Introduction to Probability
Chapter 1: Introduction

Dr. Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

1. Motivation
2. Introduction to the concepts of probability
3. Assigning probabilities to events
4. Continuity theorem in probability

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.
4. https://medium.com/@jrodthoughts/statistical-learning-in-artificial-intelligence-systems-e68927792175 by Jesus Rodriguez (2017)

Motivation

In artificial intelligence (AI) environment, uncertainty is a key element. Due to uncertainty the AI agent does not know the precise outcome of the given situation. Uncertainty is the typical result of random/probabilistic or partially observable environment. Statistical learning is helpful in these AI situations.

For example Bayes' theorem helps in dealing with uncertainty in the real world:

$$P(cause|effect) = P(effect)xP(effect|cause)/P(cause),$$

where $P(A|B)$ is the probability of occurrence of A given B. Replacing cause and effect with the probabilities of any state-action combination in an AI environment we arrive to the fundamentals of Bayesian learning. Many AI algorithms are based on Bayesian learning or statistical learning.

Motivation

In reliability computation of r-out-of n system the probability concepts are used where the components are assumed to have random life. r-out-of n system is a system which functions if atleast r out of its n components functions. Series and parallel system are n-out of n system and 1-out of n system, respectively.

Consider a situation where a redundant component or spare is provided to the system to increase its reliability. Then using probability concepts we can find increase in the reliability of the system.

Introduction

- Random Experiment (E): is an experiment whose outcome may not be perdicted in advance.
- Sample Space (Ω): Collection of all possible outcomes of random experiment.

Example

If E_1: Toss a coin, then $\Omega_1 = \{H, T\}$.
If E_2: Toss a coin till we get a head, then $\Omega_2 = \{H, TH, TTH, \ldots\}$.
If E_3: Lifetime of a bulb, then $\Omega_3 = [0, \infty)$.
If E_4: Radioactive particles emitted by a radioactive substance, then $\Omega_4 = \{0, 1, 2, \ldots\}$.
If E_5: Roll a pair of dice and see up face, then $\Omega_5 = \{(i, j),\ i = 1, 2, 3, 4, 5, 6;\ j = 1, 2, 3, 4, 5, 6\}$.

Introduction

- Event: is subset of sample space. Event is denoted by capital letter.
- The set of all subsets is power set for a finite sample space .

Example

In E_1 the event is the toss yield a head $A_1 = \{H\}$
In E_2 we are getting head in third toss then event is $A_2 = \{TTH\}$.
In E_3 an event is $A_3 = (0, 2)$.
In E_4 if the radioactive particles emitted is 2, then $A_4 = \{2\}$.
In E_5 if sum of number on up faces is 4, then $A_5 = \{(1,3),(2,2),(3,1)\}$.
If $E_1, E_2, \ldots$ are events in Ω and events $E_1, E_2, \ldots$ are mutually exclusive, then $\bigcup_{i=1}^{\infty} E_i \subseteq \Omega$.

Sigma Field

Suppose E is an experiment with sample space as Ω. Let f be a collection of subsets of Ω. Then f is said to be a sigma field if

1. $\Omega \in f$.
2. If $A \in f$, then $\bar{A} \in f$.
3. If $A_1, A_2 \in f$, then $A_1 \cup A_2 \in f$.

Example-Sigma Field

Example (1)

E: Toss a coin, then $\Omega = \{H, T\}$. Then
$f_1 = \{\phi, \{H\}, \{T\}, \Omega\}$ is the power set and is a sigma field.
$f_2 = \{\phi, \Omega\}$ is trivial sigma field.

Example (2)

E: Toss a two coin, then $\Omega = \{HH, HT, TH, TT\}$. Then
$f_1 = \{\phi, \Omega\}$ is trivial sigma field.
f_2=Power set of Ω, is a sigma field.

Probability

Definition

Consider a random experiment E having sample space Ω. Let f be a sigma field of subsets of Ω. Consider an event A defined on f, then $P(A)$ is a real number called a probability of event A if $P(\cdot)$ satisfies following axioms:

1. $P(\Omega) = 1, \Omega \in f$.
2. $P(A) \geq 0, A \in f$.
3. If $A_1, A_2, \ldots$ are mutually exclusive events in f, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

How to assign the probabilities

The assignment of probability is done on the basis of

1. prior experience or prior observations;
2. analysis of the experimental conditions;
3. assumptions.

Relative frequency approximation

$$P(A) = \lim_{m \to \infty} \frac{m_A}{m} = \frac{\text{number of times event A occurs}}{\text{number of times experiment was done}}.$$

For large number of trails, the approximate probability obtained is quite near to the exact probability. The disadvantage of this approach is that the experiment should be repeated and is not a one off situation.

Classical Method

Here we assume that the possible outcomes of random experiment are equally likely and their total number is finite. Then

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{\text{number of favourable cases to event } A}{\text{number of cases in } \Omega}.$$

Simple consequences of axioms

Consider the experiment E on (Ω, f).

1. $P(\phi) = 0, \ \phi = \{\} \in f$.
2. A and B are two events in f, then
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$
3. If $A, B \in f$ and $A \subseteq B$, then $P(A) \leq P(B)$.
4. Let $A_i \in f, \ i = 1, 2, \ldots, n$, then
$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{\substack{i,j=1 \\ i<j}}^{n} P(A_i \cap A_j) + \sum_{\substack{i,j,k=1 \\ i<j<k}}^{n} P(A_i \cap A_j \cap A_k)$$
$$+ \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n).$$

Example

Example

Suppose we are rolling two fair dices independently. We want to find the probability that

1. the sum of faces up is 7.
2. total sum of numbers on faces up is greater than 9.

Solution: 1.Let event A_1 denote sum of faces up is 7. Favourable cases for $A_1 = \{(1,6),(6,1),(2,5),(5,2),(3,4),(4,3)\}$. Also total number of cases are 36. Hence

$$P(A_1) = \frac{\text{Number of favourable cases}}{\text{Total number of cases}} = \frac{6}{36} = \frac{1}{6}.$$

2. Event A_2 denote sum of numbers on up-faces is greater than 9. Then $A_2 = \{(4,6),(6,4),(5,5),(5,6),(6,5),(6,6)\}$. Therefore

$$P(A_2) = \frac{6}{36} = \frac{1}{6}.$$

Example

Example

An urn contains 5 red, 2 black and 4 yellow balls. Two balls are drawn at random from the urn. Find the probability that both balls are of same colour.

