

फास्यावर फक्त 1 ते 6 या अंकांपैकी एकच येऊ शकतो, त्यामुळे 6 पेक्षा मोठा अंक येण्याची शक्यता शून्य आहे.

$$\therefore A = \{ \} \text{ म्हणजेच, } A = \emptyset$$

$$\therefore n(A) = 0$$

$$\therefore P(A) = \frac{\text{अपेक्षित परिणामांची संख्या}}{\text{एकूण परिणामांची संख्या}}$$

$$\therefore P(A) = \frac{0}{6} = 0$$

2. तीन निष्पक्ष नाणे (fair coins) फेकले कमीत कमी दोन छापा दिसण्याची शक्यता शोधा.

उत्तर: तीन नाणे टाकल्यावर, एकूण $2^3 = 8$ संभाव्य परिणाम असतात

समजा H = छापा (हेड) आणि T = काटा (टेल) असे आहे

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

$$\therefore n(S) = 8$$

A या घटनेची अट अशी आहे की "कमीत कमी दोन छापा"

$$\therefore A = \{HHH, HHT, HTH, THH\}$$

$$\therefore n(A) = 4$$

$$\therefore P(A) = \frac{\text{अपेक्षित परिणामांची संख्या}}{\text{एकूण परिणामांची संख्या}}$$

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

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

3. A आणि B या प्रयोगातील परस्पर अपसंगत (Mutually Exclusive) इव्हेंट आहेत. जर

$P(A \text{ नाही}) = 0.65$, $P(A \cup B) = 0.65$ आणि $P(B) = p$ असेल तर p चे मूल्य शोधा.

उत्तर:

$$P(\text{नाही } A) = 0.65, \text{ म्हणजेच } P(A') = 0.65$$

आणि संपूर्ण नमुना जागेची (sample space) संभाव्यता 1 असते, त्यामुळे:

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

$$P(A) + 0.65 = 1$$

$$P(A) = 1 - 0.65 = 0.35$$

∴ परस्पर अपसंगत (Mutually Exclusive) घटनांसाठी बेरीज प्रमेयानुसार,

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

$$\therefore 0.65 = 0.35 + p$$

$$\therefore p = 0.65 - 0.35$$

$$\therefore p = 0.30$$

4. A आणि B यापैकी किमान एक इव्हेंट घडण्याची शक्यता 0.6 आहे. जर A आणि B दोन्ही एकत्र घडण्याची शक्यता 0.2 असेल, तर $P(A') + P(B')$ चा शोध घ्या.

उत्तर: दिलेले $P(A \cup B) = 0.6$, $P(A \cap B) = 0.2$

संभाव्यतेचा बेरीज नियमानुसार

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

$$\therefore 0.6 = P(A) + P(B) - 0.2$$

$$\therefore 0.6 = 1 - P(\bar{A}) + 1 - P(\bar{B}) - 0.2$$

$$\therefore 0.6 = 2 - 0.2 - [P(\bar{A}) + P(\bar{B})]$$

$$\therefore [P(\bar{A}) + P(\bar{B})] = 1.8 - 0.6$$

$$\therefore [P(\bar{A}) + P(\bar{B})] = 1.2$$

5. A, B, C या तीन परस्पर अनन्य आणि संपूर्ण इव्हेंट आहेत ज्या एका यादृच्छिक प्रयोगाशी संबंधित आहेत. जर $P(B) = \frac{3}{2} P(A)$ and $P(C) = \frac{1}{2} P(B)$ असेल, तर $P(A)$ शोधा.

उत्तर : दिलेले

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

$$P(C) = \frac{1}{2} P(B)$$

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

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

A, B आणि C या तीन परस्पर अपसंगत (mutually exclusive) आणि सर्वसमावेशक (exhaustive) इव्हेंट आहेत. म्हणून,

We have $A \cup B \cup C = S$

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

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

$$\therefore 1 = P(A) + \frac{3}{2} P(A) + \frac{3}{4} P(A)$$

$$\therefore 1 = P(A) \left[1 + \frac{3}{2} + \frac{3}{4} \right]$$

$$\therefore 1 = P(A) \left[\frac{4 + 6 + 3}{4} \right]$$

$$\therefore 1 = P(A) \left[\frac{13}{4} \right]$$

$$\therefore P(A) = \frac{4}{13}$$

6. लीप (Leap) नसलेल्या वर्षात त्रेपन्न रविवार असण्याची शक्यता किती आहे?

उत्तर : लिप नसलेले वर्ष 365 दिवसांचे असते,

365 दिवस = 52 पूर्ण आठवडे + 1 अतिरिक्त दिवस.

∴ 52 पूर्ण आठवड्यांमध्ये नक्कीच 52 रविवार असतील. पण उरलेला 1 अतिरिक्त दिवस कोणताही असू शकतो:

रविवार, सोमवार, मंगळवार, बुधवार, गुरुवार, शुक्रवार, शनिवार.

53 रविवार यायला, अतिरिक्त दिवस रविवार असायला हवा.

एकूण 7 शक्यता आहेत, पण त्यातील फक्त 1 शक्यता अशी आहे जिथे हा अतिरिक्त दिवस रविवार असेल.

$$\therefore P(53 \text{ रविवार}) = \frac{\text{अपेक्षित परिणामांची संख्या}}{\text{एकूण परिणामांची संख्या}} = \frac{1}{7}$$

7. एका भांड्यात (urn) 6 पांढरे, 4 लाल आणि 9 काळे चेंडू आहेत. यादृच्छिकरित्या काही चेंडू काढले तर खालील घटनांची संभाव्यता शोधा:

a) काढलेल्या दोन चेंडूंपैकी दोन्ही पांढरे असतील.

b) एकही लाल चेंडू नसेल.

c) प्रत्येक रंगाचा एक चेंडू असेल (एक पांढरा, एक लाल, एक काळा).

d) किमान एक चेंडू पांढरा असेल.

उत्तर: एका कलशामध्ये 6 पांढरे + 4 लाल + 9 काळे = 19 चेंडू आहेत.

यापैकी 3 चेंडू यादृच्छिकरित्या निवडले जातात.

$$n(S) = {}^{19}C_3 = 969$$

a) समजा A या घटनेची अट अशी आहे की "दोन चेंडू पांढरे असतील"

म्हणजेच 2 चेंडू पांढरे + 1 चेंडू कोणत्याही रंगाचा (19-6 पांढरे = बाकीचे 13 चेंडू) असावा.,

$$n(A) = {}^{19}C_2 \times {}^{13}C_1 =$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{{}^{6}C_2 \times {}^{13}C_1}{{}^{19}C_3} = \frac{15 \times 13}{19 \times 3 \times 17} = \frac{195}{1938}$$

b) समजा B या घटनेची अट अशी आहे की "एकही लाल चेंडू नसेल"

म्हणजेच काढलेले 3 चेंडू पांढऱ्या आणि काळ्या चेंडूंमधून असले पाहिजेत

6 पांढरे + 9 काळे = एकूण 15 चेंडू

$$\therefore n(B) = {}^{15}C_3 = 455$$

$$P(B) = \frac{n(B)}{n(S)} = \frac{{}^{15}C_3}{{}^{19}C_3} = \frac{455}{969}$$

c) समजा C या घटनेची अट अशी आहे की "प्रत्येक रंगाचा एक चेंडू असेल (एक पांढरा, एक लाल, एक काळा)".

$$\therefore n(C) = {}^6C_1 \times {}^4C_1 \times {}^9C_1$$

$$P(C) = \frac{n(C)}{n(S)} = \frac{{}^6C_1 \times {}^4C_1 \times {}^9C_1}{{}^{19}C_3} = \frac{6 \times 4 \times 9}{19 \times 3 \times 17} = \frac{72}{323}$$

d) समजा D या घटनेची अट अशी आहे की "किमान एक चेंडू पांढरा असणे"

$$P(D) = 1 - P(\text{तीन चेंडूंपैकी एकही पांढरा नाही})$$

$$= 1 - \frac{{}^{15}C_3}{{}^{19}C_3}$$

$$= 1 - \frac{286}{969}$$

$$= 1 - 0.295$$

$$= 0.705$$

8. विद्यार्थ्यांची एच.एस.सी. परीक्षा उत्तीर्ण होण्याची शक्यता $\frac{2}{3}$ आहे आणि तो एच.एस.सी. आणि आय.आय.टी. प्रवेश परीक्षा उत्तीर्ण होण्याची शक्यता $\frac{14}{15}$ आहे. तो किमान एक परीक्षा उत्तीर्ण होण्याची शक्यता $\frac{4}{5}$ आहे. तो आय.आय.टी. प्रवेश परीक्षा उत्तीर्ण होण्याची शक्यता किती आहे?

उत्तर:

समजा A या घटनेची अट अशी आहे की " विद्यार्थी एच.एस.सी. परीक्षा उत्तीर्ण होईल.

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

समजा B या घटनेची अट अशी आहे की " विद्यार्थी आय.आय.टी. परीक्षा उत्तीर्ण होईल.

$P(B)$ म्हणजे I.I.T. प्रवेश परीक्षा उत्तीर्ण होण्याची शक्यता

आता, $P(\text{तो किमान एक परीक्षा उत्तीर्ण होतो})$

$$= P(A \text{ किंवा } B \text{ किंवा दोन्ही})$$

$$= P(A \cup B) = \frac{4}{5}$$

एच.एस.सी. आणि आय.आय.टी. प्रवेश परीक्षा उत्तीर्ण होण्याची शक्यता $\frac{14}{45}$ आहे.

$$P(\text{दोन्ही परीक्षा उत्तीर्ण होण्याची शक्यता}) = P(A \cap B) = \frac{14}{45}$$

$P(\text{तो कमीत कमी एक परीक्षा उत्तीर्ण होतो})$

$$= P(A \text{ or } B \text{ or both})$$

$$= P(A \cup B) = \frac{4}{5}$$

आता, संभाव्यतेचा बेरीज नियमानुसार

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

$$\frac{4}{5} = \frac{2}{3} + P(B) - \frac{14}{45}$$

$$\therefore P(B) = \frac{4}{5} - \frac{2}{3} + \frac{14}{45} = \frac{20}{45}$$

$$\therefore P(B) = \frac{4}{9}$$

9. खोलीत 3 इलेक्ट्रिक बल्ब आहेत. 15 इलेक्ट्रिक बल्बच्या संग्रहातून ज्यापैकी केवळ 10 चांगले आहेत, त्यातून 3 बल्ब यादृच्छिकपणे निवडले जातात आणि लावले जातात. खोली किमान एका बल्बने प्रकाशित होण्याची शक्यता शोधा.

उत्तर :

एकूण 15 इलेक्ट्रिक बल्बच्या संग्रहातून 3 बल्ब निवडण्याच्या इव्हेंट ांची संख्या:

$$n(S) = {}^{15}C_3$$

10 बल्ब चांगले आहेत म्हणजेच 5 बल्ब खराब आहेत.

समजा A या घटनेची अट अशी आहे की ' सर्व 3 बल्ब खराब आहेत'

$$\therefore n(A) = {}^5C_3$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{{}^5C_3}{{}^{15}C_3} = \frac{2}{91}$$

आता, A' हि अशी इव्हेंट आहे की 'किमान 3 बल्बपैकी एक चांगला आहे'

$$P(A') = 1 - P(A) = 1 - \frac{2}{91} = \frac{89}{91}$$

10. सहा बाजू असलेले दोन निरपेक्ष फासे (six-face unbiased dice) टाकले. दाखवलेल्या संख्यांची बेरीज 7 किंवा गुणाकार 12 असण्याची शक्यता शोधा.

उत्तर: दोन निरपेक्ष फासे टाकले असता नमुना जागा

$$S = \left\{ \begin{array}{l} (1,1), (1,2), (1,3), (1,4), (1,5), (1,6) \\ (2,1), (2,2), (2,3), (2,4), (2,5), (2,6) \\ (3,1), (3,2), (3,3), (3,4), (3,5), (3,6) \\ (4,1), (4,2), (4,3), (4,4), (4,5), (4,6) \\ (5,1), (5,2), (5,3), (5,4), (5,5), (5,6) \\ (6,1), (6,2), (6,3), (6,4), (6,5), (6,6) \end{array} \right\}$$

$$\therefore n(S) = 36$$

समजा इव्हेंट A = 'दाखवलेल्या संख्यांची बेरीज 7 आहे'

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

$$\therefore n(A) = 6$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{6}{36} = \frac{1}{6}$$

समजा इव्हेंट B = 'संख्येचा गुणाकार 12 आहे'

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

$$\therefore n(B) = 4$$

$$P(B) = \frac{n(B)}{n(S)} = \frac{4}{36} = \frac{1}{9}$$

आता इव्हेंट $A \cap B$ = 'संख्यांची बेरीज 7 आहे आणि गुणाकार 12 आहे'

$$= \{(3,4), (4,3)\}$$

$$\therefore n(A \cap B) = 2$$

$$\therefore P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{2}{36} = \frac{1}{18}$$

आणि इव्हेंट $E = A \cup B$ = 'संख्यांची बेरीज 7 आहे किंवा गुणाकार 12 आहे'

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

$$= \frac{1}{6} + \frac{1}{9} - \frac{1}{18}$$

$$= \frac{3+2-1}{18}$$

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

11. एका कलशात 10 लाल, 5 पांढरे आणि 5 काळे चेंडू असतात. दोन चेंडू यादृच्छिकपणे काढले जातात. ते वेगवेगळ्या रंगांचे असण्याची शक्यता शोधा.

उत्तर : एका कलशामध्ये 10 लाल + 5 पांढरे + 5 काळे = 20 चेंडू आहेत

20 चेंडूंपैकी 2 चेंडू निवडण्याचे एकूण मार्ग = ${}^{20}C_2 = 190$

समजा R या घटनेची अट अशी आहे की " दोन्ही चेंडू लाल आहेत"

समजा W या घटनेची अट अशी आहे की " दोन्ही चेंडू पांढरे आहेत"

समजा B या घटनेची अट अशी आहे की " दोन्ही चेंडू काळे आहेत"

$$n(R) = {}^{10}C_2 \quad \therefore P(R) = \frac{n(R)}{n(S)} = \frac{{}^{10}C_2}{{}^{20}C_2} = \frac{45}{190}$$

$$n(W) = {}^{5}C_2 \quad \therefore P(R) = \frac{n(W)}{n(S)} = \frac{{}^{5}C_2}{{}^{20}C_2} = \frac{10}{190}$$

$$n(B) = 5C_2 \quad \therefore P(R) = \frac{n(B)}{n(S)} = \frac{5C_2}{20C_2} = \frac{10}{190}$$

समजा E या घटनेची अट अशी आहे की" दोन्ही काढलेले चेंडू एकाच रंगाचे आहेत"

R, W आणि हे B परस्पर अपसंगत (mutually exclusive) आहेत.

$$\therefore P(E) = (R \cup W \cup B) = (R) + P(W) + P(B)$$

$$= \frac{45}{190} + \frac{10}{190} + \frac{10}{190}$$

$$= \frac{65}{190} = \frac{13}{38}$$

आणि E' या घटनेची अट अशी आहे की" दोन्ही काढलेले चेंडू एकाच रंगाचे नाहीत"

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

$$= 1 - \frac{13}{38}$$

$$= \frac{25}{38}$$

सरावासाठी उदाहरणे

1. 52 पत्त्यांच्या चांगल्या शफलड गठ्यातून दोन पत्ते काढले गेले आहेत. तर खालील अटींची संभाव्यता

(Probability) काढा a)दोघेही राजे आहेत b) एक राजा आणि इतर राणी आहे.

2. जर $P(A) = 0.6, P(B) = 0.5$ and $P(A \cap B) = 0.2$ तर खालील

a) $P(A')$ b) $P(A \cap B')$ c) $P(B \cap A')$ संभाव्यता (Probability) काढा.

3.एका पिशवीत एकूण 100 मार्बल्स आहेत. 60 निळे 30 हिरवे, आणि 10 पिवळे, आहेत. एका मुलाला फक्त निळे किंवा हिरवे मार्बल्स निवडायचे आहेत. म्हणून, मुलाला मान्य असलेले मार्बल्स निवडण्याची संभाव्यता (Probability)किती आहे?

1.3.1 व्हेन आकृती (Venn diagram):

व्हेन आकृती (Venn diagram) ही संभाव्यता (Probability) मध्ये विविध घटनांमधील (Events) संबंधांचे दृश्यरूपात (Visual Representation) प्रदर्शन करण्यासाठी वापरली जाणारी एक प्रभावी साधन आहे. यामध्ये, प्रत्येक इव्हेंट एका वर्तुळाद्वारे (Circle) दर्शविली जाते, आणि वर्तुळे एकमेकांवर आच्छादित (Overlapping) होऊन त्यांच्या संयोग (Union) आणि छेद (Intersection) क्षेत्रांचे प्रतिनिधित्व करतात. व्हेन आकृतीच्या मदतीने, आपण विविध घटनांच्या संभाव्यता सहजपणे काढू शकतो.

समजा A आणि B या दोन इव्हेंट व्हेन आकृती (Venn diagram) मध्ये असतील तर जर दोन इव्हेंट A आणि B परस्पर अपसंगत (Mutually exclusive) असतील, म्हणजेच त्यांचा छेद (Intersection) रिक्त असेल ($A \cap B = \emptyset$), तर व्हेन आकृतीमध्ये त्यांच्या वर्तुळे एकमेकांवर आच्छादित (overlap) होत नाहीत. याचा अर्थ, A आणि B एकत्रितपणे कधीही घडू शकत नाहीत.

जर A आणि B या इव्हेंट परस्पर अपूर्ण नसतील, म्हणजेच त्यांचा काही समान भाग असेल ($A \cap B \neq \emptyset$), तर व्हेन आकृतीमध्ये त्यांच्या वर्तुळे एकमेकांवर आच्छादित होतात. आच्छादन क्षेत्र (Overlapping Region) A आणि B च्या छेदाचे ($A \cap B$) प्रतिनिधित्व करते, म्हणजेच दोन्ही इव्हेंट एकत्रितपणे घडण्याची शक्यता. व्हेन आकृतीमध्ये, नमुना जागेच्या (Sample Space) आत पण दोन्ही वर्तुळांच्या बाहेरील क्षेत्र A आणि B च्या संयोगाच्या पूरकाचे (Complement of $A \cup B$) प्रतिनिधित्व करते. याचा अर्थ, ना A ना B घडण्याची शक्यता व्हेन आकृतीच्या मदतीने, आपण विविध घटनांमधील संबंध आणि त्यांच्या संभाव्यता अधिक स्पष्टपणे समजू शकतो.व्हेन आकृती (Venn Diagram) मध्ये, प्रत्येक विभागलेल्या क्षेत्रात आपण खालीलपैकी कोणतीही माहिती समाविष्ट करू शकतो:

घटनेचे परिणाम (Outcomes of the event)

घटनेतील परिणामांची संख्या (Number of outcomes in the event)

घटनेची संभाव्यता (Probability of the event)

a. **सर्वसमावेशक संच (Universal Set) U आयत (Rectangle)** म्हणून दर्शविला जातो, आणि संच A त्याच्या आत एक वर्तुळ (Circle) म्हणून प्रदर्शित केला जातो. हे वर्तुळ संच A मधील सर्व घटकांचे प्रतिनिधित्व करते. आयताच्या बाहेरील क्षेत्र U च्या बाहेरील घटकांचे प्रतिनिधित्व करते.



Figure 1.5: सर्वसमावेशक संच (Universal Set)

b. **पूरक संच (Compliment of set)**

A' हा संच A चा पूरक (Complement) आहे. त्यामध्ये A मध्ये नसलेले सर्व घटक आहेत

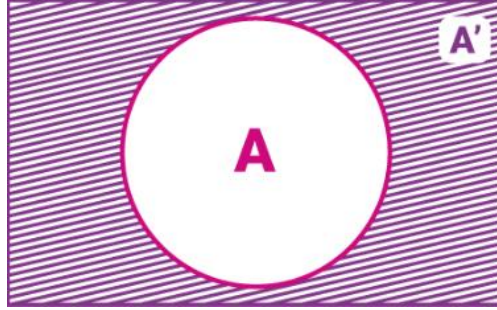


Figure 1.6: पूरक (Complement)

c. **छेदनबिंदू (Intersection of sets)**

समजा A आणि B हे कोणतेही दोन संच असतील तर त्यांचा छेदनबिंदू (Intersection) जो $A \cap B$ आहे तो खालीलप्रमाणे दर्शविला जातो. त्यामध्ये A आणि B मधील सर्व सामान्य घटक असतात.

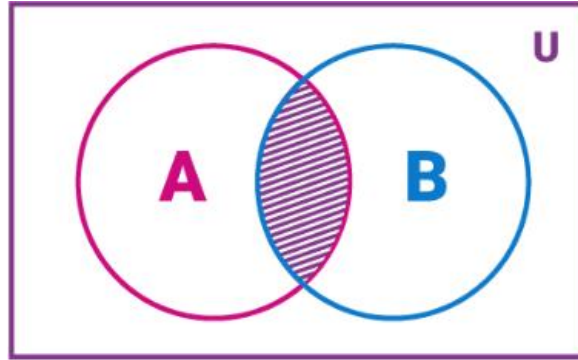


Figure 1.7: छेदनबिंदू (Intersection)

$A \cup B$ हा संच A आणि B च्या एकत्रित घटकांचे प्रतिनिधित्व करतो. तो खालीलप्रमाणे दर्शविला जातो

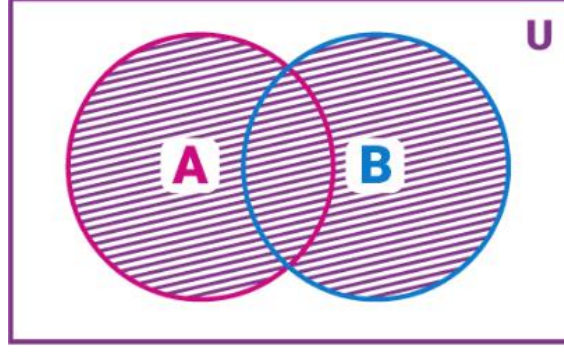


Figure 1.8

d. संचांच्या संयोगाच्या पूरक (Complement of Union of sets)
 $(A \cup B)'$ हे घटक A मध्ये किंवा B मध्ये नसलेले घटक दर्शवितात.

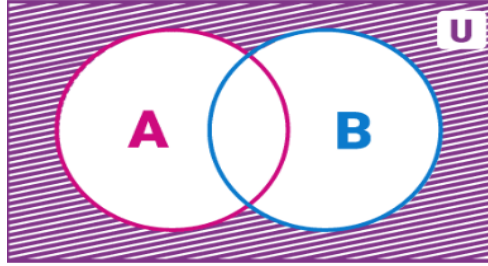


Figure 1.9: संचांच्या संयोगाच्या पूरक (Complement of Union of sets)

e. संचांच्या छेदनबिंदूचे पूरक
 $(A \cap B)'$ संच A आणि B मध्ये सामान्य नसलेल्या सार्वत्रिक घटकांचे घटक दर्शविते.

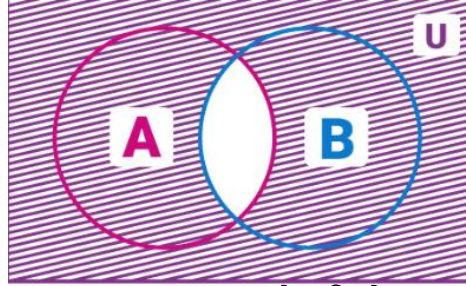


Figure 1.10: संचांच्या छेदनबिंदूचे पूरक

f. दोन संचांमधील फरक (A-B)

A-B संच A चे घटक दर्शवितो जे संच B मध्ये नाहीत. खालील व्हेन आकृतीमध्ये आकृती: 1.11 छायांकित प्रदेश A-B आहे.

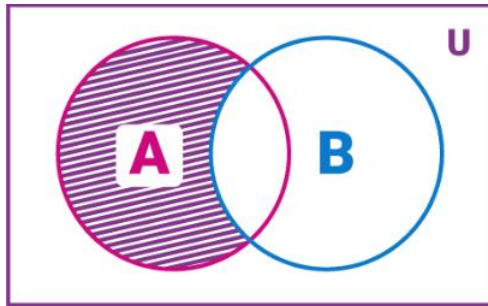


Figure 1.11

g. दोन संचातील सममितीय फरक (Symmetric difference between two sets)

संच $A \ominus B$ (A symmetric difference B) म्हणजेच सममितीय फरक. हा संच A किंवा B यामध्ये असलेल्या पण दोन्हीमध्ये नसलेल्या घटकांचे प्रतिनिधित्व करतो. व्हेन आकृतीत $A \ominus B$ हा संच A आणि B च्या वर्तुळांमध्ये असलेल्या पण त्यांच्या छेदात $(A \cap B)$ नसलेल्या भागाने दर्शविला जातो.

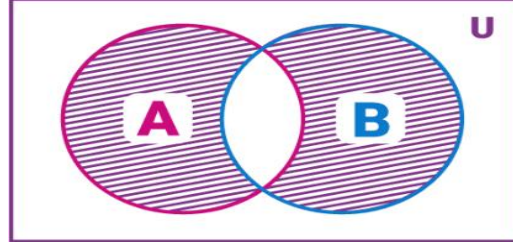


Figure 1.12: दोन संचातील सममितीय फरक (Symmetric difference between two sets)

व्हेन आकृतीवरील उदाहरणे

एका वर्गात एकूण 100 विद्यार्थी आहेत. त्यापैकी 70 विद्यार्थ्यांना गणित आवडते, 60 विद्यार्थ्यांना भौतिकशास्त्र आवडते. 40 विद्यार्थ्यांना दोन्ही आवडतात (गणित आणि भौतिकशास्त्र दोन्ही). व्हेन आकृती वापरून "फक्त गणित आवडणाऱ्या विद्यार्थ्यांची शक्यता (probability) शोधा.

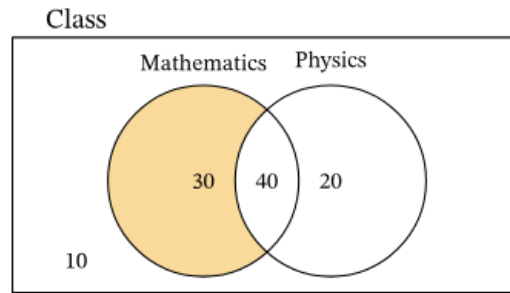


Figure 1.13

व्हेन आकृती आकृती 1.13 वरून आपल्याला असे दिसून येते की 30 विद्यार्थ्यांना गणित आवडते पण भौतिकशास्त्र आवडत नाही यादृच्छिक निवडलेल्या विद्यार्थ्याला गणित आवडते पण भौतिकशास्त्र आवडत नाही

$$\begin{aligned} \text{संभाव्यता सूत्रानुसार, } P(A) &= \frac{\text{अपेक्षित परिणामांची संख्या}}{\text{एकूण परिणामांची संख्या}} \\ &= \frac{30}{100} \\ &= 0.3 \end{aligned}$$

उदाहरण: अण्णा 1 ते 10 मधील संख्येचा विचार करतात. खालील प्रश्नांची उत्तरे देण्यासाठी व्हेन आकृती काढा.

1. संख्या दोनच्या पटीत असण्याची संभाव्यता किती आहे?
2. संख्या तीनच्या पटीत असण्याची संभाव्यता किती आहे?
3. संख्या दोन किंवा तीनच्या पटीत असण्याची संभाव्यता किती आहे?
4. संख्या दोन आणि तीनच्या पटीत असण्याची संभाव्यता किती आहे?
5. संख्या दोन किंवा तीनच्या पटीत नसण्याची संभाव्यता किती

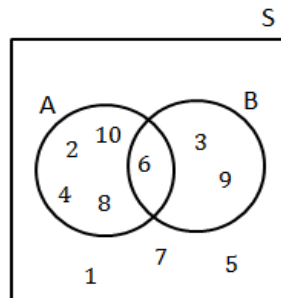


Figure 1.14

व्हेन आकृती आकृती: 1.14 दाखवते की 1 ते 10 पर्यंतच्या सर्व संख्या असलेल्या सॅम्पल स्पेस आहे.

$$\therefore U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

समजा A या घटनेची अट अशी आहे की, "सगळे अंक २ च्या पटीत आहेत"

$$\therefore A = \{2, 4, 6, 8, 10\}$$

समजा B या घटनेची अट अशी आहे की, "सगळे अंक 3 च्या पटीत आहेत"

$$\therefore B = \{3, 6, 9\}.$$

2 आणि 3 दोन्हीच्या पटीत येणाऱ्या संख्या = $A \cap B = \{6\}$

किंवा 3 यांच्या पटीत येणाऱ्या एकूण संख्या = $\{2, 3, 4, 6, 8, 9, 10\}$

1. निवडलेली संख्या २ च्या पटीत असण्याची संभाव्यता

$$P(A) = \frac{A \text{ मधील एकूण संख्या}}{U \text{ मधील एकूण परिणामांची संख्या}} = \frac{5}{10} = \frac{1}{2} = 0.5$$

2. निवडलेली संख्या ३ च्या पटीत असण्याची शक्यता

$$P(B) = \frac{B \text{ मधील एकूण संख्या}}{U \text{ मधील एकूण परिणामांची संख्या}} = \frac{3}{10} = 0.3$$

3. निवडलेली संख्या 2 किंवा 3 च्या पटीत असण्याची शक्यता

$$\begin{aligned} \therefore P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= \frac{1}{2} + \frac{3}{10} - \frac{1}{10} \\ &= \frac{7}{10} \end{aligned}$$

4. निवडलेली संख्या 2 आणि 3 दोन्हीच्या पटीत असण्याची शक्यता

$$P(A \cap B) = \frac{1}{10}$$

5. निवडलेली संख्या 2 किंवा 3 च्या पटीत नसण्याची शक्यता

$$\begin{aligned} P(A \cup B)' &= 1 - P(A \cup B) \\ P(A \cup B)' &= 1 - \frac{7}{10} = \frac{3}{10} \end{aligned}$$

उदाहरण: एका महाविद्यालयीन वसतिगृहात 75 विद्यार्थी आहेत, त्यापैकी 15 विद्यार्थ्यांना चहा प्यायला आवडतो, 40 विद्यार्थ्यांना कॉफी प्यायला आवडते आणि 20 विद्यार्थ्यांना चहा किंवा कॉफी दोन्ही आवडत नाही. या वसतिगृहातील दोन विद्यार्थी एका कॅन्टीनमध्ये येतात. दोघेही एकच पेय ऑर्डर करतील याची शक्यता शोधा.

उत्तर

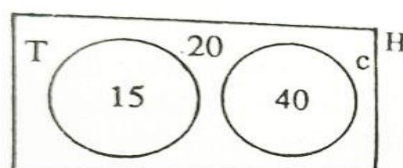


Figure 1.15

समजा T हा फक्त चहा पिण्यास आवडत असलेल्या विद्यार्थ्यांचा संच आहे आणि C हा फक्त कॉफी पिण्यास आवडत असलेल्या विद्यार्थ्यांचा संच आहे म्हणूनच, असा कोणताही विद्यार्थी नाही ज्याला दोन्ही पेय आवडतात.

समजा A या घटनेची अट अशी आहे की, "दोन्ही विद्यार्थी चहा ऑर्डर करतील".

समजा B या घटनेची अट अशी आहे की, "दोन्ही विद्यार्थी कॉफी ऑर्डर करतील"

$A \cup B$ = दोन्ही विद्यार्थी एकच पेय ऑर्डर करतील

A आणि हे B परस्पर अपसंगत (mutually exclusive) आहेत.

$$\therefore P(A \cup B) = P(A) + P(B) \dots \dots \dots (1)$$

निकांलांची एकूण संख्या = 75_{C_2}

A मधील एकूण निकालांची (Outcomes) संख्या = $15C_2$

B मधील एकूण निकालांची (Outcomes) संख्या = $40C_2$

$$P(A) = \frac{n(A)}{n(S)} = \frac{15C_2}{75C_2} = \frac{105}{2775}$$

$$P(B) = \frac{n(B)}{n(S)} = \frac{40C_2}{75C_2} = \frac{780}{2775}$$

∴ समीकरण (1) वरून

$$\therefore P(A \cup B) = \frac{105}{2775} + \frac{780}{2775} = \frac{885}{2775} = \frac{59}{185}$$

उदाहरण: एका विद्यापीठात प्रवेश घेणाऱ्या 100 विद्यार्थ्यांच्या नमुन्यात, एका प्रश्नावलीतून असे दिसून आले की त्यापैकी 45 जणांनी इंग्रजीचा सरावासाठी उदाहरणे केला, 40 जणांनी फ्रेंचचा सरावासाठी उदाहरणे केला, 35 जणांनी जर्मनचा सरावासाठी उदाहरणे केला, 20 जणांनी इंग्रजी आणि फ्रेंचचा सरावासाठी उदाहरणे केला, 23 जणांनी इंग्रजी आणि जर्मनचा सरावासाठी उदाहरणे केला, 19 जणांनी फ्रेंच आणि जर्मनचा सरावासाठी उदाहरणे केला आणि 12 जणांनी तिन्ही भाषांचा सरावासाठी उदाहरणे केला. व्हेन आकृती वापरून, यादृच्छिकपणे निवडलेल्या व्यक्तीने तीनपैकी फक्त एकाच भाषेचा सरावासाठी उदाहरणे केला असेल याची शक्यता शोधा.

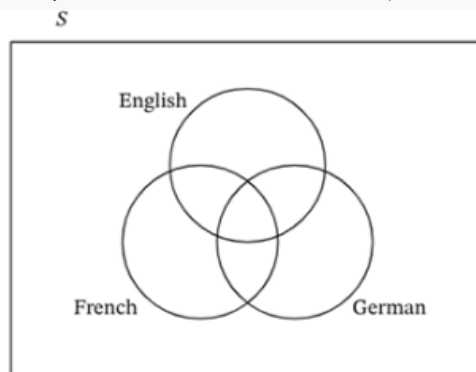


Figure 1.16

आता आपण येथे मध्यभा पासून सुरुवात करू जिथे तिन्ही वर्तुळे एकमेकांवर आदळतात(overlap). म्हणजे तीन घटनांचा छेदनबिंदू (intersection). हा प्रदेश अशा विद्यार्थ्यांचे प्रतिनिधित्व करतो ज्यांनी तिन्ही भाषा शिकल्या आहेत. असे 12 विद्यार्थी आहेत.

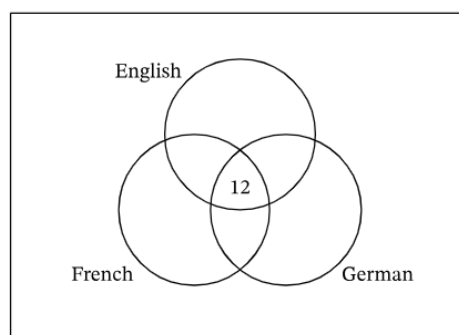


Figure 1.17

तसेच आपल्याला माहिती आहे की 20 विद्यार्थ्यांनी इंग्रजी आणि फ्रेंच दोन्ही भाषा शिकल्या. परंतु 12 विद्यार्थ्यांनी तिन्ही भाषा शिकल्या म्हणजेच त्यांनी जर्मन भाषा देखील शिकली.

यावरून आपण असा निष्कर्ष काढतो की इंग्रजी आणि फ्रेंच दोन्ही भाषा शिकणाऱ्या पण जर्मन न शिकणाऱ्या विद्यार्थ्यांची संख्या = $20 - 12 = 8$ आहे

इंग्रजी आणि जर्मन शिकणाऱ्या आणि फ्रेंच न शिकणाऱ्या विद्यार्थ्यांची संख्या = $23 - 12 = 11$ आहेत.

फ्रेंच आणि जर्मन शिकणाऱ्या आणि इंग्रजी न शिकणाऱ्या विद्यार्थ्यांची संख्या = $19 - 12 = 7$ आहेत.

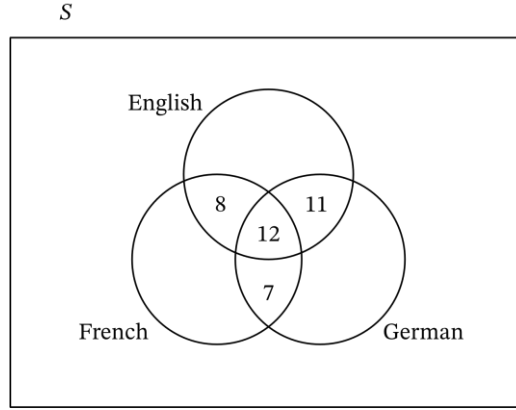


Figure 1.18

आपल्याला माहित आहे की 45 विद्यार्थ्यांनी इंग्रजी भाषा शिकली परंतु वरील व्हेन आकृतीवरून दिसून येते की $8+12+11=31$ विद्यार्थ्यांनी जर्मन किंवा फ्रेंच भाषा देखील शिकली. म्हणून ज्या विद्यार्थ्यांनी फक्त इंग्रजी भाषा शिकली परंतु फ्रेंच किंवा जर्मन भाषा शिकली नाही त्या विद्यार्थ्यांची संख्या $45 - 31 = 14$ आहे. तसेच, ज्या विद्यार्थ्यांनी फक्त फ्रेंच भाषा शिकली परंतु इंग्रजी किंवा जर्मन भाषा शिकली नाही त्या विद्यार्थ्यांची संख्या $= 40 - (8 + 12 + 7) = 13$ आहे. ज्यांनी फक्त जर्मन भाषा शिकली परंतु इंग्रजी किंवा फ्रेंच भाषा शिकली नाही त्या विद्यार्थ्यांची संख्या $= 35 - (7 + 12 + 11) = 5$

$\therefore$ आपली वेन आकृती खालीलप्रमाणे होईल

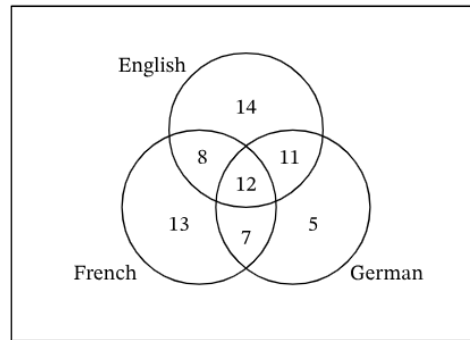


Figure 1.19

म्हणून व्हेन आकृतीमध्ये फक्त एकाच भाषेचा सरावासाठी उदाहरणे करणाऱ्या विद्यार्थ्यांना हिरव्या रंगाने छायांकित केले आहे.

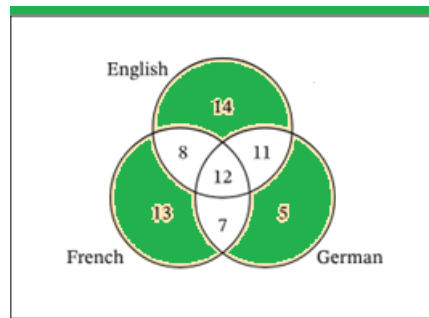


Figure 1.20

फक्त एकाच भाषा शिकलेल्या विद्यार्थ्यांची संख्या $13+14+5=32$ आहे.

$$\begin{aligned}
 \therefore \text{फक्त एकच भाषा शिकलेल्या विद्यार्थ्यांची संभाव्यता} &= \frac{\text{फक्त एकच भाषा शिकलेल्या विद्यार्थ्यांची संख्या}}{\text{एकूण विद्यार्थ्यांची संख्या}} \\
 &= \frac{32}{100} \\
 &= 0.32
 \end{aligned}$$

1.4 कंडिशनल प्रोबॅबिलिटी, मल्टिप्लिकेशन थेरम आणि इन्डिपेन्डेंट इव्हेंट

(Conditional probability, Multiplication theorem and Independent events):

1.4.1 कंडिशनल प्रोबॅबिलिटी: कंडिशनल प्रोबॅबिलिटी म्हणजे इव्हेंट B आधीच घडली आहे हे गृहित धरून इव्हेंट A घडण्याची संभाव्यता. ही संभाव्यता $P(A/B)$ या स्वरूपात दर्शविली जाते, आणि इव्हेंट B घडली असताना इव्हेंट A घडण्याची संभाव्यता" असे वाचले जाते.

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, P(B) \neq 0$$

इव्हेंट A आधीच घडली आहे हे गृहित धरून इव्हेंट B घडण्याची संभाव्यता खालीलप्रमाणे व्यक्त केली जाते

$$P(B/A) = \frac{P(A \cap B)}{P(A)}, P(A) \neq 0$$

जर A आणि B या दोन इव्हेंट एखाद्या यादृच्छिक प्रयोगाशी (random experiment) संबंधित असतील, तर त्यांची संयुक्त संभाव्यता (joint probability) खालीलप्रमाणे दिली जाते:

$$P(A \cap B) = P(A) \cdot P(B/A) \text{ if } P(A) \neq 0$$

$$P(A \cap B) = P(B) \cdot P(A/B) \text{ if } P(B) \neq 0$$

याठिकाणी $P(A/B)$ आणि $P(B/A)$ सशर्त संभाव्यता आहेत.

जर A आणि B या इव्हेंट स्वतंत्र (Independent) असतील, म्हणजेच A घडली तरी B घडण्यावर कोणताही परिणाम होत नसेल, तर

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

कारण स्वतंत्र घटनांसाठी, $P(B | A) = P(B)$ असते.

जर $A_1, A_2, A_3, \dots, A_n$ या सर्व इव्हेंट स्वतंत्र (Independent Events) असतील आणि त्या एखाद्या यादृच्छिक प्रयोगाशी (Random Experiment) संबंधित असतील, तर त्यांच्या सर्वांच्या एकत्र घडण्याची (Joint probability) शक्यता पुढील सूत्राने दिली जाते.

$$P(A_1 \cap A_2 \cap A_3 \dots \dots \dots \cap A_n) = P(A_1) \times P(A_2) \times P(A_3) \times \dots \dots \dots \times P(A_n)$$

कंडिशनल प्रोबॅबिलिटी वरील उदाहरणे

उदाहरण: जर A आणि B या दोन इव्हेंट असतील ज्यासाठी $P(A) = 0.5$, $P(B) = 0.6$ आणि

$$P(A \cup B) = 0.8 \text{ दिले असतील, तर } P(A/B) \text{ आणि } P(B/A) \text{ शोधा.}$$

उत्तर: आपल्याला माहित आहे कि संभाव्यतेचा बेरीज नियमानुसार,

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

$$\begin{aligned}
 \therefore P(A \cap B) &= P(A) + P(B) - P(A \cup B) \\
 &= 0.5 + 0.6 - 0.8
 \end{aligned}$$

$$\therefore P(A \cap B) = 0.3$$

$$\therefore P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.6} = 0.5$$

$$\therefore P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{0.3}{0.5} = 0.6$$

उदाहरण: बॅगमध्ये 10 पांढरे आणि 15 काळे बॉल असतात. दोन बॉल एकामागून एक सलगपणे पण बदल न करता (परत न ठेवता). अनुक्रमे काढले जातात. प्रथम पांढरा आणि दुसरा काळा आहे याची संभाव्यता काय आहे?

उत्तर: समजा A हा पहिल्या वेळेस पांढरा चेंडू मिळण्याची इव्हेंट आहे,

आणि B हा दुसऱ्या वेळेस काळा चेंडू मिळण्याची इव्हेंट आहे.

आवश्यक संभाव्यता (Required probability)

= पहिल्या वेळेस पांढरा चेंडू मिळण्याची संभाव्यता $\times$ दुसऱ्या वेळेस काळा चेंडू मिळण्याची संभाव्यता.

= $P(A \text{ and } B)$

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

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

$$\text{आता, } P(A) = \frac{{}^{10}C_1}{{}^{25}C_1} = \frac{10}{25} = \frac{2}{5}$$

$P(B|A)$ = पहिल्या ड्रॉमध्ये एक पांढरा चेंडू आधीच निवडला गेला असल्यास, दुसऱ्या ड्रॉमध्ये काळा चेंडू मिळण्याची संभाव्यता

$$P(B|A) = \frac{{}^{10}C_1}{{}^{24}C_1} = \frac{10}{24} = \frac{5}{8}$$

$$\therefore \text{आवश्यक संभाव्यता} = P(A \cap B) = P(A) \cdot P(B|A)$$

$$= \frac{2}{5} \times \frac{5}{8} = \frac{1}{4}$$

उदाहरण: एक नाणे दोनदा फेकले गेले आणि चार संभाव्य परिणाम समान संभाव्य आहेत असे गृहीत धरले जाते. जर A या घटनेची अट अशी आहे की 'छापा आणि काटा दोन्ही दिसले आहेत' आणि B या घटनेची अट अशी आहे की 'जास्तीत जास्त एक काटा दिसून येतो. तर $P(A)$, $P(B)$, $P(A/B)$ आणि $P(B/A)$ शोधा.