Total number of balls are 11. Two balls are drawn out of 11 balls in $\binom{11}{2}$ ways. Let event E_1 denote that both balls are of same colour. If balls are red the number of ways of choosing them are $\binom{5}{2}$. If balls are black the number of ways of choosing them are $\binom{2}{2}$. If balls are yellow the number of ways of choosing them are $\binom{4}{2}$. Therefore number of favourable cases to E_1 is $\binom{5}{2}+\binom{2}{2}+\binom{4}{2}$. Hence required probability is

$$P(E_1) = \frac{\binom{5}{2}+\binom{2}{2}+\binom{4}{2}}{\binom{11}{2}} = \frac{17}{55}.$$

Example

Example

Four persons A, B, C, and D take turns (in the sequence $A, B, C, D, A, B, C, D, A, \ldots$) in tossing a biased coin.The biased coin has probability $3/4$ of head up. The first person to get a tail wins. We want to determine the probability that B wins. The probability of getting a tail in tossing the coin is $p = 1/4$ and $q = 1 - p$. Then required probability is

$$\begin{aligned} P(B \text{ wins}) &= qp + q^5 p + q^9 p + \cdots \\ &= pq(1 + q^4 + q^8 + \cdots) \\ &= \frac{pq}{1 - q^4} \\ &= 0.274. \end{aligned}$$

Definition

(a) A sequence of events $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$ are said to be monotonically increasing if, for all n, $A_n \subseteq A_{n+1}$

(b) A sequence of events $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$ are said to be monotonically decreasing if, for all n, $A_n \supseteq A_{n+1}$

Definition (Limit of sequence)

(a) For monotonically increasing sequence of events $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$ the limit of sequence of events is defined as

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$$

(b) For monotonically decreasing sequence of events $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$ the limit of sequence of events is defined as

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$$

Example

Example

Let $A_{n-1} = \{\omega : 0 < \omega < \frac{n-2}{n-1}\},\ n = 2, 3, \ldots$. Then $A_1 = \phi,\ A_2 = (0, \frac{1}{2}),\ A_3 = (0, \frac{2}{3}), \ldots$. Hence sequence of events $\{A_n\}_{n=1}^{\infty}$ is monotonically increasing. The limit of sequence is

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$$
$$= \{\omega : 0 < \omega < 1\}.$$

Continuity theorem in probability

Let (Ω, f, P) be a probability model.

(a) If $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$, be monotonically increasing sequence of events, then

$$P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P\left(A_n\right)$$

(b) If $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$, be monotonically decreasing sequence of events, then

$$P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P\left(A_n\right)$$

Summary

To analyze the the algoritmns and computer systems, computer scientists need powerful tools. Many of the tools require the foundation in the probability theory. Hence we require to study the concepts of probability. The Russian mathemematican Kolmogorov (1903-1987) provided foundational work of the modern probability theory. In this chapter we introduced the concepts of random experiment, the sample space and the mathematical definition of probability. Followed by the assignment of the probabilities to the events. Some simple consequences of the axioms of the definition of the probability were also discussed. Finally we discussed the continuity theorem of probability. Various examples are provided to understand the concepts.

Introduction to Probability
Chapter 2: Conditional Probability

Dr. Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

1. Conditional Probability
2. Independence
3. Theorem of total probability
4. Bayes' theorem

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.

Conditional Probability

Consider a family having two children, then $\Omega = \{BB, GB, BG, GG\}$, $n(\Omega) = 4$. Consider the event A: both the children are girls. Then $P(A) = 1/4$.
If some information in the form of "B: at least one of the children is a girl" is known. Then reduced sample space is $B = \{GB, BG, GG\}$. Then the probability of A given the condition B is $P(A|B) = 1/3$.
Note that $P(A|B) \geq P(A)$.
$P(A|B) = \frac{n(A\cap B)}{n(B)} = \frac{n(A\cap B)/n(\Omega)}{n(B)/n(\Omega)} = \frac{P(A\cap B)}{P(B)}$, provided $P(B) > 0$.

Definition

Let probability model be (Ω, f, P). Then the conditional probability of $A \in f$ given B is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \; P(B) > 0.$$

Multiplication rule

The probability that n events $A_1, A_2, \ldots, A_n \in f$ occur in a sequence is

$$P(\cap_{i=1}^{n} A_i) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cdots \cap A_{n-1}),$$

provided $P(A_1) > 0,\ P(A_1 \cap A_2) > 0, \ldots, P(A_1 \cap A_2 \cdots \cap A_{n-1}) > 0.$

Example

A bag contains 5 red, 5 white and 4 blue balls. If someone draws 3 balls one by one without replacement, then the probability that three balls will be drawn in the sequence red-white-blue is

$$\begin{aligned} P(R_1 \cap W_2 \cap B_3) &= P(B_3|W_2 \cap R_1)P(W_2|R_1)P(R_1) \\ &= P(R_1)P(W_2|R_1)P(B_3|W_2 \cap R_1) \\ &= \frac{5}{14} \times \frac{5}{13} \times \frac{4}{12}. \end{aligned}$$

Here note that $R_1,\ W_2,\ B_3$ are dependent events.

Example

Example

A bag contains 5 red, 5 white and 4 blue balls. If someone draws 3 balls one by one with replacement, then the probability that three balls will be drawn in the sequence red-white-blue is

$$\begin{aligned} P(R_1 \cap W_2 \cap B_3) &= P(R_1)P(W_2)P(B_3) \\ &= \frac{5}{14} \times \frac{5}{14} \times \frac{4}{14}. \end{aligned}$$

Here note that $R_1,\ W_2,\ B_3$ are independent events.

Example

Example

In a war game, submarine S_1 targets S_2, and both S_2 and S_3 target S_1. The probabilities of S_1, S_2 and S_3 hitting their targets are $1/2$, $2/3$ and $1/3$ respectively. They shoot simultaneously. We want to determine the conditional probability that S_2 hits the target and S_3 does not given that S_1 is hit. Here the required probability is

$$\begin{aligned} P(S_2 \cap \bar{S_3} | S_1 \text{ is hit}) &= \frac{P(S_2)P(\bar{S_3})}{P(S_2 \cup S_3)} \\ &= \frac{\frac{2}{3} \times \frac{2}{3}}{\frac{2}{3} + \frac{1}{3} - \frac{2}{3} \times \frac{1}{3}} \\ &= \frac{4}{7} \end{aligned}$$

Independence

Two events are independent if the occurrence of one does not effect the occurrence or nonoccurrence of the other.

Definition

Events A and B are independent if $P(A|B) = P(A)$. Hence $P(A \cap B) = P(A)P(B)$ and also $P(B|A) = P(B)$.

- If A and B are independent, then A and $\bar{B}$ are independent.
- If A and B are independent, then $\bar{A}$ and B are independent.
- If A and B are independent, then $\bar{A}$ and $\bar{B}$ are independent.

Example

Example

Suppose that $P(A) = 0.4$, $P(B) = 0.5$, and A and B are independent events. Determine $P(A^c \cup B^c)$. Note that

$$\begin{aligned} P(A^c \cup B^c) &= P(A^c) + P(B^c) + P(A^c \cap B^c) \\ &= 1 - P(A) + 1 - P(B) - P(A^c)P(B^c) \\ &= 1 - 0.4 + 1 - 0.5 - (1 - 0.4)(1 - 0.5) \\ &= 0.8, \end{aligned}$$

it follows since A and B are independent, then $\bar{A}$ and $\bar{B}$ are also independent.

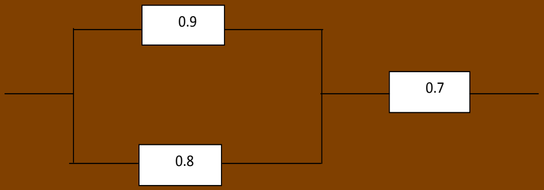
Definition

The n events $A_1, \ldots, A_n$ are mutually independent if and only if the probability of intersection of any $2, 3, \ldots, n$ of these sets is product of their respective probabilities, i.e., for $r = 2, 3, \ldots, n$,

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_r}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_r}).$$

Example

Consider the following electronic system (see diagram), which shows the probabilities of the system components operating properly (i.e, the reliability of the components). Assume that each component operates independently. Find the system reliability, i.e., the probability that the entire system operates?



Solution: Since components are mutually independent. Hence

$$\text{System reliability} = (1 - (1 - 0.9)(1 - 0.8)) \times 0.7 = 0.686$$

Example

Consider the experiment of rolling two fair dice repeatedly and independently until a total of 5 or a total of 7 appears. We want to determine the probability that a total of 5 is rolled before a total of 7 is rolled.

Solution: Let event A_i denote that 5 is rolled before a 7 in the ith trail and experiment terminates.

$$P(A_i) = P\left(\left\{\bigcap_{j=1}^{i-1}(5\cup 7)^c\right\}\cap 5\right) =^{\text{indep}} \left(\frac{26}{36}\right)^{i-1} \times \frac{4}{36}$$

$$\text{Now } P(5 \text{ before } 7) = P\left(\bigcup_{i=1}^{\infty} A_i\right) =^{\text{disjoint}} \sum_{i=1}^{\infty} P(A_i)$$

$$= \sum_{i=1}^{\infty}\left(\frac{26}{36}\right)^{i-1} \times \frac{4}{36} = \frac{4/36}{1-\frac{26}{36}} = \frac{2}{5}.$$

Theorem of total probability

Events $E_1, \ldots, E_n$ are mutually exclusive and exhustive, and event A is caused by happening of $E_1, \ldots, E_n$, then

$$P(A) = \sum_{i=1}^{n} P(A|E_i)P(E_i),$$

here $P(E_i) > 0, \; i = 1, 2, \ldots, n.$

Bayes' theorem

Events $E_1, \ldots, E_n$ are mutually exclusive and exhustive, and event A is caused by happening of $E_1, \ldots, E_n$, then for $i = 1, 2, \ldots, n$

$$P(E_i|A) = \frac{P(A|E_i)P(E_i)}{\sum_{j=1}^{n} P(A|E_j)P(E_j)},$$

here $P(A) > 0$ and $P(E_i) > 0, \; i = 1, 2, \ldots, n.$

Example

In a town there are 200 car drivers, 500 two-wheeler drivers and 20 bus drivers. There is a probability 0.01, 0.03 and 0.15 respectively for an accident involving car, two-wheeler and bus. One of the drivers meets with an accident, what is the probability that he/she was driving a car?

Solution: Let event A,B,C denote the events that the chosen driver drives a car, a two-wheeler, a bus, respectively. Let event E denote an accident. Here $P(A) = \frac{200}{200+500+20} = \frac{20}{72}$, $P(B) = \frac{50}{72}$, $P(C) = \frac{2}{72}$. Also $P(E|A) = 0.01$, $P(E|B) = 0.03$ and $P(E|C) = 0.15$. Now the required probability, using Bayes' theorem, is

$$\begin{aligned} P(A|E) &= \frac{P(A)P(E|A)}{P(A)P(E|A) + P(B)P(E|B) + P(C)P(E|C)} \\ &= \frac{\frac{20}{72} \times 0.01}{\frac{20}{72} \times 0.01 + \frac{50}{72} \times 0.03 + \frac{2}{72} \times 0.15} \\ &= 0.1 \end{aligned}$$

Example

Consider the experiment of rolling two fair dice repeatedly and independently until a total of 5 or a total of 7 appears. We want to determine the probability that a total of 5 is rolled before a total of 7 is rolled.

Solution: Let event A denote that 5 is rolled before a 7. Then $A = \bigcup_{i=1}^{n}(A \cap B_i)$, where B_i is the event that the game terminates in ith roll. Now the required probability, using theorem of total probability, is

$$\begin{aligned} P(A) &= \sum_{i=1}^{\infty} P(A \cap B_i) \\ &= \sum_{i=1}^{\infty} \left(\frac{26}{36}\right)^{i-1} \times \frac{4}{36} \\ &= \frac{4/36}{1-\frac{26}{36}} = \frac{2}{5}. \end{aligned}$$

Summary

Since there may be some information available about the outcome of the trail in a given experiment, hence we introduced the concept of the conditional probability. Also if this information is irrelevant to the event under considration from there comes the definiton of the independence of the events. These definitions can be used to find the reliabilities of the series-parallel or parallel-series structures. Hence examples are provided for the same. In the last theorem of total probability and Bayes' theorem were presented.

Introduction to Probability
Chapter 3: Random Variable

Dr. Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

1. Random variable
2. Probability mass function (PMF)
3. Probability density function (PDF)
4. Cummulative distribution function (CDF)

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.

Random Variable

Definition (Random Variable)

Let (Ω, f, P) be probability model. A real valued function X defined on sample space Ω is a random variable if for all $x \in \mathbb{R}$, $\{\omega : X(\omega) \leq x\} \in f$, i.e. $X^{-1}((-\infty, x]) \in f$. i.e. $X^{-1}((-\infty, x])$ is an event.

Example

E: Toss a coin, then $\Omega = \{H, T\}$. Suppose X counts the number of heads. Consider $f = \{\phi, \{H\}, \{T\}, \Omega\}$. Then

$$\mathrm{X}^{-1}((-\infty, x]) = \begin{cases} \phi, & x < 0 \\ \{T\}, & 0 \leq x < 1 \\ \{H, T\}, & x \geq 1 \end{cases}$$

is in f, i.e., $X^{-1}((-\infty, x])$ an event. Hence X is a random variable.

Random Variable

Example

E: Toss a coin two times, then $\Omega = \{HH, HT, TH, TT\}$. Suppose X counts the number of heads. Consider f as power set of Ω. Then

$$X^{-1}((-\infty, x]) = \begin{cases} \phi, & x < 0 \\ \{TT\}, & 0 \le x < 1 \\ \{TT, HT, TH\}, & 1 \le x < 2 \\ \Omega, & x \ge 2 \end{cases}$$

is an event. Hence X is a random variable.

Discrete Random Variable

- X is a discrete type random variable if we can associate a number $p_X(x) = P(X = x)$ with each outcome x in the range space R_X such that $p_X(x) \geq 0$ and $\sum_{x \in R_X} p_X(x) = 1$.
- The pair $(x, p_X(x)),\ x \in R_X$ is called probability distribution and $p_X(x)$ is called the probability mass function (PMF).