उत्तर: येथे नमुना जागा = $S = \{HH, HT, TH, TT\}$

$$A = \{HT, TH\}, \text{ and } B = \{HH, HT, TH\} \quad A \cap B = \{HT, TH\}$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{2}{4} = \frac{1}{2}$$

$$P(B) = \frac{n(B)}{n(S)} = \frac{3}{4}$$

$$P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{2}{4} = \frac{1}{2}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2}{3}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/2}{1/2} = 1$$

उदाहरण: X एखाद्या समस्येचे 90% समाधान करू शकतो आणि Y 70% समस्या सोडवू शकतो. एका पुस्तकातून कोणतीही समस्या निवडली असता, किमान एकाने ती सोडवली असेल, याची शक्यता किती?

उत्तर: समजा A घटनेची अट अशी आहे कि, "X समस्या सोडवेल"

$$\therefore P(A) = \frac{90}{100} = \frac{9}{10} = 0.9$$

समजा B घटनेची अट अशी आहे कि, "Y समस्या सोडवेल"

$$\therefore P(B) = \frac{70}{100} = \frac{7}{10} = 0.7$$

येथे A आणि B स्वतंत्र (Independent) इव्हेंट आहेत.

$$\begin{aligned}
\therefore \text{आवश्यक संभाव्यता (Required probability)} &= P(A \cap B) = 1 - P(\bar{A}) \times P(\bar{B}) \\
&= 1 - (1 - 0.9)(1 - 0.7) \\
&= 1 - (0.1 \times 0.3) \\
&= 1 - 0.03 \\
&= 0.97
\end{aligned}$$

उदाहरण: 52 पत्त्यांच्या पॅकमधून सलग दोन पत्ते काढली जातात. पहिले कार्ड राजा (किंग) आणि दुसरे राणी ((क्विन) असण्याची शक्यता शोधा.

- a) जर पहिले कार्ड बदलले असेल
- b) जर पहिले कार्ड बदलले नाही

उत्तर: A) एकूण 52 पत्त्यांपैकी 4 राजा (किंग) असतात.

पहिल्या ड्रॉमध्ये किंग येण्याची शक्यता

$$P(\text{पहिले कार्ड राजा}) = \frac{4}{52} = \frac{1}{13}$$

पत्ता (कार्ड) परत ठेवला आहे, त्यामुळे दुसऱ्या ड्रॉमध्ये पुन्हा 52 पत्ते असतील. त्यापैकी 4 क्वीन असतात. दुसऱ्या ड्रॉमध्ये क्वीन येण्याची शक्यता

$$P(\text{दुसरे कार्ड राणी}) = \frac{4}{52} = \frac{1}{13}$$

या दोन इव्हेंट शर्तीय (Conditional) असल्यामुळे आणि गुणाकार सिद्धांत (Multiplication Theorem of Probability) यावर आधारित असल्यामुळे,

$$\begin{aligned}
P(\text{पहिले कार्ड राजा आणि दुसरे कार्ड राणी}) &= \frac{1}{13} \times \frac{1}{13} \\
&= \frac{1}{169}
\end{aligned}$$

$$B) P(\text{पहिले कार्ड राजा}) = \frac{4}{52} = \frac{1}{13}$$

$$P(\text{दुसरे कार्ड राणी}) = \frac{4}{51}$$

$$\therefore P(\text{पहिले कार्ड राजा आणि दुसरे कार्ड राणी}) = P(\text{पहिले कार्ड राजा}) \times P(\text{दुसरे कार्ड राणी})$$

$$\begin{aligned}
&= \frac{1}{13} \times \frac{4}{51} \\
&= \frac{4}{663}
\end{aligned}$$

उदाहरण: एक 52 पत्त्यांचा (कार्डचा) गट्टा आहे आणि त्यामधून एक कार्ड यादृच्छिकरित्या (randomly) निवडले आहे. आपल्याला खालील A आणि B इव्हेंट स्वतंत्र आहेत का ते सिद्ध करायचे आहे.

A: निवडलेले कार्ड स्पेड नसणे

B: निवडलेले कार्ड किंग असणे

उत्तर: P (A): स्पेड नसलेल्या कार्डांची शक्यता (♥, ♦, ♣)

एकूण 52 पत्त्यांपैकी स्पेडचे 13 कार्ड असतात.

त्यामुळे स्पेड नसलेल्या कार्डांची संख्या = 52 - 13 = 39

$$\text{म्हणून, } P(A) = \frac{n(A)}{n(S)} = \frac{39}{52} = \frac{3}{4}$$

P(B): किंग असलेल्या कार्डांची शक्यता

एकूण 4 किंग असतात (♠, ♥, ♦, ♣)

$$\text{त्यामुळे, } P(B) = \frac{n(B)}{n(S)} = \frac{4}{52} = \frac{1}{13}$$

P(A ∩ B): स्पेड नसलेल्या किंगची शक्यता

एकूण 4 किंग आहेत, त्यापैकी 1 स्पेड आहे.

त्यामुळे स्पेड नसलेल्या किंगची संख्या = 4-1=3

$$\text{म्हणून } P(A \cap B) = \frac{3}{52}$$

A आणि B इव्हेंट स्वतंत्र असतील तर $P(A \cap B) = P(A) \times P(B)$

उजवीकडील बाजू (RHS) = $P(A) \times P(B)$)------(1)

$$= \frac{3}{4} \times \frac{1}{13}$$

$$= \frac{3}{52}$$

डावीकडील बाजू (LHS) = $P(A \cap B) = \frac{3}{52}$ ----- (2)

समीकरण (1) आणि (2) वरून

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

म्हणून, A आणि B इव्हेंट स्वतंत्र आहेत.

उदाहरण: एका पती-पत्नी एका कार्यालयातील दोन रिक्त पदांसाठी मुलाखतीसाठी उपस्थित होते. पती-पत्नी निवडून येण्याची शक्यता $1/7$ आहे आणि पत्नीच्या निवडीची शक्यता $1/5$ आहे. दोघांचीही निवड झाली आहे A आणि B त्यापैकी फक्त एकाची वर निवड झाली आहे याची शक्यता शोधा.

उत्तर: A: पती निवडला जाण्याची इव्हेंट

B: पत्नी निवडली जाण्याची इव्हेंट

A' (A complement): पती निवडला गेला नाही

B' (B complement): पत्नी निवडली गेली नाही

$$\text{आता } P(A) = \frac{1}{7}$$

$$\therefore P(A') = 1 - P(A) = 1 - \frac{1}{7} = \frac{6}{7}$$

$$\text{आणि } P(B) = \frac{1}{5}$$

$$\therefore P(B') = 1 - P(B) = 1 - \frac{1}{5} = \frac{4}{5}$$

दोघांचीही निवड झाली आहे.

A आणि B दोन्ही इव्हेंट स्वतंत्र आहेत.

$$\therefore P(A \cap B) = P(A) \times P(B) = \frac{1}{7} \times \frac{1}{5} = \frac{1}{35}$$

फक्त एकच जण निवडला जाईल याची शक्यता:

फक्त एकच जण निवडला जाणे म्हणजे एक तर पती निवडला जाईल आणि पत्नी निवडली जाणार नाही किंवा पत्नी निवडली जाईल आणि पती निवडला जाणार नाही.

$A \cap B'$: पती निवडला गेला आणि पत्नी निवडली नाही

$$P(A \cap B') = P(A) \times P(B') = \frac{1}{7} \times \frac{4}{5} = \frac{4}{35}$$

$A' \cap B$: पत्नी निवडली गेली आणि पती निवडला नाही

$$P(A' \cap B) = P(A') \times P(B) = \frac{6}{7} \times \frac{1}{5} = \frac{6}{35}$$

समजा E अशी इव्हेंट आहे कि, "दोघांपैकी एकाची निवड होईल."

$$E = (A \cap B') \cup (A' \cap B)$$

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

$$= \frac{4}{35} + \frac{6}{35}$$

$$= \frac{10}{35} = \frac{2}{7}$$

उदाहरण: A लक्ष्यावर (टार्गेट) गोळीबार करू शकतो याची शक्यता $5/7$ आहे आणि B त्याच लक्ष्यावर (टार्गेट) गोळीबार (हिट) करू शकतो याची शक्यता $3/5$ आहे. A आणि B स्वतंत्रपणे गोळीबार(हिट) करतात. A) लक्ष्यावर(टार्गेट) अजिबात गोळीबार (हिट) झाला नाही आणि B) त्यापैकी किमान एकाने लक्ष्यावर (टार्गेट) गोळीबार (हिट) केला आहे याची शक्यता शोधा.

उत्तर: A ने टार्गेट हिट करण्याची शक्यता $P(A)$

$$P(A) = \frac{5}{7} \quad \therefore P(A^c) = 1 - P(A) = 1 - \frac{5}{7} = \frac{2}{7}$$

B ने टार्गेट हिट करण्याची शक्यता $P(B)$

$$P(B) = \frac{3}{5} \quad \therefore P(B^c) = 1 - P(B) = 1 - \frac{3}{5} = \frac{2}{5}$$

A आणि B हे स्वतंत्ररीत्या शूट करतात. A^c आणि B^c ह्या सुद्धा स्वतंत्र इव्हेंट आहेत.

समजा E_1 अशी इव्हेंट आहे कि, "टार्गेट अजिबात हिट झाले नाही." $= A^c \cap B^c$

$$\therefore P(E_1) = P(A^c \cap B^c) = P(A^c) \times P(B^c) = \frac{2}{7} \times \frac{2}{5} = \frac{4}{35}$$

समजा E_2 अशी इव्हेंट आहे कि, "किमान एकाने तरी टार्गेट हिट केले आहे."

E_2 हि इव्हेंट E_1 ची पूरक (Complementary) इव्हेंट आहे.

$$\therefore P(E_2) = 1 - P(E_1) = 1 - \frac{4}{35} = \frac{31}{35}$$

सरावासाठी उदाहरणे:

1. E आणि F अशा इव्हेंट आहेत की $P(E) = 0.6$ $P(F) = 0.6$ आणि $P(E \cap F) = 0.2$ तर

$P(E/F)$ and $P(F/E)$ शोधा.

2. जर $P(A) = 0.8, P(B) = 0.5$ आणि $P\left(\frac{B}{A}\right) = 0.4$ तर i) $P(A \cap B)$ ii) $P\left(\frac{A}{B}\right)$

iii) $P(A \cup B)$ शोधा.

3. एका नाण्याला तीन वेळा फेकले असता. खालील घटनांसाठी $P(E/F)$ निश्चित करा:

E: तिसऱ्यांदा फेकल्यावर छापा (Head) येणे आणि F: पहिल्या दोन वेळा फेकल्यावर छापा येणे.

E: किमान दोन वेळा छापा येणे आणि F: जास्तीत जास्त दोन वेळा छापा येणे.

E: जास्तीत जास्त दोन वेळा काटा (Tail) येणे आणि F: किमान एकदा तरी काटा येणे.

4. आई, वडील आणि मुलगा कुटुंबाच्या फोटोसाठी अनियमित क्रमाने रांगेत उभे राहतात.

E आणि F अशा इव्हेंट आहेत की, E: मुलगा एका टोकाला असणे F: वडील मधोमध असणे तर $P(E/F)$ शोधा.

5. एक काळा आणि लाल असे दोन फासे (Dice) फेकले आहेत. जर काळ्या फासाचा निकाल 5 असेल तर 9 पेक्षा जास्त बेरीज मिळण्याची सशर्त संभाव्यता (Conditional probability) शोधा. जर लाल फासाचा निकाल 4 पेक्षा कमी असेल तर 8 बेरीज मिळण्याची सशर्त संभाव्यता शोधा.

6. समजा, जन्माला येणारे मूल मुलगा किंवा मुलगी होण्याची समान शक्यता आहे. जर एखाद्या कुटुंबात दोन मुले असतील, तर खालील परिस्थितीत दोन्ही मुली असण्याची सशर्त संभाव्यता शोधा –

a) धाकटी मुलगी आहे.

b) किमान एक मुलगी आहे.

7. एका समस्येचे निराकरण तीन व्यक्तींना (P, Q, R) करायचे आहे, ज्यांची अनुक्रमे समस्या सोडवण्याची शक्यता $2/7, 4/7$ आणि $4/9$ आहे. तर, ती समस्या सोडवली जाईल याची संभाव्यता किती?

1.5 बेयस सिद्धांत आणि सशर्त संभाव्यता (Bayes' Theorem and Conditional Probability)

बेयसचा सिद्धांत एखाद्या घटनेची सशर्त संभाव्यता (Conditional Probability) शोधण्यासाठी वापरला जातो. याचा उपयोग पूर्वीच्या माहितीनुसार एखादी इव्हेंट घडण्याची संभाव्यता शोधण्यासाठी केला जातो. बेयसचा सिद्धांत (Bayes' Theorem, Bayes' Rule, किंवा Bayes' Law) इव्हेंट A घडण्याची सशर्त संभाव्यता शोधतो, जेव्हा इव्हेंट B आधीच घडली आहे.

बेयस सिद्धांताचा सामान्य स्वरूप:

$$P(A/B) = \frac{P(B/A) \cdot P(A)}{P(B)}$$

P (A) = इव्हेंट A घडण्याची संभाव्यता (Prior Probability).

P (B) = इव्हेंट B घडण्याची संभाव्यता (Total Probability).

P(A/B) = B आधीच घडले असताना, A घडण्याची संभाव्यता (Posterior Probability).

P(B/A) = A घडले असताना, B घडण्याची संभाव्यता (Likelihood).

बेयसच्या सिद्धांताचे उदाहरण:

समजा, तीन पिशव्यांमध्ये पांढरे आणि काळे मार्बल आहेत. एक पांढरा मार्बल काढला आहे. तो पहिल्या पिशवीतून आला असेल, याची संभाव्यता शोधण्यासाठी आपण बेस प्रमेयाचा उपयोग करू शकतो.

बेयस प्रमेयाचे जनरल (General) सूत्र पुढीलप्रमाणे आहे.

समजा, $E_1, E_2, E_3, \dots, E_n$ ही नमुना अवकाश S S शी संबंधित घटनांची एक संच आहे, जिथे प्रत्येक इव्हेंट $E_1, E_2, E_3, \dots, E_n$ ला शून्याहून अधिक संभाव्यता आहे आणि त्या S च्या विभाजनास तयार करतात. A ही S शी संबंधित कोणतीही इव्हेंट असल्यास, बेयसच्या सिद्धांतानुसार:

$$P(E_i/A) = \frac{P(A/E_i)P(E_i)}{P(A/E_1)P(E_1) + P(A/E_2)P(E_2) + \dots + P(A/E_k)P(E_k)}$$

1. समजा, पहिल्या भांड्यामध्ये 3 पांढरे आणि 4 लाल चेंडू आहेत, आणि दुसऱ्या भांड्यामध्ये 5 पांढरे आणि 6 लाल चेंडू आहेत. यापैकी एक चेंडू यादृच्छिकरित्या कोणत्याही एका भांड्यातून निवडला जातो आणि तो पांढरा असतो. तर, तो चेंडू पहिल्या भांड्या मधून निवडला गेला असेल, याची संभाव्यता शोधा.

उत्तर: समजा E_1 , आणि E_2 अनुक्रमे, पहिले भांडे आणि दुसरे भांडे निवडले जाण्याच्या इव्हेंट दर्शवतात.

$$\therefore P(E_1) = P(E_2) = \frac{1}{2}$$

निवडलेल्या भांड्या मधून काढलेला एक चेंडू पांढरा असतो अशी इव्हेंट A मानू मग पहिल्या भांड्या मधून पांढरा चेंडू निवडला जाण्याची संभाव्यता

$$= P(A/E_1) = \frac{\text{पहिल्या भांड्या मधील पांढऱ्या चेंडूंची संख्या}}{\text{पहिल्या भांड्या मधील एकूण चेंडूंची संख्या}} = \frac{3C_1}{7C_1} = \frac{3}{7}$$

दुसऱ्या मधून पांढरा चेंडू निवडला जाण्याची संभाव्यता

$$= P(A/E_2) = \frac{\text{दुसऱ्या भांड्या मधील पांढऱ्या चेंडूंची संख्या}}{\text{दुसऱ्या भांड्या मधील एकूण चेंडूंची संख्या}} = \frac{5C_1}{11C_1} = \frac{5}{11}$$

यादृच्छिकपणे निवडलेला पांढरा चेंडू पहिल्या मधून आलेला असेल, याची संभाव्यता.

$$\begin{aligned} = P(E_1/A) &= \frac{P(A/E_1)P(E_1)}{P(A/E_1)P(E_1) + P(A/E_2)P(E_2)} \\ &= \frac{\frac{3}{7} \times \frac{1}{2}}{\left(\frac{3}{7} \times \frac{1}{2}\right) + \left(\frac{5}{11} \times \frac{1}{2}\right)} \end{aligned}$$

$$\begin{aligned}
&= \frac{\frac{3}{7} \times \frac{1}{2}}{\frac{1}{2} \left(\frac{3}{7} + \frac{5}{11} \right)} \\
&= \frac{\frac{3}{7}}{\frac{68}{77}} \\
&= \frac{3}{7} \times \frac{77}{68} \\
&= \frac{33}{68}
\end{aligned}$$

2. समजा, बोल्ट फॅक्टरीमध्ये मशीन X, Y आणि Z एकूण उत्पादनाच्या अनुक्रमे 25%, 35% आणि 40% बोल्ट तयार करतात. त्यांच्या उत्पादनांपैकी 5%, 4% आणि 2% बोल्ट दोषपूर्ण असतात. एक बोल्ट यादृच्छिकरित्या निवडला जातो आणि तो दोषपूर्ण आढळतो.

तर, तो मशीन X, Y किंवा Z

उत्तर: समजा E_1, E_2 and E_3 या अनुक्रमे मशीन X, Y आणि Z ने तयार केलेल्या बोल्टच्या इव्हेंट दर्शवतात.

A ही इव्हेंट आहे की, "निवडलेला बोल्ट दोषपूर्ण आहे."

$$P(E_1) = \frac{25}{100} = 0.25$$

$$P(E_2) = \frac{35}{100} = 0.35$$

$$P(E_3) = \frac{40}{100} = 0.40$$

मशीन X ने तयार केलेला बोल्ट दोषपूर्ण असेल, याची संभाव्यता:

$$P(A/E_1) = \frac{5}{100} = 0.05$$

$$P(A/E_2) = \frac{4}{100} = 0.04$$

$$P(A/E_3) = \frac{2}{100} = 0.02$$

यादृच्छिकरित्या निवडलेला दोषपूर्ण बोल्ट मशीन X ने तयार केला असेल, याची संभाव्यता.

$$\begin{aligned}
P(E_1/A) &= \frac{P(A/E_1)P(E_1)}{P(A/E_1)P(E_1) + P(A/E_2)P(E_2) + P(A/E_3)P(E_3)} \\
&= \frac{0.05 \times 0.25}{0.05 \times 0.25 + 0.04 \times 0.35 + 0.02 \times 0.40} \\
&= \frac{0.0125}{0.0125 + 0.0140 + 0.008} \\
&= \frac{0.0125}{0.0345} = \frac{25}{69}
\end{aligned}$$

यादृच्छिकरित्या निवडलेला दोषपूर्ण बोल्ट मशीन Y ने तयार केला असेल, याची संभाव्यता

$$\begin{aligned}
P(E_2/A) &= \frac{P(A/E_2)P(E_2)}{P(A/E_1)P(E_1) + P(A/E_2)P(E_2) + P(A/E_3)P(E_3)} \\
&= \frac{0.35 \times 0.04}{0.05 \times 0.25 + 0.04 \times 0.35 + 0.02 \times 0.40} \\
&= \frac{0.0140}{0.0345} \\
&= \frac{140}{345} \\
&= \frac{28}{69}
\end{aligned}$$

यादृच्छिकरित्या निवडलेला दोषपूर्ण बोल्ट मशीन Z ने तयार केला असेल, याची संभाव्यता

$$\begin{aligned} P(E_3/A) &= 1 - [P(E_1/A) + P(E_2/A)] \\ &= 1 - \left[\frac{25}{69} + \frac{28}{69} \right] \\ &= 1 - \frac{53}{69} \\ &= \frac{16}{69} \end{aligned}$$

- एका फॅक्टरीत दोन मशीन आहेत. प्रायोगिक निरीक्षणांनुसार, मशीन I आणि मशीन II अनुक्रमे 30% आणि 70% उत्पादन करतात. तसेच, मशीन I मधील 5% आणि मशीन II मधील 1% उत्पादन दोषपूर्ण असते. जर यादृच्छिकरित्या निवडलेला एक घटक दोषपूर्ण असेल, तर तो मशीन II ने तयार केला असेल, याची संभाव्यता किती?
- एका नव्याने बांधलेल्या उड्डाणपुलाच्या कोसळण्याची शक्यता असते, भले त्याची रचना दोषपूर्ण असो किंवा नसो. रचना दोषपूर्ण असेल, तर ती होण्याची शक्यता 5% आहे. जर रचना दोषपूर्ण असेल, तर उड्डाणपुलाच्या कोसळण्याची शक्यता 95% आहे. जर रचना योग्य असेल, तरीसुद्धा उड्डाणपुल कोसळण्याची शक्यता 30% आहे. जर उड्डाणपुल कोसळला असेल, तर तो दोषपूर्ण रचनेमुळे कोसळला असेल, याची किती शक्यता आहे?
- एका पिशवीत 6 लाल आणि 5 निळे चेंडू आहेत, आणि दुसऱ्या पिशवीत 5 लाल आणि 8 निळे चेंडू आहेत. पहिल्या पिशवीतून एक चेंडू घेतला जातो आणि त्याचा रंग न पाहता दुसऱ्या पिशवीत ठेवला जातो. नंतर दुसऱ्या पिशवीतून एक चेंडू निवडला जातो. तो चेंडू निळा असेल, याची किती शक्यता आहे?
- X, Y आणि Z यांच्या व्यवस्थापक होण्याच्या शक्यता अनुक्रमे 4:2:3 आहेत. जर X व्यवस्थापक झाला, तर बोनस योजना लागू होण्याची शक्यता 0.3 आहे. जर Y व्यवस्थापक झाला, तर बोनस योजना लागू होण्याची शक्यता 0.5 आहे. जर Z व्यवस्थापक झाला, तर बोनस योजना लागू होण्याची शक्यता 0.8 आहे. जर बोनस योजना लागू झाली असेल, तर X व्यवस्थापक झाला असेल, याची किती शक्यता आहे?
- E_1 आणि E_2 या दोन इव्हेंट समान शक्यता असलेल्या, परस्पर वर्ज्य आणि सर्वसमावेशक आहेत. जर $P(A/E_1) = 0.2$, $P(A/E_2) = 0.3$. तर $P(E_1/A)$ शोधा.
- तीन पिशव्या आहेत, प्रत्येकी 100 चेंडू आहेत. पहिल्या पिशवीत 75 लाल आणि 25 निळे चेंडू आहेत. दुसऱ्या पिशवीत 60 लाल आणि 40 निळे चेंडू आहेत. तिसऱ्या पिशवीत 45 लाल आणि 55 निळे चेंडू आहेत. जर यादृच्छिकरित्या एक पिशवी निवडली गेली आणि त्यातून एक चेंडू घेतला गेला, तर तो चेंडू लाल असेल, याची किती शक्यता आहे?

युनिट-2

रँडम व्हेरिअबल आणि प्रोबॅबिलिटी डिस्ट्रीब्यूशन

(Random variable and Probability distribution)

विषय निष्पत्ती (Course Outcome):

CO2: डेटाचा प्रसार मिळविण्यासाठी यादृच्छिक चल आणि संभाव्यता वितरणाच्या प्रकारांच्या संकल्पना वापरा.

घटक निष्पत्ती (Theory Learning Outcome):

1. यादृच्छिक चलाचे प्रकार वेगळे करा.
2. संभाव्यता वितरणावर दिलेल्या समस्या सोडवा.
3. यादृच्छिक चलाचे अपेक्षित मूल्य काढा.
4. यादृच्छिक चलाचे भिन्नता आणि मानक विचलन काढा.
5. दिलेल्या समस्येवर लागू करावयाच्या वितरणाचा प्रकार ओळखा.
6. दिलेले वितरण लागू करून संभाव्यता काढा.

2.1 रँडम व्हेरिअबल व्याख्या आणि प्रकार: (Random variable: Definition and types)

2.1.1 रँडम व्हेरिअबल (Random Variable): रँडम व्हेरिअबल म्हणजे एक व्हेरिअबल जो चाचणी, प्रयोग किंवा घटनेचे परिणाम दर्शवतो. ही एक विशिष्ट संख्या असते जी चाचणी, प्रयोग किंवा घटनेची पुनरावृत्ती करताना प्रत्येक वेळी वेगळी असते. विशेषतः स्टॅटिस्टिक्स आणि प्रोबॅबिलिटीमध्ये रँडम व्हेरिअबल म्हणजे एक व्हेरिअबल ज्याचे रँडम व्हेल्यू प्रयोगाच्या परिणामावर अवलंबून असते. ते सहसा मोठ्या अक्षर X किंवा Y इत्यादींनी दर्शविले जाते.

रँडम व्हेरिअबल्सचे दोन प्रकार आहेत:

a. डिस्क्रीट रँडम व्हेरिअबल (Discrete random variable): ज्या रँडम व्हेरिअबलमध्ये मोजता येण्याजोग्या आयसोलेटेड व्हॅल्यूज (पूर्णांक) गृहीत धरता येतात त्याला डिस्क्रीट रँडम व्हेरिअबल म्हणतात.

उदाहरण

- i. कुटुंबातील मुलांची संख्या
- ii. राज्यातील जिल्ह्यांची संख्या
- iii. वर्गातील विद्यार्थ्यांची संख्या
- iv. फुटबॉल सामन्यात केलेल्या गोलची संख्या.

b. कंटीन्यूअस रँडम व्हेरिअबल (Continuous random variable): कोणताही रँडम व्हेरिअबलमध्ये ज्याला दिलेल्या अंतरालात (इंटरव्हल) सर्व वास्तविक मूल्ये स्वीकारता येतात, त्याला कंटीन्यूअस रँडम व्हेरिअबल म्हणतात.

उदाहरण

- i. व्यक्तीचे वय
- ii. विद्युत घटकाचे आयुष्य (तासांमध्ये)
- iii. किमी/तासात वाहनाचा वेग.
- iv. दिलेल्या दिवशी तापमान

2.2 प्रोबॅबिलिटी मास फंक्शन (Probability of Mass Function (p.m.f))

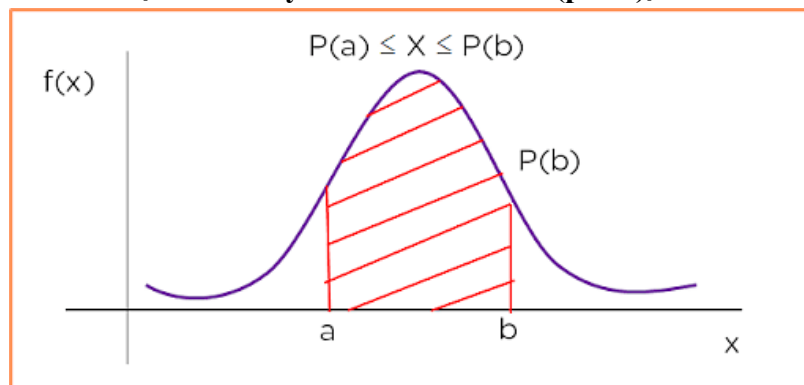


Figure 2.1: प्रोबॅबिलिटी मास फंक्शन (Probability of Mass Function (p.m.f))

प्रोबॅबिलिटी मास फंक्शन (PMF) हे एक गणितीय कार्य आहे जे एका विशिष्ट व्हेल्यू घेणाऱ्या एका डिस्क्रीट रँडम चलाच्या प्रोबॅबिलिटीचे वर्णन करते. डिस्क्रीट रँडम चलांसाठी प्रोबॅबिलिटी सिद्धांतातील ही एक महत्वाची संकल्पना आहे.

2.2.1 व्याख्या (Definition)(डेफिनेशन): X हे नमुना स्पेस S वर परिभाषित केलेले एक वेगळे रँडम व्हेरिएबल आहे आणि प्रत्येक डिस्क्रीट रँडम व्हेरिएबल $P_i = P[X=x_i] \quad i=1,2,3,\dots,n$ असा क्रमांक दिल्यास X ची श्रेणी $\{X_1, X_2, X_3, \dots, X_n\}$ असेल

$$0 \leq P_i \leq 1 \quad i=1,2,3,\dots,n$$

$$\sum_{i=1}^n P_i$$

मग P या कार्याला प्रोबॅबिलिटी मास फंक्शन (PMF) म्हणतात.

1) निःपक्षपाती(फेअर) डाय रोल करणे:

चला, सहा बाजू असलेला फासा फिरवण्याचे उदाहरण पाहू. फासावर दिसणारी संख्या X दर्शवूया. पॉसिबल आऊटकम्स 1,2,3,4,5,6 आहेत आणि प्रत्येक परिणाम घडण्याची शक्यता समान आहे (कारण फासा निःपक्षपाती(फेअर) आहे).

या डिस्क्रीट रँडम व्हेरिएबल X साठी PMF आहे:

$$P(X = x) = \frac{1}{6} \quad x = 1,2,3,4,5,6$$

तर फासेच्या रोलसाठी पीएमएफ आहे:

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

$$P(X = 4) = \frac{1}{6}, P(X = 5) = \frac{1}{6}, P(X = 6) = \frac{1}{6}$$

येथे, प्रोबॅबिलिटीची बेरीज अशी आहे:

$$P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) = \frac{1}{6} \times 6 = 1$$

2) नाणे फ्लिप करणे:

समजा तुम्ही योग्य नाणे फ्लिप केले आणि X ला परिणाम दर्शवू द्या: हेडसाठी $X=1$ आणि टेलसाठी $X=0$. या परिस्थितीसाठी PMF आहे:

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

तर, हेडसची प्रोबॅबिलिटी $\frac{1}{2}$ आहे आणि टेलची प्रोबॅबिलिटी देखील $\frac{1}{2}$ आहे.

उदाहरण

1) प्रोबॅबिलिटी मास फंक्शन खालील प्रमाणे दिले आहे: $f(x) = a \cdot 2^x$ $x = 0, 1, 2$ साठी, नंतर a ची किंमत शोधा.

उत्तर: " a " चे मूल्य शोधण्यासाठी आम्ही PMF प्रॉपर्टी वापरतो.

$$\begin{aligned} P(X \in E) &= \sum_{x \in E} f(x) = 1 \\ f(0) + f(1) + f(2) &= 1 \\ a \times 0^2 + a \times 1^2 + a \times 2^2 &= 1 \\ a \times (0^2 + 1^2 + 2^2) &= 1 \\ a \times (0 + 1 + 4) &= 1 \\ 5a &= 1 \\ a &= \frac{1}{5} \end{aligned}$$

2) खालील प्रोबॅबिलिटी मास फंक्शन सारणीवरून CDF निश्चित करा.

X	1	2	3	4
P(X = x)	0.2	0.05	0.25	0.5

उत्तर: $P(X < 3)$ साठी CDF शोधण्यासाठी आपण PMF चा गुणधर्म वापरतो.

$$\begin{aligned} P(X \in E) &= \sum_{x \in E} f(x) \\ P(X < 3) &= P(X = 1) + P(X = 2) \\ P(X < 3) &= 0.2 + 0.05 \\ P(X < 3) &= 0.25 \\ \text{तर, } P(X < 3) \text{ चा CDF} &= 0.25 \end{aligned}$$

3) खालील तक्ता रँडम व्हेरिअबल X साठी PMF दर्शवितो. p ची किंमत शोधा.

X	2	4	6
P(X = x)	p	p + 0.5	0.2

उत्तर: p ची किंमत शोधण्यासाठी आपण सूत्र वापरतो

$$\begin{aligned} P(X \in E) &= \sum_{x \in E} f(x) = 1 \\ p + (p + 0.5) + 0.2 &= 1 \\ 2p + 0.7 &= 1 \\ 2p &= 1 - 0.7 \\ 2p &= 0.3 \\ p &= \frac{0.3}{2} \\ p &= 0.15 \end{aligned}$$

सरावासाठी उदाहरणे

1) जर PMF $P(x) = 2x^2 + 5x - 1$ असेल आणि $x = 0, 1$ आणि 2 असेल तर q ची किंमत शोधा.

2) प्रोबॅबिलिटी मास फंक्शन टेबलवरून b ची किंमत निश्चित करा.

X	1	2	3	4
P(X = x)	b	$2b^2$	$3b$	0.3

लक्षात ठेवण्यासारखे महत्वाचे मुद्दे:

- PMF फक्त डिस्क्रीट रँडम व्हेरिअबल्सना लागू होते (जे मर्यादित किंवा मोजता येण्याजोग्या असीम मूल्यांचा संच घेऊ शकतात).

- प्रोबॅबिलिटीच्या सर्व पॉसिबल आऊट कम्सच्या प्रोबॅबिलिटीची बेरीज 1 च्या बरोबरीची असणे आवश्यक आहे.
- PMF डिस्क्रीट रँडम व्हेरिएबलसाठी प्रत्येक विशिष्ट परिणामाची प्रोबॅबिलिटी मोजण्याचा एक मार्ग प्रदान करते.

2.2.1 प्रोबॅबिलिटी डिस्ट्रीब्यूशन फंक्शन (Probability Distribution Function) (p.d.f)

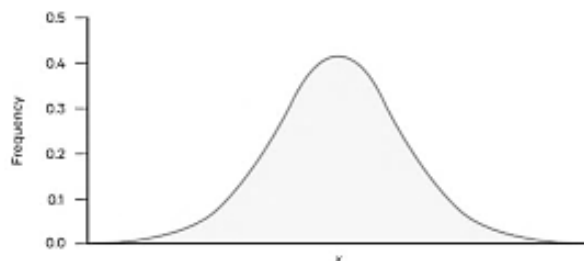


Figure 2.2: प्रोबॅबिलिटी डिस्ट्रीब्यूशन फंक्शन (Probability Distribution Function) (p.d.f)

प्रोबॅबिलिटी डिस्ट्रीब्यूशन फंक्शन (पीडीएफ), विशेषतः सतत रँडम व्हेरिएबल्ससाठी, हे एक फंक्शन आहे जे रँडम व्हेरिएबल्स ने विशिष्ट मूल्य घेण्याची सापेक्ष शक्यता प्रदान करते. पीडीएफ रँडम व्हेरिएबल्सच्या पॉसिबल मूल्यांवर प्रोबॅबिलिटी चे वितरण वर्णन करते. सतत रँडम व्हेरिएबल्ससाठी प्रोबॅबिलिटी सिद्धांतातील ही एक महत्त्वाची संकल्पना आहे.

2.2.2 व्याख्या(Definition)(डेफिनेशन): कोणत्याही रँडम व्हेरिएबल्स X साठी, जिथे त्याचे मूल्य ' x ' बिंदूवर मूल्यांकन केले जाते, तेव्हा प्रोबॅबिलिटी डिस्ट्रीब्यूशन फंक्शन प्रोबॅबिलिटी देते की X ने x च्या बरोबरीचे मूल्य घेतले आहे.

आपण प्रोबॅबिलिटी डिस्ट्रीब्यूशन असे दर्शवितो की, $F(x) = P(X \leq x)$

एक सतत रँडम व्हेरिएबल X साठी, प्रोबॅबिलिटी डिस्ट्रीब्यूशन फंक्शन (पीडीएफ), जे $F(x)$ म्हणून दर्शविले जाते, त्याचे खालील गुणधर्म आहेत:

अ-नकारात्मकता(Non-negativity):

सर्व x साठी $F(x) \geq 0$

पीडीएफ नेहमी x च्या सर्व मूल्यांसाठी शून्यापेक्षा जास्त किंवा समान असते.

एकूण प्रोबॅबिलिटी 1 आहे (Total Probability is 1):

पीडीएफच्या वक्राखालील एकूण क्षेत्रफळ 1 च्या बरोबरीचे असले पाहिजे, जे एकूण प्रोबॅबिलिटी दर्शवते:

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

याचा अर्थ असा की रँडम व्हेरिएबल X ने संभाव्य मूल्यांच्या संपूर्ण श्रेणीतील कोणतेही मूल्य घेण्याची प्रोबॅबिलिटी 1 आहे.

2.2.3 प्रोबॅबिलिटी ओवर एंटरव्हल (Probability over an Interval):

पीडीएफच्या वक्र अंतर्गत a आणि b मधील क्षेत्र हे रँडम व्हेरिएबल X विशिष्ट अंतराल $[a, b]$ मध्ये येण्याची शक्यता दर्शवते. हे मोजण्यासाठी PDF हे $[a, b]$ मध्यांतरावर एकत्रित केले आहे

$$P(a \leq x \leq b) = \int_a^b F(x)dx$$

प्रोबॅबिलिटी डिस्ट्रीब्यूशन कार्यासाठी वेगवेगळी सूत्रे खाली सूचीबद्ध आहेत:

2.2.4 डिसक्रीट रँडम व्हेरिएबल चे प्रोबॅबिलिटी डिस्ट्रीब्यूशन (Probability Distribution of a Discrete Random Variable): डिसक्रीट रँडम व्हेरिएबल हे असे व्हेरिएबल आहे जे 0, 1, 2, 3 ही भिन्न मोजण्यायोग्य मूल्ये घेते. एका स्वतंत्र रँडम व्हेरिएबलच्या प्रोबॅबिलिटी डिस्ट्रीब्यूशन सूत्र असे आहे,

प्रोबेबिलिटी डिस्ट्रिब्यूशन कार्य:

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

2.2.5 कंटीन्यूअस व्हेरिअबल चे प्रोबेबिलिटी डिस्ट्रिब्यूशन (Probability Distribution of a Continuous Random Variable):

कंटीन्यूअस रँडम व्हेरिअबल हे एक व्हेरिअबल आहे, जे अनंतपणे अनेक व्हॅल्यूज घेते. कंटीन्यूअस रँडम व्हेरिअबलच्या प्रोबेबिलिटी डिस्ट्रिब्यूशन सूत्र असे आहे,

प्रोबेबिलिटी डिस्ट्रिब्यूशन कार्य: $F(x) = P(X \leq x)$

याव्यतिरिक्त, प्रोबेबिलिटी डेंसिटी फंक्शन (PDF) हे CDF चे डेरीवेटीव्ह आहे:

$$f(x) = \frac{d}{dx} (F(x)), \quad \text{where } F(x) = \int_a^b f(u) du$$

उदाहरण

1) समजा आपण दोन फासे फेकले. फासाच्या बेरजेसाठी प्रोबेबिलिटी सारणी बनवा. शक्यता आहेत: 2,3,4,5,6,7,8,9,10,11,12.

उत्तर: प्रोबेबिलिटी डिस्ट्रिब्यूशन टेबल

X	2	3	4	5	6	7	8	9	10	11	12
P(x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

2) रँडमपणे नवडलेल्या शहर ब्लॉकवरील घरांमध्ये राहणाऱ्या वृद्धांची संख्या खालील प्रोबेबिलिटी डिस्ट्रिब्यूशनद्वारे वर्णन केली आहे.

उत्तर: प्रोबेबिलिटी डिस्ट्रिब्यूशन टेबल

नंबर ऑफ ओल्ड मेन (x)	3	4	5	6
प्रोबेबिलिटी P(x)	0.50	0.25	0.10	?

3) रँडमपणे निवडलेल्या घरात 6 किंवा त्याहून अधिक वृद्ध लोक राहतात याची शक्यता किती आहे?

उत्तर: सर्व $p(\text{प्रोबेबिलिटी})$ ची बेरीज १ च्या बरोबरीची आहे

एका घरात सहा किंवा त्याहून अधिक वृद्ध लोक राहतात याची शक्यता किती आहे,

$$= 1 - (0.50 + 0.25 + 0.10) = 0.15$$

अशा प्रकारे, एका घरात सहा किंवा त्याहून अधिक वृद्ध लोक राहत असल्याची शक्यता 0.15 इतकी आहे.

महत्त्वाचे मुद्दे:

- पीडीएफ हा कंटीन्यूअस रँडम व्हेरिअबलसाठी वापरला जातो, तर डिस्क्रीट रँडम व्हेरिअबलसाठी, आपण प्रोबेबिलिटी मास फंक्शन (PMF) वापरतो.
- कंटीन्यूअस रँडम व्हेरिअबलसाठी एका बिंदूवर प्रोबेबिलिटी शून्य असते, कारण एका बिंदूवर वक्र अंतर्गत क्षेत्र शून्य असते. त्याऐवजी प्रोबेबिलिटी अंतराने आढळतात.

2.2.6 क्युमुलेटीव डिस्ट्रिब्यूशन फंक्शन (Cumulative Distribution Function) (c.d.f.)

क्युमुलेटीव डिस्ट्रिब्यूशन फंक्शन (CDF) हे एक गणितीय कार्य आहे जे चल एखाद्या विशिष्ट संख्येपेक्षा कमी किंवा समान व्हॅल्यू घेईल याची प्रोबेबिलिटी प्रदान करते.

X च्या सर्वात लहान शक्य व्हॅल्यूसाठी CDF 0 पासून सुरू होते आणि X च्या सर्वात मोठ्या शक्य व्हॅल्यूच्या जवळ येताच 1 पर्यंत वाढते. हे एक कमी न होणारे कार्य आहे जे रँडम व्हेरिअबलसलाच्या डिस्ट्रिब्यूशनचे संपूर्ण वर्णन प्रदान करते.

उदाहरणार्थ, जर तुम्ही 80 च्या चाचणी गुणांसाठी CDF पाहत असाल आणि ते तुम्हाला 0.75 देते, तर याचा अर्थ असा की रँडम विद्यार्थ्याचा गुण 80 किंवा त्यापेक्षा कमी असण्याची 75% शक्यता आहे.

थोडक्यात, CDF तुम्हाला पुढे जाताना प्रोबॅबिलिटी सारांशित करून रँडम व्हेल्यू एका विशिष्ट श्रेणीत असण्याची शक्यता समजून घेण्यास मदत करते.

व्याख्या(Definition)(डेफिनेशन)::

रँडम व्हेरिएबल X चे क्युमुलेटीव डिस्ट्रिब्यूशन फंक्शन $F(x)$ असे परिभाषित केले आहे:

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

क्युमुलेटीव डिस्ट्रिब्यूशन फंक्शन (CDF) साठी वेगवेगळी सूत्रे खाली सूचीबद्ध आहेत:

डिस्क्रीट रँडम व्हेरिएबलसाठी क्युमुलेटीव डिस्ट्रिब्यूशन फंक्शन: X हा cdf $F(x)$ असलेला एक रँडम व्हेरिएबल मानूया, नंतर $P[a < X \leq b] = F(b) - F(a)$.

कंटीन्यूअस रँडम व्हेरिएबलसाठी क्युमुलेटीव डिस्ट्रिब्यूशन फंक्शन: X हा cdf $F(x)$ असलेला एक कंटीन्यूअस व्हेरिएबल मानूया, नंतर

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

उदाहरण

1) रँडम व्हेरिएबल X चा विचार करा जे दोन नाण्यांच्या फ्लिपमध्ये मिळवलेल्या हेडची संख्या दर्शविते. X चा CDF $F(x)$ काढा.

उत्तर: X ची पॉसिबल मूल्ये 0, 1 आणि 2 आहेत. प्रोबॅबिलिटी आहेत:

$$P(X = 0) = \frac{1}{4} \quad (\text{हेड नाही})$$

$$P(X = 1) = \frac{1}{2} \quad (\text{एक हेड})$$

$$P(X = 2) = \frac{1}{4} \quad (\text{दोन हेड})$$

आणि CDF $F(x)$ आहेत:

$$F(x) = \begin{cases} 0 & \text{जर } x < 0 \\ \frac{1}{4} & \text{जर } 0 \leq x < 1 \\ \frac{3}{4} & \text{जर } 1 \leq x < 2 \\ 1 & \text{जर } x \geq 2 \end{cases}$$

2) X ला $0 \leq x \leq 1$ साठी PDF $f(x) = 2x$ सह सतत रँडम व्हेरिएबल समजा. X चा CDF $F(x)$ शोधा.