Example

Let we toss the coin three times. Then sample space is $S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$; where H, T denote head and tail, respectively. Let X denote number of heads in tossing coin three times. Then $R_X = \{0, 1, 2, 3\}$. The probability distribution is

x	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Continuous Random Variable

- Let X be continuous type random variable if we can associate a function called probability density function (PDF) $f(x)$ such that

  1. $f(x) \geq 0$, for all x
  2. $\int_{-\infty}^{\infty} f(x)dx = 1$, and
  3. $P(a \leq X \leq b) = \int_a^b f(x)dx$, for $-\infty < a < b < \infty$.

Example

Let X be the time to failure (in days) of an electronic device has the following PDF

$$f(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases}$$

Here for $0 < a < b$, $P(a < X < b) = \int_a^b \lambda e^{-\lambda x} dx = e^{-\lambda a} - e^{-\lambda b}$. Then the electronic device has life more than 10 days but less than 30 days is

$$P(10 < X < 30) = e^{-10\lambda} - e^{-30\lambda}.$$

Cumulative Distribution Function (CDF)

- For random variable X, the CDF is defined as $F_X(x) = P(X \leq x)$.
- For discrete case: for random variable X, the CDF is defined as $F_X(x) = P(X \leq x) = \sum_{x_i \leq x} p_X(x_i)$, where $p(x_i)$ is the PMF. Note that $P(X = a) = F(a) - F(a-)$.
- For continuous case: for random variable X, the CDF is defined as $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du$, where $f(x)$ is PDF. Note that $\frac{d}{dx}F(x) = f(x)$ and $P(X = a) = 0$.

Example

Example (1)

Let the random variable X denote the number of defects in a device having the CDF as

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

Then $P(X = 1) = F(1) - F(1-) = 1 - \frac{1}{2} = \frac{1}{2}$ and $P(X = 0) = F(0) - F(0-) = \frac{1}{2} - 0 = \frac{1}{2}$. Also $P\left(0 \leq X \leq \frac{3}{2}\right) = P\left(X \leq \frac{3}{2}\right) - P(X < 0) = F\left(\frac{3}{2}\right) - F(0-) = 1 - 0 = 1$.

Example (2)

A pen drive has either 2GB, 4GB, 8GB, 16GB or 64GB of memory. Let Y denote the amount of memory in a purchased pen drive with probability mass function as $p(2) = 0.05$, $p(4) = 0.10$, $p(8) = 0.35$, $p(16) = k$, $p(64) = 0.10$, where k is a constant and $p(x) = P(Y = x)$. We want to determine the value of k and the CDF of Y. Since $p(2) + p(4) + p(8) + p(4) + p(16) + p(64) = 1$, therefore $k = 0.4$. The CDF of rv Y is

$$F(x) = \begin{cases} 0, & x < 2 \\ 0.05, & 2 \leq x < 4 \\ 0.15, & 4 \leq x < 8 \\ 0.50, & 8 \leq x < 16 \\ 0.90, & 16 \leq x < 64 \\ 1, & x \geq 64 \end{cases}$$

Example (3)

Let X denote the number of trials to get a success has the following PMF

$$p(x) = q^{x-1}p,\ \ x = 1, 2, \ldots;\ \ p + q = 1.$$

Then for any positive integer x,

$$F(x) = \sum_{y \le x} p(y) = \sum_{y=1}^{x} q^{y-1}p = p\sum_{y=0}^{x-1} q^{y} = p \times \frac{1-q^{x}}{1-q} = 1 - q^{x}.$$

Hence the CDF of X is

$$F(x) \quad = \quad \begin{cases} 0, & x < 1 \\ 1 - q^{[x]}, & x \ge 1 \end{cases}$$

where $[x]$ is the largest integer $\le x$ (e.g., $[4.7] = 4$).

Example (4)

Let magnitude X denote the dynamic load on a bridge (in Newtons). The PDF of X is

$$f(x) = \begin{cases} x - \frac{7}{4}, & 0 \leq x \leq 4 \\ 0 & \text{otherwise} \end{cases}$$

For any number $0 \leq x \leq 4$,

$$F(x) = \int_{-\infty}^{x} f(y)dy = \int_{0}^{x} \left(y - \frac{7}{4} \right) dy = \frac{x^2}{2} - \frac{7}{4}x.$$

Then the CDF of X is

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{2} - \frac{7}{4}x, & 0 \leq x < 4 \\ 1, & x \geq 4 \end{cases}$$

Properties of CDF

- $0 \leq F_X(x) \leq 1, \forall x.$
- $F_X(x)$ is non-decreasing in x.
- $F_X(x)$ is right continuous.
- $\lim_{x \to \infty} F_X(x) = 1$ and $\lim_{x \to -\infty} F_X(x) = 0.$

Summary

This chapter introduced the concept of random variable. Then the discrete and continuous random variables are defined. The probability mass function (PMF), the probability density function (PDF) and the cummulative distribution function (CDF) are introduced after that.

Introduction to Probability
Chapter 4: Expectation

Dr Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

1. Expectation
2. Moments
3. Moment Generating Function (MGF)
4. Chebyshev's Inequality

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.

Expectation

Let X be discrete type random variable with probability distribution $(x, p(x)),\ x = x_1, x_2, \ldots$. If

$$\sum_{x=x_1}^{\infty} |x| p(x) < \infty$$

then expected value of X exist. And the expected value of X or mean of X is defined as

$$\mu = E(X) = \sum_{x=x_1}^{\infty} x p(x).$$

The variance of X is

$$\sigma^2 = E(X - \mu)^2 = \sum_{x=x_1}^{\infty} (x - \mu)^2 p(x) = E(X^2) - (E(X))^2.$$

Expectation

Let X be continuous type random variable with PDF $f(x)$. If

$$\int_{-\infty}^{\infty} |x| f(x) dx < \infty$$

then expected value of X exist. And the expected value of X or mean of X is defined as

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

The variance of X is

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

Example

A random variable X has the following PMF

x	-2	-1	0	1	2	3
p(x)	0.1	k	0.2	k	0.3	k

Find (i) the value of k. (ii) $E(X)$ and $Var(X)$.
(i) Since $0.1 + k + 0.2 + k + 0.3 + k = 1 \Rightarrow k = 0.133$.
(ii) $E(X) = -2 \times 0.1 - 1 \times k + 0 + 1 \times k + 2 \times 0.3 + 3 \times k = 0.799$
$E(X^2) = 4 \times 0.1 + 1 \times k + 0 + 1 \times k + 4 \times 0.3 + 9 \times k = 3.063$
$Var(X) = E(X^2) - (E(X))^2 = 2.42$.

Example

Let X be the time to failure (in days) of an electronic device has the following PDF

$$f(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0, \ \lambda > 0 \end{cases}$$

Then the mean life or average life of electronic device is

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_{0}^{\infty} x \lambda e^{-\lambda x} dx \\ &= \frac{\lambda}{(\lambda)^2} \\ &= \frac{1}{\lambda}. \end{aligned}$$

Moments

- Let X be discrete type random variable with probability distribution $(x, p(x)),\ x = x_1, x_2, \ldots,$ then the origin moment or rth moment about origin is $\mu_r' = E(X^r) = \sum_{x=x_1}^{\infty} x^r p(x),\ r = 1, 2, \ldots.$ The central moment or rth moment about mean is $\mu_r = E(X-\mu)^r = \sum_{x=x_1}^{\infty} (x-\mu)^r p(x),\ r = 1, 2, \ldots.$
- Let X be continuous type random variable with PDF $f(x)$,then the origin moment or rth moment about origin is $\mu_r' = E(X^r) = \int_{-\infty}^{\infty} x^r f(x) dx,\ r = 1, 2, \ldots.$ The central moment or rth moment about mean is $\mu_r = E(X-\mu)^r = \int_{-\infty}^{\infty} (x-\mu)^r f(x) dx,\ r = 1, 2, \ldots.$

Relationship between central moment and origin moment

$$\begin{aligned}\mu_r &= E(X-\mu)^r \\ &= \sum_{j=0}^{r} \binom{r}{j} (-1)^j \mu^j \mu'_{r-j}\end{aligned}$$

Moment Generating Function (MGF)

Let X be random variable, then the Moment Generating Function (MGF) is defined as $M_X(t) = E(e^{tX})$. Note that

$$M_X(t) = E\left(1 + tX + \frac{t^2X^2}{2!} + \cdots\right)$$
$$= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \cdots$$

Hence $\frac{d^r}{dt^r} M_X(t)|_{t=0} = E(X^r) = \mu'_r$.

Example

Example

The random variable X has PMF
$P(X=x)=\binom{n}{x}p^x\ q^{n-x},\ x=0,1,\ldots,n,\ p+q=1,\ 0<p<1.$ MGF is

$$\begin{aligned} M(t) &= E(e^{tX}) \\ &= \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x\ q^{n-x} \\ &= \sum_{x=0}^{n}\binom{n}{x}(pe^t)^x\ q^{n-x} \\ &= \left(q+pe^t\right)^n \end{aligned}$$

Here $\mu=E(X)=\frac{d}{dt}M_X(t)|_{t=0}=np;\ \mu_2'=E(X^2)=\frac{d^2}{dt^2}M_X(t)|_{t=0}=np\,(1+(n-1)p)$ and $\mu_2=Var(X)=\mu_2'-\mu^2=npq$.

Example

Let X be the time to failure (in days) of an electronic device has the following PDF

$$f(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0, \ \lambda > 0 \end{cases}$$

MGF is

$$\begin{aligned} M(t) &= E(e^{tX}) \\ &= \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx \\ &= \lambda \int_0^\infty e^{-(\lambda - t)x} dx \\ &= \left(1 - \frac{t}{\lambda}\right)^{-1}, \end{aligned}$$

for $t < \lambda$. Here $\mu = E(X) = \frac{d}{dt} M_X(t)|_{t=0} = \frac{1}{\lambda}$;
$\mu_2' = E(X^2) = \frac{d^2}{dt^2} M_X(t)|_{t=0} = \frac{2}{\lambda^2}$ and $\mu_2 = Var(X) = \mu_2' - \mu^2 = \frac{1}{\lambda^2}$.

Example

Example

Let X be a discrete random variable with MGF $M_X(t) = \frac{e^{2t}}{6} + \frac{e^{3t}}{2} + \frac{e^{4t}}{3}$. Determine Var($X$). Note that

$$E(X) = \frac{d}{dt} M_X(t)|_{t=0} = \left(\frac{2e^{2t}}{6} + \frac{3e^{3t}}{2} + \frac{4e^{4t}}{3}\right)|_{t=0} = \frac{19}{6}$$

$$E(X^2) = \frac{d^2}{dt^2} M_X(t)|_{t=0} = \left(\frac{4e^{2t}}{6} + \frac{9e^{3t}}{2} + \frac{16e^{4t}}{3}\right)|_{t=0} = \frac{63}{6}$$

Hence

$$\begin{aligned} Var(X) &= E(X^2) - (E(X))^2 \\ &= \frac{17}{36}. \end{aligned}$$

Let X be a random variable, then for $a, b \in \mathbb{R}$,

1. $E(aX + b) = aE(X) + b$
2. $Var(aX + b) = a^2 \; Var(X)$

Example

Let random variable X has mean μ and variance σ^2. Then the random variable $Z = \frac{X-\mu}{\sigma}$ has the mean

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{E(X)-\mu}{\sigma} = 0,$$

and variance

$$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2} Var(X) = 1.$$

Here $Z = \frac{X-\mu}{\sigma}$ is called the standarized random variable.

Chebyshev's Inequality

Let X be a general random variable with mean μ and variance σ^2. Then, for any $k > 0$,

$$P\left(|X - \mu| \geq k\right) \leq \frac{\sigma^2}{k^2}.$$

or, equivalently,

$$P\left(|X - \mu| \leq k\right) \geq 1 - \frac{\sigma^2}{k^2}.$$

Example

Suppose we have sampled the weights of Ponies in the local animal shelter and found that our sample has mean 40 and variance 9 pounds. Using Chebyshev's inequality, find the lower bound of the probability that the weight of Ponies is between 34 pounds to 46 pounds.

Solution: Let X be the weight of Ponies. Here mean $\mu = 40$ and variance $\sigma^2 = 9$. Then, using Chebyshev's Inequality, we have

$$\begin{aligned} P(34 < X < 46) &= P(|X - 40| < 6) \\ &\geq 1 - \frac{\sigma^2}{6^2} = 1 - \frac{9}{36} = \frac{27}{36}. \end{aligned}$$

Example

Let X be a random variable with mean μ and variance σ^2, then using Chebyshev's inequality following values may easily be observed

$$P(|X-\mu| \leq 2\sigma) \geq 1 - \frac{1}{2^2} = 0.75$$

,i.e., 75% chances is that the observed value of X lies between two standard deviations of the mean.

$$P(|X-\mu| \leq 3\sigma) \geq 1 - \frac{1}{3^2} = 0.89$$

,i.e., 89% chances is that the observed value of X lies between three standard deviations of the mean.

$$P(|X-\mu| \leq 4\sigma) \geq 1 - \frac{1}{4^2} = 0.94$$

,i.e., 94% chances is that the observed value of X lies between four standard deviations of the mean.

Summary

This chapter introduced the concept of expectation. The moments were introduced, the mean and variance are particular cases of that. The moment generating function (MGF) was introduced and examples are presented for finding mean and variance using MGF. In last the Chebyshev's inequality was presented.