उत्तर: CDF शोधण्यासाठी PDF इंटीग्रेट करा:

$$F(x) = \int_{-\infty}^x f(t) dt = \int_0^x 2t dt = x^2 \text{ for } 0 \leq x \leq 1$$

अशाप्रकारे, CDF असे आहे

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ x^2 & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x > 1 \end{cases}$$

3) $x \geq 0$ साठी CDF $F(x) = 1 - e^{-x}$ दिल्यास X 1 आणि 2 च्या दरम्यान असण्याची प्रोबॅबिलिटी शोधा.

उत्तर: X हा 1 आणि 2 च्या दरम्यान असण्याची प्रोबॅबिलिटी अशी आहे

$$P(1 \leq X \leq 2) = F(2) - F(1)$$

गणना करा: $(2) = 1 - e^{-2}$ आणि $F(1) = 1 - e^{-1}$

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

सरावासाठी उदाहरणे

1) X हा खालील वितरणासह मिश्रित रँडम व्हेरिअबल मानावा: $P(X=0) = 0.2$, $P(X=1) = 0.3$ आणि सतत भाग $[2, 3]$ वर एकसमानपणे वितरित केला आहे. CDF $F(x)$ शोधा.

2) $-1 \leq x \leq 1$ साठी $f(x) = \frac{3}{4}(1 - x^2)$ ही PDF दिली तर CDF $F(x)$ शोधा.

उदाहरण

1) नाणे टाकणे

समजा तुमच्याकडे एक फेअर नाणे आहे आणि तुम्ही ते हेड्सवर उतरते की टेलवर, यावर पैज लावत आहात. जर आपण निकाल 1 विचारात घेतला तर: जर हेड्सवर, तुम्ही 20 रुपये जिंकलात आणि निकाल 2: जर टेलवर, तर तुम्ही 20 रुपये गमावाल.

हेड्सची संभाव्यता $P(H) = \frac{1}{2}$ आणि टेलवर, तुम्ही 20 रुपये गमावाल.

तुमच्या जिंकण्याचे अपेक्षित मूल्य असे मोजता येते:

$$E(X) = 20 \times \frac{1}{2} + (-20) \times \frac{1}{2} = 10 + (-10) = 0$$

म्हणून, पैजचे अपेक्षित मूल्य रु.0 आहे. याचा अर्थ, सरासरी, तुम्हाला दीर्घकाळात पैसे मिळत नाहीत किंवा गमावले जात नाहीत, हा एक फेअर गेम आहे.

2) लॉटरी तिकीट

समजा तुम्ही 1 रुपये किमतीचे लॉटरी तिकीट खरेदी करता आणि त्याचे पॉसिबल परिणाम असे आहेत:

जर आपण जिंकण्याचा विचार केला तर: 100 रुपये (0.01 च्या प्रोबॅबिलिटीसह) आणि पराभव: 0 रुपये (0.99 च्या प्रोबॅबिलिटीसह)

लॉटरी तिकीटाचे एक्स्पेक्टेड व्हॅल्यू असे आहे:

$$E(X) = 100 \times 0.01 + 0 \times 0.99 = 1$$

तर, या तिकीटाचे अपेक्षित मूल्य एक्स्पेक्टेड व्हॅल्यू 1 रुपये आहे, याचा अर्थ असा की सरासरी, तुम्हाला बरोबरी मिळण्याची अपेक्षा आहे. तथापि, हे या वस्तुस्थितीचे कारण नाही की, प्रत्यक्षात, तुम्ही बहुतेक वेळा तुमचे 1 रुपये गमावाल.

3) X च्या खालील प्रोबॅबिलिटी डिस्ट्रीब्यूशनसाठी

X	0	1	2	3	4
PX=x	k	2k	4k	2k	k

i) k ची किंमत शोधा,

ii) $P(X \geq 2)$, $P(X < 3)$, $P(X \leq 1)$

उत्तर: i) $P(X = x)$ साठी, आपल्याकडे आहे

$$\sum_{x=0}^4 p(x) = 1$$

$$\therefore k + 2k + 4k + 2k + k = 1$$

$$\therefore 10k = 1$$

$$\therefore k = \frac{1}{10}$$

ii) X चे प्रोबॅबिलिटी डिस्ट्रीब्यूशन खालील तक्त्यामध्ये दाखवल्याप्रमाणे आहे:

X	0	1	2	3	4
PX=x	110	2110=210	4110=410	110=210	110

$$\begin{aligned}
 P(X \geq 2) &= p(2) + p(3) + p(4) \\
 &= \frac{4}{10} + \frac{2}{10} + \frac{1}{10} \\
 &= \frac{7}{10}
 \end{aligned}$$

$$\begin{aligned}
 P(X < 3) &= p(0) + p(1) + p(2) \\
 &= \frac{1}{10} + \frac{2}{10} + \frac{4}{10} \\
 &= \frac{7}{10}
 \end{aligned}$$

$$\begin{aligned}
 P(X \leq 1) &= p(0) + p(1) \\
 &= \frac{1}{10} + \frac{2}{10} \\
 &= \frac{3}{10}
 \end{aligned}$$

4) रँडम व्हेरिअबल X मध्ये खालील प्रोबॅबिलिटी डिस्ट्रीब्यूशन असते:

X	0	1	2	3	4	5	6
PX=x	k	3k	5k	7k	9k	11k	13k

i) k $P(X \geq 2)$, $P(0 < X < 4)$ शोधा

ii) x चा cdf मिळवा.

उत्तर:

i) $P(X)$ साठी, आपल्याकडे

$$\sum_{i=0}^6 P(x) = 1$$

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

$$\therefore 49k = 1$$

$$\therefore k = \frac{1}{49}$$

आता X चे प्रोबॅबिलिटी डिस्ट्रीब्यूशनखालील तक्त्यामध्ये दाखवल्याप्रमाणे आहे:

X	0	1	2	3	4	5	6
PX=x	149	349	549	749	949	1149	1349

टेबलावरून,

$$\begin{aligned}
 P(X \geq 2) &= p(2) + p(3) + p(4) + p(5) + p(6) \\
 &= \frac{5}{49} + \frac{7}{49} + \frac{9}{49} + \frac{11}{49} + \frac{13}{49} \\
 &= \frac{45}{49}
 \end{aligned}$$

$$\begin{aligned}
 P(0 < X < 4) &= p(1) + p(2) + p(3) \\
 &= \frac{3}{49} + \frac{5}{49} + \frac{7}{49} = \frac{15}{49}
 \end{aligned}$$

ii) X चा Cdf:

व्याख्येनुसार,

$$F(x_i) = P_1 + P_2 + P_3 + \dots + P_n \quad i = 1, 2, 3, \dots, n$$

$$\therefore F(x_0) = P_0 = \frac{1}{49}$$

$$\therefore F(x_1) = P_0 + P_1 = \frac{1}{49} + \frac{3}{49} = \frac{4}{49}$$

$$\therefore F(x_2) = P_0 + P_1 + P_2 = \frac{4}{49} + \frac{5}{49} = \frac{9}{49}$$

$$\therefore F(x_3) = P_0 + P_1 + P_2 + P_3 = \frac{9}{49} + \frac{7}{49} = \frac{16}{49}$$

$$\therefore F(x_4) = P_0 + P_1 + P_2 + P_3 + P_4 = \frac{16}{49} + \frac{9}{49} = \frac{25}{49}$$

$$\therefore F(x_5) = P_0 + P_1 + P_2 + P_3 + P_4 + P_5 = \frac{25}{49} + \frac{11}{49} = \frac{36}{49}$$

$$\therefore F(x_6) = P_0 + P_1 + P_2 + P_3 + P_4 + P_5 + P_6 = \frac{36}{49} + \frac{13}{49} = \frac{49}{49} = 1$$

अशाप्रकारे, X चा cdf खालील तक्त्यामध्ये दाखवल्याप्रमाणे आहे,

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

सरावासाठी उदाहरणे

- 33 नाण्यांच्या टॉसमधून मिळणाऱ्या हेडच्या संख्येसाठी एक्स्पेक्टेड व्हॅल्यू किती आहे?
- एका मानक डेकवरून सहा कार्डे विकली जातात, क्रमाची पर्वा न करता. जर एक किंवा अधिक एसेस आले तर खेळाडूला 45 रुपये जिंकतात. अन्यथा तो किंवा ती 30 रुपये गमावते. खेळाडूची एक्स्पेक्टेड व्हॅल्यू किती आहे?
- सोळा फासे फिरवले जातात आणि क्रम न लावलेले निकाल नोंदवले जातात. जर किमान पाच षटकार आले तर खेळाडूला 70 रुपये जिंकतात, अन्यथा खेळाडूंना 4 रुपये द्यावे लागतात. खेळाडूंना एक्स्पेक्टेड व्हॅल्यू किती आहे?

2.3 व्हेरिएन्स (Variance):

व्हेरिएन्स ही एक स्टॅटिस्टिक्सकल माप आहे जी डेटा पॉइंट्सच्या संचाचा प्रसार किंवा विसर्जन दर्शवते. ते डेटासेटमधील डेटा पॉइंट्स सरासरी (सरासरी) मूल्यापेक्षा किती वेगळे आहेत हे दर्शवते.

उच्च व्हेरिएन्स दर्शवते की डेटा पॉइंट्स सरासरीभोवती विस्तृतपणे पसरलेले आहेत, तर कमी व्हेरिएन्स दर्शवते की ते सरासरीभोवती जवळून क्लस्टर केलेले आहेत. व्हेरिएन्स डेटासेट मधील परिवर्तनशीलता समजून घेण्यास मदत करते. वरील उदाहरणात, 3.7 चे व्हेरिएन्स सूचित करते की डेटा पॉइंट्स सरासरीपासून काही प्रमाणात पसरलेले आहेत. वास्तविक-जगातील अनुप्रयोगांमध्ये, जोखीम मूल्यांकन करण्यासाठी वित्त मध्ये, सुसंगतता मोजण्यासाठी गुणवत्ता नियंत्रण मध्ये आणि इतर अनेक क्षेत्रात परिवर्तनशीलतेचे विश्लेषण करण्यासाठी व्हेरिएन्स वापरली जाते.

व्हेरिएन्सचा फॉर्म्युला

1. पोपुलेशन व्हेरिएन्स (Population Variance) (σ^2):

पोपुलेशन व्हेरिएन्सचा फॉर्म्युला:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{N}$$

जेथे:

- σ^2 हा पोपुलेशन व्हेरिएन्स आहे.

- x_i हा पोपुलेशनतील प्रत्येक डेटा बिंदू दर्शवतो.
- μ हा पोपुलेशनचा सरासरी आहे.
- N हा पोपुलेशनतील डेटा बिंदूंची एकूण संख्या आहे.

2. सॅम्पल व्हेरिएन्स (Sample Variance) (s^2):

सॅम्पल व्हेरिएन्सचा फॉर्म्युला:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

जेथे:

- S^2 हा सॅम्पल व्हेरिएन्स आहे.
- x_i हा सॅम्पलतील प्रत्येक डेटा बिंदू दर्शवतो.
- $\bar{x}$ हा सॅम्पलचा मध्य आहे.
- n हा सॅम्पलतील डेटा बिंदूंची एकूण संख्या आहे.

2.3.1 एक्स्पेक्टेड व्हॅल्यू आणि व्हेरिएन्स यांच्यातील संबंध:

एक्स्पेक्टेड व्हॅल्यू वापरून खालील प्रकारे व्हेरिएन्स देखील व्यक्त केली जाऊ शकते:

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

एक्स्पेक्टेड व्हॅल्यू आणि व्हेरिएन्स यांच्यातील संबंध अधिक स्पष्टपणे दर्शविण्यासाठी हे सूत्र विस्तृत केले जाऊ शकते:

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

लिनिअरटीचे एक्स्पेक्शन वापर करून, हे होते:

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

म्हणून, रँडम व्हेरिएबल X चा फरक X च्या वर्गाच्या एक्स्पेक्टेड व्हॅल्यू आणि X च्या एक्स्पेक्टेड व्हॅल्यूच्या वर्गातील फरक म्हणून मोजता येतो:

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

उदाहरण

1) नमुना डेटाच्या व्हेरिएन्सची गणना करा: 7, 11, 15, 19, 24

उत्तर: आमच्याकडे डेटा आहे 7, 11, 15, 19, 24

डेटाचा मीन शोधा.

$$\begin{aligned}\bar{x} &= (7 + 11 + 15 + 19 + 24)/5 \\ &= 76/5 \\ &= 15.2\end{aligned}$$

व्हेरिएन्सचे सूत्र वापरून आपल्याला मिळते,

$$\begin{aligned}s^2 &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \\ &= \frac{(67.24 + 17.64 + 0.04 + 14.44 + 77.44)}{(5 - 1)} \\ &= 176.8/4 \\ &= 44.2\end{aligned}$$

2) जर डेटाचा फरक 12 असेल आणि सरासरीपासून डेटाच्या वर्ग फरकांची बेरीज 156 असेल तर निरीक्षणांची संख्या काढा.

उत्तर: आमच्याकडे,

$$\begin{aligned}(x_i - \mu)^2 &= 156 \\ \sigma^2 &= 12\end{aligned}$$

व्हेरिएन्सचे सूत्र वापरून आपल्याला मिळते,

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{N}$$

$$12 = 156/n$$

$$n = 156/12$$

$$n = 13$$

2.3.2 स्टॅंडर्ड डेव्हिएशन (Standard Deviation):

जरीस्टॅंडर्ड डेव्हिएशनची कल्पना डिस्क्रिट आणि कंटिन्युअस रँडम व्हेरिएबल्ससाठी सारखीच असली तरी, आपण ज्या पद्धतींनी त्याची गणना करतो त्या व्हेरिएबलच्या प्रकारानुसार काही प्रमाणात बदलतात. कंटिन्युअस रँडम व्हेरिएबलच्या स्टॅंडर्ड विचलन आणि डिस्क्रिट रँडम व्हेरिएबलमधील मुख्य फरक खाली दिले आहेत.

1. डिस्क्रिट रँडम व्हेरिएबलचे स्टॅंडर्ड डेव्हिएशन (Standard deviation of discrete random variable):

डिस्क्रिट रँडम व्हेरिएबलसाठी, स्टॅंडर्ड विचलन हे मोजते की रँडम व्हेरिएबलचे वैयक्तिक परिणाम एक्स्पेक्टेड व्हॅल्यू (सरासरी) पासून किती विचलित होतात, जिथे परिणाम वेगळे, मोजण्यायोग्य मूल्ये असतात.

स्टॅंडर्ड डेव्हिएशन मोजण्यासाठी पायऱ्या:

a) एक्स्पेक्टेड व्हॅल्यू(सरासरी) शोधा:

$$E(x) = \sum_{i=1}^n x_i \cdot P(x_i)$$

जिथे x_i रँडम व्हेरिएबल विशिष्ट संभाव्य मूल्ये X आहेत आणि (x_i) ही प्रत्येक मूल्याची संभाव्यता आहे x_i

b) $E(X^2)$ ची गणना करा:

$$E(X^2) = \sum_{i=1}^n x_i^2 \cdot P(x_i)$$

c) व्हेरिएन्स शोधा: सूत्र वापरा:

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

d) स्टॅंडर्ड डेव्हिएशन शोधा: शेवटी, स्टॅंडर्ड डेव्हिएशन हे व्हेरिएन्सचे वर्गमूळ आहे:

$$\sigma_X = \sqrt{Var(X)}$$

2. कंटिन्युअस रँडम व्हेरिएबलचे स्टॅंडर्ड डेव्हिएशन (Standard deviation of continuous random variable):

कंटिन्युअस रँडम व्हेरिएबलचे स्टॅंडर्ड डेव्हिएशन एक्स्पेक्टेड व्हॅल्यूभोवती असलेल्या मूल्यांच्या श्रेणीचे मोजमाप करते, परंतु पॉसिबल आउटकमस अगणित आणि अंतहीन असतात. विशिष्ट श्रेणीतील कोणतेही मूल्य (जसे की 0 आणि 1 मधील वास्तविक संख्या, किंवा वास्तविक रेषेवरील कोणताही मध्यांतर) कंटिन्युअस रँडम व्हेरिएबलला नियुक्त केले जाऊ शकते.

स्टॅंडर्ड डेव्हिएशन मोजण्यासाठी पायऱ्या:

a) एक्स्पेक्टेड व्हॅल्यू शोधा (सरासरी): प्रोबॅबिलिटी डेन्सिटी फंक्शन $f(x)$ असलेल्या कंटिन्युअस रँडम व्हेरिएबल X साठी एक्स्पेक्टेड व्हॅल्यू खालीलप्रमाणे दिले आहे:

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

जिथे $f(x)$ हे व्हेरिएबलचे प्रोबॅबिलिटी डेन्सिटी फंक्शन आहे.

b) $E(X^2)$ शोधा: X^2 चे एक्स्पेक्टेड व्हॅल्यू आहे:

$$E(x^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx$$

b) व्हेरिएन्स शोधा: व्हेरिएन्स आहे:

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

d) स्टँडर्ड डेव्हिएशन शोधा: शेवटी, स्टँडर्ड डेव्हिएशन हे व्हेरिएन्सचे वर्गमूळ आहे:

$$\sigma_X = \sqrt{\text{Var}(X)}$$

उदाहरणे:

1) 52 पत्त्यांच्या चांगल्या प्रकारे हलवलेल्या पॅकमधून दोन पत्ते सलग काढली जातात. एक्काच्या संख्येचा मीन आणि व्हेरियन्स शोधा.

उत्तर: X हा एक्काची संख्या दर्शविणारा एक रँडम व्हेरिएबल समजा. X ची पॉसिबल आऊट कम्स 0,1,2 आहेत.

$$P(X = x = 0) = p(\text{एक्का नसून आणि एक्का नसून})$$

$$P(X = 0) = p(\text{एक्का नसून}) \times p(\text{एक्का नसून})$$

$$P(X = 0) = \frac{48}{52} \times \frac{48}{52}$$

$$P(X = 0) = \frac{144}{169}$$

$$P(X = x = 1) = p(\text{एक्का आणि एक्का नसून किंवा एक्का नसून आणि एक्का})$$

$$P(X = 1) = p(\text{एक्का आणि एक्का नसून}) + p(\text{एक्का नसून आणि एक्का})$$

$$P(X = 1) = p(\text{एक्का}) \times p(\text{एक्का नसून}) + p(\text{एक्का नसून}) \times p(\text{एक्का})$$

$$P(X = 1) = \frac{4}{52} \times \frac{48}{52} + \frac{48}{52} \times \frac{4}{52}$$

$$P(X = 1) = \frac{24}{169}$$

$$P(X = x = 2) = p(\text{एक्का आणि एक्का})$$

$$P(X = 2) = p(\text{एक्का}) \times p(\text{एक्का})$$

$$P(X = 2) = \frac{4}{52} \times \frac{4}{52}$$

$$P(X = 2) = \frac{1}{169}$$

म्हणून प्रोबॅबिलिटी डिस्ट्रीब्यूशन असे दिले आहे,

X	0	1	2
P(x)	$\frac{144}{169}$	$\frac{24}{169}$	$\frac{1}{169}$

$$E(X) = \sum_{i=0}^2 x_i \cdot P(x_i) = \mu = 0 \times \frac{144}{169} + 1 \times \frac{24}{169} + 2 \times \frac{1}{169} = \frac{26}{169}$$

$$E(X^2) = \sum_{i=0}^2 (x_i)^2 \cdot P(x_i) = 0^2 \times \frac{144}{169} + 1^2 \times \frac{24}{169} + 2^2 \times \frac{1}{169} = \frac{28}{169}$$

$$\text{Var}(X) = E[X^2] - [E(X)]^2 = \frac{28}{169} - \left(\frac{26}{169}\right)^2 = \frac{24}{169}$$

2) जर तुम्ही सहा बाजू असलेला फासा रोल केला तर पॉसिबल आऊटकम्स 1,2,3,4,5 आणि 6 असतील, प्रत्येकाची प्रोबॅबिलिटी 1/6 असेल. शोधा i) एक्स्पेक्टेड व्हॅल्यू ii) व्हेरिएन्स iii) स्टँडर्ड डेव्हिएशन

उत्तर: येथे पॉसिबल आऊटची प्रोबॅबिलिटी आहे,

येथे पॉसिबल आऊटची प्रोबॅबिलिटी आहे,

$P(X = 1) = \frac{1}{6}, P(X = 2) = \frac{1}{6}, P(X = 3) = \frac{1}{6}, P(X = 4) = \frac{1}{6}, P(X = 5) = \frac{1}{6}, P(X = 6) = \frac{1}{6}$
एक्स्पेक्टेड व्हॅल्यू

$$E(x) = \sum_{i=1}^n x_i \cdot P(x_i) = 3.5$$

$$\text{व्हेरिएन्स, } Var(X) = E[X^2] - [E(X)]^2 = 2.92.$$

स्टॅण्डर्ड डेव्हिएशन, $\sigma_x = \sqrt{Var(X)}$ अंदाजे 1.71 आहे.

3) 0 आणि 1 मधील युनिफॉर्म डिस्ट्रीब्यूशनचा विचार करा, जिथे प्रोबॅबिलिटी डेन्सिटी फंक्शन

$f(x) = 1$ साठी $0 \leq x \leq 1$ साठी आहे.

उत्तर: एक्स्पेक्टेड व्हॅल्यू

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

$$\text{व्हेरिएन्स, } Var(X) = E[X^2] - [E(X)]^2 = \frac{1}{12}$$

$$\text{स्टॅण्डर्ड डेव्हिएशन, } \sigma_x = \sqrt{Var(X)} = \sqrt{\frac{1}{12}} \approx 0.29$$

2.4 बर्नोली डिस्ट्रीब्यूशन (Bernoulli Distribution):

बर्नोली डिस्ट्रीब्यूशन हे अशा परिस्थितींमध्ये प्रोबॅबिलिटी मोजण्यासाठी एक मूलभूत साधन म्हणून परिभाषित केले आहे जिथे फक्त दोन पर्याय असतात (म्हणजेच बायनरी सिच्युएशन), जसे की उत्तीर्ण होणे किंवा अपयशी होणे, जिंकणे किंवा हरणे, किंवा सरळ हो किंवा नाही. बर्नोली डिस्ट्रीब्यूशन हे नाणे उलटण्यासारखे असू शकते. बायनरी परिस्थितींमध्ये फक्त दोन शक्यता असतात: यश किंवा अपयश. उदाहरणार्थ, नाणे उलटवताना, ते हेड्सवर उतरू शकते, जे यश(सक्सेस) दर्शवते; किंवा टेल, जे अपयश(फेलियर) दर्शवते. हेड्स मिळविण्याची शक्यता p आहे आणि टेल मिळण्याची शक्यता $1-p$ किंवा q आहे.

2.4.1 व्याख्या(Definition)(डेफिनेशन):

बर्नोली डिस्ट्रीब्यूशनात फक्त दोन पॉसिबल परिणामांसह एक रँडम व्हेरिएबल समाविष्ट आहे. जेव्हा चल 1 च्या बरोबरीचा असतो, तेव्हा ते प्रोबॅबिलिटी p सह यश दर्शवते तर जेव्हा चल 0 च्या बरोबरीचा असतो, तेव्हा ते फेलियर दर्शवते. सक्सेसची पॉसिबल q आहे, जी $1-p$ म्हणून मोजली जाते.

2.4.2 बर्नोली डिस्ट्रीब्यूशन संबंधित संज्ञा (Terminologies associated with Bernoulli Distribution)

1. **सक्सेस आणि फेल्युअर:** बर्नोली ट्रायलमध्ये, निकालांना सक्सेस (1) आणि **फेल्युअर** (0) असे संबोधले जाते.
2. **प्रोबॅबिलिटी ऑफ सक्सेस (p):** एकाच बर्नोली ट्रायलमध्ये सक्सेस मिळण्याची शक्यता 0 आणि 1 च्या दरम्यान असते.
3. **प्रोबॅबिलिटी ऑफ फेल्युअर (q):** फेलियरची शक्यता सक्सेस मिळण्याच्या प्रोबॅबिलिटीला पूरक असते आणि ती q म्हणून दर्शविली जाते, जिथे $q = 1 - p$
4. **रँडम व्हेरिएबल (X):** बर्नोली डिस्ट्रीब्यूशनच्या संदर्भात, X हा व्हेरिएबल दर्शवितो जो 1 किंवा 0 व्हॅल्यू घेऊ शकतो, जो सक्सेस होणाऱ्यांची संख्या दर्शवितो.
5. **बर्नोली ट्रायल:** फक्त दोन पॉसिबल आऊटकम्ससह एक इंडिव्हिज्युअल एक्सपेरिमेंट किंवा ट्रायल.
6. **बर्नोली पॅरामीटर:** हे बर्नोली डिस्ट्रीब्यूशनात सक्सेस मिळण्याची शक्यता (p) दर्शवते.
7. **एक्स्पेक्टेड व्हॅल्यू (सरासरी):** बर्नोली ट्रायलच्या मालिकेतील सरासरी किंवा सरासरी निकाल बहुतेकदा $E[X]$ म्हणून दर्शविला जातो.
8. **व्हेरिएन्स:** रँडम व्हेरिएबल X ची व्हॅल्यू किती बदलतात हे शोधण्यासाठी एक उपाय.

9. **बायनॉमियल डिस्ट्रीब्यूशन:** बर्नौली डिस्ट्रीब्यूशन हे बर्नौली डिस्ट्रीब्यूशनासाठी एक बिल्डिंग ब्लॉक आहे जे बर्नौली ट्रायलच्या निश्चित संख्येतील यशांची संख्या वर्णन करते.
10. **जॉमेट्रिक डिस्ट्रीब्यूशन:** बर्नौलीपासून घेतलेले आणखी एक डिस्ट्रीब्यूशन, जे पहिले सक्सेस मिळविण्यासाठी आवश्यक असलेल्या ट्रायलच्या संख्येचे मॉडेल करते.
11. **निगेटिव्ह बायनॉमियल डिस्ट्रीब्यूशन:** हे डिस्ट्रीब्यूशन विशिष्ट संख्येतील सक्सेस मिळविण्यासाठी आवश्यक असलेल्या ट्रायलच्या संख्येचे वर्णन करते. हे बर्नौली ट्रायलमधून देखील घेतले जाते.

2.4.3 बर्नौली डिस्ट्रीब्यूशन सूत्र (Formula of Bernoulli Distribution)

बर्नौली डिस्ट्रीब्यूशन सूत्र दोन पॉसिबल आऊटकमसच्या प्रोबॅबिलिटीचे वर्णन करण्यासाठी वापरले जाते. सक्सेस आणि फेलियर. ते $X \sim \text{बर्नौली}(p)$ म्हणून दर्शविले जाऊ शकते, ज्यामध्ये पॅरामीटर p सक्सेसची प्रोबॅबिलिटी दर्शवितो. बर्नौली डिस्ट्रीब्यूशन, प्रोबॅबिलिटी डिस्ट्रीब्यूशन फंक्शन (PDF) आणि क्युमुलेटीव डिस्ट्रीब्यूशन फंक्शन (CDF) द्वारे दिले जाऊ शकते.

I. बर्नौली डिस्ट्रीब्यूशनासाठी प्रोबॅबिलिटी डिस्ट्रीब्यूशन फंक्शन (Probability Distribution Function for Bernoulli Distribution)

PDF एका डिस्क्रीट रँडम व्हेरिअबलची स्पेसिफिक व्हॅल्यू घेण्याची प्रोबॅबिलिटी मोजते. बर्नौली डिस्ट्रीब्यूशनसाठी, ते सक्सेस (1) किंवा फेलियर (0) मिळण्याची प्रोबॅबिलिटी शोधण्यात मदत करते.

$$P(X = x) = p^x(1 - p)^{1-x}, x = 0, 1; \quad 0 < p < 1$$

येथे,

- फक्त 0 किंवा 1 असू शकते.
- $P(X = x)$ म्हणून दर्शविलेले PDF, रँडम व्हेरिअबल X हे स्पेसिफिक व्हॅल्यू x च्या बरोबरीचे असण्याची प्रोबॅबिलिटी मोजते.
- जर $x=1$ (सक्सेस), तर प्रोबॅबिलिटी p आहे.
- जर $x=0$ (फेलियर), तर प्रोबॅबिलिटी q आहे, जी p ची पूरक आहे आणि $(1 - p)$ म्हणून देखील लिहिता येते.

II. बर्नौली डिस्ट्रीब्यूशन साठी क्युमुलेटीव डिस्ट्रीब्यूशन फंक्शन (Cumulative Distribution Function for Bernoulli Distribution)

CDF एका विशिष्ट व्हॅल्यूपेक्षा कमी किंवा समान असलेल्या रँडम व्हेरिअबलच्या प्रोबॅबिलिटीची गणना करते. बर्नौली डिस्ट्रीब्यूशनमध्ये, ते आपल्याला 1 किंवा 0 पेक्षा कमी किंवा इकल व्हॅल्यू मिळण्याची शक्यता सांगते.

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < 0 \\ 1 - p, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

येथे,

- CDF, जो $P(X \leq x)$ म्हणून दर्शविला जातो, तो X ने x पेक्षा कमी किंवा इकल व्हॅल्यू घेण्याची प्रोबॅबिलिटीची निश्चित करतो.
- जर $x < 0$ असेल, तर प्रोबॅबिलिटी 0 असेल.
- जर $x = 0$ आणि 1 च्या दरम्यान असेल, तर प्रोबॅबिलिटी $1 - p$ असेल.
- जर $x \geq 1$ असेल, तर प्रोबॅबिलिटी 1 असेल.

2.4.4 बर्नौली डिस्ट्रीब्यूशनचे मीन आणि डिफरन्स (Mean and Variance of Bernoulli Distribution):

I. बर्नौली डिस्ट्रीब्यूशनचे मीन (Mean (μ) of Bernoulli Distribution)

जेव्हा आपण मोठ्या संख्येने स्वतंत्र बर्नौली ट्रायल करतो तेव्हा सरासरी, ज्याला एक्स्पेक्टेड व्हॅल्यू असेही म्हणतात, एवरेज आऊटकम दर्शवते. सोप्या भाषेत सांगायचे तर, जर हा प्रयोग अनेक वेळा केला गेला, तर एवरेज एक्स्पेक्टेड एवरेज निकाल प्रदान करते. विशेषतः बर्नौली वितरणासाठी, एवरेज सक्सेसची प्रोबॅबिलिटी दर्शवते (p).

उदाहरण

जर सक्सेसची प्रोबॅबिलिटी 0.5 असेल, तर अपेक्षित निकाल 0.5 असेल.

$$E[X] = \mu = p$$

बर्नौली रँडम व्हेरिअबल X साठी $P(X = 1) = p$ (सक्सेसची प्रोबॅबिलिटी), आणि

$P(X = 0) = q = 1 - p$ (फेलियरची प्रोबॅबिलिटी). एक्स्पेक्टेड व्हॅल्यू (मीन) $E[X]$ म्हणून गणना केली जाते:

$$E[X] = [P(X = 1) \times 1] + [P(X = 0) \times 0]$$

$$E[X] = (p \times 1) + (q \times 0)$$

$$E[X] = \mu = p$$

II. बर्नौली डिस्ट्रिब्युशनचे व्हेरियन्स (Variance of Bernoulli Distribution):

व्हेरियन्स म्हणजे सरासरीपासून संभाव्य परिणाम किती पसरले आहेत याचे मोजमाप. बर्नौली डिस्ट्रिब्युशनसाठी, व्हेरियन्सचे सूत्र $p(1 - p)$ आहे. व्हेरियन्स ही सक्सेसची प्रोबॅबिलिटी (p) आणि फेलियरची प्रोबॅबिलिटी ($q = 1 - p$) यावर अवलंबून असते. या दोन प्रोबॅबिलिटीचा गुणाकार करून विस्तार किंवा एक्स्पेक्टेड व्हॅल्यूपासून विचलन निश्चित केले जाते. जेव्हा p 0.5 च्या जवळ असेल तेव्हा विस्तार मोठा असेल. दुसरीकडे, जेव्हा p 0 किंवा 1 च्या जवळ असेल तेव्हा प्रसार लहान असेल.

$$\text{Var}(X) = \sigma^2 = p(1 - p) \text{ or } pq$$

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

$E[X^2]$ च्या गुणधर्माचा वापर करून, आमच्याकडे आहे,

$$E[X^2] = (1^2 \times p) + (0^2 \times q)$$

$$E[X^2] = p$$

$$E[X^2] = (1 \times p)^2 + (0 \times q)^2$$

$$E[X^2] = p^2$$

म्हणून, ही मूल्ये पुन्हा व्हेरियन्स सूत्रात बदलणे:

$$\text{Var}(X) = p - p^2$$

$$\text{Var}(X) = \sigma^2 = p(1 - p) \text{ or } pq$$

हे आपल्याला बर्नौली डिस्ट्रिब्युशनच्या व्हेरियन्सचे सूत्र देते, जे $p(1 - p)$ किंवा pq आहे.

2.4.5 बर्नौली वितरणाचे गुणधर्म (Properties of Bernoulli Distribution):

- बायनरी आउटकम (Binary Outcomes):** बर्नौली वितरण केवळ दोन संभाव्य आउटकमांसह परिस्थितींचे मॉडेल बनवते, जसे की सक्सेस (1) आणि फेलियर (0).
- कॉन्स्टंट प्रोबॅबिलिटी (Constant Probability):** बर्नौली डिस्ट्रिब्युशनातील सर्व चाचण्यांमध्ये सक्सेसची प्रोबॅबिलिटी (p) सुसंगत राहते.
- इंडिपेंडेंट ट्रायल्स (Independent Trials):** प्रत्येक ट्रायल्सचा निकाल त्यानंतरच्या ट्रायल्समुळे प्रभावित होत नाही.
- कॉम्प्लिमेंटरी प्रोबॅबिलिटी (Complementary Probability):** फेलियरची प्रोबॅबिलिटी (q) सक्सेसची प्रोबॅबिलिटी (p) १ मधून वजा करून शोधता येते; $q = 1 - p$.
- इंडिपेंडेंट प्रोबॅबिलिटी डिस्ट्रिब्युशन (Discrete Probability Distribution):** बर्नौली डिस्ट्रिब्युशन मर्यादित संख्येने भिन्न आउटकम असतात ज्यामुळे ते एक स्वतंत्र वितरण बनते.
- एक्स्पेक्टेड व्हॅल्यू (मीन) (Expected Value):** बर्नौली डिस्ट्रिब्युशनचा सरासरी सक्सेसच्या प्रोबॅबिलिटी इतका असतो (p), जो एवरेज निकाल दर्शवतो.
- व्हेरियन्स (Variance):** व्हेरियन्स (pq) बर्नौली डिस्ट्रिब्युशनातील आउटकम सरासरीपासून कसे विचलित होतात हे मोजते.

2.4.6 बर्नौली डिस्ट्रिब्युशनचा ग्राफ (Bernoulli Distribution Graph):

बर्नौली डिस्ट्रिब्युशनचा ग्राफ हा फक्त दोन बार असलेला एक साधा बार चार्ट आहे.

1. एक बार निकाल "1" दर्शवितो (सामान्यतः सक्सेस दर्शवितो), आणि त्याची उंची सक्सेसच्या प्रोबॅबिलिटीशी जुळते, "p."
2. दुसरा बार निकाल "0" दर्शवितो (सामान्यतः फेलियर दर्शवितो), आणि त्याची उंची फेलियर ची प्रोबॅबिलिटी दर्शवते, "q," जिथे $q = 1 - p$

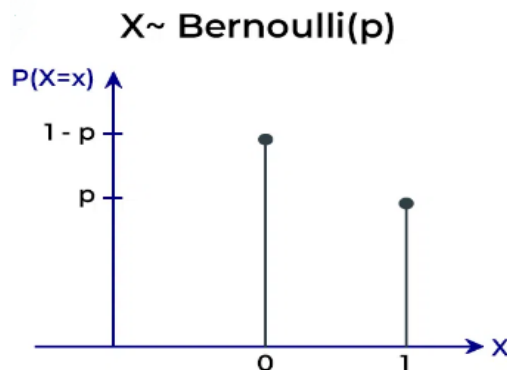


Figure 2.3: बर्नौली डिस्ट्रिब्युशनचा ग्राफ (Bernoulli Distribution Graph):

2.4.7 बर्नौली ट्रायल्स (Bernoulli Trial)

बर्नौली ट्रायल्सचे वर्णन असे केले जाते की ज्यांचे दोन पॉसिबल आउटकम आहेत, सक्सेस आणि फेल्युअर, हो आणि नाही, किंवा हेड आणि टेल. प्रत्येक ट्रायल्समध्ये, सक्सेसची शक्यता (p) आणि फेल्युअरची शक्यता (q) असते. ट्रायल्सचे नाव स्विस गणितज्ञ जेम्स बर्नौली यांच्या नावावरून ठेवण्यात आले आहे ज्यांनी प्रोबॅबिलिटीच्या क्षेत्रात महत्त्वपूर्ण योगदान दिले. बर्नौली ट्रायल्सची विशिष्ट उदाहरणे म्हणजे-

- नाणे उलटवणे (हेड किंवा टेल)
- आज पाऊस पडेल की नाही हे ठरवणे (हो किंवा नाही)
- नवजात बाळाचे (मुलगी किंवा मुलगा) लिंग निश्चित करणे.

2.4.8 बर्नौली ट्रायल्सच्या अटी (Conditions of Bernoulli Trials)

1. ट्रायल्सची निश्चित आणि मर्यादित संख्या आहे.
2. प्रत्येक ट्रायल्स दुसऱ्या ट्रायल्सपासून इंडिपेंडेंट आहे म्हणजेच एका ट्रायल्सचा निकाल इतरांवर परिणाम करत नाही.
3. प्रत्येक ट्रायल्समध्ये, फक्त दोन पॉसिबल आउटकम आहेत, सक्सेस किंवा फेल्युअर. या निकालांची प्रोबॅबिलिटी सर्व ट्रायल्समध्ये सुसंगत राहते.

उदाहरण

1) निःपक्षपाती नाणे उलथवून हेड्स (यश) मिळण्याची प्रोबॅबिलिटी शोधा.

उत्तर:

नाणे टाकण्याचा निकाल X दर्शवितो.

जर हेड्स असतील तर $X = 1$, जर टेल असतील तर $X = 0$.

निःपक्षपाती नाण्यासाठी p (सक्सेसची प्रोबॅबिलिटी) 0.5 आहे आणि

q (फेल्युअरची प्रोबॅबिलिटी) = $1 - 0.5$ आहे

$$p = 0.5.$$

2) जर एखाद्या विद्यार्थ्याला बहुपर्यायी परीक्षेत उत्तीर्ण होण्याची प्रोबॅबिलिटी 60% असेल, तर तो नापास होण्याची प्रोबॅबिलिटी किती आहे?

उत्तर:

$P(X = 0) = q$ (फेल्युअर होण्याची प्रोबॅबिलिटी)

दिलेले, $p = 0.6$ (सक्सेस होण्याची प्रोबॅबिलिटी)

$$q = 1 - p$$

$$q = 1 - 0.6$$

$$q = 0.4$$

3) जर तुम्ही एका स्टँडर्ड डेकवरून कार्ड काढले तर नॉन-एक्का कार्ड (एक्कानसणं) काढण्याची प्रोबॅबिलिटी किती आहे?

उत्तर:

$P(X = 0) = q$ (फेल्युअरची प्रोबॅबिलिटी)

दिले, $p = 1/13$ म्हणजेच, सक्सेसची प्रोबॅबिलिटी, एक्का काढणे

$$q = 1 - p$$

$$q = 1 - \frac{1}{13}$$

$$q = \frac{12}{13}$$

4) जेम्स 10 चेंडूंच्या बॅगमधून 5 वेळा चेंडू काढतो ज्यामध्ये 5 लाल आणि 5 हिरवे चेंडू बदलून घेतले जातात. हे बर्नोली ट्रायल्सचे उदाहरण आहे का ते तपासा.

उत्तर: बर्नोली ट्रायल्सच्या सर्व अटी पूर्ण झाल्या आहेत का ते आपल्याला तपासावे लागेल.

जेम्स 5 वेळा चेंडू काढतो, म्हणजेच ट्रायल्सची संख्या 5 आहे जी मर्यादित आहे.

बॉल बदलला असल्याने, प्रत्येक ट्रायल्स इंडिपेंडेंट असते.

फक्त दोनच पॉसिबल आउटकम आहेत - लाल आणि हिरवा

लाल चेंडू काढण्याची प्रोबॅबिलिटी = हिरवा चेंडू काढण्याची प्रोबॅबिलिटी $= \frac{15}{10} = \frac{1}{2}$

याचा अर्थ बर्नोली ट्रायल्सच्या सर्व अटी पूर्ण झाल्या आहेत.

म्हणून आपण म्हणतो की, दिलेले उदाहरण बर्नोली प्रयोग आहे.

5) एका फुटबॉल खेळाडूने प्रत्येक शॉटवर गोल मिळण्याची प्रोबॅबिलिटी 0.6 च्या 7 इंडिपेंडेंट फ्री शॉट्स मारले. ट्रायल्सची संख्या आणि प्रत्येक शॉटमध्ये गोल न मिळण्याची प्रोबॅबिलिटी निश्चित करा.

उत्तर: दिलेला केस बर्नोली ट्रायल्सचे उदाहरण आहे कारण ट्रायल्सची संख्या 7 आहे जी मर्यादित आहे.

फक्त दोनच संभाव्य परिणाम आहेत - गोल मिळणे आणि गोल न मिळणे.

प्रत्येक ट्रायल्स दुसऱ्यापासून इंडिपेंडेंट असते आणि गोल मिळण्याची शक्यता $= 0.6 =$ सक्सेसची प्रोबॅबिलिटी

$\Rightarrow$ गोल न मिळण्याची प्रोबॅबिलिटी = फेलियरची प्रोबॅबिलिटी $= 1 - 0.6 = 0.4$

म्हणून, ट्रायल्सची संख्या 7 आहे आणि प्रत्येक शॉटमध्ये गोल न मिळण्याची प्रोबॅबिलिटी 0.4 आहे.

2.4.9 युनिफॉर्म डिस्ट्रीब्यूशन (Uniform Distribution):

युनिफॉर्म डिस्ट्रीब्यूशन म्हणजे प्रोबॅबिलिटी डिस्ट्रीब्यूशन जे एका विशिष्ट श्रेणीतील सर्व निकालांची समान शक्यता दर्शवते. म्हणजेच प्रत्येक निकाल येण्याची प्रोबॅबिलिटी समान असते. निःपक्षपाती डायच्या साध्या रोलशी व्यवहार करत असो किंवा सतत मध्यांतरातून रँडम नंबर निवडत असो, युनिफॉर्म डिस्ट्रीब्यूशन रँडमनेस समजून घेण्यासाठी एक सरळ पण शक्तिशाली मॉडेल प्रदान करते.

युनिफॉर्म डिस्ट्रीब्यूशनचे दोन प्रकार आहेत: डिस्ट्रिक्ट युनिफॉर्म डिस्ट्रीब्यूशन आणि कंटीन्यूअस युनिफॉर्म डिस्ट्रीब्यूशन (प्राथमिक सांख्यिकीमध्ये सर्वात सामान्य प्रकार). ते रँडम व्हेरिएबल, मीन आणि व्हेरिएन्सचे घनता कार्य परिभाषित करते.

युनिफॉर्म डिस्ट्रीब्यूशन म्हणजे असे डिस्ट्रीब्यूशन ज्यामध्ये समान प्रोबॅबिलिटी असलेल्या घटनांमुळे स्थिर प्रोबॅबिलिटी असते. त्याला आयताकृती वितरण (कंटीन्यूअस युनिफॉर्म डिस्ट्रीब्यूशन) असेही म्हणतात. त्याचे दोन

पॅरामीटर्स **a** आणि **b** आहेत: **a** = किमान(Minimum) आणि **b** = कमाल(Maximum). डिस्ट्रीब्यूशन **U (a, b)** असे लिहिले आहे.

व्याख्या(Definition)(डेफिनेशन): युनिफॉर्म डिस्ट्रीब्यूशन म्हणजे प्रोबॅबिलिटी डिस्ट्रीब्यूशन एक प्रकार आहे जिथे प्रत्येक पॉसिबल निकालाची समान प्रोबॅबिलिटी असते. याचा अर्थ असा की दिलेल्या श्रेणीतील सर्व मूल्ये पाहिली जाण्याची शक्यता समान असते.

युनिफॉर्म डिस्ट्रीब्यूशनचा ग्राफ(Graph of Uniform Distribution):

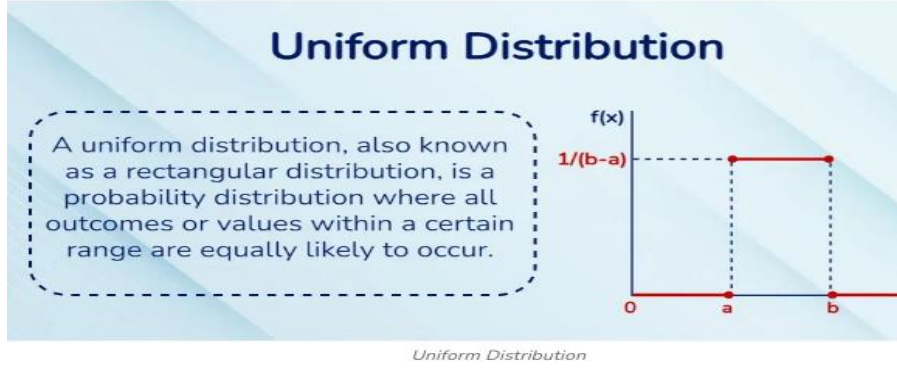


Figure 2.4: युनिफॉर्म डिस्ट्रीब्यूशनचा ग्राफ(Graph of Uniform Distribution)

आपण दोन उदाहरणांनी सहज समजू शकतो

- निःपक्षपाती डाय रोल करणे (डिस्क्रीट युनिफॉर्म डिस्ट्रीब्यूशन)
सहा बाजू असलेला निःपक्षपाती डाय रोल करताना, प्रत्येक फेस (1,2,3,4,5,6) वरच्या दिशेने उतरण्याची प्रोबॅबिलिटी 1/6 इतकी असते. हे डिस्क्रीट युनिफॉर्म डिस्ट्रीब्यूशनचे एक उत्कृष्ट उदाहरण आहे.
- ग्राफ(आलेख): फेअर डाय रोल करण्याचे प्रोबॅबिलिटी मास फंक्शन (PMF)
या उदाहरणातील ग्राफ x-अक्षावर प्रत्येक पॉसिबल आउटकम (1 ते 6) आणि y-अक्षावर प्रोबॅबिलिटी (1/6) दर्शविल.

■ रँडम नंबर जनरेशन (कंटीन्यूअस युनिफॉर्म डिस्ट्रीब्यूशन)(Random Number Generation (Continuous Uniform Distribution))

समजा एक रँडम नंबर जनरेटर 0 आणि 1 दरम्यान एक रियल व्हॅल्यू तयार करण्यासाठी प्रोग्राम केलेला आहे, या श्रेणीतील प्रत्येक संख्या समान प्रोबॅबिलिटी असलेली आहे. हे सतत एकसमान डिस्ट्रीब्यूशनचे उदाहरण आहे.

● युनिफॉर्म डिस्ट्रीब्यूशन फॉर्मूला(Uniform Distribution Formula)

एक रँडम व्हेरिएबल **X** हा $a < X < b$ या अंतराने एकसमानपणे डिस्ट्रीब्यूट केला जातो असे म्हटले जाते. युनिफॉर्म डिस्ट्रीब्यूशन फॉर्मूला,

प्रोबॅबिलिटी डेन्सिटी फंक्शन (pdf)	$f(x) = \frac{1}{(b-a)}, a \leq x \leq b$
मीन (μ)	$\int_a^b x \cdot f(x) dx = \frac{1}{b-a} \left(\frac{x^2}{2} \right)_a^b = \frac{(a+b)}{2}$
व्हेरियन्स (σ^2)	$\begin{aligned} \int_a^b x^2 \cdot f(x) dx &= Var(X) = E[X^2] - [E(X)]^2 \\ &= \int_a^b x^2 \cdot \frac{1}{b-a} dx - \left(\frac{a+b}{2} \right)^2 \\ &= \frac{(b-a)^2}{12} \end{aligned}$

स्टँडर्ड डेव्हिएशन (σ)	$\sqrt{\frac{(b-a)^2}{12}}$
क्युमुलेटीव डिस्ट्रिब्यूशन फंक्शन (cdf)	$= \frac{(x-a)}{(b-a)}$ for $x \in [a, b]$
मेडियन	$= \frac{(a+b)}{2}$
कंडीशनल प्रोबॅबिलिटी साठी = $P(c < x < d)$	$= (d-c) \times f(x)$ $= \frac{(d-c)}{(b-a)}$

युनिफॉर्म डिस्ट्रिब्यूशनचे टाईप्स (Types of Uniform Distribution):

1. कंटिन्यूअस युनिफॉर्म डिस्ट्रिब्यूशन (Continuous Uniform Distribution): कंटिन्यूअस युनिफॉर्म डिस्ट्रिब्यूशन म्हणजे असे डिस्ट्रिब्यूशन ज्यामध्ये विशिष्ट श्रेणीमध्ये परिभाषित केलेल्या अनंत संख्येच्या मूल्यांची संख्या असते. त्यात रेक्टांग्युलर आकाराचा आलेख असतो ज्याला रेक्टांग्युलर डिस्ट्रिब्यूशन म्हणतात. ते कंटिन्यूअस स्वरूपाच्या मूल्यांवर कार्य करते. उदाहरण: रँडम व्हॅल्यू जनरेटर
2. डिस्ट्रिक्ट युनिफॉर्म डिस्ट्रिब्यूशन (Discrete Uniform Distribution): डिस्ट्रिक्ट युनिफॉर्म प्रोबॅबिलिटी डिस्ट्रिब्यूशन म्हणजे असे डिस्ट्रिब्यूशन ज्यामध्ये एका विशिष्ट श्रेणीमध्ये मर्यादित संख्येने मूल्ये परिभाषित केली जातात. त्याच्या आलेखामध्ये प्रत्येक मर्यादित मूल्यासाठी विविध उभ्या रेषा असतात. ते अशा मूल्यांवर कार्य करते जे वेगळे स्वरूपाचे असतात. उदाहरण: एक फासा गुंडाळलेला असतो.

या प्रकारांची सविस्तर चर्चा खालीलप्रमाणे करूया.

• कंटिन्यूअस युनिफॉर्म डिस्ट्रिब्यूशन किंवा रेक्टांग्युलर डिस्ट्रिब्यूशन (Continuous Uniform Distributions or Rectangular Distributions)

कंटिन्यूअस युनिफॉर्म डिस्ट्रिब्यूशन, ज्याला रेक्टांग्युलर डिस्ट्रिब्यूशन असेही म्हणतात, हे प्रोबॅबिलिटी डिस्ट्रिब्यूशन आहे जिथे प्रोबॅबिलिटी मास फंक्शन (पीडीएफ) एका विशिष्ट अंतरामध्ये स्थिर असते आणि इतरत्र शून्य असते. याचा अर्थ असा की मध्यांतरातील सर्व परिणाम समान संभाव्य असतात. कंटिन्यूअस युनिफॉर्म डिस्ट्रिब्यूशन परिभाषित अंतरांमध्ये रँडमनेस समजून घेण्यासाठी आणि मॉडेलिंग करण्यासाठी एक साधी पण शक्तिशाली चौकट प्रदान करते, ज्यामुळे ते प्रोबॅबिलिटी सिद्धांत आणि लागू सांख्यिकीमध्ये आवश्यक साधने बनतात.

• प्रोबॅबिलिटी डेंसिटी फंक्शन (Probability Density Function) (PDF)

कंटिन्यूअस युनिफॉर्म डिस्ट्रिब्यूशनचे प्रोबॅबिलिटी डेंसिटी फंक्शन (पीडीएफ) एका विशिष्ट अंतरालमध्ये येणाऱ्या रँडम व्हॅल्यूची प्रोबॅबिलिटी परिभाषित करते. मध्यांतर $[a, b]$ वर कंटिन्यूअस युनिफॉर्म डिस्ट्रिब्यूशनसाठी, पीडीएफ खालीलप्रमाणे दिले जाते:

$$f(x) = \frac{1}{(b-a)} \quad a \leq x \leq b \text{ साठी}$$

आणि अन्यथा $f(x) = 0$.

• क्युमुलेटीव डिस्ट्रिब्यूशन फंक्शन (Cumulative Distribution Function) (CDF)

कंटिन्यूअस युनिफॉर्म डिस्ट्रिब्यूशनचे क्युमुलेटीव डिस्ट्रिब्यूशन फंक्शन (CDF) प्रोबॅबिलिटी देते की रँडम व्हॅल्यू एका विशिष्ट मूल्यापेक्षा कमी किंवा समान आहे. $[a, b]$ वर सतत एकसमान डिस्ट्रिब्यूशनसाठी, CDF अशी परिभाषित केली आहे:

$$F(x) = \frac{(x-a)}{(b-a)} \quad a \leq x \leq b \text{ साठी}$$

आणि $F(x) = 0$, $x < a$ साठी, $F(x) = 1$, $x > b$ साठी.

डिस्क्रीट आणि कंटिन्यूअस युनिफॉर्म डिस्ट्रिब्युशन यातील फरक:

खालील तक्ता कंटिन्यूअस आणि डिस्क्रीट एकसमान वितरणातील फरक दर्शवितो:

Aspect	डिस्क्रीट युनिफॉर्म डिस्ट्रिब्युशन	कंटिन्यूअस युनिफॉर्म डिस्ट्रिब्युशन
आऊटकम	पॉसिबल आऊटकमची मर्यादित किंवा मोजण्यायोग्य संख्या	सततच्या अंतराने असंख्य पॉसिबल आऊटकम
प्रोबॅबिलिटी	प्रत्येक परिणामाची एक विशिष्ट, समान प्रोबॅबिलिटी असते	संपूर्ण मध्यांतरात प्रोबॅबिलिटी डेंसिटी स्थिर असते, परंतु कोणत्याही विशिष्ट एकाच निकालाची प्रोबॅबिलिटी 0 असते.
उदाहरण	निःपक्षपाती डाय रोल करणे (6 पॉसिबल आऊटकम) - डेकवरील कार्ड काढणे (52 पॉसिबल आऊटकम)	0 ते 1 दरम्यान रँडम व्हॅल्यू निर्माण करणे - बस स्टॉपवर पोहोचण्याची वेळ (0 ते 10 मिनिटांदरम्यान)
ग्राफ	प्रत्येक निकालाचे समान उंचीसह प्रतिनिधित्व करणारे वैयक्तिक बार असलेले प्रोबॅबिलिटी मास फंक्शन (PMF)	प्रोबॅबिलिटी डेंसिटी फंक्शन (PDF) ज्यामध्ये मध्यांतरावर एक सपाट, सतत रेषा असते जी सर्व बिंदूवर समान शक्यता दर्शवते.

उदाहरण1) एका रँडम व्हेरिएबल X चे युनिफॉर्म डिस्ट्रिब्युशन $(-2, 2)$ वर असते,(i) ज्यासाठी $P(X > k) = \frac{1}{2}$ आहे असा k शोधा (ii) $P(X > k) = \frac{1}{2}$ चे मूल्यांकन करा (iii) $P[|X - 1| < 1]$ **उत्तर:**

$$(i) X = f(x) = \frac{1}{(b-a)} = \frac{1}{(2-(-2))} = \frac{1}{4}$$

$$P(X > k) = 1 - P(X \leq k) = 1 - \int_{-2}^k f(x) dx = 1 - \frac{1}{4} \int_{-2}^k dx = 1 - \frac{(k+2)}{4} = \frac{1}{2}$$

सोडवून आपल्याला मिळते $k = 0$

$$(ii) P(X < 1) = \int_{-2}^1 f(x) dx = \frac{1}{4} \int_{-2}^1 dx = \frac{3}{4}$$

$$(iii) P[|X - 1| < 1] = P[1 - 1 < X < 1 + 1] =$$

$$P[0 < X < 2] = \int_0^2 f(x) dx = \frac{1}{4} \int_0^2 dx = \frac{1}{4}$$

2) जर X हे $(-1, 4)$ मध्ये एकसारखे वितरित केले असेल तर

(i) सरासरी (ii) फरक (iii) प्रमाण विचलन (iv) मध्यक

उत्तर: येथे, $a = -1$ आणि $b = 4$

$$i) \text{ मीन } (\mu) = \frac{(4-1)}{2} = 1.5$$

$$(ii) \text{ व्हेरियन्स } (\sigma^2) = \frac{(4+1)^2}{12} = 2.08$$

$$(iii) \text{ स्टँडर्ड डिव्हिएशन } (\sigma) = \sqrt{2.08} = 1.443$$

$$(iv) \text{ मेडियन } = \frac{(4-1)}{2} = 1.5$$

3) जर पारंपारिक पत्त्यांच्या डेकमध्ये चार सूट असलेल्या पत्त्यांच्या डेकमध्ये 52 पत्ते असतील: हार्ट्स, स्पेड, क्लब आणि डायमंड्स. प्रत्येक सूटमध्ये 13 पत्ते असतात ज्यापैकी 3 पत्ते फेस कार्ड असतात. नवीन डेक पत्त्यांची संख्या वगळून तयार केला जातो. तर सुधारित डेकमधून हार्ट कार्ड मिळण्याची शक्यता किती आहे?

उत्तर: प्रश्नात, दिलेली पत्त्यांची संख्या मर्यादित आहे म्हणून ती एक डिस्क्रीट युनिफॉर्म डिस्ट्रिब्युशन आहे.

डिस्क्रीट युनिफॉर्म डिस्ट्रिब्युशनतील प्रोबॅबिलिटीचे सूत्र आहे $P(X) = \frac{1}{n}$

सुधारित डेकमध्ये हार्ट मिळण्याची शक्यता $= \frac{1}{4} = 0.25$

4) रँडम व्हेरिएबल X साठी युनिफॉर्म डिस्ट्रीब्यूशन प्रोबॅबिलिटी डेंसिटी फंक्शन वापरून, $(0, 20)$, मध्ये $P(3 < X < 16)$ शोधा.

उत्तर: येथे, $a = 0, b = 20$

$$f(x) = \frac{1}{(20 - 0)} = \frac{1}{20}$$

$$P(3 < X < 16) = (16 - 3) \times \left(\frac{1}{20}\right) = \frac{13}{20}$$

5) रँडम व्हेरिएबल X चे $(-5, 6)$ वर युनिफॉर्म डिस्ट्रीब्यूशन असते, $x = 3$ साठी क्युमुलेटीव डिस्ट्रीब्यूशन फंक्शन शोधा.

उत्तर: येथे, $a = -5, b = 6, x = 3$

$$CDF = \frac{(3 - (-5))}{(6 - (-5))} = \frac{8}{11}$$

सरावासाठी उदाहरणे

1. एक रँडम व्हेरिएबल X हा मध्यांतर $[2, 10]$ वर युनिफॉर्म डिस्ट्रीब्यूशन नंतर येतो. X , 4 आणि 8 च्या दरम्यान असण्याची प्रोबॅबिलिटी शोधा.
2. समजा एक कंटिन्यूअस रँडम व्हेरिएबल Y मध्यांतर $[0, 5]$ वर युनिफॉर्मली डिस्ट्रीब्यूट केला गेला आहे.
 - a. एक्स्पेक्टेड व्हॅल्यू $E(Y)$ काढा.
 - b. व्हेरियंस $Var(Y)$ काढा.
3. एक रँडम व्हेरिएबल Z हा $[3, 15]$ वर युनिफॉर्मली डिस्ट्रीब्यूट केला गेला आहे. Z चा CDF काढा आणि Z , 9 पेक्षा कमी किंवा समान असल्याची प्रोबॅबिलिटी शोधण्यासाठी त्याचा वापर करा.

2.4.10 एक्सपोनेन्शियल डिस्ट्रीब्यूशन (Exponential distribution)

एखाद्या फंक्शनचे एक्सपोनेन्शियल डिस्ट्रीब्यूशन शोधण्यासाठी एक्सपोनेन्शियल डिस्ट्रीब्यूशन सूत्र वापरले जाते. एक्सपोनेन्शियल डिस्ट्रीब्यूशन म्हणजे अशी प्रक्रिया ज्यामध्ये घटना स्थिर सरासरी दराने स्वतंत्रपणे आणि सतत घडते. एक्सपोनेन्शियल डिस्ट्रीब्यूशन बहुतेकदा मेमरी लेस डिस्ट्रीब्यूशन म्हणून ओळखले जाते कारण याचा अर्थ असा होतो की भूतकाळातील माहितीचा भविष्यातील प्रोबॅबिलिटीवर कोणताही परिणाम होत नाही.

व्याख्या (Definition)(डेफिनेशन):

एक्सपोनेन्शियल डिस्ट्रीब्यूशनसाठी, व्हेरिएबल कंटिन्यूअस आणि इंडिपेंडेंट असणे आवश्यक आहे. एक्सपोनेन्शियल डिस्ट्रीब्यूशन सूत्र खालीलप्रमाणे दिले आहे:

$$f(x) = me^{-mx}$$

or

$$f(x) = \frac{1}{\mu} e^{\left(-\frac{1}{\mu}\right)x}$$

जेथे, $m = f(x) = \frac{1}{\mu} =$ दर(रेट) पॅरामीटर किंवा क्षय(डिके) पॅरामीटर.

$\mu =$ घटनांमधील सरासरी वेळ

उदाहरण

1) एक पोस्टल क्लर्क त्यांच्या ग्राहकांसोबत सरासरी 4 मिनिटे घालवतो. वेळेचे एक्सपोनेन्शियल डिस्ट्रीब्यूशन असते. एक्सपोनेन्शियल फंक्शन सूत्र वापरून $x = 5$ वर फंक्शनचे मूल्य शोधा.

उत्तर: दिले $\mu = 4, x = 5$ म्हणून $m = \frac{1}{\mu} = \frac{1}{4} = 0.25$

$$f(x) = me^{-mx}$$

$$f(x) = 0.25 e^{(-0.25)5}$$

$$f(x) = 0.072$$

तर, फंक्शनचे मूल्य येथे $x = 5$ is 0.072.

2) एका व्यक्तीला एका काउंटरवर सरासरी 10 मिनिटे लागतात. वेळेचे एक्सपोनेन्शियल डिस्ट्रीब्यूशन असते. एक्सपोनेन्शियल फंक्शन सूत्र वापरून $x = 7$ वर फंक्शनचे मूल्य शोधा.

उत्तर:

दिले $\mu = 10, x = 7$ म्हणून $m = \frac{1}{\mu} = \frac{1}{10} = 0.1$

$$f(x) = me^{-mx}$$

$$f(x) = 0.1 e^{(-0.1)7}$$

$$f(x) = 0.04966$$

तर, फंक्शनचे मूल्य येथे $x = 7$ is 0.04966.

3) विद्यार्थ्याला कोणतीही समस्या सोडवण्यासाठी लागणारा वेळ हा एक्सपोनेन्शियल डिस्ट्रीब्यूशन नुसार येतो ज्याचा सरासरी वेळ 8 मिनिटे असतो. तो समस्या सोडवण्यासाठी 5 मिनिटे लागण्याची प्रोबॅबिलिटी किती असेल?

उत्तर: दिले $\mu = 8, x = 5$ म्हणून $m = \frac{1}{\mu} = \frac{1}{8} = 0.125$

$$f(x) = me^{-mx}$$

$$f(x) = 0.125 e^{(-0.125)5}$$

$$f(x) = 0.04966$$

तर, फंक्शनचे मूल्य येथे $x = 5$ is 0.04966.

4) एका बँकेत दर तासाला सरासरी 6 ग्राहक येतात. बँकेत सलग दोन ग्राहक येण्यामधील वेळ 10 मिनिटे असण्याची प्रोबॅबिलिटी किती आहे?

उत्तर: दिले $\mu = 6, x = 10$ म्हणून $m = \frac{1}{\mu} = \frac{1}{6} = 0.166$

$$f(x) = me^{-mx}$$

$$f(x) = 0.166 e^{(-0.166)10}$$

$$f(x) = 0.035$$

तर, फंक्शनचे मूल्य येथे $x = 10$ is 0.035.

5) प्लंबरला समस्या सोडवण्यासाठी लागणारा वेळ हा एक्सपोनेन्शियल डिस्ट्रीब्यूशन नुसार येतो ज्याचा सरासरी वेळ 12 मिनिटे असतो. समस्या सोडवण्यासाठी त्याला 5 मिनिटे लागण्याची प्रोबॅबिलिटी किती असेल?

उत्तर: दिले $\mu = 12, x = 5$ म्हणून $m = \frac{1}{\mu} = \frac{1}{12} = 0.083$

$$f(x) = me^{-mx}$$

$$f(x) = 0.083 e^{(-0.083)12}$$

$$f(x) = 0.226$$

तर, फंक्शनचे मूल्य येथे $x = 5$ is 0.226

सरावासाठी उदाहरणे

1) एका व्यक्तीला साधारणपणे दर तासाला 2 फोन येतात. वेळेचे एक्सपोनेन्शियल डिस्ट्रीब्यूशन असते. पुढच्या तासात फोन कॉल येणाऱ्या फंक्शनचे मूल्य शोधा.

- 2) एक दुकानदार त्यांच्या ग्राहकांसोबत सरासरी 8 मिनिटे घालवतो. वेळेचे एक्सपोनेन्शियल डिस्ट्रीब्यूशन असते. एक्सपोनेन्शियल फंक्शन सूत्र वापरून $x = 7$ वर फंक्शनचे मूल्य शोधा.
- 3) एक ग्राहक काउंटरवर सरासरी 20 मिनिटे घालवतो. वेळेचे एक्सपोनेन्शियल डिस्ट्रीब्यूशन असते. एक्सपोनेन्शियल फंक्शन सूत्र वापरून $x = 9$ वर फंक्शनचे मूल्य शोधा
- 4) वर्धापनदिन कार्ड खरेदी करण्यासाठी जोडीदारांना आठ मिनिटांच्या सरासरी वेळेच्या एक्सपोनेन्शियल डिस्ट्रीब्यूशन द्वारे मॉडेल करता येते. वितरण लिहा, प्रोबॅबिलिटी डेंसिटी फंक्शन सांगा आणि डिस्ट्रीब्यूशनचा आलेख तयार करा.
- 5) प्रवाशांनी त्यांचे विमान तिकीट खरेदी करण्याचे दिवस पुढे किती आहेत हे एक्सपोनेन्शियल डिस्ट्रीब्यूशन द्वारे मॉडेल करता येते ज्याचा सरासरी वेळ 15 दिवसांचा असतो. प्रवासी दहा दिवसांपेक्षा कमी वेळेपूर्वी तिकीट खरेदी करेल याची प्रोबॅबिलिटी शोधा. सर्व प्रवाशांपैकी निम्मे प्रवासी किती दिवस वाट पाहतात?

युनिट 3

नेटवर्क एनालिसिस

(Network analysis)

विषय निष्पत्ती (Course Outcome):

CO3: नातेसंबंध आणि नेटवर्क संरचनेच्या जटिल नमुनाचा अंदाज घेण्यासाठी नेटवर्क विश्लेषण लागू करा.

घटक निष्पत्ती (Theory Learning Outcome):

1. दिलेल्या समस्येचे नेटवर्क डायग्राम तयार करा.
2. दिलेल्या समस्येवर आधारित क्रिटिकल पाथची गणना करा.
3. पीईआरटी अल्गोरिदम वापरून समस्या सोडवा.
4. पीईआरटी वापरून अपेक्षित प्रोजेक्ट पूर्ण होण्याच्या तारखेची गणना करा.

परिचय(Introduction)

नेटवर्क हे बांधकाम, देखभाल, फॅब्रिकेशन, खरेदी, संगणक प्रणाली इन्स्टंटेशन, संशोधन आणि विकास नियोजन इत्यादी क्षेत्रातील मोठ्या प्रकल्पांचे नियोजन आणि वेळापत्रक तयार करण्यासाठी वापरले जाणारे तंत्र आहे. अशा अनेक ऑपरेशन्स रिसर्च परिस्थिती आहेत ज्यांचे मॉडेलिंग आणि नेटवर्क म्हणून निराकरण केले जाऊ शकते.

नेटवर्क विश्लेषण अनेक नावांनी ओळखले जाते PERT (प्रोग्राम इव्हॅल्युएशन अँड रिव्ह्यू टेक्निक), CPM (क्रिटिकल पाथ मेथड), PEP (प्रोग्राम इव्हॅल्युएशन प्रोसिजर), हा धडा दोन अल्गोरिदम सादर करेल.

1. पीईआरटी (PERT)
2. सीपीएम (CPM)

3.1 नेटवर्क आकृतीचे बांधकाम (Network diagram construction) (बाण आकृती)

नेटवर्क डायग्राम हे तार्किक आणि अनुक्रमिकपणे जोडलेल्या ऍक्टिव्हिटी आणि प्रोजेक्टच्या इव्हेंटचे ग्राफिकल प्रतिनिधित्व आहे. नेटवर्कला एरो डायग्राम देखील म्हणतात. पीईआरटी (प्रोग्राम इव्हॅल्युएशन रिव्ह्यू टेक्निक) आणि (क्रिटिकल पाथ मेथड) ही दोन सर्वात मोठ्या प्रमाणावर लागू केलेली तंत्रे आहेत.

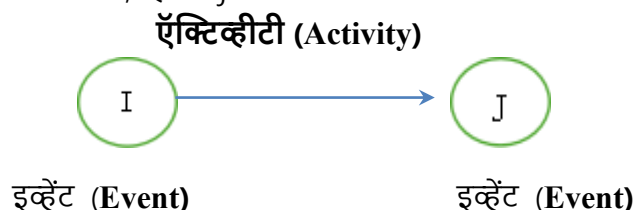
3.1.1 नेटवर्कमध्ये सामान्यतः वापरल्या जाणाऱ्या काही संज्ञा खालीलप्रमाणे आहेत

1. क्रियाकलाप(ऍक्टिव्हिटी) (Activity)

हा एखाद्या प्रोजेक्टचा भौतिकदृष्ट्या ओळखण्यायोग्य भाग आहे जो त्याच्या अंमलबजावणीसाठी वेळ आणि संसाधने वापरतो. नेटवर्कमध्ये, ऍक्टिव्हिटी बाणाद्वारे दर्शविला जातो, ज्याची टेल प्रारंभ दर्शवते आणि हेड ऍक्टिव्हिटी समाप्तीचे प्रतिनिधित्व करते. बाणाची लांबी, आकार आणि दिशा यांचा ऍक्टिव्हिटीच्या आकाराशी कोणताही संबंध नाही.

2. इव्हेंट (Event)

हे एखाद्या ऍक्टिव्हिटीच्या सुरुवातीचे आणि अंतिम बिंदूचे प्रतिनिधित्व करते. कार्यक्रमाला वेळ लागत नाही. हे सामान्यतः क्रमांकित वर्तुळाद्वारे दर्शविले जाते. हेड इव्हेंटला j^{th} इव्हेंट म्हणतात ज्याला i^{th} इव्हेंट म्हटल्या जाणाऱ्या टेल इव्हेंटपेक्षा नेहमीच जास्त संख्या असते, म्हणजे $j > i$

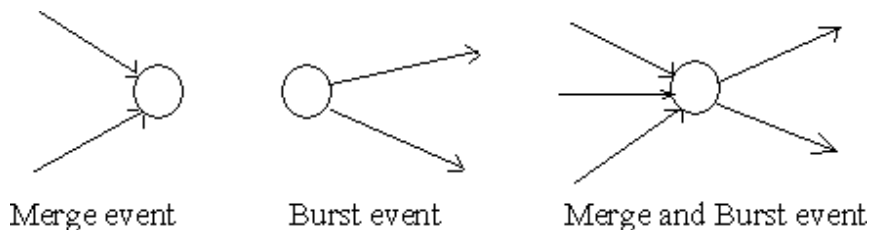


घटनांचे तीन प्रकार आहेत

i.मर्ज इव्हेंट (इव्हेंट)(Merge event): जेव्हा एकापेक्षा जास्त ऍक्टिव्हिटी येतात आणि इव्हेंट(घटनां)मध्ये सामील होतात. अशा इव्हेंटला मर्ज इव्हेंट (घटनां) म्हणून ओळखले जाते.

ii. **बर्स्ट इव्हेंट (Burst Event):** जेव्हा एकापेक्षा जास्त ऍक्टिव्हिटी इव्हेंट सोडतात तेव्हा अशा घटनेला बर्स्ट इव्हेंट म्हणून ओळखले जाते

iii. **मर्ज आणि बर्स्ट इव्हेंट (Merge and Burst event):** मर्ज आणि बर्स्ट इव्हेंट: एखादी गतिविधी मर्ज इव्हेंट आणि बर्स्ट इव्हेंट असू शकते कारण काही ऍक्टिव्हिटीच्या संदर्भात ती एक मर्ज इव्हेंट असू शकते आणि काही इतर ऍक्टिव्हिटीच्या संदर्भात ती बर्स्ट इव्हेंट असू शकते (खालील चित्रात दाखवल्याप्रमाणे)



3. मार्ग (पाथ) (Path)

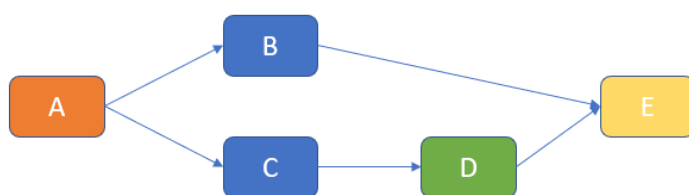
ऍक्टिव्हिटीची एक अखंड साखळी, सुरुवातीच्या इव्हेंटला इतर इव्हेंटशी (घटनेशी) जोडणारे बाण याला पाथ म्हणतात.

4. नेटवर्क आकृती (Network Diagram)

हे तार्किक आणि अनुक्रमिकपणे जोडलेले बाण (ऍक्टिव्हिटी) आणि नोड्स (इव्हेंट) चे ग्राफिकल प्रतिनिधित्व आहे जे एखाद्या प्रोजेक्टच्या ऍक्टिव्हिटी आणि इव्हेंटचे प्रतिनिधित्व करते.

5. **समवर्ती ऍक्टिव्हिटी (काँकुरेंट ऍक्टिव्हिटी) (Concurrent Activity):** समवर्ती ऍक्टिव्हिटी म्हणून ओळखले जातात जे एकाच वेळी पूर्ण केले जाऊ शकतात. हे लक्षात घेतले जाऊ शकते की एखादी ऍक्टिव्हिटी एखाद्या इव्हेंटची प्रिडीसेसर किंवा सक्सेसर असू शकते किंवा ती एक किंवा अधिक इतर ऍक्टिव्हिटीसोबत असू शकते.

6. **पूर्ववर्ती क्रियाकलाप (प्रिडीसेसर ऍक्टिव्हिटी) (Predecessor Activity):** तयार केलेल्या नेटवर्क आकृती मध्ये, विशिष्ट ऍक्टिव्हिटी सुरू करण्यापूर्वी पूर्ण करणे आवश्यक असलेल्या ऍक्टिव्हिटींना पूर्ववर्ती ऍक्टिव्हिटी (प्रिडीसेसर ऍक्टिव्हिटी) म्हणतात. आकृती 2.1 मध्ये इव्हेंट A इव्हेंट B चा पूर्ववर्ती आहे, इव्हेंट B इव्हेंट E चा पूर्ववर्ती आहे.



Activity List	Predecessor
A	Nil
B	A
C	A
D	C
E	B, D

7. उत्तराधिकारी क्रियाकलाप (सक्सेसर ऍक्टिव्हिटी) (Successor Activity)

तयार केलेल्या आकृतीमध्ये कोणत्याही विशिष्ट ऍक्टिव्हिटीचे अनुसरण करणे आवश्यक असलेल्या ऍक्टिव्हिटींना उत्तराधिकारी ऍक्टिव्हिटी (सक्सेसर ऍक्टिव्हिटी) म्हणतात

8. डमी ऍक्टिव्हिटी (ऍक्टिव्हिटी)

एक ऍक्टिव्हिटी जी एका ऍक्टिव्हिटी ची दुसऱ्यावर अवलंबित्व दर्शवते, परंतु कोणत्याही संसाधनाचा वापर करत नाही त्याला डमी ऍक्टिव्हिटी म्हणतात. डमी सामान्यतः ठिपके असलेल्या रेषा बाणांनी दर्शविल्या जातात जे खाली आकृती 3.1 मध्ये स्पष्ट केले आहे.

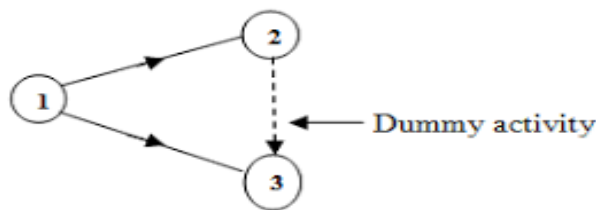


Figure 3.1: डमी ऍक्टिव्हिटी (Dummy Activity)

9. परस्परावलंबन (इंटरडिपेंडेंसी) (Interdependency)

प्रोजेक्ट मध्ये अनेक ऍक्टिव्हिटी असतात या सर्व ऍक्टिव्हिटी एकमेकांशी संबंधित असतात. संपूर्ण प्रोजेक्ट ची अनेक लहान-लहान (ऍक्टिव्हिटी)क्रियाकलापांमध्ये विभागणी केली पाहिजे आणि त्यांचे अवलंबित्व शोधले पाहिजे

10. नेटवर्क डायग्राममधील त्रुटी

a) लूपिंग (Looping): काहीवेळा सदोष नेटवर्क क्रमामुळे आकृती 3.2 मध्ये वर्णन केलेली स्थिती उद्भवते. येथे D, E, F ऍक्टिव्हिटी एक लूप तयार करतात F पूर्ण होईपर्यंत ऍक्टिव्हिटी D सुरू होऊ शकत नाही, जे E च्या पूर्णतेवर अवलंबून असते परंतु E D च्या पूर्णतेवर अवलंबून असते. अशा प्रकारे नेटवर्क पुढे जाऊ शकत नाही.

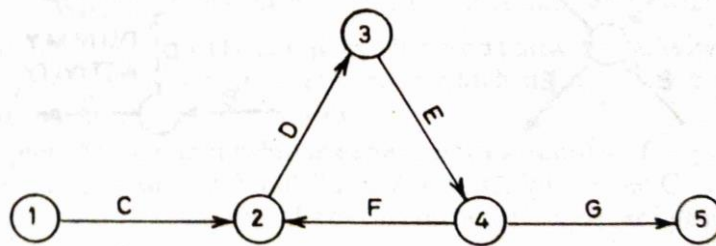


Figure 3.2: लूपिंग (looping)

b) ड्यंगलिंग (Dangling): ही अशी परिस्थिती आहे जिथे अंतिम (ऍक्टिव्हिटी) क्रियाकलापांव्यतिरिक्त इतर (ऍक्टिव्हिटी)क्रियाकलापांमध्ये कोणतेही सक्सेसर (Successor) ऍक्टिव्हिटी नसतात. हे टाळण्यासाठी संपूर्ण प्रोजेक्ट ची सुरुवात एकच आणि एकच शेवटची क्रिया असावी.

11.नोड रिप्रेसेंटेशन (Node representation): प्रोजेक्ट मॅनेजमेंट नेटवर्क विश्लेषणामध्ये, "नोड" एखाद्या प्रोजेक्टमधील विशिष्ट ऍक्टिव्हिटी किंवा इव्हेंटचे प्रतिनिधित्व करतो आणि नोड्समधील संबंध त्या (ऍक्टिव्हिटी)क्रियाकलापांमधील अवलंबित्व किंवा अनुक्रम दर्शवितो, त्यांना नेटवर्क आकृतीवर जोडणाऱ्या बाणांनी चित्रित केले आहे; मूलतः एक नोड एखाद्या क्रियाकलापाचा प्रारंभ किंवा समाप्ती बिंदू दर्शवितो आणि नोड्सला जोडणाऱ्या रेषा ऍक्टिव्हिटी कोणत्या क्रमाने पूर्ण करणे आवश्यक आहे हे स्पष्ट करतात.

महत्त्वाचे मुद्दे

- 1. व्हिज्युअल प्रतिनिधित्व (रिप्रेसेंटेशन):** नेटवर्क डायग्रामवर नोड्स सामान्यतः सर्कल किंवा बॉक्स म्हणून चित्रित केले जातात, ज्यामुळे (ऍक्टिव्हिटी)क्रियाकलापांचा प्रवाह आणि त्यांचे अवलंबित्व पाहणे सोपे होते.
- 2. ऍक्टिव्हिटी प्रतिनिधित्व (रिप्रेसेंटेशन):** नोड्स सहसा नेटवर्क आकृतीमध्ये वर्तुळे किंवा चौकोन म्हणून दर्शवले जातात, ज्यामुळे ऍक्टिव्हिटीचा प्रवाह आणि त्यांचे परस्पर अवलंबित्व सहजपणे पाहता येते.
- 3. डिपेंडेंसी एरॉस (अवलंबित्व बाण) (Dependency arrows):** नोड्सला जोडणारे बाण क्रियाकलापांमधील तार्किक क्रम दर्शवतात, कोणती ऍक्टिव्हिटी दुसरी सुरू होण्यापूर्वी पूर्ण करणे आवश्यक आहे हे दर्शविते.
- 4. संबंधांचे प्रकार (टाईप्स ऑफ रिलेशनशिप्स):** नोड्समध्ये वेगवेगळ्या प्रकारचे संबंध दर्शवले जाऊ शकतात, ज्यामध्ये "फिनिश-टू-स्टार्ट" (सर्वात सामान्य), "स्टार्ट-टू-स्टार्ट", "फिनिश-टू-फिनिश" आणि "स्टार्ट-टू-फिनिश" यांचा समावेश होतो, जे प्रकल्पाच्या विशिष्ट गरजांवर अवलंबून असते.

3.1.2 नेटवर्क डायग्रामचे बांधकाम (कन्स्ट्रक्शन ऑफ नेटवर्क डायग्रॅम) (ए-ओ-ए नेटवर्क)

नेटवर्क नियम

1. प्रारंभिक नोडला फक्त बाहेर जाणारे बाण असतात. नेटवर्कमध्ये फक्त एकच प्रारंभिक नोड असावा.
2. कोणतीही इव्हेंट तेव्हाच घडू शकते जेव्हा तिच्यापर्यंत पोहोचवणाऱ्या सर्व क्रिया पूर्ण झालेल्या असतात.
3. कोणतीही इव्हेंट दोनदा घडू शकत नाही, म्हणजेच यापूर्वी घडलेल्या घटनेकडे परत जाणारा कोणताही नेटवर्क मार्ग असू शकत नाही.
4. अंतिम नोड वगळता कोणताही मार्ग बंद होऊ नये. नेटवर्कमध्ये फक्त एकच अंतिम नोड असावा.
5. कोणतीही कृती तेव्हाच सुरू होऊ शकते जेव्हा तिच्या मागील इव्हेंट (पूर्व इव्हेंट) घडलेली असेल.
6. प्रत्येक बाण एकाच परिस्थितीचे प्रतिनिधित्व करावा. प्रकल्पातील क्रियांच्या संख्येशी बाणांच्या संख्येची समानता असावी.
7. नेटवर्कचे प्रतिनिधित्व असे असावे की प्रत्येक ऍक्टिव्हिटी अंतिम उद्दिष्टापर्यंत पोहोचण्यासाठी पूर्ण होईल.
8. सर्व मर्यादा आणि परस्परावलंबन योग्य डमी वापरून नेटवर्कवर योग्यरित्या दर्शविले जावे.
9. टवर्कचे लॉजिक नेहमी राखले गेले पाहिजे, म्हणजे अॅरो हेड खरी नियंत्रण परिस्थिती दर्शवण्यासाठी योग्य मार्ग दाखवतात.
10. वेळेचा प्रवाह डावीकडून उजवीकडे दाखवण्याचा नेहमीचा सराव आहे.

3.1.3 नेटवर्कची ग्राफिकल मार्गदर्शक तत्त्वे (ग्राफिकल गाईडलाईन्स ऑफ नेटवर्क) (Graphical guidelines of network)

1. बाण(Arrow) वेक्टर (Vector)नाहीत. ते त्याच्या लांबीद्वारे कालावधी (किंवा वेळ) दर्शवण्यासाठी कधीही वापरले जात नाहीत.
2. जेव्हा शक्य असेल तेव्हा सर्व बाण समान आकाराचे असावेत.
3. मसुदा तयार करण्याच्या सोयीनुसार बाणांमधील कोन निवडला पाहिजे. नोड्स सोडताना किंवा जोडणाऱ्या बाणांमधील कोन शक्य तितका मोठा असावा जेणेकरून आकृतीमध्ये इतर संबंधित माहिती जोडण्यासाठी अधिक जागा उपलब्ध होईल.
4. शक्यतो सरळ बाण वापरावेत.
5. बाण एकमेकांना ओलांडू नयेत. जर ओलांडणे अपरिहार्य असेल, तर त्या परिस्थितीत बाण तुटलेला दाखवून तो दुसऱ्या बाणावरून पूल करून जावा.

3.1.4 इव्हेंट्स क्रमांकासाठी फुलकर्सन (Fulkerson) नियम

नियम 1:सुरुवातीची इव्हेंट, ज्यामध्ये सर्व बाहेर जाणारे बाण असतात आणि कोणताही येणारा बाण नसतो, तिला क्रमांक "1" दिला जातो.

नियम 2: नोड '1' मधून बाहेर जाणारे सर्व बाण काढून टाका. यामुळे किमान एक नवीन प्रारंभिक इव्हेंट तयार होईल.

नियम 3: या प्रारंभिक घटनांना अनुक्रमे 2, 3... असे क्रमांक द्या.

नियम 4:या क्रमांकित घटनांमधून बाहेर जाणारे सर्व बाण काढून टाका, ज्यामुळे अधिक प्रारंभिक इव्हेंट तयार होतील. या नवीन घटनांना पुढील अनुक्रमिक क्रमांक द्या.

नियम 5:संपूर्ण प्रक्रिया अंतिम किंवा टर्मिनल नोडपर्यंत सुरू ठेवा, ज्यामध्ये सर्व बाण आत येतात आणि कोणताही बाण बाहेर जात नाही, आणि त्याला क्रमांक दिला जातो.

सोडवलेली उदाहरणे

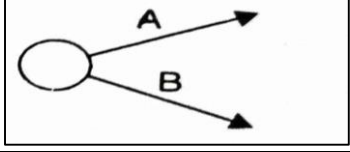
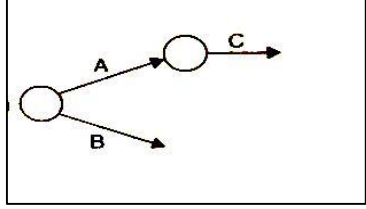
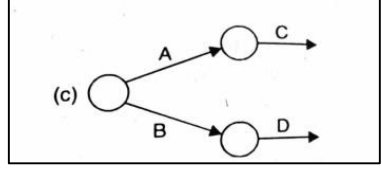
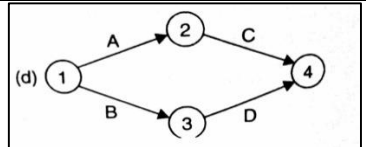
उदाहरण 1. A, B, C आणि D असे लेबल असलेल्या आणि खालीलप्रमाणे संबंधित चार ऍक्टिव्हिटी असलेल्या प्रोजेक्ट साठी नेटवर्क काढा.

i.क्रिया A आणि क्रिया B एकाच वेळी केली जाऊ शकतात.

ii.क्रिया A ही क्रिया C ची इमिडिएट प्रिडीसेसर आहे, आणि त्याचप्रमाणे B आणि D यांच्यातही समान संबंध आहे.

iii. क्रिया C आणि D पूर्ण झाल्यावर प्रोजेक्ट ची पूर्तता होते.