Introduction to Probability
Chapter:5 Discrete Distributions

Dr Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

1. Binomial Trail
2. Binomial Distribution
3. Geometric Distribution
4. Negative Binomial Distribution
5. Hypergeometric Distibution

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.

Bernoulli trail $X \sim Ber(p)$

- Consider a trail in which we can have success with probability p and failure with probability q, such that $p + q = 1$
- Let X counts the number of successess. Then X can take values $0, 1$.
- PMF is $P(X = 1) = p$ and $P(X = 0) = q$.
- MGF is $M(t) = q + pe^t$.
- $E(X) = p$ and $Var(X) = E(X^2) - (E(X))^2 = p - p^2 = pq$

Binomial distribution $X \sim Bin(n, p)$

- Let n independent bernoulli trails are conducted.
- Let X counts the number of successess in n trails. Then X can take values $0, 1, \ldots, n$.
- PMF is $P(X = x) = \binom{n}{x}\ p^x\ q^{n-x},\ x = 0, 1, \ldots, n;\ 0,$ otherwise.
- MGF is $M(t) = (q + pe^t)^n$.
- $E(X) = np$; $E(X^2) = n(n-1)p^2 + np$; and $Var(X) = npq$.

Binomial distribution

Example

Suppose that we are throwing pair of dice 10 times and asking for the probability of double five. Here the probability of getting double five is $p = \frac{1}{36}$. If X denote the number of times getting a double five from rolling two dices. Then $X \sim Bin(n, p),\ n = 10,\ p = \frac{1}{36}$ Then Required probability $P(X > 0) = 1 - P(X = 0) = 1 - \binom{10}{0}\left(\frac{1}{36}\right)^0\left(\frac{35}{36}\right)^{10}$

Binomial distribution

Example

Suppose that a football player makes 30% of his shot attempts. If the players shoots 10 shots in a game, and X is the number of shots made. Assuming that the shots are independent, find the mean of random variable X. Clearly $X \sim Bin(n, p),\ n = 10,\ p = 0.30$. Then mean of random variable X is

$$E(X) = np = 10 \times 0.30 = 3.$$

Geometric distribution $X \sim Geo(p)$

- Let independent bernoulli trails are conducted till we get a success.
- Let X counts the number of trails to get a success. Then X can take values $1, 2, \ldots \infty$.
- PMF is $P(X = x) = q^{x-1}p,\ x = 1, 2, \ldots \infty;\ 0,$ otherwise, $p + q = 1$.
- MGF is $M(t) = \frac{pe^t}{1-qe^t}$.
- $E(X) = \frac{1}{p}$; and $Var(X) = \frac{q}{p^2}$.

Geometric distribution

Example

In each trial the probability of success is 0.25, in how many trails we expect first success?

Solution: Let X denote the number of trails to have first success. $X \sim Geo(p)$, where $p = 0.25$.Then expected number of trials are $E(X) = \frac{1}{p} = 4$.

Geometric distribution

- Geometric distribution has memoryless property, i.e.,

$$P(X > x + y | X > y) = P(X > x)$$

Example

Let getting a '1' is a success in roll of a die. We will roll a fair die until we observe success. Let we have already rolled the die ten times without a success. The probability that more than two additional tosses are required to have a success is

$$\begin{aligned} P(X > 12 | X > 10) &= P(X > 2) \\ &= 1 - P(X = 1) - P(X = 2) \\ &= 1 - \frac{1}{6} - \frac{1}{6} \times \frac{5}{6} \\ &= \frac{25}{36}. \end{aligned}$$

Negative Binomial distribution $X \sim NB(r, p)$

- Let independent bernoulli trails are conducted till we get r successes.
- Let X counts the number of trails to get r success. Then X can take values $r, r+1, \ldots \infty$.
- PMF is $P(X = x) = \binom{x-1}{r-1}\ p^r q^{x-r},\ x = r, r+1, \ldots \infty;\ 0,\ $ otherwise.
- MGF is $M(t) = \left(\frac{pe^t}{1-qe^t}\right)^r$.
- $E(X) = \frac{r}{p}$; and $Var(X) = \frac{rq}{p^2}$.

Negative Binomial distribution

Example

In ODI cricket match series between two teams A and B, the team who wins three games will be the winner. Suppose that the team A has probability 0.60 of winning over team B. We want to find the probability that team A will win the series in 5 games. Let X counts the number of games required by team A to win the series. Then $X \sim NB(r, p)$, $r = 3, \; p = 0.60$. Now the required probability is $P(X = 5) = \binom{5-1}{3-1}$ $(0.60)^3(1 - 0.60)^{5-3}$.

Negative Binomial distribution

Example

A factory produces components for computers. Let 5% of components are defective. We need to find 3 non-defective components for our 3 new computers. Components are tested until 3 non-defectives are found. What is the probability that more than 5 components will be tested?

Solution: Let X be the number of components tested for getting r non-defectives. Then $X \sim NB(r,p)$, $r=3$, $p=0.95, q=1-p$. Then required probability is

$$\begin{aligned} P(X>5) &= 1-P(X\leq 5) \\ &= 1-[P(X=3)+P(X=4)+P(X=5)] \\ &= 1-[p^3+3p^3q+6p^3q^2] \\ &= 0.0012. \end{aligned}$$

Hypergeometric distribution

A box contains N balls. M are drawn at random, marked and returned to the box. Next n balls are drawn at random from box and then the marked balls are counted. Let X denote the number of marked balls. Then PMF is

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, \ \max(0, M+n-N) \leq x \leq \min(M, n).$$

$E(X) = \frac{n}{N}M, \ Var(X) = \frac{nM}{N^2(N-1)}(N-M)(N-n).$

Summary

The discrete distributions were presented in this chapter. These distributions has wide use in the field of engineering, science and management. Particulaly we intoduced the binomial, geometric, negative Binomial and hypergeometric distribution.

Introduction to Probability
Chapter 6 Poisson Process

Dr Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

1. Poisson Process
2. Properties of Poisson Process
3. Poisson Distribution
4. Binomial Approximation to Poisson Distribution

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.

small $'o'$

- $f = o(h)$ if $Lim_{h \to 0} \frac{f(h)}{h} = 0$
- Example $f(x) = x^2$. Here f is $o(h)$.
- If $f_i = o(h)$, then $f = \sum_{i=1}^{n} a_i f_i = o(h)$
- ,i.e., linear combination of $o(h)$ functions is $o(h)$.
- Example $f_1(x) = x^2$ and $f_2(x) = x^5$. Both f_1 and f_2 are $o(h)$. Hence $f_1 - f_2$ is also $o(h)$.

Poisson Process

Let events are occuring in time and $N(t)$ counts the number of events occuring in the interval $(0, t]$. Then $N(t)$ is said to follow Poisson Process with rate λ, i.e., $N(t) \sim PP(\lambda)$ if following assumptions holds:

1. $N(t)$ has independent increment, i.e., events occuring in disjoint time interval are independent.
2. $N(t)$ has stationary increment,i.e., the distribution of $N(t)$ depends on length of interval not on where it is situated.
3. If h is small

$$P(N(h) = 1) = \lambda h + o(h)$$
$$P(N(h) \geq 2) = o(h).$$

Using assumption (3) $P(N(h) = 0) = 1 - \lambda h + o(h)$.