उत्तर:

i	क्रिया A आणि क्रिया B एकाच वेळी केली जाऊ शकतात	
ii	प्रिडीसेसर (Predecessor) आहे, आणि त्याचप्रमाणे B आणि D यांच्यातही समान संबंध आहे.	
iii	क्रिया D ही क्रिया B ची इमिडिएट सक्सेसर (Successor) आहे.	
iv	क्रिया C आणि D पूर्ण झाल्यावर प्रोजेक्ट ची पूर्तता होते	

उदाहरण 2. A, B आणि C या सात क्रिया एकाच वेळी चालणाऱ्या प्रोजेक्ट चे नेटवर्क काढा. ऍक्टिव्हिटी (Activity) पूर्ववर्ती (प्रिडिसेसर -predecessor) संबंध खालीलप्रमाणे आहेत:

ऍक्टिव्हिटी

इमिडिएट प्रिडिसेसर

D

A

E

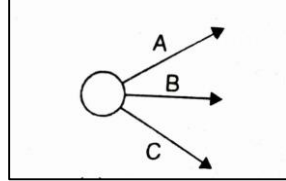
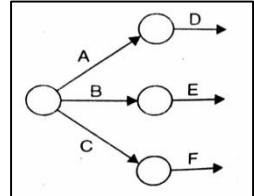
B

F

C

ऍक्टिव्हिटी G हे प्रोजेक्ट चे शेवटचे ऑपरेशन आहे आणि ते D, E आणि F चे सक्सेसर देखील आहे.

उत्तर:

i	ऍक्टिव्हिटी A, B आणि C समांतर चालतात	
ii	ऍक्टिव्हिटी D, E आणि F हे अनुक्रमे ऍक्टिव्हिटी A, B आणि C यांचे इमिडिएट सक्सेसर आहेत.	

iii	शेवटचा एक्टिव्हिटी G हा D, E आणि F यांचा इमिडिएट सक्सेसर आहे.	
-----	---	--

उदाहरण 3. प्र प्रोजेक्ट मध्ये खालील संबंधांसह, A ते F पर्यंत नियुक्त केलेल्या सहा एक्टिव्हिटी चा समावेश आहे:

- A हे केले जाणारे पहिले काम आहे.
- B आणि C एकाच वेळी केले जाऊ शकते आणि A चे अनुसरण करणे आवश्यक आहे.
- B च्या आधी D असणे आवश्यक आहे
- E हे C नंतर यायलाच हवे, पण ते B पूर्ण होईपर्यंत सुरू होऊ शकत नाही.
- शेवटची प्रक्रिया F ही दोन्ही पूर्ण होण्यावर अवलंबून आहे. नेटवर्क आकृती काढा.

उत्तर:

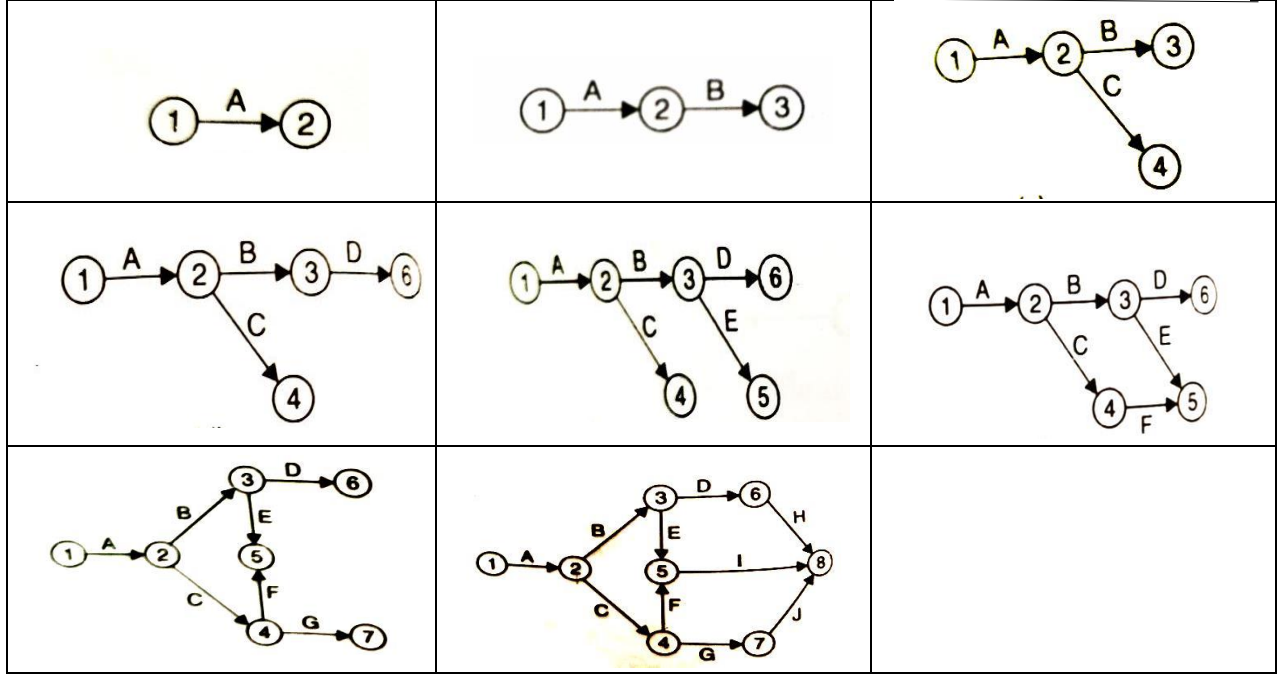
i	A हे केले जाणारे पहिले काम आहे	
ii	B आणि C एकाच वेळी केले जाऊ शकते आणि A चे अनुसरण करणे आवश्यक आहे.	
iii	B च्या आधी D असणे आवश्यक आहे	
iv	E हे C नंतर यायलाच हवे, पण ते B पूर्ण होईपर्यंत सुरू होऊ शकत नाही.	
v	शेवटची प्रक्रिया F ही दोन्ही पूर्ण होण्यावर अवलंबून आहे. नेटवर्क आकृती काढा.	

उदाहरण 4. इमारतीच्या देखभाल प्रोजेक्ट मध्ये दहा कामे समाविष्ट आहेत. प्रिडिसेसर संबंध त्याच्या नोड क्रमांकांद्वारे ओळखले जातात, जसे खाली दर्शविले आहे.

काम	क्रिया	काम	क्रिया
A	(1, 2)	F	(4, 5)
B	(2, 3)	G	(4, 7)
C	(2, 4)	H	(5, 8)
D	(3, 6)	I	(6, 8)
E	(3, 5)	J	(7, 8)

प्रोजेक्ट साठी नेटवर्क आकृती काढा.

उत्तर:



सरावासाठी उदाहरणे

1. सात ऍक्टिव्हिटी असलेल्या प्रोजेक्ट चे नेटवर्क आकृती काढा. ऍक्टिव्हिटी A, B आणि C एकाच वेळी चालतात. ऍक्टिव्हिटी चे प्रिडीसेसर संबंध खालीलप्रमाणे आहेत:

ऍक्टिव्हिटी

इमिडिएट प्रिडीसेसर

D

A

E

B

F

C

2. खालील आंतर-संबंधांसह 9 ऍक्टिव्हिटी असलेल्या प्रोजेक्ट साठी नेटवर्क आकृती काढा:

- C, D नंतर येते पण F पूर्वी येते.
- C हे B नंतर येते पण H पूर्वी येते
- G हे F नंतर येते पण I पूर्वी येते.
- E हे A नंतर येते पण I पूर्वी येते.
- D हे A नंतर येते.
- H आणि I एकाच वेळी समाप्त होतात.
- A आणि B एकाच वेळी सुरू होतात

3. दहा घटनांचा समावेश असलेल्या प्रोजेक्ट योजनेमध्ये पुढीलप्रमाणे (प्रिडीसेसर) पूर्ववर्ती संबंध आहेत.

कार्यक्रम	तत्काळ प्रिडिसेसर	कार्यक्रम	तत्काळ प्रिडिसेसर
1	-	6	3, 5
2	1	7	3, 4
3	2	8	3, 7
4	2	9	7
5	2	10	3,6,8,9

प्रोजेक्ट योजनेसाठी नेटवर्क आकृती काढा.

4. एका प्रोजेक्ट मध्ये चौदा उपक्रम A ते M या क्रि एक्टिव्हिटी मधील संबंध खाली दिले आहेत.

- A हे पहिले ऑपरेशन आहे.
 - B आणि C समांतर केले जाऊ शकतात आणि ते A चे त्वरित सक्सेसर आहेत
 - D, E आणि F हे B नंतर येतात.
 - G हे E नंतर येते
 - H हे D नंतर येते, परंतु ते E पूर्ण होईपर्यंत सुरू होऊ शकत नाही
 - I आणि J हे G नंतर येतात
 - F आणि J हे K पूर्वी येतात
 - H आणि I हे K पूर्वी येतात
 - M हे L आणि K नंतर येते.
 - शेवटची प्रक्रिया N, M आणि C नंतर येते
- नेटवर्क आकृतिबंध तयार करा

3.2 क्रिटिकल पाथ च्या निर्धार (Determination of Critical Path) (डिटर्मिनेशन ऑफ क्रिटिकल पाथ)

क्रिटिकल पाथ काय आहे?

बहुतेक इमिडिएट प्रोजेक्ट मध्ये, परस्परसंबंधित किंवा अवलंबित कार्यांचे किमान दोन किंवा अधिक अनुक्रम आहेत जे समांतरपणे कार्यान्वित करणे आवश्यक आहे. प्रोजेक्ट चा क्रिटिकल पाथ हा प्रोजेक्ट पूर्ण करण्यासाठी कार्यान्वित करणे आवश्यक असलेल्या कार्यांचा सर्वात लांब क्रम आहे.

क्रिटिकल पाथ वरील कार्यांना क्रिटिकल एक्टिव्हिटी म्हणतात कारण त्यांना उशीर झाल्यास, संपूर्ण प्रोजेक्ट पूर्ण होण्यास उशीर होईल, नॉन क्रिटिकल कार्य अनुक्रमांमधील क्रि एक्टिव्हिटी च्या विपरीत ज्याचा एकूण प्रोजेक्ट वेळेवर परिणाम न होता विलंब होऊ शकतो.

क्रिटिकल पाथ शोधण्यासाठी, प्रोजेक्ट मॅनेजर क्रिटिकल पाथ पद्धत वापरतात (सी पी एम)

3.2.1 क्रिटिकल पाथ पद्धतीची व्याख्या

क्रिटिकल पाथ मेथड ही एक तंत्र आहे जी प्रोजेक्ट मॅनेजर नवीन प्रोजेक्ट शेड्यूल तयार करण्यासाठी आणि प्रोजेक्टच्या एकूण वेळेचा अंदाज घेण्यासाठी वापरली जाते. हे क्रिटिकल पाथ विश्लेषण देखील ओळखले जाते.

सीपीएम (CPM) एक नेटवर्क आकृतिबंध वापरून सर्व कार्यांच्या अनुक्रमाचे प्रतिनिधित्व करते, जे प्रोजेक्ट पूर्ण करण्यासाठी आवश्यक असते. प्रत्येक कार्याचा मार्ग त्याच्या कालावधीने परिभाषित केला जातो आणि त्यावर गणना करून महत्वाचा मार्ग (क्रिटिकल पाथ) ओळखला जातो, जो प्रोजेक्ट च्या एकूण कालावधीचा निर्धार करतो

3.2.2 CPM चार्ट काढण्याची पद्धत

- प्रोजेक्ट पूर्ण करण्यासाठी आवश्यक असलेल्या सर्व एक्टिव्हिटी ची यादी
- एक्टिव्हिटी चा क्रम शोधा
- नेटवर्क आकृती काढा
- प्रत्येक एक्टिव्हिटी पूर्ण होण्यासाठी लागणारा वेळ शोधा.

5. एक्टिव्हिटी मधील अवलंबित्व शोधा

6. गंभीर मार्ग (क्रिटिकल पाथ) शोधा

7. प्रोजेक्ट च्या प्रगतीनुसार नेटवर्क आकृतिबंध अद्यतनित करा (अपडेट करा)

CPM नियोजित क्रियाकलापांच्या प्रकल्पाच्या समाप्तीपर्यंतच्या सर्वात लांब मार्गाची गणना करते आणि सर्वात लवकर आणि नवीनतम मार्गाची गणना करते की प्रत्येक एक्टिव्हिटी प्रोजेक्ट लांब न करता सुरू आणि पूर्ण करू शकतो.

CPM मध्ये वापरल्या जाणाऱ्या काही संज्ञा(terms)

1.अर्लीएस्ट स्टार्ट टाईम: एक्टिव्हिटी सुरू होऊ शकणाऱ्या सर्वात आधीच्या वेळेस, प्रारंभिक वेळ (अर्लीएस्ट स्टार्ट टाईम) म्हणतात.

2.अर्लीएस्ट कंप्लेशन टाइम: ही क्रियाकलापाच्या सर्वात लवकर सुरू होण्याची वेळ आणि त्या एक्टिव्हिटी ची कालावधी यांची बेरीज आहे.

3.लेटेस्ट स्टार्ट टाइम: ही एक्टिव्हिटी च्या शेवटच्या पूर्ण होण्याच्या वेळेतून त्या एक्टिव्हिटी च्या कालावधीची वजाबाकी आहे.

4. लेटेस्ट कंप्लेशन टाइम: ही त्या इव्हेंट ची शेवटची घडण्याची वेळ आहे ज्यावर एक्टिव्हिटी समाप्त होतो.

5.टोटल फ्लोट: एखाद्या क्रियाकलापासाठी अनुमत जास्तीत जास्त वेळ आणि त्याचा अंदाजित कालावधी यांच्यातील फरकाला टोटल फ्लोट म्हणतात.

6.फ्री फ्लोट: एखाद्या एक्टिव्हिटी च्या पूर्ण होण्याच्या वेळेत होऊ शकणाऱ्या विलंबाचा कालावधी, जो पुढील एक्टिव्हिटी च्या सुरुवातीवर परिणाम करत नाही, याला फ्री फ्लोट म्हणतात

7.इंडिपेंडेंट फ्लोट: क्तिव्हिटी सुरू होण्यास किती वेळ उशीर केला जाऊ शकतो, हे गृहीत धरून कोणत्याही तत्काळ पुढील एक्टिव्हिटी च्या लवकरात लवकर सुरू होण्याच्या वेळेवर परिणाम न करता, अगोदरची क्रिया त्याच्या नवीनतम समाप्तीच्या वेळेस संपली आहे.

8 वर्क-ब्रेकडाऊन स्ट्रक्चर: हे प्रोजेक्ट अंमलबजावणीसाठी आवश्यक असलेल्या सर्व क्रियाकलापांची व्याख्या करते.

9. स्लॅक इव्हेंट: नवीनतम इव्हेंट आणि सर्वात आधीच्या इव्हेंटच्या वेळेतील फरक म्हणून त्याची व्याख्या केली जाते **वापरलेली नोटेशन्स खालील प्रमाणे आहेत**

(i, j) = टेल इव्हेंट i आणि हेड इव्हेंट j सह एक्टिव्हिटी.

E_i = घटनेची सर्वात जुनी इव्हेंट i

L_j = j इव्हेंटची नवीनतम परवानगीयोग्य इव्हेंट

D_{ij} = एक्टिव्हिटी पूर्ण होण्याची अंदाजे वेळ (i, j)

$(Es)_{ij}$ = एक्टिव्हिटी ची सुरुवातीची सर्वात लवकर वेळ (i, j)

$(Ef)_{ij}$ = एक्टिव्हिटी ची सर्वात लवकर पूर्ण वेळ (i, j)

$(Ls)_{ij}$ = एक्टिव्हिटी ची नवीनतम सुरुवातीची वेळ (i, j)

$(Lf)_{ij}$ = क्रियाकलापाची नवीनतम पूर्ण वेळ (i, j)

प्रक्रिया खालीलप्रमाणे आहे

1.लवकरात लवकर वेळेचे निर्धारण(डिटर्मिनेशन ऑफ अर्लीएस्ट पाथ) (E_j)): फॉरवर्ड पास गणना

टप्पा1: गणना स्टार्ट नोडपासून सुरू होते आणि शेवटच्या नोडकडे जाते. सुलभतेसाठी, प्रारंभिक प्रोजेक्ट इव्हेंटसाठी शून्याचा सर्वात जुनी इव्हेंट वेळ गृहीत धरून फॉरवर्ड पासची गणना सुरू होते.

टप्पा2: i) एक्टिव्हिटी ची सुरुवातीची वेळ (i, j) ही टेल एंड इव्हेंटची सर्वात जुनी इव्हेंट वेळ आहे म्हणजे

$(Es)_{ij} = E_i$

ii) एक्टिव्हिटी ची सर्वात जुनी समाप्ती वेळ (i, j) ही सर्वात सुरुवातीची वेळ + एक्टिव्हिटी वेळ आहे म्हणजे $(Ef)_{ij} = (Es)_{ij} + D_{ij}$ or $(Ef)_{ij} = E_i + D_{ij}$

iii) इव्हेंट j साठी सर्वात लवकर इव्हेंट वेळ ही त्या इव्हेंटमध्ये समाप्त होणाऱ्या सर्व क्रियाकलापांच्या लवकरात लवकर समाप्त होण्याच्या वेळेची कमाल आहे म्हणजे $E_j = \max [(Ef)_{ij}]$. (i, j) च्या

सर्व तात्काळ पूर्ववर्ती साठी किंवा $E_j = \max [E_i + D_{ij}]$

2. बॅकवर्ड पास गणना (नवीनतम स्वीकार्य वेळेसाठी)(फॉर लेटेस्ट अलोकेबल टाईम)

टप्पा 1 : इव्हेंट समाप्त करण्यासाठी $E = L$ असे गृहीत धरा. लक्षात ठेवा की सर्व E ची गणना फॉरवर्ड पास गणनेद्वारे केली गेली आहे

टप्पा 2: एक्टिव्हिटी साठी नवीनतम समाप्ती वेळ (i, j) इव्हेंट j च्या नवीनतम इव्हेंट वेळेच्या समान आहे

म्हणजे $(Lf)_{ij} = L_j$

टप्पा 3: एक्टिव्हिटी ची नवीनतम सुरुवात वेळ $(i, j) = (i, j)$ ची नवीनतम पूर्णता वेळ - एक्टिव्हिटी वेळ किंवा $(Ls)_{ij} = (Lf)_{ij} - D_{ij}$ किंवा $(Ls)_{ij} = L_j - D_{ij}$

टप्पा 4: इव्हेंट ' i ' साठी नवीनतम इव्हेंट वेळ ही त्या इव्हेंटपासून उद्भवलेल्या सर्व एक्टिव्हिटी च्या नवीनतम प्रारंभ वेळेची किमान आहे म्हणजे $L_i = \min [(Ls)_{ij} \text{ for all immediate successor of } (i, j)] = \min [(Lf)_{ij} - D_{ij}] = \min [L_j - D_{ij}]$

3. फ्लोट्स आणि स्लॅक वेळा निश्चित करणे

फ्लोट्सचे तीन प्रकार आहेत

टोटल (एकूण) फ्लोट: एकूण प्रोजेक्ट कालावधी वेळेवर परिणाम न करता लवकरात लवकर अपेक्षित पूर्ण होण्याच्या वेळेच्या पलीकडे एक्टिव्हिटी पूर्ण होण्यासाठी किती वेळ विलंब होऊ शकतो.

गणितानुसार,

$$(Tf)_{ij} = (Ls)_{ij} - (Es)_{ij} \text{ किंवा } (Tf)_{ij} = (L_j - D_{ij}) - E_i$$

फ्री फ्लोट: एखाद्या एक्टिव्हिटी च्या पूर्ण होण्यास झालेला विलंब लवकरात लवकर पूर्ण होण्याच्या वेळेपेक्षा किती काळपर्यंत वाढवता येऊ शकतो, ज्यामुळे पुढील एक्टिव्हिटी च्या लवकरात लवकर सुरुवातीच्या वेळेवर कोणताही परिणाम होत नाही

गणितानुसार,

$$(Ff)_{ij} = (E_j - E_i) - D_{ij}$$

इंडिपेंडेंट(स्वतंत्र) फ्लोट: एक्टिव्हिटी सुरू होण्यास किती वेळ उशीर केला जाऊ शकतो, हे गृहीत धरून कोणत्याही तात्काळ पुढील एक्टिव्हिटी च्या लवकरात लवकर सुरू होण्याच्या वेळेवर परिणाम न करता, अगोदरची क्रिया त्याच्या नवीनतम समाप्तीच्या वेळेस संपली आहे.

गणितानुसार,

$$(If)_{ij} = (E_j - L_i) - D_{ij} \text{ नकारात्मक स्वतंत्र फ्लोट नेहमी शून्य म्हणून घेतले जाते.}$$

इव्हेंट स्लॅक: नवीनतम इव्हेंट आणि सर्वात आधीच्या इव्हेंटच्या वेळेतील फरक म्हणून त्याची व्याख्या केली जाते

$$\text{हेड इव्हेंट स्लॅक} = L_j - E_j, \text{ टेल इव्हेंट स्लॅक} = L_i - E_i$$

4. क्रिटिकल चा पाथ निर्धार:

क्रिटिकल इव्हेंट (इव्हेंट): शून्य स्लॅक वेळा असलेल्या घटनांना क्रिटिकल इव्हेंट म्हणतात. दुस-या शब्दात इव्हेंट i ला क्रिटिकल म्हटले जाते जर $E_i = L_i$

क्रिटिकल एक्टिव्हिटी: झीरो टोटल फ्लोट असलेल्या एक्टिव्हिटी ना क्रिटिकल एक्टिव्हिटी म्हणून ओळखले जातात. दुस-या शब्दात सांगायचे तर, एखाद्या प्रोजेक्ट च्या प्रारंभास उशीर झाल्यास संपूर्ण प्रकल्पाच्या पूर्ण होण्याच्या तारखेस आणखी विलंब झाल्यास ती गंभीर असल्याचे म्हटले जाते.

क्रिटिकल पाथ: नेटवर्कमधील गंभीर ऍक्टिव्हिटी च्या क्रमाला क्रिटिकल पाथ म्हणतात. क्रिटिकल पाथ हा नेटवर्कमधील सुरुवातीच्या इव्हेंटपासून शेवटच्या इव्हेंटपर्यंतचा सर्वात लांब मार्ग आहे आणि प्रोजेक्ट पूर्ण करण्यासाठी लागणारा किमान वेळ परिभाषित करतो.

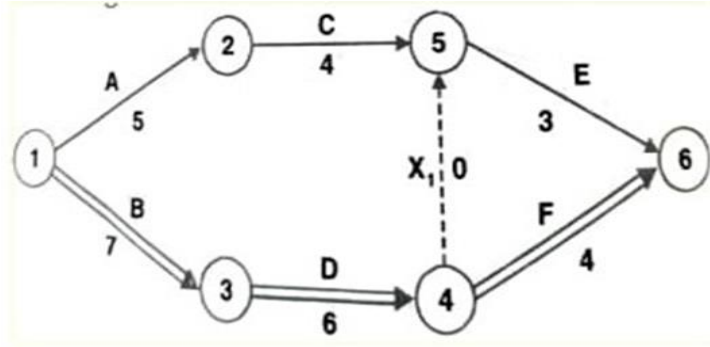
सोडवलेली उदाहरणे

उदाहरण 1: खालील उदाहरणासाठी

1. नेटवर्क आकृती काढा
2. क्रिटिकल पाथ आणि प्रोजेक्ट पूर्ण होण्याची वेळ शोधा
3. क्रिटिकल आणि नॉन - क्रिटिकल ऍक्टिव्हिटी ओळखा

ऍक्टिव्हिटी	प्रीसिडींग ऍक्टिव्हिटी	कालावधी (दिनांक)
A	-	5
B	-	7
C	A	4
D	B	6
E	C & D	3
F	D	4

उत्तर: नेटवर्क डायग्राम



नेटवर्कमधील मार्ग

नेटवर्क डायग्राममधील मार्ग म्हणजे पहिल्या इव्हेंटशी (म्हणजे प्रकल्पाची सुरुवात) शेवटच्या इव्हेंटशी (म्हणजे प्रोजेक्ट पूर्ण होणे) क्रियाकलापांचा प्रवाह.

या नेटवर्क डायग्राममध्ये तीन मार्ग आहेत:

- 1) A-C-E कालावधी=5+4+3=12 दिवस
- 2) B-D-X₁-E कालावधी=7+6+0+3=16 दिवस
- 3) B-D-F कालावधी=7+6+4=17 दिवस

क्रिटिकल पाथ : नेटवर्क आकृतीमध्ये क्रिटिकल पाथ हा सर्वात लांब मार्ग आहे. हा जास्तीत जास्त कालावधीचा मार्ग आहे (म्हणजे कमाल लांबी)

क्रिटिकल पाथ = B-D-F

आणि प्रोजेक्ट पूर्ण होण्याची वेळ = 17 दिवस

सब-क्रिटिकल पाथ

सब-क्रिटिकल पाथ हा नेटवर्क डायग्राममधील दुसरा सर्वात लांब मार्ग आहे,

सब-क्रिटिकल B-D-X₁-E कालावधी=7+6+0+3=16 दिवस

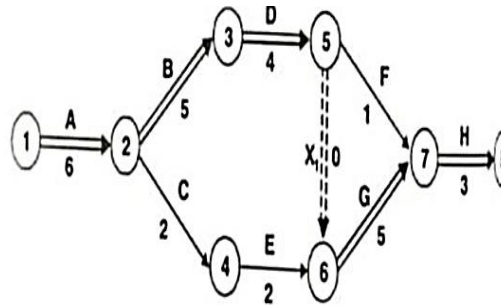
उदाहरण 2: खालील आकृतीसाठी

- a. नेटवर्क आकृती काढा

- b. क्रिटिकल पाथ आणि पूर्ण होण्याची वेळ शोधा
c. क्रिटिकल आणि नॉन-क्रिटिकल ऍक्टिव्हिटी ओळखा

ऍक्टिव्हिटी	प्रीसिडींग ऍक्टिव्हिटी	कालावधी (दिनांक)
A	-	6
B	A	5
C	A	2
D	B	4
E	C	2
F	D	1
G	D & E	5
H	F & G	3

उत्तर: नेटवर्क डायग्राम



नेटवर्कमधील मार्ग

या नेटवर्क डायग्राममध्ये तीन मार्ग आहेत.

1. A-B-D-F-H कालावधी=19 दिवस
2. A-B-D-X-G-H कालावधी=23 दिवस
3. A-C-E-G-H कालावधी=18 दिवस

क्रिटिकल पाथ

क्रिटिकल पाथ = A-B-D-X-G-H आणि प्रोजेक्ट पूर्ण होण्याची वेळ = 23 दिवस

क्रिटिकल ऍक्टिव्हिटी

उपक्रम A, B, D, X, G आणि H

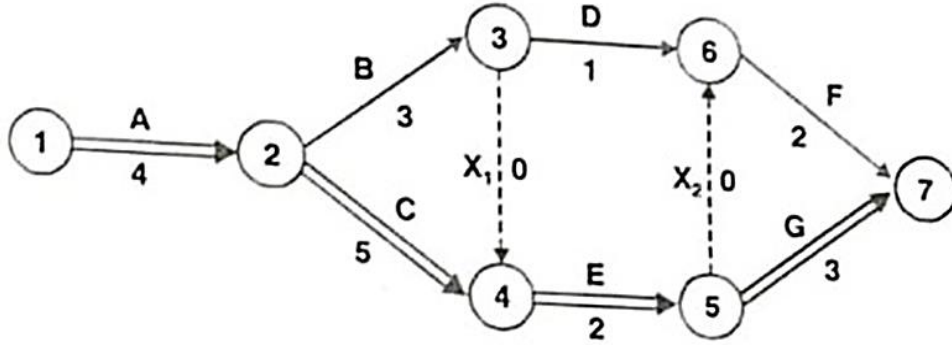
क्रिटिकल नसलेल्या ऍक्टिव्हिटी C, E, आणि F

उदाहरण 3: खालील चित्रासाठी:

- a. नेटवर्क आकृती काढा
- b. क्रिटिकल पाथ आणि प्रोजेक्ट पूर्ण होण्याची वेळ शोधा
- c. क्रिटिकल आणि नॉन-क्रिटिकल ऍक्टिव्हिटी ओळखा

ऍक्टिव्हिटी	प्रीसिडींग ऍक्टिव्हिटी	कालावधी (दिनांक)
A	-	4
B	A	3
C	A	5
D	B	1
E	B&C	2
F	D&E	2
G	E	3

उत्तर: a. नेटवर्क डायग्राम



b. नेटवर्कमधील पथ

या चित्रात तीन मार्ग आहेत

1. A-B-D-F कालावधी = 10 दिवस
2. A-B-X1-E-X2-F कालावधी = 11 दिवस
3. A-B- X1-E-G कालावधी = 12 दिवस
4. A-C-E-X2-F कालावधी = 13 दिवस
5. A-C-E-G कालावधी = 14 दिवस

क्रिटिकल पाथ

क्रिटिकल पाथ = A-C-E-G

आणि प्रोजेक्ट पूर्ण होण्याची वेळ = 14 दिवस

क्रिटिकल ऍक्टिव्हिटी

उपक्रम A, C, E आणि G

नॉन-क्रिटिकल ऍक्टिव्हिटी B, D, F, X1 आणि X2

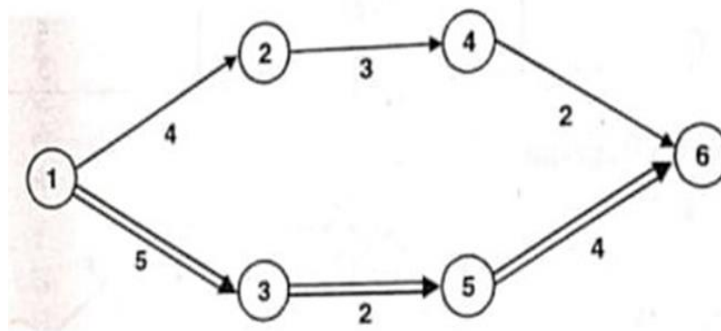
उदाहरण 4: खालील चित्रासाठी:

- a. नेटवर्क आकृती काढा
- b. क्रिटिकल पाथ आणि प्रोजेक्ट पूर्ण होण्याची वेळ शोधा
- c. क्रिटिकल आणि नॉन-क्रिटिकल ऍक्टिव्हिटी ओळखा

एॅक्टिव्हिटी	कालावधी (दिनांक)
1-2	4
1-3	5
2-4	3
3-5	2
4-6	2
5-6	4

उत्तर: या उपक्रम नोड संबंध स्वरूपात दिले आहेत

नेटवर्क डायग्राम



नेटवर्कमधील मार्ग

1. 1-2-4-6 कालावधी = 9 दिवस
2. 1-3-5-6 कालावधी = 11 दिवस

क्रिटिकल पाथ

क्रिटिकल पाथ = 1-3-5-6

आणि प्रोजेक्ट पूर्ण होण्याची वेळ = 11 दिवस

क्रिटिकल ऍक्टिव्हिटी

उपक्रम 1-3, 3-5, 5-6

नॉन-क्रिटिकल ऍक्टिव्हिटी 1-2, 2-4, 4-6

3.3 प्रोजेक्ट मूल्यांकन आणि पुनरावलोकन तंत्र(प्रोजेक्ट एव्हल्यूएशन अँड रिव्ह्यू टेक्निक) project evaluation and review technique (PERT Algorithm) (पीईआरटी अल्गोरिदम)

ही एक प्रक्रिया आहे ज्याद्वारे एखाद्या प्रोजेक्ट च्या ऍक्टिव्हिटी ना त्याच्या योग्य क्रम आणि वेळेनुसार प्रस्तुत केले जाते. हे एक शेड्यूलिंग तंत्र आहे ज्याचा वापर प्रोजेक्ट मध्ये शेड्यूल, आयोजित आणि कार्ये एकत्रित करण्यासाठी केला जातो. पीईआरटी ही मुळात व्यवस्थापन नियोजन आणि नियंत्रणाची यंत्रणा आहे जी एखाद्या विशिष्ट प्रोजेक्ट साठी ब्लूप्रिंट प्रदान करते. प्रकल्पातील सर्व प्राथमिक घटक किंवा इव्हेंट शेवटी PERT द्वारे ओळखल्या गेल्या आहेत.

या तंत्रात, एक पीईआरटी चार्ट तयार केला जातो जो प्र प्रोजेक्ट मधील सर्व निर्दिष्ट कार्यांचे वेळापत्रक दर्शवतो. पीईआरटी चार्टमधील कार्ये किंवा घटनांचे अहवाल स्तर काही प्रमाणात वर्क ब्रेकडाउन स्ट्रक्चरमध्ये परिभाषित केल्याप्रमाणे आहेत.

3.3.1 पीईआरटी चार्ट म्हणजे काय?

पीईआरटी चार्ट हे प्रोजेक्ट मॅनेजमेंट टूल आहे जे कामांची योजना आखण्यासाठी आणि शेड्यूल करण्यासाठी वापरले जाते, प्रोजेक्ट ऍक्टिव्हिटी चा क्रम आणि वेळ स्पष्ट करते. पीईआरटी चार्टचा वापर प्रोजेक्ट मधील कार्ये शेड्यूल(वेळापत्रक), आयोजन आणि समन्वय साधण्यासाठी केला जातो. पीईआरटी चार्टचा उद्देश गंभीर मार्ग निश्चित करणे आहे, ज्यामध्ये महत्त्वपूर्ण ऍक्टिव्हिटी चा समावेश आहे ज्या वेळापत्रकानुसार पूर्ण केल्या पाहिजेत. हा तक्ता प्रोजेक्ट नियोजन ऍक्टिव्हिटी मध्ये व्युत्पन्न केलेल्या माहितीच्या मदतीने तयार केला जातो जसे की प्रयत्नांचा अंदाज, सॉफ्टवेअर डेव्हलपमेंटसाठी योग्य प्रक्रिया मॉडेलची निवड आणि उपकार्यामध्ये कार्यांचे विघटन.

पीईआरटी चार्टमध्ये काय असते?

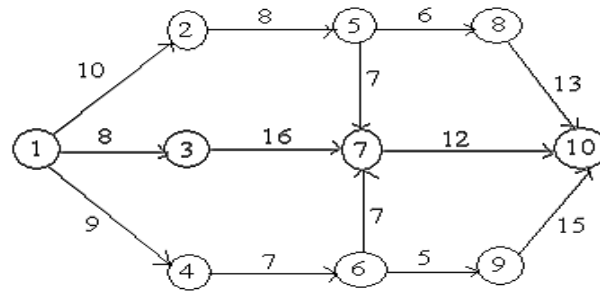
येथे PERT चार्टचे मुख्य घटक आहेत:

- **नोड्स:** हे कार्य किंवा टप्पे दर्शवते. प्रत्येक नोड कार्याचे नाव दर्शवतो आणि कार्याचा कालावधी देखील दर्शवू शकतो.

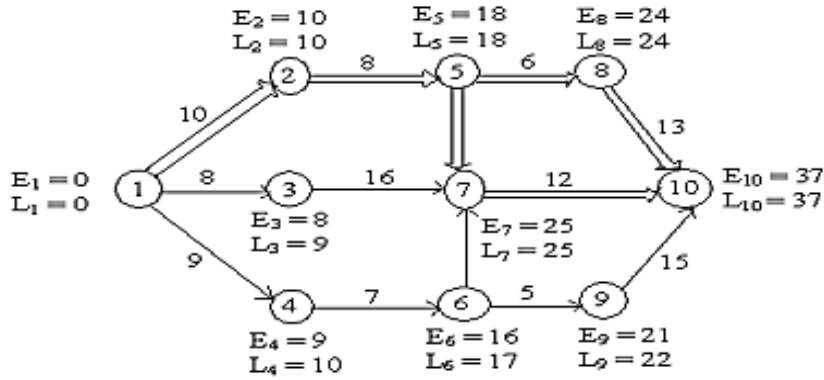
- **बाण(एरो):** हे कार्याची दिशा किंवा क्रम आणि त्यांच्यामधील अवलंबित्व देखील सूचित करते. समजा A ते B पर्यंत एक अरे असेल तर टास्क B च्या आधी टास्क A पूर्ण करणे आवश्यक आहे.
- **टाईम एस्टिमेशन(Time estimation) वेळेचा अंदाज:** हे कार्य पूर्ण करण्यासाठी कालावधीचा अंदाज लावते.
- **क्रिटिकल पाथ:** क्रिटिकल पाथ हा प्रोजेक्ट मॅनेजमेंटमधील सर्वात मोठा मार्ग आहे ज्याचा परिणाम नेहमी प्रोजेक्ट पूर्ण करण्यासाठी सर्वात कमी वेळेत होतो.
- **माईल स्टोन्स (Mile stones)(महत्त्वपूर्ण टप्पा):** हा प्रोजेक्ट टाइमलाइनमधील महत्त्वाचा मुद्दा आहे जो महत्त्वाच्या इव्हेंट किंवा मुदतीचे प्रतिनिधित्व करतो.

CPM ची सोडवलेली उदाहरणे

उदाहरण 1. सर्व नोड पॉइंट्सच्या संदर्भात अर्ली स्टार्ट (लवकर प्रारंभ) आणि लेट स्टार्ट (उशीरा प्रारंभ) निश्चित करा आणि खालील नेटवर्कसाठी क्रिटिकल पाथ ओळखा.



उत्तर : नेटवर्कमध्ये प्रत्येक नोडसाठी E आणि L ची गणना दर्शविली आहे



नेटवर्क टेबल

एक्टिव्हिटी (i, j)	नॉर्मल टाईम (D_{ij})	अर्लीएस्ट टाईम		लेटेस्ट टाईम		फ्लोट टाईम ($L_i - D_{ij} - E_i$)
		स्टार्ट (E_i)	फिनीश ($E_i + D_{ij}$)	स्टार्ट ($L_i - D_{ij}$)	फिनीश (L_i)	
1, 2)	10	0	10	0	10	0
(1, 3)	8	0	8	1	9	1
(1, 4)	9	0	9	1	10	1
(2, 5)	8	10	18	10	18	0
(4, 6)	7	9	16	10	17	1

(3, 7)	16	8	24	9	25	1
(5, 7)	7	18	25	18	25	0
(6, 7)	7	16	23	18	25	2
(5, 8)	6	18	24	18	24	0
(6, 9)	5	16	21	17	22	1
(7, 10)	12	25	37	25	37	0
(8, 10)	13	24	37	24	37	0
(9,10)	15	21	36	22	37	1

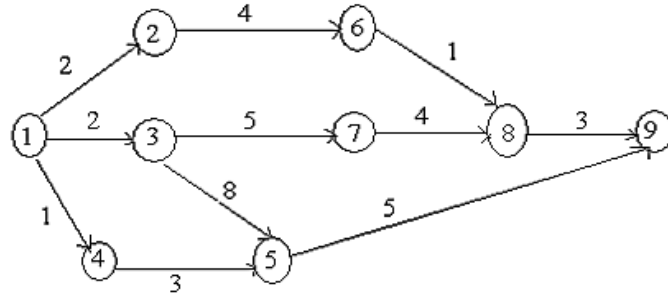
टेबलवरून, क्रिटिकल नोड्स (1, 2), (2, 5), (5, 7), (5, 8), (7, 10) आणि (8, 10) आहेत.

टेबलवरून, दोन संभाव्य क्रिटिकल पाथ आहेत

i. $1 \rightarrow 2 \rightarrow 5 \rightarrow 8 \rightarrow 10$

ii. $1 \rightarrow 2 \rightarrow 5 \rightarrow 7 \rightarrow 10$

उदाहरण 2. क्रिटिकल पाथ शोधा आणि खालील नेटवर्कसाठी स्लॅक टाईम काढा

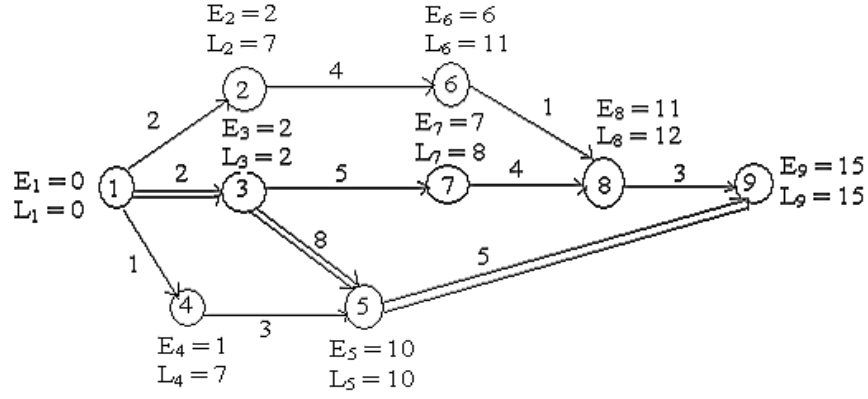


उत्तर:

सर्वात अर्लीएस्ट टाईम आणि लेटेस्ट टाईम खालील प्रमाणे आहे

एक्टिव्हिटी (i, j)	नॉर्मल टाईम (D_{ij})	अर्लीएस्ट टाईम		लेटेस्ट टाईम		फ्लोट टाईम ($L_i - D_{ij} - E_i$)
		स्टार्ट (E_i)	फिनीश ($E_i + D_{ij}$)	स्टार्ट ($L_i - D_{ij}$)	फिनीश (L_i)	
(1, 2)	2	0	2	5	7	5
(1, 3)	2	0	2	0	2	0
(1, 4)	1	0	1	6	7	6
(2, 6)	4	2	6	7	11	5
(3, 7)	5	2	7	3	8	1
(3, 5)	8	2	10	2	10	0
(4, 5)	3	1	4	7	10	6
(5, 9)	5	10	15	10	15	0
(6, 8)	1	6	7	11	12	5
(7, 8)	4	7	11	8	12	1
(8, 9)	3	11	14	12	15	1

वरील सारणीतून, क्रिटिकल नोड्स (1, 3), (3, 5) आणि (5, 9) एक्टिव्हिटी आहेत.