Poisson Process

Statement: Under the assumptions (1), (2) and (3) of Poisson Process,

$$P(N(t)=n)=\frac{e^{-\lambda t}(\lambda t)^n}{n!},\ n=0,1,\ldots$$

Example

Suppose that in a bank customers are arriving according to the Poisson process with rate $\lambda = 2$ customers per hour. The probability that from 10:00 am to 10:30 am no customer arrive is

$$\begin{aligned} P(N(1/2)=0) &= e^{-2\times\frac{1}{2}} \\ &= e^{-1}, \end{aligned}$$

here $N(t)$ is number of customer arriving in the interval $(0,t)$.

Properties of Poisson Process

- Sum of independent Poisson Process is a Poisson Process. That is if $N_1(t) \sim PP(\lambda_1)$ and $N_2(t) \sim PP(\lambda_2)$; and $N_1(t)$ and $N_2(t)$ are independent, then $N_1(t) + N_2(t) \sim PP(\lambda_1 + \lambda_2)$.
- Suppose each time an event occurs is classified as type-I or type-II. Each event is classified as type-I with probability p and is classified as type-II with probability $q = 1 - p$. Let $N(t) \sim PP(\lambda)$, where $N(t)$ is number of events occuring in the interval $(0, t]$. Let $N_1(t)$ and $N_2(t)$ denote, respectively, the number of type-I and type-II event occuring in the interval $(0, t]$. Here $N(t) = N_1(t) + N_2(t)$. Then $N_1(t) \sim PP(\lambda p)$ and $N_2(t) \sim PP(\lambda q)$; also $N_1(t)$ and $N_2(t)$ are independent.

Example

Example

A radioactive source emits particles (either with redish or with white light) at a rate of 6 per minute in accordance with Poisson process. Particles that are emitted with redish light has a probability $1/3$ and those emitted with white light has probability $2/3$. Find the probability that 5 particles emit with white light in 7 minute period.

Solution: $N(t)$ is the number of particles emitted with white light in interval $(0, t]$. Here $N(t) \sim PP(6 \times \frac{2}{3})$. Hence required probability

$$P(N(7) = 5) = \frac{e^{-6\times\frac{2}{3}\times 7}(6 \times \frac{2}{3} \times 7)^5}{5!}$$

Poisson Distribution

- Fix $\lambda t = \mu$ then the Poisson Process $N(t) = Y$ becomes Poisson distribution such that PMF is,

$$P(Y = n) = \frac{e^{-\mu}(\mu)^n}{n!}, \ n = 0, 1, \ldots$$

- MGF is $M(t) = e^{\mu(e^t - 1)}$.
- $E(X) = \mu$ and $Var(X) = \mu$.

Binomial Approximation to Poisson Distribution

- $Lim_{n \to \infty; np=\lambda} B(n, p) = Poiss(\lambda)$
- If n is large and p is small then Binomial distribution can be approximated by Poisson distribution.

Example

Consider a situation where due to certain infection persons are dying with probability 0.001. Consider that the population has 10000 persons in totallity. Find the probability that more than 5 persons will die due to this infection. Consider X denote the number of persons dying due to infection. $X \sim Bin(n, p)$, where $n = 10000$ and $p = 0.001$. X can be approximated by Possion distribution, i.e., $X \sim Poiss(\lambda)$, here $\lambda = np = 10$. Hence required probability $P(X > 5) = 1 - P(X \leq 5) = 1 - \sum_{i=0}^{5} \frac{e^{-10}(10)^i}{i!}$

Summary

Poisson process was introduced in this chapter. Properties of Poisson process were presented with illustrated examples. Then Poisson distribution was introduced. In last Binomial approximation to Poisson distibution was studied.

Introduction to Probability
Chapter 7 Continuous Distributions

Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

1. Exponential distibution
2. Gamma distibution
3. Normal distribution
4. Standard normal distribution

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.

Exponential distribution $X \sim Exp(\lambda)$

- PDF is

$$f(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0,\ \lambda > 0 \end{cases}$$

- CDF is

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

- MGF is $M(t) = \left(1 - \frac{t}{\lambda}\right)^{-1},\ t < \lambda.$
- $E(X) = \frac{1}{\lambda}$; and $Var(X) = \frac{1}{\lambda^2}.$

Relationship between $Exp(\lambda)$ and Poisson Process

Let $N(t)$ counts the number of events occuring in $(0, t]$, i.e., $N(t) \sim PP(\lambda)$, then

$$P(N(t) = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, \; n = 0, 1, \ldots$$

Let T_1 denote the time of occurance of first event, i.e., time of first arrival of event. Then

$$\begin{aligned} T_1 > t &\equiv N(t) = 0 \\ \Rightarrow \; P(T_1 > t) &= P(N(t) = 0) \\ &= e^{-\lambda t}, \; t > 0, \; \lambda > 0. \end{aligned}$$

Hence $T_1 \sim Exp(\lambda)$

Example

Suppose that on average 30 programme per hour queued to be proceesed in accordance with Poisson process. What is the probability that the server will wait more than 5 minutes before the first programme arrives?

Solution: Let $N(t)$ is the number of programme arriving in $(0, t]$ such that $N(t) \sim PP(\lambda)$

$$\lambda = 30 \text{ per hour} = \frac{30}{60} \text{ per min} = \frac{1}{2} \text{ per min}$$

Let T_1 denote the time for first arrival of programme at server. Here $T_1 \sim exp(\lambda)$. Hence the required probability is

$$\begin{aligned} P(T_1 > 5) &= P(N(5) = 0) \\ &= \frac{e^{-\lambda t}(\lambda t)^0}{0!} \\ &= e^{-\frac{5}{2}} \end{aligned}$$

Memoryless Property of Exponential distribution

- Exponential distribution has memoryless property, i.e., for all $t > 0, \ s > 0$

$$P(X > t + s | X > t) = P(X > s)$$

Gamma distribution $X \sim Gamma(\lambda, r)$

PDF is

$$f(x) = \begin{cases} 0, & x < 0 \\ \frac{\lambda^r}{\Gamma(r)} e^{-\lambda x} x^{r-1}, & x \geq 0,\ \lambda > 0,\ r > 0 \end{cases}$$

here $\Gamma(r) = \int_0^\infty e^{-x} x^{r-1} dx$. Note that $\Gamma(r) = (r-1)\Gamma(r-1)$ and $\Gamma(1) = 1$

MGF is $M(t) = \left(1 - \frac{t}{\lambda}\right)^{-r},\ t < \lambda$. $E(X) = \frac{r}{\lambda}$; and $Var(X) = \frac{r}{\lambda^2}$.