क्रिटिकल पाथ $1 \rightarrow 3 \rightarrow 5 \rightarrow 9$ आहे

उदाहरण 3: प्रकल्पाचे खालील टाइम शेड्यूल (वेळापत्रक) आहे

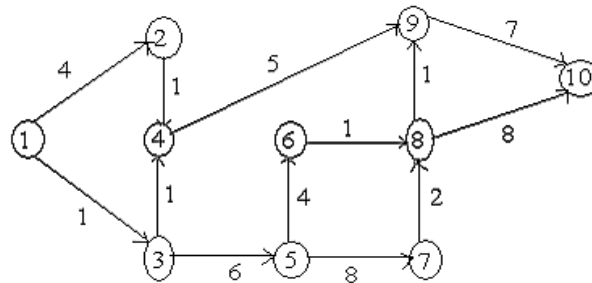
एॅक्टिव्हिटी	कालावधी (आठवड्यांत)	एॅक्टिव्हिटी	कालावधी (आठवड्यांत)
(1 - 2)	4	(5 - 7)	8
(1 - 3)	1	(6 - 8)	1
(2 - 4)	1	(7 - 8)	2
(3 - 4)	1	(8 - 9)	1
(3 - 5)	6	(8 - 10)	8
(4 - 9)	5	(9 - 10)	7
(5 - 6)	4		

नेटवर्क तयार करा आणि गणना करा

1. प्रत्येक इव्हेंट साठी TE आणि TL
2. प्रत्येक एॅक्टिव्हिटी साठी प्लोट
3. क्रिटिकल पाथ आणि त्याचा कालावधी

उत्तर:

नेटवर्क आहे



इव्हेंट क्रमांक: 1 2 3 4 5 6 7 8 9 10

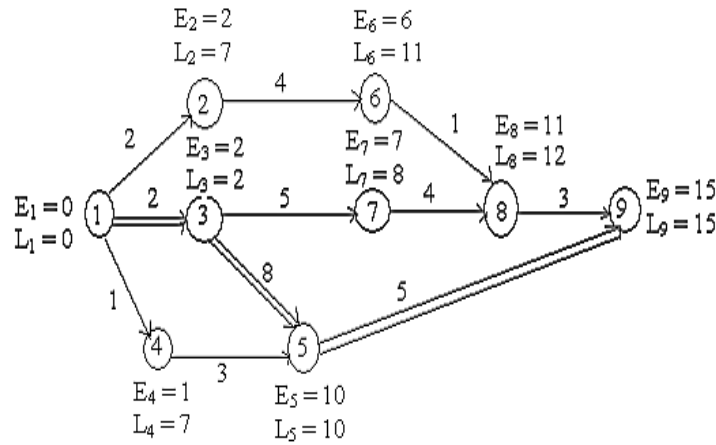
TE: 0 4 1 5 7 11 15 17 18 25

TL: 0 12 1 13 7 16 15 17 18 25

फ्लोट = TL (हेड इव्हेंट) – TE (टेल इव्हेंट) – कालावधी

एक्टिव्हिटी	कालावधी	T _E (टेल इव्हेंट)	T _L (हेड इव्हेंट)	फ्लोट
(1 – 2)	4	0	12	8
(1 – 3)	1	0	1	0
(2 – 4)	1	4	13	8
(3 – 4)	1	1	13	11
(3 – 5)	6	1	7	0
(4 – 9)	5	5	18	8
(5 – 6)	4	7	16	5
(5 – 7)	8	7	15	0
(6 – 8)	1	11	17	5
(7 – 8)	2	15	17	0
(8 – 9)	1	17	18	0
(8 – 10)	8	17	25	0
(9 – 10)	7	18	25	0

परिणामी नेटवर्क क्रिटिकल पाथ दाखवते



दोन क्रिटिकल पाथ आहेत

i. 1 → 3 → 5 → 7 → 8 → 9 → 10

ii. 1 → 3 → 5 → 7 → 8 → 10

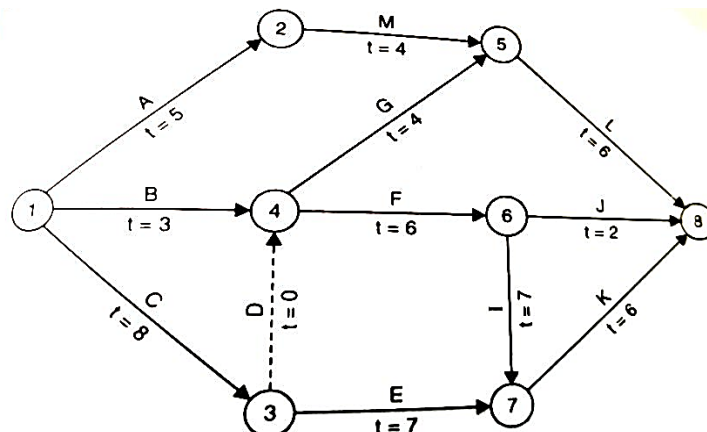
सरावासाठी उदाहरणे

उदाहरण 1: एका विशिष्ट प्रोजेक्ट चे नेटवर्क आकृतीमध्ये दर्शविले आहे, ज्यामध्ये विविध एक्टिव्हिटी चे अंदाजित कालावधी दिले आहेत.

खालील गोष्टी निर्धारित करा:

- प्रत्येक इव्हेंट ची लवकरात लवकर आणि उशिरात उशिरा वेळ (Earliest Event Time आणि Latest Event Time).
- प्रत्येक क्रियाकलापाची अर्लीएस्ट स्टार्ट आणि लेटेस्ट स्टार्ट व समाप्ती वेळ (Earliest आणि Latest Start आणि Finish Time).

- c) प्रत्येक ऍक्टिव्हिटी साठी टोटल फ्लोट (Total Float) आणि फ्री (मुक्त) फ्लोट (Free Float).
d) नेटवर्कसाठी क्रिटिकल पाथ (Critical Path).



पीईआरटी द्वारे विश्लेषणाचा मुख्य उद्देश विशिष्ट तारखेच्या आत विशिष्ट **इव्हेंट** ची पूर्तता शोधणे आहे. पीईआरटी दृष्टीकोन अनिश्चितता लक्षात घेते. तीन **टाईम** (वेळ) मूल्ये प्रत्येक ऍक्टिव्हिटी शी संबंधित आहेत

a. ऑप्टिमिस्टिक टाईम (Optimistic time)

ही सर्वात कमी वेळ आहे ज्यामध्ये ऍक्टिव्हिटी पूर्ण केला जाऊ शकतो. हे गृहीत धरते की सर्वकाही खूप चांगले होते. हे t_o द्वारे दर्शविले जाते

b. मोस्ट लैकली टाईम (Most likely time)

क्रियाकलापाला लागणाऱ्या सामान्य वेळेचा हा अंदाज आहे. हे सामान्य विलंब गृहीत धरते. जर एखादा आलेख पूर्ण होण्याच्या वेळेत आणि त्या कालावधीत पूर्ण होण्याची वारंवारता प्लॉट केला असेल, तर बहुधा वेळ इव्हेंट घडण्याची सर्वोच्च वारंवारता दर्शविल. हे t_m द्वारे दर्शविले जाते.

c. पेसिमिस्टिक टाईम (Pessimistic time)

व काही चुकीचे झाल्यास क्रियाकलापास लागू शकणारा सर्वात जास्त वेळ हे दर्शवते. आशावादी अंदाजानुसार, हे मूल्य असे असू शकते की शंभरपैकी एक किंवा वीसपैकी एक या मूल्यापेक्षा जास्त वेळ घेईल. हे t_p द्वारे दर्शविले जाते.

3.3.2 PERT गणनेमध्ये, टक्के अपेक्षित मूल्य मिळविण्यासाठी सर्व मूल्ये वापरली जातात.

1. **एक्सपेक्टेड टाईम (Expected time) अपेक्षित वेळ** - एखाद्या ऍक्टिव्हिटी मोठ्या संख्येने पुनरावृत्ती झाल्यास सरासरी वेळ लागतो आणि ऍक्टिव्हिटी वेळ बीटा **डिस्ट्रिब्युशन** (वितरण) चे अनुसरण करते या गृहीतावर आधारित आहे, हे $t_e = (t_o + 4t_m + t_p) / 6$ द्वारे दिले जाते.

2. ऍक्टिव्हिटी साठी व्हेरियन्स (परिवर्तन) (Variance) खालील सूत्राने दिले जाते: $\sigma^2 = [(t_p - t_o) / 6]^2$

PERT अल्गोरिदम खालील चरणांद्वारे सोडवता येतो:

चरण 1: PERT नेटवर्कमध्ये, सर्वाधिक संभाव्य वेळ (Most Likely Time) ला t_m , आशावादी वेळ (Optimistic Time) ला t_o , आणि निराशावादी वेळ (Pessimistic Time) ला t_p असे दर्शवा.

चरण 2: ऍक्टिव्हिटी कालावधीसाठी β -वितरण (Beta Distribution) वापरा. खालील सूत्राचा वापर करून प्रत्येक ऍक्टिव्हिटी साठी अपेक्षित वेळ t_e शोधा: $t_e = (t_p + 4t_m + t_o) / 6$.

चरण 3: लवकरात लवकर सुरुवातीची वेळ (Earliest Start Time) आणि उशिरात उशिरा पूर्ण होण्याची वेळ (Latest Finish Time) यांची गणना करा. एकूण फ्लोट (Total Float) हे या दोघांमधील फरक म्हणून मोजा.

चरण 4: महत्वाच्या क्रियाकलापांचा (Critical Activities) आणि महत्वाचा मार्ग (Critical Path) ओळखा, तसेच प्रोजेक्ट च्या अपेक्षित पूर्ण होण्याच्या तारखेची गणना करा.

चरण 5: t_p आणि t_o च्या मूल्यांचा वापर करून प्रत्येक क्रियाकलापाच्या कालावधीचे परिवर्तन (Variance, σ^2) खालील सूत्राने शोधा:

$$\sigma^2 = ((t_p - t_o) / 6)$$

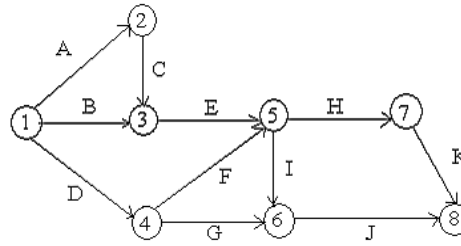
चरण 6: खालील सूत्राचा वापर करून z_0 गणना करा:

$z_0 =$ देय तारीख - पूर्ण होण्याची अपेक्षित तारीख $\sqrt{\text{प्रोजेक्ट भिन्नता}}$

चरण 7: स्टँडर्ड नॉर्मल टेबल (प्रमाणित सामान्य तक्ता) (Standard Normal Table) वापरून $P(z \leq z_0)$ चा शोध घ्या, जो प्रोजेक्ट नियोजित वेळेत पूर्ण होण्याची शक्यता दर्शवतो. येथे z हे $N(0,1)$ सामान्य वितरणाचे अनुसरण करते.

सोडवलेली उदाहरणे

उदाहरण 1: प्रोजेक्ट साठी



टास्क (कार्य)	A	B	C	D	E	F	G	H	I	J	K
लिस्ट टाईम (कमीत कमी वेळ)	4	5	8	2	4	6	8	5	3	5	6
ग्रेटेस्ट टाईम (सर्वात मोठा वेळ)	8	10	12	7	10	15	16	9	7	11	13
मोस्ट लाईकली टाईम (सर्वात संभाव्य वेळ)	5	7	11	3	7	9	12	6	5	8	9

प्रत्येक इव्हेंटसाठी **अर्लिअस्ट** आणि **लेटेस्ट एक्सपेक्टेड टाईम** शोधा आणि नेटवर्कमधील क्रिटिकल पाथ देखील शोधा.

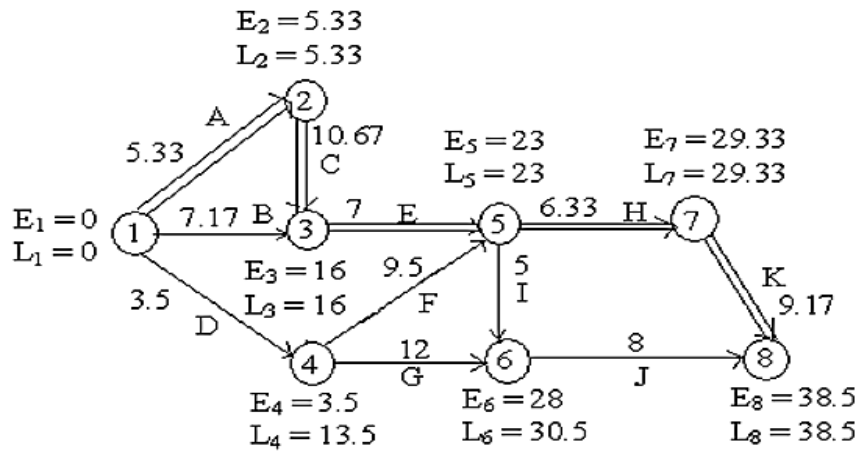
उत्तर:

टास्क	लिस्ट टाईम (t_o)	ग्रेटेस्ट टाईम (t_p)	मोस्ट लाईकली टाईम (t_m)	एक्सपेक्टेड टाईम ($t_o + t_p + 4t_m$)/6
A	4	8	5	5.33
B	5	10	7	7.17
C	8	12	11	10.67
D	2	7	3	3.5
E	4	10	7	7
F	6	15	9	9.5

G	8	16	12	12
H	5	9	6	6.33
I	3	7	5	5
J	5	11	8	8
K	6	13	9	9.17

टास्क	एक्सपेक्टेड टाईम (t_e)	स्टार्ट		फिनीश		टोटल प्लोट (संपूर्ण विलंब कालावधी)
		अर्लिस्ट	लेटेस्ट	अर्लिस्ट	लेटेस्ट	
A	5.33	0	0	5.33	5.33	0
B	7.17	0	8.83	7.17	16	8.83
C	10.67	5.33	5.33	16	16	0
D	3.5	0	10	3.5	13.5	10
E	7	16	16	23	23	0
F	9.5	3.5	13.5	13	23	10
G	12	3.5	18.5	15.5	30.5	15
H	6.33	23	23	29.33	29.33	0
I	5	23	25.5	28	30.5	2.5
J	8	28	30.5	36	38.5	2.5
K	9.17	29.33	29.33	31.5	38.5	0

नेटवर्क



क्रिटिकल पाथ आहे: $A \rightarrow C \rightarrow E \rightarrow H \rightarrow K$

उदाहरण 2: एका प्रकल्पाच्या खालील वैशिष्ट्ये आहेत.

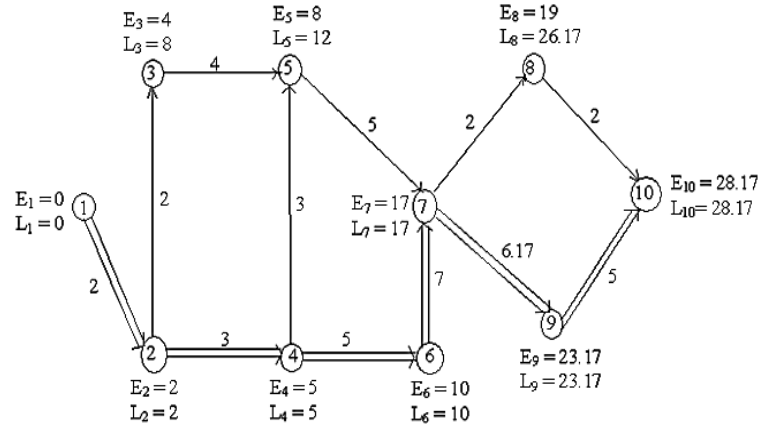
एक्टिव्हिटी	मोस्ट ऑप्टिमिस्टिक टाईम (a)	मोस्ट पॅसिमिस्टिक टाईम (b)	मोस्ट लाईकली टाईम (m)
(1 – 2)	1	5	1.5
(2 – 3)	1	3	2
(2 – 4)	1	5	3
(3 – 5)	3	5	4
(4 – 5)	2	4	3
(4 – 6)	3	7	5
(5 – 7)	4	6	5
(6 – 7)	6	8	7
(7 – 8)	2	6	4
(7 – 9)	5	8	6
(8 – 10)	1	3	2
(9 – 10)	3	7	5

पीईआरटी नेटवर्क तयार करा. प्रत्येक इव्हेंटसाठी क्रिटिकल पाथ आणि व्हेरियन्स (भिन्नता) शोधा

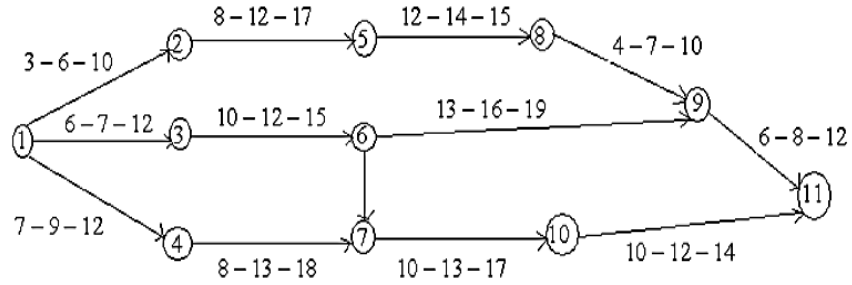
उत्तर:

एक्टिव्हिटी	(a)	(b)	(m)	(4m)	$t_e (a + b + 4m)/6$	$\frac{v}{[(b - a) / 6]^2}$
(1 – 2)	1	5	1.5	6	2	4/9
(2 – 3)	1	3	2	8	2	1/9
(2 – 4)	1	5	3	12	3	4/9
(3 – 5)	3	5	4	16	4	1/9
(4 – 5)	2	4	3	12	3	1/9
(4 – 6)	3	7	5	20	5	4/9
(5 – 7)	4	6	5	20	5	1/9
(6 – 7)	6	8	7	28	7	1/9
(7 – 8)	2	6	4	16	4	4/9
(7 – 9)	5	8	6	24	6.17	1/4
(8 – 10)	1	3	2	8	2	1/9
(9 – 10)	3	7	5	20	5	4/9

नेटवर्क खालीलप्रमाणे आहे



उदाहरण 3: प्रत्येक एक्टिव्हिटी साठी व्हेरियन्स आणि एक्सपेक्टेड टाईम ची गणना करा.



उत्तर:

एक्टिव्हिटी	(t _o)	(t _m)	(t _p)	$t_e(t_o + t_p + 4t_m)/6$	$v [(t_p - t_o) / 6]^2$
(1 - 2)	3	6	10	6.2	1.36
(1 - 3)	6	7	12	7.7	1.00
(1 - 4)	7	9	12	9.2	0.69
(2 - 3)	0	0	0	0.0	0.00
(2 - 5)	8	12	17	12.2	2.25
(3 - 6)	10	12	15	12.2	0.69
(4 - 7)	8	13	19	13.2	3.36
(5 - 8)	12	14	15	13.9	0.25
(6 - 7)	8	9	10	9.0	0.11
(6 - 9)	13	16	19	16.0	1.00
(8 - 9)	4	7	10	7.0	1.00
(7 - 10)	10	13	17	13.2	1.36
(9 - 11)	6	8	12	8.4	1.00
(10 - 11)	10	12	14	12.0	0.66

सरावासाठी उदाहरणे

1. खालील डेटासाठी, नेटवर्क काढा. अर्लिअस्ट एक्सपेक्टेड टाईम आणि लेटेस्ट एक्सपेक्टेड टाईम ची गणना केल्यानंतर क्रिटिकल पाथ, स्लॅक टाईम शोधा

एॅक्टिव्हिटी	कालावधी	एॅक्टिव्हिटी	कालावधी
(1 – 2)	5	(5 – 9)	3
(1 – 3)	8	(6 – 10)	5
(2 – 4)	6	(7 – 10)	4
(2 – 5)	4	(8 – 11)	9
(2 – 6)	4	(9 – 12)	2
(3 – 7)	5	(10 – 12)	4
(3 – 8)	3	(11 – 13)	1
(4 – 9)	1	(12 – 13)	7

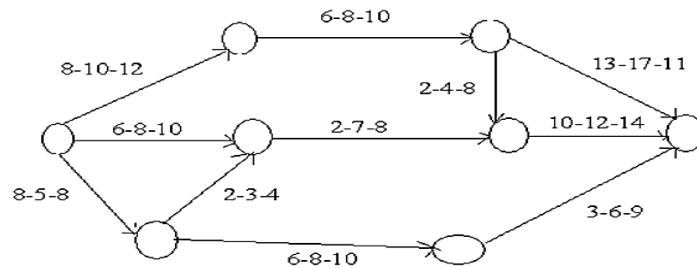
2. प्रोजेक्ट शेड्यूलमध्ये खालील वैशिष्ट्ये आहेत

एॅक्टिव्हिटी	मोस्ट ऑप्टिमिस्टिक टाईम	एॅक्टिव्हिटी	मोस्ट ऑप्टिमिस्टिक टाईम
(1 – 2)	1	2	3
(2 – 3)	1	2	3
(2 – 4)	1	3	5
(3 – 5)	3	4	5
(4 – 5)	2	5	4
(4 – 6)	3	5	7
(5 – 7)	4	5	6
(6 – 7)	6	7	8
(7 – 8)	2	4	6
(7 – 9)	4	6	8
(8 – 10)	1	2	3
(9 – 10)	3	5	7

पीईआरटी नेटवर्क तयार करा आणि शोधा

- अर्लिअस्ट पॉसिबल टाईम
- लेटेस्ट अलोव्हेबल टाईम
- स्लॅक व्हॅल्यू (मूल्ये)
- क्रिटिकल पाथ

3. प्रत्येक एॅक्टिव्हिटी साठी व्हेरियन्स एक्सपेक्टेड टाईम ची गणना करा



3.3.3 प्रोबॅबिलिटी ऑफ मीटिंग शेड्युल डेट (Probability of meeting schedule date) (बैठक वेळापत्रक तारीख संभाव्यता)

क्रिटिकल पाथ आणि सर्व ॲक्टिव्हिटीजच्या घडण्याच्या वेळा मोजल्यानंतर, पुढील म्हणजे नियोजित तारखेला किंवा त्यापूर्वी एखादी विशिष्ट इव्हेंट घडण्याची शक्यता शोधणे.

पूर्ण होण्याची नियोजित वेळ पूर्ण करण्याची संभाव्यता शोधण्याची प्रक्रिया

(नॉर्मल डिस्ट्रिब्युशन कर्व) सामान्य वितरण वक्राच्या गुणधर्मांचा आणि टक्केवारी संभाव्यता तक्त्याचा वापर फक्त तेव्हाच करता येतो जेव्हा विशिष्ट नेटवर्कमधून मिळालेला यादृच्छिक वक्र (रँडम कर्व) प्रमाणित (सामान्यीकृत) स्वरूपात रूपांतरित केला जातो.

हे अंतिम इव्हेंट चे मूल्य सामान्य वितरण वक्राच्या मोडेल मूल्याशी जुळवून प्राप्त केले जाते. निर्धारित वेळेत पूर्ण होण्याची संभाव्यता शोधण्याची प्रक्रिया खालीलप्रमाणे आहे.

टप्पा 1: नेटवर्कसाठी महत्त्वपूर्ण मार्गास अनुरूप असलेले प्रमाणित विचलन (σ) खालील संबंधाचा वापर करून निर्धारित करा.

$$\sigma = \sqrt{\text{sum of variances along critical path}}$$

$$\text{म्हणजे, } \sigma = \sqrt{\sum \sigma_{ij}^2}$$

जेथे σ_{ij}^2 = क्रिटिकल पाथ वर i-j ॲक्टिव्हिटी साठी व्हेरियन्स $\sigma_{ij}^2 = \left(\frac{t_p - t_o}{6}\right)^2$

टप्पा 2: शेड्यूलड टाइम कम्लिशन (निर्धारित पूर्णता वेळ) (t_s) आणि अर्लिअस्ट एक्सपेक्टेड टाइम (लवकरात लवकर अपेक्षित पूर्णता वेळ) (t_e) ज्ञात असताना, वेळेतील अंतर ($t_s - t_e$) शोधा आणि ते संभाव्यता घटक Z च्या स्वरूपात खालील संबंधाने व्यक्त करा.

$$Z = \frac{t_s - t_e}{\sigma} \quad \text{किंवा} \quad Z = \frac{t_s - t_e}{\sqrt{\sum \sigma^2}}$$

प्रोबॅबिलिटी फॅक्टर (संभाव्यता घटक) Z हा सकारात्मक, शून्य किंवा नकारात्मक असू शकतो.

जेव्हा Z पॉझिटिव्ह असतो (t_s t_e च्या उजवीकडे), प्रोजेक्ट वेळेत पूर्ण होण्याची शक्यता 50% पेक्षा जास्त असते.

जेव्हा Z शून्य असतो (t_s सोबत t_e जुळतो) प्रोजेक्ट वेळेत पूर्ण होण्याची शक्यता पन्नास- पन्नास असते.

जेव्हा Z ऋण असेल (t_s t_e च्या डावीकडे), प्रोजेक्ट वेळेत पूर्ण होण्याची शक्यता 50% पेक्षा कमी असते.

टप्पा 3: नॉर्मल व्हेरिएट टेबल (सामान्य विचलित सारणी) च्या संदर्भात परसेंटेज प्रोबॅबिलिटी (टक्केवारी संभाव्यता) शोधा

उदाहरण 1: एका प्रोजेक्ट ला गंभीर मार्गावर 15 महिने लागतील, 3 महिन्यांचे मानक विचलन असेल. अ) 15 महिन्यांत ब) 18 महिने आणि क) 12 महिन्यांत प्रोजेक्ट पूर्ण होण्याची संभाव्यता किती आहे?

उत्तर:

संभाव्यता घटक Z खालील सूत्राने दिला जातो:

$$Z = \frac{t_s - t_e}{\sigma}$$

जेथे, $t_e = 15$ महिने $\sigma = 3$ महिने

a) दिलेले आहे: $t_s = 15$ महिने $\therefore Z = \frac{15-15}{3} = 0$

तक्त्यानुसार, $Z = 0$ असताना, संभाव्यता = 50%

b) दिलेले आहे: $t_s = 18$ महिने $\therefore Z = \frac{18-15}{3} = 1$

तक्त्यानुसार, $Z = 1$ साठी संभाव्यता 84.13% आहे.

c) दिलेले आहे: $t_s = 12$ महिने $\therefore Z = \frac{12-15}{3} = -1$

तक्त्यानुसार, $Z = -1$ साठी संभाव्यता 15.87% आहे.

उदाहरण 2: पीईआरटी गणनेमध्ये 16 च्या फरकासह प्रकल्पाची लांबी 50 आठवडे मिळते. अ) 95% b) 75% c) 40% संभाव्यतेसह प्रोजेक्ट पूर्ण होण्यासाठी तुम्ही किती आठवड्यांत सोडाल?

उत्तर:

मानक विचलन स्टँडर्ड डेव्हिएशन $\sigma = \sqrt{16} = 4$

a) 95% संभाव्यतेसाठी $Z=1.65$

$$Z = \frac{t_s - t_e}{\sigma} \therefore t_s = \sigma Z + t_e = 4 \times 1.65 + 50 = 56.6 \approx 57 \text{ आठवडे}$$

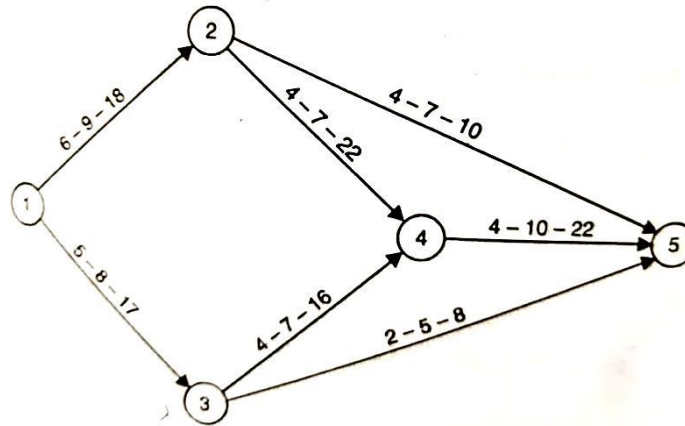
b) 75% संभाव्यतेसाठी, $Z=0.69$

$$\therefore t_s = \sigma Z + t_e = 4 \times 0.69 + 50 = 52.76 \approx 53 \text{ आठवडे}$$

c) 40% संभाव्यतेसाठी, $Z=-0.25$

$$\therefore t_s = \sigma Z + t_e = 4 \times (-0.25) + 50 = 49 \text{ आठवडे}$$

उदाहरण 3: खाली दर्शविलेल्या नेटवर्कसाठी प्रत्येक क्रियाकलापासाठी वेळेचा अंदाज (दिवसांमध्ये) नमूद केला आहे. 35 दिवसात प्रोजेक्ट पूर्ण होण्याची शक्यता निश्चित करा.



आम्ही प्रथम प्रत्येक कृतीसाठी अपेक्षित पूर्णता वेळ (t_e) आणि विभेदन (σ^2) खालील सूत्रांचा वापर करून निर्धारित करतो.

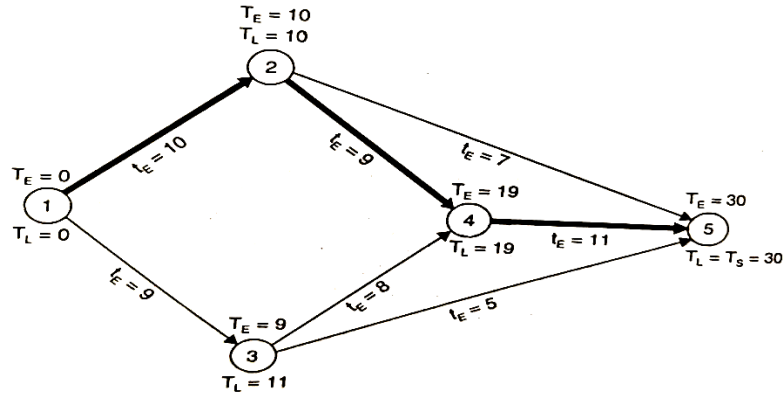
$$t_e = \frac{t_o + 4t_m + t_p}{6} \quad \text{and} \quad \sigma^2 = \left(\frac{t_p - t_o}{6} \right)^2$$

एक्टिव्हिटी	t_o	t_m	t_p	t_e	σ^2
1-2	6	9	18	10	4
1-3	5	8	17	9	4
2-4	4	7	22	9	9
3-4	4	7	16	8	4
4-5	4	10	22	11	9
2-5	4	7	10	7	1
3-5	2	5	8	5	1

वेगवेगळ्या घटनांसाठी अपेक्षित वेळ (t_e), मध्य वेळ (t_m) आणि शिल्लक वेळ (Slack) यांची गणना खालील तक्त्यात मांडली जाते. शेवटच्या घटनेसाठी, t_m ला t_e च्या बरोबरीने घेतले जाते.

खालील तक्त्यामध्ये आपल्याला आढळते की इव्हेंट 1, 2, 4 आणि 5 मध्ये स्लॅक कमीत कमी आहे. म्हणून 1-2-4-5 हा क्रिटिकल पाथ आहे. क्रिटिकल पाथ जाड रेषांनी चिन्हांकित केला आहे.

इव्हेंट नंबर	अर्लिअस्ट एक्सपेक्टेड टाईम				लेटेस्ट ओकारन्स टाईम				स्लॅक S
	प्रीडीसेसर इव्हेंट (i)	t_e^{ij}	t_e^j	t_e	सक्सेसर इव्हेंट (j)	t_m^{ij}	t_m^j	t_m	
1	-	-	0	0	2 3	10 9	<u>0</u> 2	0	0
2	1	10	<u>10</u>	10	4 5	9 7	<u>10</u>	10	0
3	1	9	<u>9</u>	9	4 5	8 5	<u>11</u>	11	2
4	2 3	9 8	<u>19</u> 17	19	5	11	<u>19</u>	19	0
5	2 3	7 5	17 14	30			30	30	0



आता मानक विचलन $\sigma = \sqrt{\sum \sigma_{ij}^2}$ द्वारे दिले जाते

$$= \sigma_{1-2}^2 + \sigma_{2-4}^2 + \sigma_{4-5}^2$$

$$= 4 + 9 + 9 = 22$$

$$\sigma = \sqrt{22} = 4.69$$

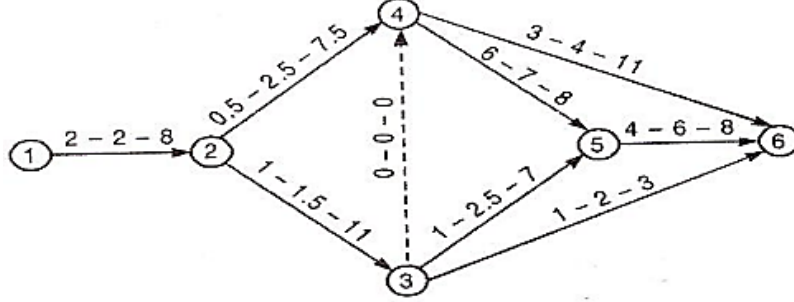
$$\therefore Z = \frac{t_s - t_e}{\sigma} \quad \text{इथे } t_s = 35 \text{ दिवस}$$

$\therefore$ प्रोजेक्ट 35 दिवसांत पूर्ण होण्याची 85.7% शक्यता आहे.

सरावासाठी उदाहरणे

- जर एखाद्या प्रकल्पाचा महत्त्वाचा मार्ग 20 महिन्यांचा असेल आणि त्यासोबत 4 महिन्यांचा मानक विचलन असेल, तर प्रोजेक्ट a) 20 महिन्यांत b) 18 महिन्यांत c) 24 महिन्यांत पूर्ण होण्याची शक्यता किती आहे?

2. पीईआरटी गणना 9 च्या फरकासह 60 आठवड्यांच्या प्रकल्पाची लांबी देते. 0.99 च्या संभाव्यतेसह प्रोजेक्ट किती आठवड्यांमध्ये पूर्ण होण्याची अपेक्षा कराल? (प्रोजेक्टची हीच लांबी आहे जी तुम्हाला 1% वेळा ओलांडण्याची अपेक्षा आहे जर प्रकल्पाची पुनरावृत्ती एकाच पद्धतीने केली गेली असेल).
3. खाली दर्शविलेल्या नेटवर्कचा विचार करा. क्रियाकलापांचे तीन वेळा अंदाज बाणांसह दिले आहेत. क्रिटिकल पाथ निश्चित करा. 20 दिवसांत प्रोजेक्ट पूर्ण होण्याची शक्यता किती आहे?



युनिट 4

बायव्हेरिएट फ्रिक्वेन्सी डिसट्रीब्युशन अँड कोरिलेशन (Bivariate frequency distribution and correlation)

विषय निष्पत्ती (Course Outcome):

CO4: बायव्हेरिएट फ्रिक्वेन्सी डिसट्रीब्युशन अँड कोरिलेशनचा वापर डेटा विश्लेषणासाठी करा.

घटक निष्पत्ती (Theory Learning Outcome):

1. कोरिलेशन आणि त्याचे प्रकार समजावून सांगा.
2. कार्ल पिअर्सनचा कोरिलेशन कोइफिशियंटचा वापर करून दोन व्हेरियेबलस् मधील लिनियर रिलेशनचे मॅग्निट्यूड मोजा.
3. स्पीअरमनचा रँक कोरिलेशन कोइफिशियंटचा वापर करून रँक मोजा.
4. दिलेल्या फ्रिक्वेन्सी डिसट्रीब्युशनचा प्रकार ठरवा.

परिचय: इंट्रोडक्शन: बायव्हेरिएट फ्रिक्वेन्सी डिसट्रीब्युशन (Introduction: Bivariate frequency distribution)

स्टॅटिस्टिक्स मध्ये बायव्हेरिएट फ्रिक्वेन्सी डिसट्रीब्युशन दोन व्हेरियेबलस् मधील रिलेशन दर्शवते आणि त्यांचे व्हॅल्यूज फ्रिक्वेन्सी डिसट्रीब्युशन मध्ये आयोजन करून विश्लेषित करते. हे दोन भिन्न डेटा संच कसे संबंधित आहेत आणि विशिष्ट संयोजनांचे मूल्य किती वेळा येतात याचे विश्लेषण करण्यास मदत करते. अशा प्रकारचे डिसट्रीब्युशन अर्थशास्त्र, जैवविज्ञान, आणि सामाजिक शास्त्र अशा क्षेत्रांमध्ये महत्त्वाचे आहे, जिथे व्हेरियेबलस् मधील संबंध डेटा विश्लेषणासाठी अत्यंत महत्त्वाचे असतात.

उदाहरणार्थ, अर्थशास्त्रामध्ये, बायव्हेरिएट डिसट्रीब्युशन जमा आणि खर्च यांमधील रिलेशन दर्शवू शकते किंवा सप्लाय आणि डिमांड यांमधील परस्परसंबंध विश्लेषित करू शकते.

4.1 कोरिलेशन (Correlation):

कोरिलेशन हे स्टॅटिस्टिक्सचे एक माप आहे जे दोन व्हेरियेबलस् मधील रिलेशनची ताकद आणि दिशा व्यक्त करते. बायव्हेरिएट डिसट्रीब्युशन मध्ये जर एका व्हेरियेबल मध्ये होणारा बदल त्या दुसऱ्या व्हेरियेबलमध्ये होणाऱ्या बदलासोबत असतो, म्हणजे एक बदल दुसऱ्या बदलासोबत वाढतो किंवा कमी होतो, तर त्या दोन व्हेरियेबलस् मध्ये रिलेशन असते उदाहरणार्थ, व्यायाम आणि शारीरिक तंदुरुस्ती, मागणी आणि पुरवठा, जमा आणि खर्च.

4.1.1 टाईप्स ऑफ कोरिलेशन (Types of Correlation):

- **पॉझिटिव्ह कोरिलेशन (Positive Correlation):** जर एका व्हेरियेबल मध्ये वाढ किंवा घट दुसऱ्या व्हेरियेबल मध्ये समांतर वाढ किंवा घट आणते, तर त्या दोन व्हेरियेबलस् मध्ये पॉझिटिव्ह कोरिलेशन असते. उदाहरणार्थ, जसे-जसे तापमान वाढते, तसतसे एसी विक्री वाढते.
- **नेगेटिव्ह कोरिलेशन (Negative Correlation):** जर एका व्हेरियेबल मध्ये वाढ किंवा घट दुसऱ्या व्हेरियेबल मध्ये उलट घट किंवा वाढ आणते, तर त्या दोन व्हेरियेबलस् मध्ये नेगेटिव्ह कोरिलेशन असते. उदाहरणार्थ, जास्त वेळ सोशल मीडियावर घालविल्यास, कामाची उत्पादकता कमी होऊ शकते.
- **लिनियर कोरिलेशन (Linear Correlation):** जर सर्व डेटा बिंदू एक सरळ रेषेसोबत जवळजवळ असतील, तर कोरिलेशन लिनियर असते. याचा अर्थ, दोन व्हेरियेबलस् मधील रिलेशन लिनियर इक्वेशन ने दर्शविला जातो.
- **नॉन- लिनियर कोरिलेशन (Non-Linear Correlation):** नॉन- लिनियर कोरिलेशन मध्ये, डेटा बिंदू सरळ रेषेसोबत नाही तर कोणत्याही वक्र रेषेसोबत असतात. याचा अर्थ दोन्ही व्हेरियेबलस् मधील रिलेशन अधिक जटिल असते आणि साध्या रेषीय मॉडेलने याला व्यक्त करता येत नाही.
- **सिम्पल कोरिलेशन (Simple Correlation):** दोन व्हेरियेबलस् मधील कोरिलेशनला सिम्पल कोरिलेशन म्हणतात. उदाहरणार्थ, उंची आणि वजन यांमधील रिलेशन.

- **मल्टिपल कोरिलेशन (Multiple Correlation):** मल्टिपल कोरिलेशन तीन किंवा अधिक व्हेरियेबलस् मधील रिलेशन दर्शवितो. उदाहरणार्थ, विद्यार्थ्यांच्या शैक्षणिक कार्यक्षमतेचे अंदाज लावताना, अभ्यास वेळ, मागील गुण, आणि उपस्थिती हे सर्व घटक एकत्रितपणे एक मल्टिपल कोरिलेशन तयार करतात.

4.2 मॅथेमॅटिकल मेथड ऑफ कोरिलेशन कोइफिशियंट (Mathematical Method of correlation coefficient):

a) स्कॅटर डायग्राम (Scatter Diagram):

स्कॅटर डायग्राम हा दोन व्हेरियेबलस् मधील रिलेशन दर्शविणारा एक ग्राफिकल प्रतिनिधित्व आहे. यामध्ये प्रत्येक पॉईंट हे दोन व्हेरियेबलस् साठी मूल्यांची जोडी दर्शवितो. याचा उपयोग कोरिलेशन पाहण्यासाठी आणि डेटा समजून घेण्यासाठी केला जातो.

b) कार्ल पिअर्सनचा कोरिलेशन कोइफिशियंट (Karl Pearson's Coefficient of Correlation):

कार्ल पीअर्सनचा कोरिलेशन कोइफिशियंट (r) दोन व्हेरियेबलस् मधील रेषीय संबंधाची तीव्रता किंवा प्रमाण मोजतो. हे दोन व्हेरियेबलस् च्या कोरिलेशन ची मोजणी करतो आणि कोरिलेशनचा प्रकार (पॉझिटिव्ह किंवा नेगेटिव्ह) समजून घेण्यास मदत करतो. कोरिलेशन कोइफिशियंट ची गणना करण्यासाठी, संबंधित डेटा आणि विविध सूत्रांचा वापर केला जातो.

c) स्पीअरमनचा रँक कोरिलेशन कोइफिशियंट (Spearman's Rank Coefficient of Correlation):

स्पीअरमनचा रँक कोरिलेशन कोइफिशियंट दोन व्हेरियेबलस् मधील रँक किंवा स्थानानुसार संबंध मोजतो. यामध्ये प्रत्येक डेटा पॉईंटला एक रँक दिला जातो आणि त्यावर आधारित कोरिलेशन मोजला जातो. याचा उपयोग डेटाच्या क्रमाने केलेल्या विश्लेषणासाठी केला जातो.

या तीन पद्धतींनी कोरिलेशन चे मापन केले जाते, ज्यामुळे आपण विविध परिस्थितींमध्ये डेटा आणि त्यांच्यातील संबंध अधिक स्पष्टपणे समजू शकतो.

मुख्य संकल्पनांचे सारांश:

- **कोरिलेशन (Correlation):** दोन व्हेरियेबलस् मधील संबंध मोजतो.
 - पॉझिटिव्ह कोरिलेशन: एका व्हेरियेबल मध्ये वाढ, दुसऱ्या मध्ये देखील वाढ.
 - नेगेटिव्ह कोरिलेशन: एका व्हेरियेबल मध्ये वाढ, दुसऱ्या मध्ये घट.
 - लिनिअर कोरिलेशन: सरळ रेषेचा संबंध.
 - नॉन- लिनिअर कोरिलेशन: वक्र रेषेचा संबंध.
 - सिम्पल कोरिलेशन: दोन व्हेरियेबलस् मधील संबंध.
 - मल्टिपल कोरिलेशन: तीन किंवा अधिक व्हेरियेबलस् मधील संबंध.

सहसंबंध मोजण्याच्या पद्धती:

1. कार्ल पिअर्सनचा कोरिलेशन कोइफिशियंट : रेषीय संबंध मोजतो.
2. स्पीअरमनचा रँक कोरिलेशन कोइफिशियंट: मोनोटोनिक संबंध मोजतो.

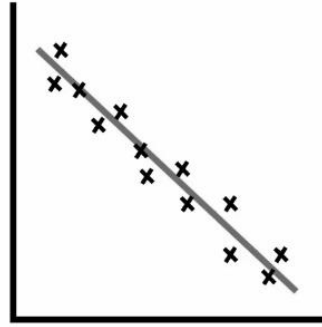
स्कॅटर डायग्राम (Scatter Diagram): स्कॅटर डायग्राम हे दोन व्हेरियेबलस् मधील संबंधाचे ग्राफिकल प्रतिनिधित्व आहे. हे स्टॅटिस्टिक्स मध्ये डेटा दृश्यमान करण्यासाठी आणि कोरिलेशन प्रभावीपणे समजून घेण्यासाठी आवश्यक उपकरण आहे. यामध्ये प्रत्येक पॉईंट हे दोन व्हेरियेबलस् साठी मूल्यांची जोडी म्हणून ग्राफवर चिन्हित केले जाते. X-अक्षावर स्वतंत्र व्हेरियेबल (Independent Variable) आणि Y-अक्षावर आश्रित व्हेरियेबल (Dependent Variable) दाखवले जातात. (Figure. 4.1 पहा)

पॉझिटिव्ह कोरिलेशन



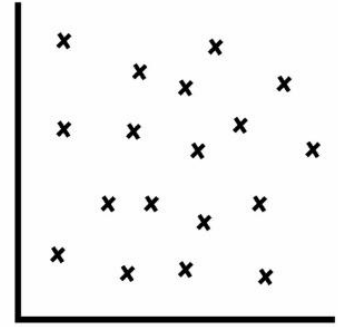
Positive
Correlation

नेगेटिव्ह कोरिलेशन



Negative
Correlation

कोरिलेशन नाही



No
Correlation

Figure 4.1: स्कॅटर डायग्राम (Scatter Diagram)

4.2.1 कार्ल पिअर्सनचा कोरिलेशन कोइफिशियंट (Karl Pearson's Coefficient of Correlation):

कार्ल पिअर्सनचा कोरिलेशन कोइफिशियंट (r) दोन व्हेरियेबलस् मधील लिनिअर कोरिलेशनची तीव्रता किंवा प्रमाण मोजतो. याचे सूत्र असे आहे:

$$r(x, y) = r = \frac{cov(x, y)}{\sigma_x \sigma_y}$$

बायव्हेरिएट फ्रिकेन्सी डिसट्रीब्युशन मध्ये, जर (x_i, y_i) मूल्ये $(x_1, y_1) (x_2, y_2) \dots (x_n, y_n)$ घेतली जातात, तर,

$$cov(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

जिथे, $\bar{x}$ आणि $\bar{y}$ हे अनुक्रमे x आणि y साठी सरासरी आहेत.
त्याचप्रमाणे,

$$\sigma_x = \sqrt{\sum_{i=1}^n \frac{1}{n} (x_i - \bar{x})^2}, \quad \sigma_y = \sqrt{\sum_{i=1}^n \frac{1}{n} (y_i - \bar{y})^2}$$

हे x आणि y साठी स्टॅण्डर्ड डेव्हिएशन्स आहेत

$$\begin{aligned} \therefore r(x, y) &= \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}} \\ &= \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sqrt{\sum_{i=1}^n x_i^2 - n \bar{x}^2} \sqrt{\sum_{i=1}^n y_i^2 - n \bar{y}^2}} \end{aligned}$$

उदाहरणे सोडवण्यासाठी खालील सूत्र देखील वापरता येईल:

$$r(x, y) = \frac{cov(x, y)}{\sigma_x \sigma_y} = \frac{\sum xy - n \bar{x} \bar{y}}{\sqrt{(\sum x^2 - n \bar{x}^2)(\sum y^2 - n \bar{y}^2)}}$$

जर x_i आणि y_i यांचे मूल्ये मोठी असतील, तर आम्ही मूल्यांचे रूपांतरण वापरू शकतो, ज्यासाठी मूल्यांमध्ये स्थान आणि प्रमाण बदल केले जातात.

समजा $u_i = \frac{x_i - a}{h}$ आणि $v_i = \frac{y_i - b}{k}$
 जिथे a आणि b नवीन उगम (Origins) आहेत आणि h आणि k प्रमाण घटक (Scale factors) आहेत,
 म्हणून

$$r_{xy} = r_{uv}$$

$$\therefore r(u, v) = \frac{\text{cov}(u, v)}{\sigma_u \sigma_v}$$

4.2.2 कोरिलेशन कोइफिशियंटचे गुणधर्म:

- $-1 \leq r \leq 1$
- जर, $r = -1$ तर पूर्ण नेगेटिव्ह कोरिलेशन असते.
- जर $r = 1$, तर पूर्ण पॉझिटिव्ह कोरिलेशन असते.
- जर $r = 0$, तर कोणतेही कोरिलेशन नाही.
- जर $-1 \leq r < 0$, , तर नेगेटिव्ह कोरिलेशन असते.
- जर $0 < r \leq 1$, तर पॉझिटिव्ह कोरिलेशन असते.

सोडवलेली उदाहरणे:

1) खालील डेटासाठी कोरिलेशन कोइफिशियंट मोजा:

X (स्टडी टाइम)	1	2	3	4	5
Y (टेस्ट स्कोर)	35	40	45	50	55

उत्तर:

x (स्टडी टाइम)	y (टेस्ट स्कोर)	x^2	y^2	xy
1	35	1	1225	35
2	40	4	1600	80
3	45	9	2025	135
4	50	16	2500	200
5	55	25	3025	275
एकूण =15	225	55	10375	725

$$\bar{x} = \frac{15}{5} = 3, \bar{y} = \frac{225}{5} = 45$$

$$r(x, y) = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{(\sum x^2 - n\bar{x}^2)(\sum y^2 - n\bar{y}^2)}}$$

$$= \frac{725 - 5(3)(45)}{\sqrt{(55 - 5(3)^2)(10375 - 5(45)^2)}}$$

$$= \frac{50}{\sqrt{(10)(250)}}$$

$$r(x, y) = \frac{50}{\sqrt{2500}} = 1$$

$$r(x, y) = 1$$

कोरिलेशन कोइफिशियंट $r = 1$, म्हणजेच दोन व्हेरियेबलस् मध्ये पूर्ण पॉझिटिव्ह कोरिलेशन आहे.

2) खालील डेटासाठी कोरिलेशन कोइफिशियंट मोजा:

आवर्स ऑफ स्लीप x	8	6	7	5	4
नंबर ऑफ आवर्स स्पेंट वर्किंग y	5	6	7	8	9

उत्तर:

x आवर्स ऑफ स्लीप	y नंबर ऑफ आवर्स स्पेंट वर्किंग	x^2	y^2	xy
8	5	64	25	40
6	6	36	36	36
7	7	49	49	49
5	8	25	64	40
4	9	16	81	36
एकूण = 30	35	190	255	201

$$\bar{x} = \frac{30}{5} = 6, \bar{y} = \frac{35}{5} = 7$$

$$r(x, y) = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{(\sum x^2 - n\bar{x}^2)(\sum y^2 - n\bar{y}^2)}}$$

$$= \frac{201 - 5(6)(7)}{\sqrt{(190 - 5(6)^2)(255 - 5(7)^2)}}$$

$$= \frac{-9}{\sqrt{100}}$$

$$r(x, y) = \frac{-9}{10} = -0.9$$

3) स्टॅटिस्टिक्स आणि कम्प्युटर प्रोग्रामिंगमध्ये १० विद्यार्थ्यांनी मिळवलेले मार्क्स दिले आहेत. कोरिलेशन कोइफिशियंट शोधा.

स्टॅटिस्टिक्स मध्ये मिळवलेले मार्क्स (x)	75	30	60	80	53	15	40	38	48	35
कम्प्युटर प्रोग्रामिंगमध्ये मिळवलेले मार्क्स (y)	45	54	91	58	63	35	43	45	44	85

उत्तर:

स्टॅटिस्टिक्स मध्ये मिळवलेले मार्क्स (x)	कंप्युटर प्रोग्रामिंगमध्ये मिळवलेले मार्क्स (y)	x^2	y^2	xy
75	45	5625	2025	3375
30	54	900	2916	1620
60	91	3600	8281	5460
80	58	6400	3364	4640
53	63	2809	3969	3339
15	35	225	1225	525
40	43	1600	1849	1720
38	45	1444	2025	1710
48	44	2304	1936	2112
35	85	1225	7225	2975
एकूण = 474	563	26132	34815	27476

$$\begin{aligned}\bar{x} &= \frac{474}{10} = 47.4, \bar{y} = \frac{563}{10} = 56.3 \\ r(x, y) &= \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{(\sum x^2 - n\bar{x}^2)(\sum y^2 - n\bar{y}^2)}} \\ &= \frac{27476 - 10(47.4)(56.3)}{\sqrt{(26132 - 10(47.4)^2)(34815 - 10(56.3)^2)}} \\ &= \frac{789.8}{\sqrt{3664.4(3118.1)}} \\ r(x, y) &= 0.2337\end{aligned}$$

4) खालील डेटा साठी r मोजा:

x	10	20	30	40	50
y	20	40	50	70	80

उत्तर:

x	y	x^2	y^2	xy
10	20	100	400	200
20	40	400	1600	800
30	50	900	2500	1500
40	70	1600	4900	2800
50	80	2500	6400	4000
एकूण = 150	260	5500	15800	9300

$$\bar{x} = \frac{150}{5} = 30, \bar{y} = \frac{260}{5} = 52$$

$$r(x, y) = \frac{cov(x, y)}{\sigma_x \sigma_y} = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{(\sum x^2 - n\bar{x}^2)(\sum y^2 - n\bar{y}^2)}}$$

$$= \frac{9300 - 5(30)(52)}{\sqrt{(5500 - 5(30)^2)(15800 - 5(52)^2)}}$$

$$r = 0.9$$

सरावासाठी उदाहरणे

1. खालील डेटा वापरून कार्ल पिअर्सनचा कोरिलेशन कोइफिशियंट मोजा.

प्राईस	10	20	30	40	50	60	70
सप्लाय (युनिट्स)	10	20	30	40	50	60	70

2. खालील डेटा वापरून कार्ल पिअर्सनचा कोरिलेशन कोइफिशियंट मोजा:

x	15	18	20	28	34
y	40	42	46	50	52

3. खालील डेटा साठी r मोजा:

x	21	23	30	54	57	58	72	78	87	90
y	60	71	72	83	110	84	100	92	113	135

4.2.3 स्पीअरमनचा रँक कोरिलेशन कोइफिशियंट (Spearman's Rank Co-Efficient of Correlation):

कधी कधी डेटा संख्यात्मक पद्धतीने मोजता येत नाही, परंतु गुणवत्तात्मक मूल्यांकन करणे शक्य असते, जसे की सुंदर, कुरूप, उत्कृष्ट, चांगले, प्रामाणिकपणा इत्यादी. अशा स्थितीत, डेटा त्या विशेष गुणधर्मानुसार क्रमांकित केला जातो, त्याऐवजी त्यावर संख्यात्मक मोजमाप घेतले जात नाही, आणि त्यामुळे सामान्य कोरिलेशन कोइफिशियंट मोजता येत नाही. याऐवजी, चार्ल्स एच. स्पीअरमन यांनी पारंपरिक कोरिलेशन कोइफिशियंटचे

एक अप्रतिम पर्याय विकसित केला, जो खालीलप्रमाणे आहे:

दिलेल्या n जोडलेल्या निरीक्षणांचा संच (x_i, y_i) साठी, जिथे $i = 1$ ते n , x निरीक्षणांना त्याच्या परिमाणानुसार क्रमांक 1, 2, ... n दिले जातात आणि त्याचप्रमाणे y निरीक्षणांना देखील क्रमांक दिले जातात. नंतर या क्रमांकांची वास्तविक संख्यात्मक मूल्यांसाठी बदल केली जातात. या पद्धतीने मोजलेला कोरिलेशन कोइफिशियंट " रँक कोरिलेशन कोइफिशियंट " किंवा " स्पीअरमनचा कोरिलेशन कोइफिशियंट " म्हणून ओळखला जातो आणि तो R ने दर्शविला जातो.

स्पीअरमनचा कोरिलेशन कोइफिशियंट $R(x, y)$ पुढीलप्रमाणे दिला जातो:

$$R(x, y) = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

इथे, d_i हे x_i आणि y_i साठी दिलेल्या रँकसाठी असलेला फरक आहे आणि n म्हणजे डेटा जोड्यांची संख्या.

टीप:

- जर x किंवा y निरीक्षणांमध्ये समान (टाय) असतील, तर त्या प्रत्येक टायसाठी त्यांचे सामायिक रँकचे सरासरी दिली जातात.
- $R(x, y)$ हा गुणांक -1 आणि 1 दरम्यान असतो.

4.2.4 रिपिटेड रँकां मधील कोरिलेशन (Correlation in case of repeated ranks):

जेव्हा दोन किंवा जास्त घटकांची समान (टाय) व्हॅल्यूज असतात, तेव्हा त्यांना वेगळ्या रँक देणे कठीण असते. अशा स्थितीत सरासरी रँक (समान रँक) प्रत्येकाला दिला जातो. यासाठी, कोरिलेशन कोइफिशियंट शोधण्यासाठी पुढील सुधारित सूत्र वापरले जाते:

$$r(x, y) = 1 - \frac{6 \left[\sum d_i^2 + \frac{m_1(m_1^2-1)}{12} + \frac{m_2(m_2^2-1)}{12} + \dots \right]}{n(n^2-1)}$$

इथे, m_1, m_2 हे त्या रिपिटेड घटकांच्या संख्यांचे मान आहेत.

सोडवलेली उदाहरणे:

1) खालील डेटा साठी रँक कोरिलेशन कोइफिशियंट मोजा:

स्टुडेंट	A	B	C	D	E	F	G	H	I	J
मॅथ्स मधील रँक	9	10	6	5	7	2	4	8	1	3
स्टॅटिस्टिक्स मधील रँक	1	2	3	4	5	6	7	8	9	10

उत्तर: इथे रँक दिल्या आहेत आणि $n = 10$

स्टुडेंट	R_1	R_2	$d = R_1 - R_2$	d^2
A	9	1	8	64
B	10	2	8	64
C	6	3	3	9
D	5	4	1	1
E	7	5	2	4
F	2	6	-4	16
G	4	7	-3	9
H	8	8	0	0
I	1	9	-8	64
J	3	10	-7	49
				$\sum d^2 = 280$

$$R = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

$$= 1 - \frac{1680}{990} = -0.697$$

2) खालील डेटा साठी रँक कोरिलेशन कोइफिशियंट मोजा:

x	35	23	47	17	10	43	9	6	28
y	30	33	45	23	8	49	12	4	31

उत्तर:

x	y	x मधील रँक	y मधील रँक	$d = R_1 - R_2$	d^2
35	30	3	5	-2	4
23	33	5	3	2	4
47	45	1	2	-1	1
17	23	6	6	0	0
10	8	7	8	-1	1
43	49	2	1	1	1
9	12	8	7	1	1
6	4	9	9	0	0
28	31	4	4	0	0
					$\sum d^2 = 12$

$$R = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6 \times 12}{9(9^2 - 1)} = 0.9$$

3) खालील डेटा साठी कोरिलेशन कोइफिशियंट मोजा जो मॅथ्समधील दोन क्विझमधील मिळवलेले मार्क्स दर्शवितो:

पहिल्या क्विझमधील मार्क्स x	6	5	8	8	7	6	10	4	9	7
दुसऱ्या क्विझमधील मार्क्स y	8	7	7	10	5	8	10	6	8	6

उत्तर:

x	4	5	6	6	7	7	8	8	9	10
x मधील रँक	1	2	3.5	3.5	5.5	5.5	7.5	7.5	9	10
y	5	6	6	7	7	8	8	8	10	10
y मधील रँक	1	2.5	2.5	4.5	4.5	7	7	7	9.5	9.5

x	y	x मधील रँक	y मधील रँक	$d = R_1 - R_2$	d^2
6	8	3.5	7	-3.5	12.25
5	7	2	4.5	-2.5	6.25
8	7	7.5	4.5	3	9
8	10	7.5	9.5	-2	4
7	5	5.5	1	4.5	20.25
6	8	3.5	7	-3.5	12.25
10	10	10	9.5	0.5	0.25
4	6	1	2.5	-1.5	2.25
9	8	9	7	2	4

7	6	5.5	2.5	3	9
					$\sum d^2 = 79.5$

$$R = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6 \times 79.5}{10(10^2 - 1)} = 0.5182$$

4. दोन विषय X आणि Y मध्ये आठ अर्जदारांचे मार्क्स खालीलप्रमाणे दिले आहेत. **कोरिलेशन कोइफिशियंट** मोजा.

अर्जदार	A	B	C	D	E	F	G	H
विषय X	15	20	28	12	40	60	20	80
विषय Y	40	30	50	30	20	10	30	60

उत्तर:

अर्जदार	विषय X	दिलेला रँक	विषय Y	दिलेला रँक	$(R_1 - R_2)^2 = D^2$
A	15	2	40	6	16
B	20	3.5	30	4	0.25
C	28	5	50	7	4
D	12	1	30	4	9
E	40	6	20	2	16
F	60	7	10	1	36
G	20	3.5	30	4	0.25
H	80	8	60	8	0
					$\sum D^2 = 81.5$

$$R = 1 - \frac{6 \left[\sum d_i^2 + \frac{m_1(m_1^2 - 1)}{12} + \frac{m_2(m_2^2 - 1)}{12} + \dots \right]}{n(n^2 - 1)}$$

येथे विषय X दोनदा 20 मध्ये पुनरावृत्ती करतो म्हणून $m_1 = 2$ आणि विषय Y 30 मध्ये 3 वेळा पुनरावृत्ती करतो म्हणून $m_2 = 2$

$$\therefore R = 1 - \frac{6 \left(81.5 + \frac{1}{12}(2^3 - 2) + \frac{1}{12}(3^3 - 3) \right)}{(8^3 - 8)}$$

$$= 0$$

दोन विषयांमध्ये मिळवलेले मार्क्स यामध्ये कोणताही **कोरिलेशन** नाही.

- 5) सौंदर्य स्पर्धेतील दहा स्पर्धकांना तीन न्यायाधीशांनी खालील प्रमाणे रँक दिले आहेत

पहिले न्यायाधीश	1	4	6	3	2	9	7	8	10	5
दुसरे न्यायाधीश	2	6	5	4	7	10	9	3	8	1
तिसरे न्यायाधीश	3	7	4	5	10	8	9	2	6	1

कोरिलेशन कोइफिशियंटची पद्धत वापरून, कोणत्या न्यायाधीशांच्या जोड्या सौंदर्याच्या सामान्य आवडाच्या जवळपास पोहोचल्या आहेत हे ठरवा.

उत्तर: समजा, R_x , R_y , R_z हे तीन न्यायाधीशांनी दिलेले रँक दर्शवितात.

R_x	R_y	R_z	$d_{xy}=R_x-R_y$	$d_{yz}=R_y-R_z$	$d_{zx}=R_z-R_x$	d_{xy}^2	d_{yz}^2	d_{zx}^2
1	2	3	-1	-1	2	1	1	4
4	6	7	-2	-1	3	4	1	9
6	5	4	1	1	-2	1	1	4
3	4	5	-1	-1	2	1	1	4
2	7	10	-5	-3	8	25	9	64
9	10	8	-1	2	-1	1	4	1
7	9	9	-2	0	2	4	0	4
8	3	2	5	1	-6	25	1	36
10	8	6	2	2	-4	4	4	16
5	1	1	4	0	-4	16	0	16
						$\sum d_{xy}^2=82$	$\sum d_{yz}^2=22$	$\sum d_{zx}^2=158$

$$R_{xy} = 1 - \frac{6\sum d_{xy}^2}{n(n^2-1)} = 1 - \frac{6(82)}{10(10^2-1)} = 0.5031$$

$$R_{yz} = 1 - \frac{6\sum d_{yz}^2}{n(n^2-1)} = 1 - \frac{6(22)}{10(10^2-1)} = 0.8667$$

$$R_{zx} = 1 - \frac{6\sum d_{zx}^2}{n(n^2-1)} = 1 - \frac{6(158)}{10(10^2-1)} = 0.0424$$

R_{yz} सर्वात जास्त आहे, म्हणून दुसऱ्या आणि तिसऱ्या न्यायाधीशांच्या जोड्या सौंदर्याच्या सामान्य आवडाच्या जवळपास पोहोचल्या आहेत.

सरावासाठी उदाहरणे

1. दोन न्यायाधीशांनी सौंदर्य स्पर्धेत ११ मुलींना खालील रँक दिले आहेत

मुली	1	2	3	4	5	6	7	8	9	10	11
न्यायाधीश A	3	4	1	2	5	10	11	7	9	8	6
न्यायाधीश B	2	4	3	1	7	9	6	11	10	5	8

रँक कोरिलेशन कोइफिशियंट मोजा.

2. खालील डेटा साठी रँक कोरिलेशन कोइफिशियंट मोजा:

x	48	33	40	9	16	16	65	24	16	57
y	13	13	24	6	15	4	20	9	6	19

3. खालील डेटा वापरून स्पीअरमनचा रँक कोरिलेशन कोइफिशियंट मोजा:

x	68	64	75	50	64	80	75	40	55	64
y	62	58	68	45	81	60	68	48	50	70

युनिट 5

रिग्रेशन एनालिसिस

(Regression Analysis)

विषय निष्पत्ती (Course Outcome):

CO5 : डेटाचा सहसंबंध निश्चित करण्यासाठी प्रतिगमन विश्लेषण लागू करा.

घटक निष्पत्ती (Theory Learning Outcome):

1. प्रतिगमन विश्लेषणाच्या संकल्पना स्पष्ट करा.
2. सहसंबंध सहगुणांक आणि प्रतिगमन सहगुणांक वापरून दोन चलांमधील संबंध शोधा.
3. डेटामधून सर्वोत्तम फिटिंग वक्र किंवा सरळ रेषा शोधण्यासाठी कमीत कमी चौरसांची पद्धत लागू करा.
4. एका चलाचे दुसऱ्या चलावरील प्रतिगमन समीकरण शोधा.

5.1 नॉन व्हेरियन्स (Non-Variance)

नॉन-व्हेरियन्स म्हणजे अशी परिस्थिती जिथे विशिष्ट वैशिष्ट्य किंवा डेटा सेटमध्ये कोणताही फरक किंवा फरक नसतो. दुसऱ्या शब्दांत, ते अशा परिस्थितीचे वर्णन करते ज्यामध्ये सर्व मूल्ये (व्हॅल्यू) स्थिर, एकसमान किंवा एकसारखी असतात. ते वेगवेगळ्या संदर्भांमध्ये लागू होऊ शकते,

1. सांख्यिकी (स्टॅटिस्टिक्स): जेव्हा डेटा पॉइंट्समध्ये कोणतेही परिवर्तनशीलता किंवा प्रसार नसतो, तेव्हा भिन्नता शून्य असते.

2. सामान्य वापर: जेव्हा गोष्टी सुसंगत किंवा अपरिवर्तित राहतात, कामगिरी, वर्तन किंवा परिणामात कोणतेही भिन्नता दर्शवत नाहीत.

• **शून्य भिन्नता (झिरो व्हेरियन्स) (Zero Variance):** जेव्हा डेटा सेटमधील सर्व मूल्ये (व्हॅल्यू) समान असतात, तेव्हा भिन्नता शून्य असते.

• **मोठे भिन्नता (लार्ज व्हेरियन्स) (Large Variance)** जेव्हा डेटा मूल्ये (व्हॅल्यू) सरासरीपासून पसरवली जातात, तेव्हा भिन्नता मोठी असते.

• **लहान भिन्नता (स्मॉल व्हेरियन्स) (Small Variance):** जेव्हा डेटा मूल्ये (व्हॅल्यू) सरासरीच्या जवळ असतात, तेव्हा भिन्नता लहान असते.

5.1.1 नॉन-व्हेरियन्सची काही सोपी उदाहरणे खालीलप्रमाणे आहेत.

उदाहरण: 5,5,5,5,5 या संख्या असलेला डेटा सेट. सर्व मूल्ये समान असल्याने फरक शून्य आहे. मूल्यांमध्ये कोणताही प्रसार किंवा फरक नाही.

उदाहरण: सर्हर सातत्याने विनंत्या अगदी त्याच वेळेत प्रक्रिया करतो (उदाहरण, विनंतीवर प्रक्रिया करण्यासाठी नेहमीच 3 सेकंद लागतात), प्रक्रियेच्या गतीमध्ये कोणताही फरक दर्शवत नाही.

उदाहरण: मानक वातावरणीय दाबाखाली शुद्ध पाण्याचा गोठणबिंदू सातत्याने 0°C असतो. जर तुम्ही नियंत्रित परिस्थितीत पाणी मोजत असाल, तर त्याचा गोठणबिंदू बदलणार नाही आणि तो स्थिर राहील.

उदाहरण: एका विशिष्ट अल्गोरिथम चालवणाऱ्या आणि समान इनपुट दिल्यावर (उदाहरण 2+2 जोडण्यासारखे गणितीय गणना) नेहमीच समान परिणाम देणाऱ्या प्रोग्राममध्ये, निकालात भिन्नता नसते ते नेहमीच 4 परत करेल.

उदाहरण: एक बास्केटबॉल खेळाडू प्रत्येक सामन्यात नेहमीच 20 गुण मिळवतो, प्रतिस्पर्धी किंवा परिस्थिती काहीही असो. कामगिरीतील ही सातत्यता फरक नसल्याचे दर्शवते.

उदाहरण: एक वाळवंट प्रदेश जिथे दररोज दुपारी तापमान नेहमीच 35°C असते, कोणताही चढ-उतार किंवा फरक नसतो, दिवसाच्या त्या विशिष्ट वेळी तेथे फरक नसतो.

उदाहरण: एखादा ग्राहक दुकानात जाताना दरवेळी एकच वस्तू (उदाहरण विशिष्ट ब्रँडची कॉफी) खरेदी करतो. त्याच्या खरेदीच्या वर्तनात कोणताही फरक नसतो.

उदाहरण: एक फूड कंपनी एक उत्पादन बनवते (जसे की चॉकलेट बार), आणि प्रत्येक बार अगदी सारखाच आकार, वजन आणि चवीचा असतो आणि बॅचमध्ये कोणताही फरक नसतो. उत्पादनाच्या वैशिष्ट्यांमध्ये कोणताही फरक नाही.

उदाहरण: एक डिजिटल घड्याळ त्याच्या आउटपुटमध्ये प्रत्येक सेकंदाला समान अचूक वेळ दाखवते (उदाहरण, त्रुटी किंवा चढउतारांशिवाय तास किंवा दिवसांसाठी 12:00:00 दाखवते). दाखवलेल्या वेळेत कोणताही फरक नाही.

उदाहरण: जर तुम्ही अशा खोलीत असाल जिथे वातानुकूलन यंत्रणा 22°C तापमान राखण्यासाठी सेट केलेली असेल आणि ती चढ-उतार न होता पूर्णपणे स्थिर राहते, तर त्या खोलीतील तापमानात कोणताही फरक नसतो.

उदाहरण: अशी कंपनी जी दर आठवड्याला एका निश्चित वेळापत्रकानुसार नेहमीच समान प्रमाणात कच्चा माल प्राप्त करते, कोणत्याही बदलाशिवाय किंवा विलंब न करता. पुरवठा साखळी कामगिरीमध्ये फरक नसलेला दिसून येतो.

या सर्व उदाहरणांमध्ये, नॉन-व्हिएरियन्सचे परिभाषित वैशिष्ट्य म्हणजे कालांतराने किंवा वेगवेगळ्या निरीक्षणांमध्ये कोणतेही विचलन, चढ-उतार किंवा फरक नाही.

5.1.2 डेटाचे मानक विचलन: (स्टॅंडर्ड डेविएशन) (Standard Deviation)

मानक विचलन (स्टॅंडर्ड डेविएशन) ही सांख्यिकीमधील एक मूलभूत संकल्पना आहे जी डेटा पॉइंट्सच्या फैलावचे मोजमाप करते जे डेटासेटमधील डेटा पॉइंट्स सरासरीपासून किती प्रमाणात विचलित होतात हे परिभाषित करते, डेटामधील परिवर्तनशीलता किंवा प्रसाराची स्पष्ट जाणीव प्रदान करते.

मानक विचलन (स्टॅंडर्ड डेविएशन) सूत्राचा वापर डेटा मूल्याचे सरासरी मूल्यापासून विचलन शोधण्यासाठी केला जातो म्हणजेच डेटा सेटमधील सर्व मूल्यांचे सरासरी मूल्यापर्यंतचे विखुरणे शोधण्यासाठी केला जातो. यादृच्छिक चलाच्या मानक विचलनाची (स्टॅंडर्ड डेविएशनची) गणना करण्यासाठी वेगवेगळे मानक विचलन सूत्र आहेत.

मानक विचलन (स्टॅंडर्ड डेविएशन) म्हणजे डेटा पॉइंटच्या सरासरी मूल्यापासून डेटा पॉइंटच्या फैलावची डिग्री. ते आपल्याला सांगते की डेटा पॉइंटचे मूल्य डेटा पॉइंटच्या सरासरी मूल्यापेक्षा कसे बदलते आणि ते आपल्याला डेटाच्या नमुन्यातील डेटा पॉइंटच्या फरकाबद्दल सांगते.

दिलेल्या डेटा सेटच्या नमुन्याचे मानक विचलन हे डेटा सेटच्या भिन्नतेचे वर्गमूळ म्हणून देखील परिभाषित केले जाते. n मूल्यांचे सरासरी विचलन (म्हणजे $x_1, x_2, \dots, x_n$) प्रत्येक मूल्याच्या फरकाच्या वर्गाची बेरीज सरासरीपासून घेऊन मोजले जाते.

$$\text{मीन डेविएशन} = (\text{Mean Deviation}) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

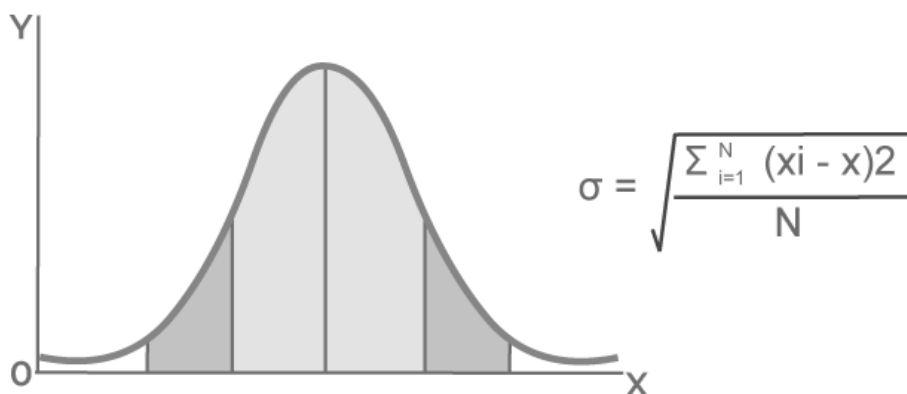


Figure 5.1: स्टॅंडर्ड डेविएशन (Standard deviation)

गटबद्ध नसलेल्या डेटाचे मानक विचलन

गटबद्ध नसलेल्या डेटासाठी, मानक विचलन तीन पद्धती वापरून मोजले जाऊ शकते,

- वास्तविक सरासरी पद्धत (एक्चुअल मीन मेथड)

- गृहीत सरासरी पद्धत (एज्युम मीन मेथड)
- चरण विचलन पद्धत (स्टेप डेविएशन मेथड)

a. वास्तविक सरासरी पद्धतीने मानक विचलन (स्टँडर्ड डेव्हिएशन बाय एक्चुअल मीन मेथड) (Standard Deviation by actual mean method)

वास्तविक सरासरी पद्धतीने मानक विचलन (स्टँडर्ड डेव्हिएशन बाय एक्चुअल मीन मेथड) दिलेल्या डेटाचा सरासरी काढण्यासाठी मूलभूत सरासरी सूत्र वापरते आणि या सरासरी मूल्याचा वापर करून आपण दिलेल्या डेटा मूल्यांचे मानक विचलन शोधतो. या पद्धतीत आपण सूत्र वापरून सरासरी काढतो,

मीन Mean $\mu = (\text{निरीक्षणांची बेरीज})/(\text{निरीक्षणांची संख्या})$

(वास्तविक सरासरी पद्धतीसाठी मानक विचलन सूत्र) स्टँडर्ड डेव्हिएशन बाय एक्चुअल मीन मेथड

$$\sigma = \sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n}}$$

उदाहरण: डेटा सेटचे स्टँडर्ड डेव्हिएशन शोधा $X = \{2, 3, 4, 5, 6\}$

उत्तर: दिलेले, $n = 5$ $x_i = \{2, 3, 4, 5, 6\}$

आपल्याला माहित आहे,

मीन Mean (μ) = (निरीक्षणांची बेरीज)/(निरीक्षणांची संख्या)

$$\mu = (2 + 3 + 4 + 5 + 6) / 5$$

$$\mu = 4$$

मानक विचलन सूत्र वापरणे (स्टँडर्ड डेव्हिएशन वापरणे)

$$\sigma = \sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n}}$$

$$\sigma^2 = 1/n[(2 - 4)^2 + (3 - 4)^2 + (4 - 4)^2 + (5 - 4)^2 + (6 - 4)^2]$$

$$\sigma^2 = 10/5 = 2$$

$$\text{अशा प्रकारे, } \sigma = \sqrt{2} = 1.414$$

b. स्टँडर्ड डेव्हिएशन बाय एज्युम मीन मेथड (Standard Deviation by Assume mean method)

x च्या खूप मोठ्या मूल्यांसाठी गटबद्ध डेटाचा मध्य शोधणे हे एक कंटाळवाणे काम आहे म्हणून आपण गृहीत धरलेली सरासरी पद्धत वापरतो जिथे आपण एक अनियंत्रित मूल्य (A) सरासरी मूल्य म्हणून गृहीत धरतो आणि नंतर सामान्य पद्धतीचा वापर करून मानक विचलनची (स्टँडर्ड डेव्हिएशनची) गणना करतो. समजा n डेटा मूल्यांच्या गटासाठी ($x_1, x_2, x_3, \dots, x_n$), गृहीत धरलेला मध्य A आहे तर विचलन (डेव्हिएशन) आहे,

$$d_i = x_i - A$$

जिथे,

x_i = डेटा मूल्ये A = गृहीत सरासरी

एज्युम मीन मेथडसाठीचे स्टँडर्ड डेव्हिएशन सूत्र

$$\sigma = \frac{\sum f_i d_i^2}{n}$$

तर,

' n ' = एकूण डेटा मूल्यांची संख्या

$$d_i = x_i - A$$

c. स्टॅंडर्ड डेव्हिएशन बाय स्टेप डेव्हिएशन मेथड (Standard Deviation by step deviation method)

आपण स्टेप डेव्हिएशन पद्धतीचा वापर करून गटबद्ध डेटाचे मानक विचलन (स्टॅंडर्ड डेव्हिएशन) देखील काढू शकतो. वरील पद्धतीप्रमाणे या पद्धतीमध्ये देखील, आपण गृहीत सरासरी म्हणून काही अनियंत्रित डेटा मूल्य निवडतो (म्हणजे A). नंतर आपण सर्व डेटा मूल्यांचे विचलन (डेव्हिएशन) मोजतो.

$$(x_1, x_2, x_3, \dots, x_n), \quad d_i = x_i - A, \quad N = \sum f$$

$$\sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2}$$

गटबद्ध वारंवारता वितरण स्वरूपात सादर केलेल्या डेटाचा विचार करताना आणि गणना आणखी कमी करण्यासाठी, आपण $u = \frac{x-A}{h}$ ठेवतो, जिथे h सामान्यतः वर्ग मध्यांतराची रुंदी म्हणून घेतले जाते आणि A सामान्यतः x पासून मध्य मूल्य असते.

$$\sigma = h * \sqrt{\frac{\sum f_i u^2}{N} - \left(\frac{\sum f_i u}{N}\right)^2}$$

d. डीस्क्रेट ग्रुप डेटाचे स्टॅंडर्ड डेव्हिएशन (Standard Deviation of Discrete grouped data)

गटबद्ध डेटामध्ये, प्रथम आम्ही एक वारंवारता सारणी बनवली आणि नंतर पुढील गणना केली गेली. स्वतंत्र गटबद्ध डेटासाठी, मानक विचलन (स्टॅंडर्ड डेव्हिएशन) तीन पद्धती वापरून देखील मोजले जाऊ शकते जे आहेत,

- वास्तविक सरासरी पद्धत (एक्चुअल मीन मेथड) (Actual Mean Method)
- गृहीत सरासरी पद्धत (एज्युम मीन मेथड) (Assumed Mean Method)
- चरण विचलन पद्धत (स्टेप डेव्हिएशन मेथड) (Step Deviation Method)

डीस्क्रेट वारंवारता वितरणावर आधारित स्टॅंडर्ड डेव्हिएशनचे सूत्र (Standard Deviation formula based on Discrete frequency Distribution)

दिलेल्या डेटा सेटसाठी जर त्याची मूल्ये n ($x_1, x_2, x_3, \dots, x_n$) असतील आणि त्यांच्याशी संबंधित वारंवारता ($f_1, f_2, f_3, \dots, f_n$) असेल तर त्याचे मानक विचलन (स्टॅंडर्ड डेव्हिएशन) सूत्र

$$\sigma = \sqrt{\frac{\sum_{i=1}^n f_i (x_i - \bar{x})^2}{n}}$$

तर,

- n म्हणजे एकूण वारंवारता ($n = f_1 + f_2 + f_3 + \dots + f_n$)
- $\bar{x}$ म्हणजे मीन
- x_i डेटा पॉइंटचे मूल्य
- f_i डेटा पॉइंटची वारंवारता

e. कंटीन्यूअस ग्रुप डेटाचे स्टॅंडर्ड डेव्हिएशन (Standard deviation of Continuous grouped data)

कंटीन्यूअस डेटासाठी आपण प्रत्येक वर्गाला त्याच्या मध्यबिंदूने (x_i म्हणून) बदलून आणि नंतर सामान्यपणे सूत्रांची गणना करून स्वतंत्र डेटा सूत्रांचा वापर करून मानक विचलन (स्टॅंडर्ड डेव्हिएशन) सहजपणे मोजू शकतो.

स्टॅंडर्ड डेव्हिएशनचे उदाहरण: (Examples on Standard deviations)

उदाहरण 1: स्टॅंडर्ड डेव्हिएशन शोधा,

x_i	5	12	15
f_i	2	3	3

उत्तर : प्रथम खालीलप्रमाणे तक्ता बनवा, जेणेकरून आपण पुढील मूल्ये सहजपणे काढू शकू.

x_i	f_i	$x_i f_i$	$x_i - \mu$	$(x_i - \mu)^2$	$f(x_i - \mu)^2$
5	2	10	-6.372	40.64	81.28
12	3	36	0.625	0.39	1.17
15	3	45	3.625	13.14	39.42
Total	8	91	-	-	121.87

$$\text{मीन Mean } (\mu) = \frac{\sum(f_i x_i)}{\sum(f_i)}$$

$$\text{मीन Mean } (\mu) = 91/8 = 11.375$$

स्टॅंडर्ड डेव्हिएशनचे सूत्रांचा वापर करून

$$\sigma = \sqrt{\frac{\sum f_i (x_i - \mu)^2}{n}}$$

$$\sigma = \sqrt{\frac{121.87}{8}}$$

$$\sigma = 3.90$$

$$\text{स्टॅंडर्ड डेव्हिएशन (Standard Derivation)}(\sigma) = 3.90$$

उदाहरण 2: खालील वितरणातून मानक विचलनाची (स्टॅंडर्ड डेव्हिएशनची) गणना करा.

क्लास	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65
फ्रिक्वेन्सी	25	35	50	90	75	60	35	25	15

उत्तर:

आपल्याकडे $h = 5$ आहे आणि आपण खालील तक्ता तयार करू

श्रेणी	मध्यम मूल्य x	वारंवारता f	$u = \frac{x - A}{h}$	u^2	fu	fu^2
20-25	22.5	25	-3	9	-75	225
25-30	27.5	35	-2	4	-70	140
30-35	32.5	50	-1	1	-50	50
35-40	37.5	90	0	0	0	0
40-45	42.5	75	1	1	75	75
45-50	47.5	60	2	4	120	240
50-55	52.5	35	3	9	105	315
55-60	57.5	25	4	16	100	400
60-65	62.5	15	5	25	75	375
Total	-	$\sum f = 410$			$\sum fu = 280$	$\sum fu^2 = 1820$

येथे $A = 37.5$

$$\bar{u} = \frac{\sum fu}{\sum f} = \frac{280}{410} = 0.68$$

मीन (Mean), $\bar{x} = A + h\bar{u} = 37.5 + 5(0.68) = 40.9$

$$\bar{x} = 40.9$$

$$\text{स्टॅंडर्ड डेव्हिएशन (Standard deviation)} = \sigma = h * \sqrt{\frac{\sum f_i u^2}{N} - \left(\frac{\sum f_i u}{N}\right)^2} = 5 \sqrt{\frac{1820}{410} - (0.68)^2}$$

$$\sigma = 5 \sqrt{4.43 - 0.4624} = 9.9495$$

स्टॅंडर्ड डेव्हिएशन (Standard deviation) = $\sigma = 9.9495$

सरावासाठी उदाहरणे

1] डेटाचे स्टॅंडर्ड डेव्हिएशन शोधा.

x	12	13	14	15	16
f	2	5	3	6	2

2] डेटाचे स्टॅंडर्ड डेव्हिएशन शोधा.

वीकली वेजेस (Rs.)	4.5-12.5	12.5-20.5	20.5-28.5	28.5-36.5	36.5-44.5
नंबर ऑफ मेन	4	24	21	18	5

5.1.3 कोवेरियन्स (Covariance)

कोवेरियन्स (Covariance) हे एक सांख्यिकीय माप आहे जे दोन यादृच्छिक चलांमधील रेखीय संबंधांच्या दिशेचे वर्णन करते. एका व्हेरिएबलमध्ये वाढ झाल्यामुळे दुसऱ्या व्हेरिएबलमध्ये वाढ किंवा घट होईल की नाही हे सूचित करते.

गणितीयदृष्ट्या, X आणि Y या दोन चलांमधील सहविभाजन खालीलप्रमाणे परिभाषित केले आहे:

$$\text{कोवेरियन्स } Cov(X, Y) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$$

- X_i and Y_i i-th निरीक्षणासाठी X आणि Y व्हेरिएबल्सची मूल्ये आहेत.
- $\bar{X}$ आणि $\bar{Y}$ अनुक्रमे X आणि Y चे मध्य (मीन)(means) आहेत.
- n ही डेटा बिंदूंची संख्या आहे.

महत्वाचे मुद्दे:

- **पॉझिटिव्ह कोवेरियन्स (Positive Covariance):** सहप्रसरण सकारात्मक असल्यास, हे सूचित करते की दोन्ही चल एकत्र वाढतात किंवा कमी होतात.
- **निगेटिव्ह कोवेरियन्स (Negative Covariance):** जर सहप्रसरण ऋणात्मक असेल, तर ते सूचित करते की जसजसे एक व्हेरिएबल वाढते, तसतसे दुसरे कमी होते.
- **शून्य कोवेरियन्स (Zero Covariance):** शून्य सहप्रसरण दर्शविते की व्हेरिएबल्समध्ये कोणताही रेखीय संबंध नाही.

तथापि, सहविभाजन मूल्यांचा अर्थ लावणे कठिण असू शकते कारण ते व्हेरिएबल्सच्या स्केलवर अवलंबून असतात. म्हणूनच सहसंबंध गुणांक बहुतेकदा वापरला जातो, कारण तो सहप्रवाह सामान्य करतो, नातेसंबंधाची ताकद आणि दिशा स्पष्ट करतो.

5.1.4 कोइफिशियंट ऑफ कोरिलेशन: (Coefficient of Correlation):

दोन व्हेरिएबल्समधील रेखीय संबंधांची तीव्रता किंवा डिग्री मोजण्यासाठी, कार्ल पियर्सनने सहसंबंध गुणांक नावाचे सूत्र विकसित केले.

$r(x, y)$ ने दर्शविलेल्या x आणि y या दोन चलांमधील सहसंबंध गुणांक म्हणून परिभाषित केले आहे

$$r(x, y) = \frac{cov(x, y)}{\sigma_x \sigma_y}$$

बायवेरिएट डिस्ट्रिब्युशन (x_i, y_i) ही मूल्ये $(x_1, y_1) (x_2, y_2) \dots (x_n, y_n)$ घेतल्यास

$$Cov(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$\bar{x}$ आणि $\bar{y}$ अनुक्रमे X आणि Y चे मध्य आहेत. त्याचप्रमाणे

$$\sigma_x = \sqrt{\sum_{i=1}^n \frac{1}{n} (x_i - \bar{x})^2} \quad \text{आणि} \quad \sigma_y = \sqrt{\sum_{i=1}^n \frac{1}{n} (y_i - \bar{y})^2}$$

जे x आणि y मालिकेसाठी स्टॅंडर्ड डेव्हिएशन आहेत. आपल्याकडे खालीलप्रमाणे साधे फॉर्म आहे,

$$Cov(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i y_i - \bar{x} \bar{y})$$

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 \quad \text{त्याचप्रमाणे} \quad \sigma_y^2 = \frac{1}{n} \sum_{i=1}^n y_i^2 - \bar{y}^2$$

$$r(x, y) = \frac{cov(x, y)}{\sigma_x \sigma_y} \quad \text{नंतर गणना केली जाऊ शकते.}$$

आपण वापरल्यास

$$u_i = X_i - A \text{ or } u_i = \frac{X_i - A}{h} \quad \text{आणि} \quad v_i = Y_i - B \text{ or } v_i = \frac{Y_i - B}{k}$$

$$Cov(u, v) = \frac{1}{n} \sum_{i=1}^n (u_i v_i - \bar{u} \bar{v})$$

$$\sigma_u^2 = \frac{1}{n} \sum_{i=1}^n u_i^2 - \bar{u}^2 \quad \text{त्याचप्रमाणे} \quad \sigma_v^2 = \frac{1}{n} \sum_{i=1}^n v_i^2 - \bar{v}^2$$

$$r(u, v) \text{ द्वारे दिले जाते } r(u, v) = \frac{cov(u, v)}{\sigma_u \sigma_v}$$

उदाहरणे सोडवण्यासाठी आपण खालील सूत्र देखील वापरू शकतो.

$$r(x, y) = \frac{cov(x, y)}{\sigma_x \sigma_y} = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{(\sum x^2 - n\bar{x}^2)(\sum y^2 - n\bar{y}^2)}}$$

काही प्रमुख मुद्दे:

- 1] हे ओळखले जाऊ शकते की $r(x, y) = r(u, v)$. $r(x, y)$ च्या तुलनेत $r(u, v)$ ची गणना करणे सोपे आहे.
- 2] सहसंबंध गुणांक 'r' नेहमी -1 आणि 1 दरम्यान असतो म्हणजेच $-1 \leq r \leq 1$.
- 3] जर $r > 0$ असेल तर सहसंबंध धनात्मक असेल आणि जर $r < 0$ असेल तर सहसंबंध ऋणात्मक असेल. जर $r = 0$, तर आपण असे म्हणतो की कॅरिएबल्स असंबंधित आहेत.
- 4] जर $r = 1$, आम्ही म्हणतो की परिपूर्ण सकारात्मक सहसंबंध आहे; जर $r = -1$ असेल तर, आपण म्हणतो की परिपूर्ण नकारात्मक सहसंबंध आहे.
- 5] आम्ही हे देखील लक्षात घेतो की, सहप्रवेश हा देखील ऑर्डरचा संयुक्त केंद्रीय क्षण (Joint Central Moment) मानला जाऊ शकतो (1, 1) पैकी (X, Y). म्हणून, आपण $\mu_{11} = cov(x, y)$ दर्शवतो.

उदाहरण 1: पॉझिटिव्ह कोव्हेरियन्स

अभ्यासाचे तास आणि परीक्षेतील गुण यांच्यातील संबंध विचारात घेऊया.

X (लर्निंग अवर्स)	Y (मार्क्स)	X ²	Y ²	XY
1	40	1	1600	40
2	45	4	2025	90
3	50	9	2500	150
4	55	16	3025	220
5	60	25	3600	300
Total=15	250	55	12750	800

$$\begin{aligned}
\bar{X} &= \frac{15}{5} = 3, \bar{Y} = \frac{250}{5} = 50 \\
r(X, Y) &= \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y} = \frac{\sum XY - n\bar{X}\bar{Y}}{\sqrt{(\sum X^2 - n\bar{X}^2)(\sum Y^2 - n\bar{Y}^2)}} \\
&= \frac{800 - 5(3)(50)}{\sqrt{(55 - 5(3)^2)(12750 - 5(50)^2)}} \\
&= \frac{50}{\sqrt{(10)(250)}} \\
r(X, Y) &= \frac{50}{\sqrt{2500}} = 1 \\
r(X, Y) &= 1
\end{aligned}$$

विश्लेषण: $r=1$, म्हणजे सरावासाठी उदाहरणलेले तास आणि चाचणी गुणांमध्ये एक परिपूर्ण पॉझिटिव्ह कोव्हेरियन्स सहसंबंध आहे सरावासाठी उदाहरणलेले तास वाढत असताना, चाचणी गुण पूर्णपणे रेषीय पद्धतीने वाढतात.

उदाहरण 1: निगेटिव्ह कोव्हेरियन्स

खालील माहितीवरून कोइफिशन्ट ऑफ कोरिलेशन काढा.

$$n = 10, \sum x = 40, \sum y = 40, \sum xy = 150, \sum x^2 = 190, \sum y^2 = 200$$

$$\bar{x} = \frac{40}{10} = 4, \bar{y} = \frac{40}{10} = 4$$

उत्तर:

$$\begin{aligned}
r(x, y) &= \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{(\sum x^2 - n\bar{x}^2)(\sum y^2 - n\bar{y}^2)}} \\
&= \frac{150 - 10(4)(4)}{\sqrt{(190 - 10(4)^2)(200 - 10(4)^2)}} \\
&= \frac{150 - 160}{\sqrt{(30)(40)}} \\
&= \frac{-10}{34.6410}
\end{aligned}$$

$$r(x, y) = -0.2850$$

विश्लेषण: $r = -1$, म्हणजे एक परिपूर्ण पॉझिटिव्ह कोव्हेरियन्स आहे.

5.1.5 कोइफिशन्ट ऑफ रिग्रेशन (Coefficient of regressions)

रेग्रेशनचा सहगुणक म्हणजे साध्या रेषीय रेग्रेशन मॉडेलमधील रेग्रेशन रेषेचा स्लोप. तो स्वतंत्र चल (X) आणि अवलंबून असलेल्या चल (Y) यांच्यातील संबंध दर्शवतो. साध्या रेषीय रेग्रेशनमध्ये, Y आणि X यांच्यातील संबंध पुढीलप्रमाणे नमूद केला जातो:

$$Y = b_0 + b_y X$$

जिथे:

- b_0 हा रेग्रेशन रेषेचा छेदबिंदू आहे (जेव्हा $X = 0$ असते तेव्हा Y चे मूल्य).
- b_y हा रेग्रेशन सहगुणक (स्लोप) आहे, जो दर्शवतो की X मध्ये एका युनिटने बदल झाल्यास Y किती बदलतो.

रेग्रेशन कोइफिशन्टचे सूत्र b_y :

रेग्रेशन कोइफिशन्टचे सूत्र b_y (स्लोप)

$$b_y = \frac{\sum(X_i - \bar{X})(Y - \bar{Y})}{\sum(X_i - \bar{X})^2}$$

जिथे:

- X_i आणि Y_i हे स्वतंत्र आणि अवलंबित चलांचे वैयक्तिक डेटा बिंदू आहेत.
- $\bar{X}$ आणि $\bar{Y}$ हे अनुक्रमे X आणि Y चे सरासरी मूल्य आहेत.

इंटरप्रेटिंग द रेग्रेशन कोइफिशियंट : (Interpreting the Regression Coefficient)

- $b_y > 0$: X आणि Y यांच्यात सकारात्मक संबंध आहे. X वाढल्यास, Y देखील वाढतो.
- $b_y < 0$: X आणि Y यांच्यात नकारात्मक संबंध आहे. X वाढल्यास, Y कमी होतो.
- $b_y = 0$: X आणि Y यांच्यात कोणताही रेषीय संबंध नाही.

5.2 लिस्ट स्क्वेअर मेथड (किमान चौरसांची पद्धत) (Method of Least Squares):

एखाद्या प्रयोगाच्या परिणामस्वरूप, समजा, चलांचे मूल्ये (x_i, y_i) नोंदवली गेली आहेत, जिथे $i = 1, 2, 3, \dots, n$. हे बिंदू आलेखावर प्लॉट केल्यास, सहसा असे दिसून येते की बहुतेक बिंदूंना एक गुळगुळीत वक्र जातो, तर काही बिंदू या वक्रापासून थोडेसे दूर असतात. ही वक्र पहिल्या श्रेणीची (पहिल्या अंशाची) असू शकते, म्हणजे सरळ रेषा $y = ax + b$, किंवा दुसऱ्या श्रेणीची पॅराबोला जसे की $y = ax^2 + bx + c$, किंवा सर्वसाधारणपणे n व्या श्रेणीची वक्र असू शकते.

सामान्य स्वरूप:

$$y = a_0x^n + a_1x^{n-1} + a_2x^{n-2} \dots \dots + a_n$$

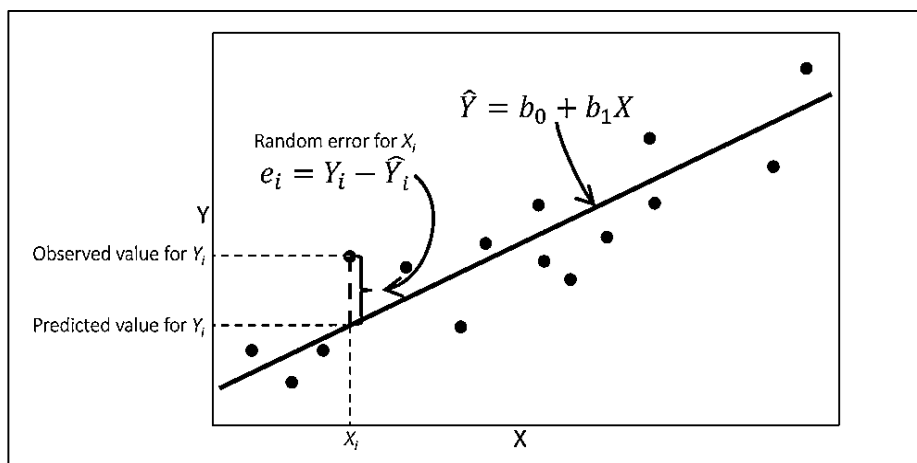


Figure 5.2: लिस्ट स्क्वेअर मेथड (Method of Least Squares)

बिंदूंच्या संचामधून जवळजवळ जाणारी वक्र समीकरण निर्धारित करण्यासाठी, आपण x आणि y यांच्यातील काही विशिष्ट संबंध गृहित धरतो. हा संबंध सरळ रेषा, दुसऱ्या श्रेणीची पॅराबोला, तिसऱ्या श्रेणीची वक्र किंवा त्याहून अधिक श्रेणीची वक्र असू शकतो. आपण अशा स्वरूपाची समीकरणे गृहित धरतो जी दिलेल्या बिंदूंना सर्वोत्तम तंदुरुस्त (best fit) ठरतील. आकृती (आकृती 5.2) मध्ये, आपण पाहतो की सरळ रेषा दिलेल्या बिंदूंच्या संचामधून जवळजवळ जात आहे. म्हणूनच, आपण सरळ रेषेचे समीकरण $y = ax + b$ (1) असे गृहित धरू शकतो. जर बिंदू (x_i, y_i) समीकरण (1) वर असल्याचे गृहित धरले तर, त्या बिंदूचा y सहनिर्देश (co-ordinate) खालीलप्रमाणे गणना करता येतो:

$$y'_i = ax + b$$

जर बिंदू खरोखर (1) वर असेल तर,

$$y_i = y'_i$$

अन्यथा $y_i - y'_i$ हे सूत्र (1) वापरून मोजलेल्या y'_i मूल्यापासून निरीक्षण केलेल्या मूल्य y_i चे विचलन दर्शविले.

लहान वर्गाच्या पद्धतीमध्ये आपण या विचलनांच्या वर्गांची बेरीज घेतो आणि कमाल किंवा किमान तत्वाचा वापर करून ही बेरीज कमी करतो. या निकषाचा वापर करून $a, b, \dots$ (1) ची मूल्ये मोजली जातात. याला लघुतम वर्ग निकष म्हणतात. वक्र (1) हा किमान वर्ग निकष वापरून कोणत्याही अंशाचा असू शकतो, आपण त्याचे समीकरण शोधू शकतो. पुढील भागात आपण दिलेल्या बिंदूंच्या संचाशी सरळ रेषा आणि दुसऱ्या अंशाचा पॅराबोला बसवण्याबद्दल चर्चा करू.

1. फिटिंग स्ट्रेट लाईन (Fitting of straight Line)

$(x_i, y_i); i = 1, 2, 3, \dots, n$ ही (x, y) ची निरीक्षण केलेली मूल्ये (व्हॅल्यू) मानू. सरळ रेषेत बसण्यासाठी, $y = ax + b$ हे किमान वर्ग निकष वापरून $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ हे निरीक्षण केलेली मूल्ये आहेत.

$$y'_1 = ax_1 + b, y'_2 = ax_2 + b, y'_3 = ax_3 + b, \dots, y'_n = ax_n + b$$

y निर्देशांकांची गणना केलेली मूल्ये (व्हॅल्यू) या गृहीतकानुसार आहेत की बिंदू $y = ax + b$ या सरळ रेषेवर आहेत. उदाहरणे सोडवण्यासाठी आपण खालील सूत्र वापरत आहोत.

$$a \sum x + nb = \sum y \dots \dots \dots (1)$$

$$a \sum x^2 + b \sum x = \sum xy \dots \dots \dots (2)$$

(1) आणि (2) सोडवून, आपण a आणि b ची मूल्ये (व्हॅल्यू) ठरवतो ज्यामुळे दिलेल्या डेटासाठी सर्वात योग्य असलेली सरळ रेषा $y = ax + b$ मिळते. कधीकधी गणनांचे आकारमान कमी करण्यासाठी आपण मूळ स्थान बदलतो. $x = X + h$ किंवा $X = x - h$ घेतो. जिथे, h हे सामान्यतः x च्या निरीक्षण केलेल्या मूल्यांच्या संचातील मध्यवर्ती स्थित मूल्य असते. आपल्याला $y = aX + b$ ही सरळ रेषा सापडते आणि नंतर X ची जागा $x - h$ ने घेते.

2. फिटिंग सेकंड डिग्री पॅराबोला (Fitting of Parabola)

आता आपण फॉर्मचा पॅराबोला बसवण्याची पद्धत स्पष्ट करू

$$y = ax^2 + bx + c$$

$(x_i, y_i); i = 1, 2, 3, \dots, n$ ही (x, y) ची निरीक्षण केलेली मूल्ये (व्हॅल्यू) मानू.

पूर्वप्रमाणे, बिंदू पॅराबोलाच्या समीकरणाची पूर्तता करतात या गृहीतकानुसार y निर्देशांकांची गणना केलेली व्हॅल्यू खालीलप्रमाणे दिली आहेत

$$y'_1 = ax_1^2 + bx_1 + c, y'_2 = ax_2^2 + bx_2 + c, \dots, y'_n = ax_n^2 + bx_n + c \dots \dots (1)$$

$$a \sum x^2 + b \sum x + nc = \sum y \dots \dots \dots (2)$$

$$a \sum x^3 + b \sum x^2 + c \sum x = \sum xy \dots \dots \dots (3)$$

$$a \sum x^4 + b \sum x^3 + c \sum x^2 = \sum x^2 y \dots \dots \dots (4)$$

समीकरण (2), (3), (4) ही तीन अज्ञात a, b, c मध्ये एकाच वेळी तीन समीकरणे आहेत. ही समीकरणे सोडवताना आपण a, b, c निश्चित करतो, जो दिलेल्या डेटासाठी सर्वोत्तम फिट पॅराबोला देतो:

पूर्वप्रमाणे, गणना कमी करण्यासाठी, आपण उत्पत्ती बदलतो $x = X + h$ किंवा $X = x - h$

$$a \sum X^2 + b \sum X + nc = \sum y \dots \dots \dots (5)$$

$$a \sum X^3 + b \sum X^2 + c \sum X = \sum Xy \dots \dots \dots (6)$$

$$a \sum X^4 + b \sum X^3 + c \sum X^2 = \sum X^2 y \dots \dots \dots (7)$$

(5), (6) आणि (7) सोडवून, आपण a, b, c शोधतो आणि पॅराबोला निश्चित करतो.

$$y = aX^2 + bX + c$$

मग आपण X ला $x - h$ ने बदलू आणि x आणि y च्या संदर्भात पॅराबोला मिळवू.

5.2.1 एक्सपोनेन्शियल रिग्रेशन (Exponential Regression)

घातांकीय प्रतिगमन (एक्सपोनेन्शियल रिग्रेशन) हा एक रेषीय प्रतिगमन प्रकार आहे. घातांकीय प्रतिगमन दोन प्रकारे व्यक्त केले जाऊ शकते. चला दोन्ही प्रकारच्या घातांकीय प्रतिगमन (एक्सपोनेन्शियल रिग्रेशन) प्रकारांची उदाहरणासह तपशीलवार चर्चा करूया. घातांकीय प्रतिगमन (एक्सपोनेन्शियल रिग्रेशन) वित्त, जीवशास्त्र,

भौतिकशास्त्र इत्यादी क्षेत्रात वापरले जाऊ शकते. घातांकीय प्रतिगमनासाठी गणितीय अभिव्यक्ती उदाहरणासह पाहू.

$$y = ae^{bx}$$

• जिथे y हा अवलंबित चल आहे

• x हा स्वतंत्र चल आहे

• a, b हे प्रतिगमन सहगुणक (रिग्रेशन कोइफिशियंट) आहेत.

घातांकीय वक्र (एक्सपोनेन्शियल कर्व) बसवताना, वरील समीकरणाच्या दोन्ही बाजूंना "ln" (पाया e सह लॉगरिदम) लागू करून आपण वरील समीकरण सरळ रेषेच्या रेषा अंतर्विच्छेद स्वरूपात (साधे रेषीय प्रतिगमन) रूपांतरित करू शकतो. $y = ae^{bx}$ दोन्ही बाजूंना ln लावल्याने आपल्याला मिळते

दोन्ही बाजूंना ln लावल्याने आपल्याला मिळते:

$$\ln(y) = \ln(ae^{bx}) \Rightarrow \ln(y) = \ln(a) + \ln(e^{bx})$$

$$\ln(y) = \ln(a) + bx$$

आपण वरील समीकरणाची तुलना $Y = A + BX$ शी करू शकतो.

जिथे $Y = \ln(y), A = \ln(a), B = b, x = X, a = e^A$ and $b = B$

सामान्य समीकरणे असतील

$$\sum Y = nA + B \sum X$$

$$\sum XY = A \sum X + B \sum X^2$$

उदाहरण: लिस्ट स्केअर मेथड वापरून स्ट्रेट लाईन काढा.

x	5	10	15	20	25
y	16	19	23	26	30

उत्तर : सरळ रेषा म्हणजे $y = ax + b$.

सामान्य समीकरणे सूत्राने दिली आहेत.

$$a \sum x + nb = \sum y \dots\dots\dots(1)$$

$$a \sum x^2 + b \sum x = \sum xy \dots\dots\dots(2)$$

आता आपल्याला खालील सारणी वापरून $\sum x, \sum x^2, \sum y, \sum xy$ ची मूल्ये (व्हॅल्यू) सापडतील,

आम्ही खालीलप्रमाणे टेबल तयार करू

x	y	x^2	xy
5	16	25	80
10	19	100	190
15	23	225	345
20	26	400	520
25	30	625	750
$\sum x = 75$	$\sum y = 114$	$\sum x^2 = 1375$	$\sum xy = 1885$

ही सर्व मूल्ये (व्हॅल्यू) समीकरण (1) आणि (2) मध्ये ठेवल्यास आपल्याला खालील सामान्य समीकरण मिळते,

$$75a + 5b = 114 \dots\dots\dots(3)$$

$$1375a + 75b = 1885 \dots\dots\dots(4)$$

दोन अज्ञातांसाठी ही दोन समीकरणे सोडवल्याने आपल्याला मिळते,

$$a = 0.7, b = 12.3,$$

मग सर्वोत्तम बसणारे सरळ रेषीय समीकरण वापरून, आपल्याकडे आहे $y = 0.7x + 12.3$

उदाहरण: लिस्ट स्केअर मेथड वापरून सेकंड डिग्री पॅराबोला चे इक्वेशन फाईंड करा

x	-3	-2	-1	0	1	2	3
y	12	4	1	2	7	15	30

उत्तर : पॅराबोला चे इक्वेशन कन्सिडर करून

$$y = ax^2 + bx + c \dots\dots(1)$$

$$a \sum x^2 + b \sum x + nc = \sum y \dots\dots\dots(2)$$

$$a \sum x^3 + b \sum x^2 + c \sum x = \sum xy \dots\dots\dots(3)$$

$$a \sum x^4 + b \sum x^3 + c \sum x^2 = \sum x^2y \dots\dots\dots(4)$$

आता आपल्याला खालील सारणी वापरून $\sum x, \sum x^2, \sum x^3, \sum x^4, \sum y, \sum xy, \sum x^2y$

आपण खालीलप्रमाणे टेबल तयार करू

x	y	x^2	x^3	x^4	xy	x^2y
-3	12	9	-27	81	-36	108
-2	4	4	-8	16	-8	16
-1	1	1	-1	1	-1	1
0	2	0	0	0	0	0
1	7	1	1	1	7	7
2	15	4	8	16	30	60
3	30	9	27	81	90	270
$\sum x = 0$	$\sum y = 71$	$\sum x^2 = 28$	$\sum x^3 = 0$	$\sum x^4 = 196$	$\sum xy = 82$	$\sum x^2y = 462$

ही सर्व मूल्ये समीकरण (2), (3), (4) मध्ये ठेवल्यास, आपल्याला मिळते

$$196a + 0b + 28c = 462 \dots (5)$$

$$0a + 28b + 0c = 82 \dots (6)$$

$$28a + 0b + 7c = 71 \dots (7)$$

तीन अज्ञातांसाठी ही तीन समीकरणे सोडवल्यास, आपल्याला मिळते

त्यामुळे $a = 2.12, b = 2.92, c = 1.66$.

उदाहरण: खालील दिलेल्या टेबल वरून एक्सपोनेन्शियल फॉर्म $y = ae^{bx}$

X	1	5	7	9	12
Y	10	15	12	15	21

उत्तर : वरील समीकरणांवरून आपल्याला $X=x, Y=\ln(y)$ माहित आहे.

x	y	X	$Y = \ln(y)$	XY	X^2
1	10	1	2.302	2.302	1
5	15	5	2.708	13.54	25
7	12	7	2.484	17.388	49
9	15	9	2.708	24.372	81
12	21	12	3.044	36.528	144

वरील तक्त्यावरून $n = 5, \sum X = 34, \sum Y = 13.246, \sum XY = 94.13, \sum X^2 = 300$

आता सामान्य समीकरणे होतात.

$$13.246 = 5A + 34B$$

$$94.13 = 34A + 300B$$

वरील समीकरण सोडवून आपण A आणि B ची किंमत मिळवू शकतो.

$$A = 2.248 \text{ and } B = 0.059$$

आपल्याला माहित असलेल्या उल्लेखित समीकरणांवरून

$$b = B \text{ and } a = e^A$$

$$a = e^{2.248} = 9.468$$

$$b = B = 0.059$$

एक्सपोनेन्शियल फॉर्म $y = ae^{bx}$

$$y = 9.468 e^{0.059x}$$

सरावासाठी उदाहरणे

1] लिस्ट स्केअर मेथड वापरून स्ट्रेट लाईन काढा.

x	-2	-1	0	1	2
y	1	2	3	3	4

2] लिस्ट स्केअर मेथड वापरून सेकंड डिग्री पॅराबोला चे इक्वेशन फाईंड करा

x	1	2	3	4	5	6	7
y	-5	-2	5	16	31	50	73

3] $y = ax^b$ चे फॉर्म एक्सप्रेस करा

x	2000	3000	4000	5000	6000
y	15	15.5	16	17	18

5.3 रिग्रेशन लाइन (Line of Regression):

रीग्रेशनची रेषा ही एक सांख्यिकीय संकल्पना आहे जी दोन चलांमधील संबंध मॉडेल करण्यासाठी वापरली जाते. हे स्कॅटर प्लॉटमधील डेटा पॉइंट्सच्या संचाद्वारे सर्वोत्कृष्ट-फिट रेषेचे प्रतिनिधित्व करते, निरीक्षण केलेली मूल्ये आणि रेषेवर आधारित अंदाजित मूल्यांमधील वर्ग फरक (त्रुटी) ची बेरीज कमी करते. सर्वात सामान्यपणे वापरली जाणारी प्रतिगमन रेषा ही कमीत कमी चौरस प्रतिगमन रेषा आहे.

साध्या रेखीय प्रतिगमनमध्ये, रेषा समीकरणाद्वारे परिभाषित केली जाते:

$$y = mx + b$$

जिथे

- y हा अवलंबित चल आहे (तुम्ही जे भाकित करण्याचा किंवा स्पष्ट करण्याचा प्रयत्न करत आहात),
- x हा स्वतंत्र चल आहे (तुम्ही भाकित करण्यासाठी वापरत असलेले चल),
- m हा रेषेचा उतार आहे (जो x मध्ये एक-युनिट बदलासाठी y किती बदलतो हे दर्शवितो),
- b हा y -इंटरसेप्ट आहे ($x=0$ असताना y चे मूल्य).
- ही रेषा निरीक्षण केलेल्या डेटा बिंदूंमधील आणि रेषेद्वारे भाकित केलेल्या मूल्यांमधील एकूण वर्ग फरक कमी करते, ज्याला बहुतेकदा किमान वर्ग त्रुटी (लिस्ट स्केअर एरर) म्हणून संबोधले जाते.

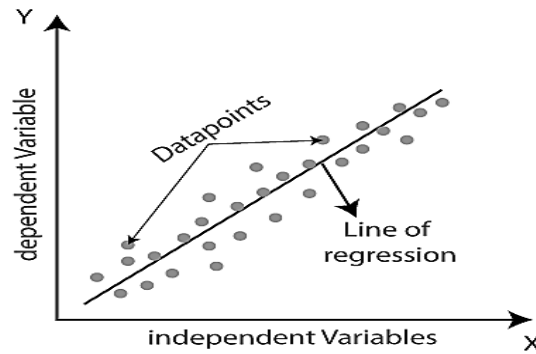


Figure 5.3: लाईन ऑफ रिग्रेशन (Line of Regression)

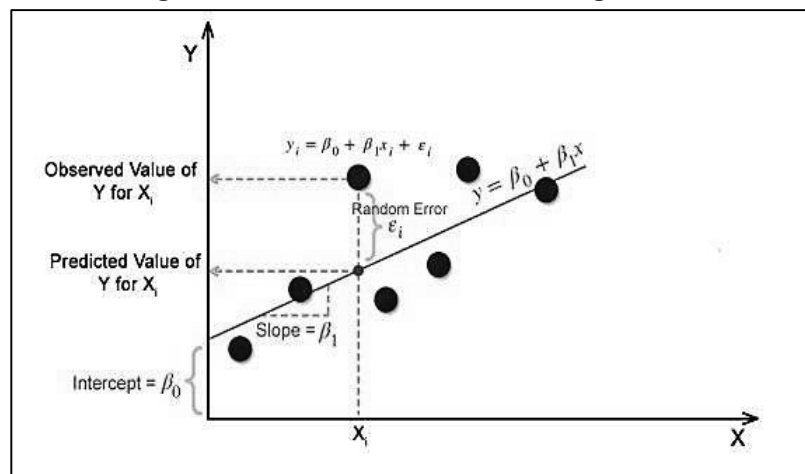


Figure 5.4: लाईन ऑफ रिग्रेशन (Line of Regression)

उदाहरणार्थ, जर आपल्याला माहित असेल की पावसाचा पीक उत्पादनावर परिणाम होतो, तर पावसाळी हंगामाच्या शेवटी पीक उत्पादनाचा अंदाज लावणे शक्य आहे. द्विवेरिएट डिस्ट्रिब्युशनमधील व्हेरिएबल्स संबंधित असल्यास, स्कॅटर डायग्राम क्लस्टरमधील बिंदू काही वक्र गोल करतात ज्याला प्रतिगमन वक्र किंवा प्रतिगमन वक्र म्हणतात. जर वक्र सरळ रेषा असेल तर तिला प्रतिगमन रेषा म्हणतात आणि अशा स्थितीत दोन चलांमधील प्रतिगमन रेषीय असते. रीग्रेशनची रेषा दुसऱ्या व्हेरिएबलच्या काही निर्दिष्ट मूल्यासाठी एका व्हेरिएबलच्या मूल्यासाठी सर्वोत्तम अंदाज देते. जर सहसंबंध परिपूर्ण नसेल (म्हणजे $r \neq \pm 1$), तर दिलेल्या बिंदूमधून अनेक रेषा काढता येतील. सर्वोत्कृष्ट तंदुरुस्त रेषा असल्याने, कमीत कमी वर्गाची पद्धत वापरून प्रतिगमन रेषा मिळविली जाते.

5.3.1 इक्वेशन ऑफ लाईन ऑफ रिग्रेशन (Equation of regression line)

x वरील y ची रिग्रेशन लाईन (The regression line of y on x is)

$$y - \bar{y} = r \frac{\sigma_y}{\sigma_x} (x - \bar{x}) = b_{yx} (x - \bar{x}) \dots \dots \dots (1)$$

y वरील x ची रिग्रेशन लाईन (The regression line of x on y is)

$$(x - \bar{x}) = r \frac{\sigma_x}{\sigma_y} (y - \bar{y}) = b_{xy} (y - \bar{y}) \dots \dots \dots (2)$$

वरील दोन समीकरणे (1) आणि (2) वरून आपल्याला मिळते,

समीकरण (1) मध्ये समाविष्ट असलेला सहगुणक b_{yx} हा x वर y चा रिग्रेशन कोइफिसिएंट म्हणून ओळखला जातो आणि समीकरण (2) मध्ये समाविष्ट असलेला सहगुणक b_{xy} हा y वर x चा रिग्रेशन कोइफिशिएंट म्हणून ओळखला जातो.

टीप: 1] (1) आणि (2) शोधण्यासाठी आपल्याला कोरिलेशन कोइफिशिएंट $r = r(x, y)$ मोजावा लागेल, जो मूळ बदल आणि प्रमाण गुणधर्म वापरून देखील निश्चित केला जाऊ शकतो.

अशा प्रकारे, $r(x, y) = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\text{cov}(u, v)}{\sigma_u \sigma_v} = r(u, v)$

$u = \frac{x-A}{h}$, $v = \frac{y-B}{k}$ नंतर

$\sigma_x = h\sigma_u$ आणि $\sigma_y = k\sigma_v$

$\sigma_u^2 = \frac{1}{n} \sum u_i^2 - \bar{u}^2$ आणि $\sigma_v^2 = \frac{1}{n} \sum v_i^2 - \bar{v}^2$

$\bar{x} = A + h\bar{u}$ आणि $\bar{y} = B + k\bar{v}$

जिथे A आणि B अनुक्रमे x आणि y पासून मध्य मूल्ये आहेत.

2] विशिष्ट मध्ये $u = x - A$, $v = y - B$ आणि $h = k = 1$ also $\sigma_x = \sigma_u$ आणि $\sigma_y = \sigma_v$

आणि $\bar{x} = A + \bar{u}$ आणि $\bar{y} = B + \bar{v}$ जिथे A आणि B अनुक्रमे x आणि y पासून मध्य मूल्ये (Mean Value) आहेत. वरील निकाल समीकरण (1) आणि (2) चे समाधान शोधण्यास मदत करतो.

3] कोरिलेशन कोइफिशियंट आणि रिग्रेशन कोइफिशियंट बीजगणितीय चिन्हे समान असतात. जर $r > 0$ असेल, तर $b_{yx} > 0$ आणि $b_{xy} > 0$. जर $r < 0$ असेल, तर आणि $b_{yx} < 0$ आणि $b_{xy} < 0$.

4] $b_{yx} b_{xy} = r^2$ कोरिलेशन कोइफिशियंट $= r = b_{yx} b_{xy}$ रिग्रेशन कोइफिशियंट म्हणजे भौमितिक मध्य.जर रिग्रेशन कोइफिशियंट धन असतील तर धन वर्गमूळ निवडा, अन्यथा ऋण.

5] रिग्रेशन रेषांमधील तीव्र कोन θ द्वारे दिलेला आहे,

$$r = \tan^{-1} \left\{ \frac{1 - r^2}{|r|} \cdot \frac{\sigma_x \sigma_y}{\sigma_x^2 \sigma_y^2} \right\}$$

6] दोन प्रतिगमन रेषेचा छेदनबिंदू $(\bar{x}, \bar{y})$ आहे.

उदाहरणे सोडवण्यासाठी आपण सूत्र खालीलप्रमाणे करतो.

पायरी 1] आपल्याला अनुक्रमे x आणि y ची सरासरी मूल्ये मिळतात.

$\bar{x} = \frac{\sum x}{n}$ आणि $\bar{y} = \frac{\sum y}{n}$ जिथे n ही निरीक्षणांची एकूण संख्या आहे.

पायरी 2] आपण खालील गोष्टी फाईंड करू

$$\sigma_x^2 = \frac{\sum x^2}{n} - \bar{x}^2 \text{ आणि } \sigma_y^2 = \frac{\sum y^2}{n} - \bar{y}^2$$

पायरी 3] आपण खालील गोष्टी फाईंड करू

$$\text{cov}(x, y) = \frac{\sum xy}{n} - \bar{x}\bar{y}$$

पायरी 4] $\text{cov}(x, y)$ नंतर आपल्याला सूत्रे वापरून b_{yx} b_{xy} हा रिग्रेशन कोइफिशियंट सापडतो.

$$b_{yx} = x \text{ वर } y \text{ चा रिग्रेशन कोइफिशियंट} = \frac{\text{cov}(x, y)}{\sigma_x^2}$$

$$b_{xy} = y \text{ वर } x \text{ चा रिग्रेशन कोइफिशियंट} = \frac{\text{cov}(x, y)}{\sigma_y^2}$$

पायरी 5] x वरील y ची रिग्रेशन लाईन

$$y - \bar{y} = b_{yx} (x - \bar{x})$$

पायरी 6] y वरील x ची रिग्रेशन लाईन

$$(x - \bar{x}) = b_{xy} (y - \bar{y})$$

लिनियर रिग्रेशनची काही उदाहरणे आहेत,

- व्यवसाय विश्लेषण
- शरीराची उंची आणि वजन विश्लेषण, इ.

5.3.2 ॲप्लिकेशन रिग्रेशन लाईनची (Application of Regression Line)

1. **भाकित विश्लेषण (Predictive Analysis):** भूतकाळातील डेटाच्या आधारे भविष्यातील मूल्यांचा अंदाज लावण्यासाठी वापरले जाते.

2. **ट्रेंड विश्लेषण (Trend Analysis):** कालांतराने ट्रेंड ओळखण्यास आणि त्यांचे विश्लेषण करण्यास मदत करते.

3. **कोरिलेशन एनालिसिस (Correlation Analysis):** चलांमधील संबंधांची ताकद आणि दिशा निश्चित करते.

4. **जोखीम व्यवस्थापन (Risk Management):** वित्त आणि आरोग्यसेवा यासारख्या विविध क्षेत्रांमध्ये जोखीमांचे मूल्यांकन आणि व्यवस्थापन करण्यात मदत करते.

उदाहरण: रिग्रेशन लाईन काढा.

x	2	3	5	7	9	10	12	15
y	2	5	8	10	12	14	15	16

उत्तर : आपण खालीलप्रमाणे टेबल तयार करू.

x	y	x^2	y^2	xy
2	2	4	4	4
3	5	9	25	15
5	8	25	64	40
7	10	49	100	70
9	12	81	144	108
10	14	100	196	140
12	15	144	225	180
15	16	225	256	240
$\sum x = 63$	$\sum y = 82$	$\sum x^2 = 637$	$\sum y^2 = 1014$	$\sum xy = 797$

जिथे n ही निरीक्षणांची एकूण संख्या आहे = 8

$$\bar{x} = \frac{\sum x}{n} = \frac{63}{8} = 7.875 \quad \text{आणि} \quad \bar{y} = \frac{\sum y}{n} = \frac{82}{8} = 10.25$$

$$\sigma_x^2 = \frac{\sum x^2}{n} - \bar{x}^2 = \frac{637}{8} - (7.875)^2 = 17.6094 \quad \text{आणि}$$

$$\sigma_y^2 = \frac{\sum y^2}{n} - \bar{y}^2 = \frac{1014}{8} - (10.25)^2 = 21.6875$$

$$\text{cov}(x, y) = \frac{\sum xy}{n} - \bar{x}\bar{y} = \frac{797}{8} - (7.875)(10.25) = 18.9063$$

$$b_{yx} = x \text{ वर } y \text{ चा रिग्रेशन कोइफिसिएंट} = \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{18.9063}{17.6094} = 1.0736$$

$$b_{xy} = y \text{ वर } x \text{ चा रिग्रेशन कोइफिसिएंट} = \frac{\text{cov}(x, y)}{\sigma_y^2} = \frac{18.9063}{21.6875} = 0.8718$$

x वरील y ची रिग्रेशन लाईन

$$(y - \bar{y}) = b_{yx} (x - \bar{x})$$

$$y - 10.25 = 1.0736 (x - 7.875)$$

$$y = 1.0736 X + 1.7954$$

y वरील x ची रिग्रेशन लाईन

$$(x - \bar{x}) = b_{xy} (y - \bar{y})$$

$$x - 7.875 = 0.8718 (y - 10.25)$$

$$x = 0.8718 y - 1.06095$$

उदाहरण: $x - 4y = 5$ आणि $x - 16y = -64$ या दिलेल्या रिग्रेशन लाईन्स आहेत.

y वर x चा रिग्रेशन कोइफिसिएंट आणि x वर y चा रिग्रेशन कोइफिसिएंट, $\bar{x}, \bar{y}$ काढा

उत्तर : येथे, y वर x चे कोणते प्रतिगमन (रिग्रेशन) समीकरण आहे आणि x वर y चे कोणते आहे हे समीकरणांकडे फक्त डोकावून पाहणे अशक्य आहे. आपण x वर y ची प्रतिगमन रेषा (रिग्रेशन लाईन्स) होण्यासाठी एक रेषा यादृच्छिकपणे निवडून प्रतिगमन सहगुणक (रिग्रेशन कोइफिसिएंट) निश्चित करतो. पुढे, ही मूल्ये स्वीकार्य आहेत का ते तपासतो.

समजा $x - 16y = -64$ हे समीकरण y वर x ची प्रतिगमन रेषा (रिग्रेशन लाईन्स) दर्शवते.

हे समीकरण $x = 16y - 64$ $b_{xy} = 16$... असे लिहिता येते..... (1)

पुढे, $x - 4y = 5$ हे समीकरण x वर y ची प्रतिगमन रेषा (रिग्रेशन लाईन्स) असेल असे समजा.

समीकरण असे लिहिता येईल $y = \frac{1}{4}x - \frac{5}{4}$ $b_{yx} = \frac{1}{4}$ (2)

(1) आणि (2) पासून, आपल्याकडे आहे

$$r^2 = b_{yx} b_{xy} = \frac{1}{4} (16) = 4 > 1$$

परिणामी, आपणचुकीच्या प्रतिगमन रेषा (रिग्रेशन लाईन्स) निवडल्या.

आता, आपण पर्यायाची देवाणघेवाण करू. समजा $x - 16y = -64$ ही y वर x ही प्रतिगमन रेषा (रिग्रेशन लाईन्स) आहे. समीकरण असे लिहिता येईल.

$$y = \frac{1}{16}x + 4 \quad \therefore b_{yx} = \frac{1}{16} \text{ (3)}$$

आणि, y वर $x - 4y = 5$ ही प्रतिगमन रेषा (रिग्रेशन लाईन्स) x म्हणून समजा.

समीकरण असे लिहिले जाऊ शकते $x = 4y + 5$ $\therefore b_{xy} = 4$ (4)

(3) आणि (4) पासून, आपल्याकडे आहे

$$r^2 = b_{yx} b_{xy} = \frac{1}{16} (4) = \frac{1}{4} < 1$$

वरील समीकरण (3) आणि (4) वरून

$b_{yx} = \frac{1}{16}$ and $b_{xy} = 4$ योग्य प्रतिगमन सहगुणक (रिग्रेशन कोइफिसिएंट) आहेत.

$$r^2 = b_{yx} b_{xy} = \frac{1}{16} (4) = \frac{1}{4} < 1$$

$$r^2 = \frac{1}{4}$$

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

आपण धन वर्गमूल निवडतो कारण प्रतिगमन सहगुणक धन असतात. आपण हे देखील लक्षात घेतो की $(\bar{x}, \bar{y})$ हा प्रतिगमन रेषांच्या छेदनबिंदूमुळे होतो. अशा प्रकारे, $(\bar{x}, \bar{y})$ दोन्ही समीकरणे पूर्ण करेल.

$$\bar{x} - 4\bar{y} = 5 \quad \text{आणि} \quad \bar{x} - 16\bar{y} = -64$$

वरील समीकरण सोडवून आपल्याला मिळते $\bar{x} = 28$ आणि $\bar{y} = \frac{23}{4}$

उदाहरण: $x = 19.13 - 0.87 y$ आणि $y = 11.64 - 0.50 x$ या दोन दिलेल्या रिग्रेशन लाईन्स आहेत , काढा

1) x चा मध्य आणि y चा मध्य

2) कोरिलेशन कोइफिशन्ट x आणि y

उत्तर : 1) x चा मध्य आणि y चा मध्य प्रतिगमन रेषांवर (रिग्रेशन लाईन्स) असल्याने, आपल्याकडे आहे

$$\bar{x} = 19.13 - 0.87 \bar{y} \quad \text{किंवा} \quad \bar{x} + 0.87 \bar{y} = 19.13$$

$$\bar{y} = 11.64 - 0.50 \bar{x} \quad \text{किंवा} \quad 0.50 \bar{x} + \bar{y} = 11.64$$

आता, आपण $\bar{x}$ आणि $\bar{y}$ ची किंमत शोधण्यासाठी वरील समीकरण सोडवतो, आपल्याला मिळते

$$\bar{x} = 15.935 \text{ आणि } \bar{y} = 3.67$$

2) x वरील y ची प्रतिगमन रेषा (रिग्रेशन लाईन्स) अशी मानावी

$$y = 11.64 - 0.50 x$$

$$b_{yx} = -0.50 \text{ कोइफिशिएंट } x$$

$$x = 19.13 - 0.87 y$$

$$b_{xy} = -0.87 \text{ कोइफिशिएंट } y$$

आपल्याकडे यासाठी सूत्र आहे

$$r = \pm \sqrt{b_{yx} \cdot b_{xy}}$$

$$r = \pm \sqrt{(-0.50)(-0.87)}$$

$$r = -0.66$$

कारण r ऋण आहे कारण b_{yx} आणि b_{xy} दोन्ही ऋण आहेत.

सरावासाठी उदाहरणे

1. खालील माहिती दिल्यास

	Variable x	Variable y
मीन	8.2	12.4
स्टँडर्ड डेविएशन	6.2	20

x आणि y मधील सहसंबंधाचा गुणांक (कोइफिशन्ट कोरिलेशन) 0.9 आहे. $y = 10$ दिल्यास x चा (रेषीय प्रतिगमन) लीनियर रिग्रेशन अंदाज काढा.

2. खालील डेटासाठी रिग्रेशन लाईन्स शोधा.

x	6	2	10	4	8
y	9	11	5	8	7

3. दोन रिग्रेशन लाईन्स समीकरण आहे

$$8x - 10y + 66 = 0 \text{ आणि } 40x - 18y = 214$$

शोधा

1) x चा मध्य आणि y चा मध्य

2) कोरिलेशन कोइफिशन्ट x आणि y

4. खालील डेटावरून x आणि y ची रिग्रेशन लाईन्स शोधा आणि $y = 5$ असताना x ची किंमत शोधा:

$$n = 25, \sum x = 125, \sum y = 100, \sum x^2 = 1650, \sum y^2 = 1500, \sum xy = 50$$

5. दोन यादृच्छिक चलांमध्ये प्रतिगमन रेषा $3x + 2y - 26 = 0$ आणि $6x + y - 31 = 0$ आहेत. कोइफिशन्ट कोरिलेशन x आणि y ची सरासरी मूल्ये शोधा. जर x चा प्रचलन 25 असेल, तर दिलेल्या डेटामधून y चे स्टँडर्ड डेविएशन शोधा.

References:

1. Higher Engineering of Mathematics by H.K. Dass, Er. Rajnish Verma [S.Chand Publication, New Delhi, 2008 ISBN: 9788121938902]
2. Fundamental of Mathematical Statistics by S.C.Gupta & V.K.Kapoor [Sultan Chand & sons Publication, New Delhi, 2008 ISBN: 978-8180545283]
3. Higher Engineering Mathematics by Grewal B. S. [Khanna publication, New Delhi, 2013. ISBN: 978-8174091955]
4. Operation Research by Er. Prem Kumar Gupta &Dr. D.S.Hira [S.Chand Publication, New Delhi,2008 ISBN:978-8121902816]
5. Probability & Statistics for Engineers by Dr. J. Ravichandran [Wiley India Pvt Ltd, ISBN: 978-8126523504]
6. Project Planning and Control with PERT and CPM by Dr B.C Punmia and K.KKhandelwal [Laxmi publications, ISBN: 81-7008-309-5]

HEAD OFFICE

Secretary,
Maharashtra State Board of Technical Education
49, Kherwadi, Bandra (East), Mumbai - 400 051
Maharashtra (INDIA)
Tel: (022)26471255 (5 -lines)
Fax: 022 - 26473980
Email: -secretary@msbte.com

Web -www.msbte.org.in

REGIONAL OFFICES:**MUMBAI**

Deputy Secretary (T),
Mumbai Sub-region,
2nd Floor, Govt. Polytechnic Building,
49, Kherwadi, Bandra (East)
Mumbai - 400 051
Phone: 022-26473253 / 54
Email: rbtemumbai@msbte.com

PUNE

Deputy Secretary (T),
M.S. Board of Technical Education,
Regional Office,
412-E, Bahirat Patil Chowk,
Shivaji Nagar, Pune
Phone: 020-25656994 / 25660319
Fax: 020-25656994
Email: rbtepn@msbte.com

NAGPUR

Deputy Secretary (T),
M.S. Board of Technical Education
Regional Office,
Mangalwari Bazar, Sadar, Nagpur - 440 001
Phone: 0712-2564836 / 2562223
Fax: 0712-2560350
Email: rbteeng@msbte.com

AURANGABAD

Deputy Secretary (T),
M.S. Board of Technical Education,
Regional Office,
Osmanpura, Aurangabad -431 001.
Phone: 0240-2334025 / 2331273
Fax: 0240-2349669
Email: rbteau@msbte.com