Particular cases:

- If $r = 1$, then $Gamma(\lambda, r) \equiv Exp(\lambda)$
- If $r = \frac{\nu}{2},\ \lambda = \frac{1}{2}$ then $Gamma(\lambda, r) \equiv \chi^2_\nu$

Relationship between $Gamma(\lambda, r)$ and Poisson Process

Let $N(t)$ counts the number of events occuring in $(0, t]$, i.e., $N(t) \sim PP(\lambda)$, then

$$P(N(t) = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!},\ \ n = 0, 1, \ldots$$

Let $S_r = \sum_{i=1}^{r} T_i$ denote the time of occurance of r events, i.e., time of r arrivals of event. Then

$$S_r > t \equiv N(t) < r$$

$$\Rightarrow\ \ P(S_r > t) = P(N(t) \leq r - 1) = e^{-\lambda t} \sum_{i=0}^{r-1} \frac{(\lambda t)^i}{i!},\ \ t > 0,\ \ \lambda > 0.$$

Hence PDF of S_r is $f_{S_r}(t) = -\frac{d}{dt} P(S_r > t) = \frac{\lambda^r}{(r-1)!} e^{-\lambda x} x^{r-1}$.
Hence $S_r \sim Gamma(\lambda, r)$

Example

Suppose that the programmes arrive at a server in according to a Possion process with $\lambda = 10$ programmes per hour. Find the probability that one has to wait more than half hour until the second programme arrive at server?

Solution: $N(t)$ is number of programmes arriving in $(0, t] \sim PP(\lambda)$.
S_r denote the time until r arrivals. Here $S_2 \sim \text{Gamma}(\lambda, r)$,
$\lambda = 10, \; r = 2$. Required probability is

$$
\begin{aligned}
P\left(S_2 > \frac{1}{2}\right) &= P\left(N\left(\frac{1}{2}\right) \leq 1\right) \\
&= P\left(N\left(\frac{1}{2}\right) = 0\right) + P\left(N\left(\frac{1}{2}\right) = 1\right) \\
&= e^{-10\times\frac{1}{2}} + e^{-10\times\frac{1}{2}}\left(10 \times \frac{1}{2}\right) \\
&= 6e^{-5}
\end{aligned}
$$

Normal distribution $X \sim N(\mu, \sigma^2)$

- PDF is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \; -\infty < x < \infty, \; -\infty < \mu < \infty, \; \sigma > 0.$$

- Probability density function is symmetric around μ, i.e., $f(\mu - x) = f(\mu + x)$.
- MGF is $M(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$.
- $E(X) = \mu$ and $Var(X) = \sigma^2$.

Standard Normal distribution $Z \sim N(0,1)$

- PDF is
$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \ -\infty < x < \infty, .$$
- Probability density function is symmetric around 0.
- CDF
$$\Phi(x) = \int_{-\infty}^{x} \phi(u) du$$
- $\Phi(x) + \Phi(-x) = 1$.
- MGF is $M(t) = e^{\frac{1}{2}t^2}$.
- $E(Z) = 0$ and $Var(Z) = 1$.
- If $Y \sim N(\mu, \sigma^2)$, then $Z = \frac{Y-\mu}{\sigma} \sim N(0,1)$.
- From Standard Normal distribution table we can find the values $\Phi(x)$. For example $\Phi(1.28) = 0.89973$.

Example

A firm manufacture devices which has random life that is normally distributed with mean 700 hours and standard deviation 50 hours. (i) Find the probability that the device survives more than 800 hours. (ii) Find the probablity that the device fails between 675 to 725 hours.

Solution: Let X be the random life of device. $X \sim N(\mu, \sigma^2)$, where $\mu = 700,\ \sigma = 50$. Then $Z = \frac{X-700}{50} \sim N(0,1)$.

$$(i)\ P(X > 800) = P(Z > 2) = 1 - \Phi(2) = 1 - 0.97725 = 0.02275.$$

$$\begin{aligned}(ii)\ P(675 < X < 725) &= P(-0.5 < Z < 0.5) = \Phi(0.5) - \Phi(-0.5)\\ &= 2\Phi(0.5) - 1 = 2 \times 0.69146 - 1 = 0.38292.\end{aligned}$$

Summary

In this chapter we presented some widely used continuous distributions. The exponential, Gamma were presented along with their relation with the Poisson process. Since the normal distribution is widely used in the statisitcal inference. Therefore we introduced the normal distribution.

Introduction to Probability
Chapter 8 Joint Distributions

Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.

Joint PMF

- Consider X and Y as discrete random variable. Then (X, Y) is two-dimensional discrete random variable with Joint probability mass function $p(x, y)$ if
  1. $p(x, y) \geq 0, \ \forall x, \ \forall y$, and
  2. $\sum_x \sum_y p(x, y) = 1$
- Marginal density of X is given by $p_X(x) = \sum_y p(x, y), \ \forall x$. Marginal density of Y is given by $p_Y(y) = \sum_x p(x, y), \ \forall x$.
- Conditional density of X given $Y = y$ is given by $p_{X|Y=y}(x) = \frac{p(x,y)}{p_Y(y)}, \ \forall x$. Conditional density of Y given $X = x$ is given by $p_{Y|X=x}(y) = \frac{p(x,y)}{p_X(x)}, \ \forall y$.

Example

A rover (a small vehicle that can move over rough ground, often used on the surface of other planets) is checked for tire wear, and headlight is checked for proper adjustment. Let X be the number of defective tires and Y be the number of headlights that needs adjustment. Consider (X, Y) having joint pmf $p_{XY}(x, y)$ as

$Y = y \downarrow X = x \rightarrow$	0	1	2
0	$2/15$	$1/15$	$1/15$
1	$3/15$	$2/15$	$1/15$
2	$2/15$	$1/15$	$2/15$

Here marginal density of X is

x	0	1	2
$p_X(x)$	$7/15$	$4/15$	$4/15$

example contd

Example

marginal density of Y is

y	0	1	2
$p_Y(y)$	$4/15$	$6/15$	$5/15$

The conditional density of X given $Y = 2$ is

x	0	1	2
$p_{X\|Y=2}(x)$	$2/5$	$1/5$	$2/5$

Joint PDF

- Consider X and Y as continuous random variables. Then (X, Y) is two-dimensional continuous random variable with Joint probability density function $f_{XY}(x, y)$ if
  1. $f_{XY}(x, y) \geq 0, \ -\infty < x < \infty, \ -\infty < y < \infty$, and
  2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dxdy = 1$
- Marginal density of X is given by $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy$. Marginal density of Y is given by $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)dx$
- Conditional density of X given $Y = y$ is given by $f_{X|Y=y}(x) = \frac{f_{XY}(x,y)}{f_Y(y)}$. Conditional density of Y given $X = x$ is given by $f_{Y|X=x}(y) = \frac{f_{XY}(x,y)}{f_X(x)}$.

Example

The front tire and back tires of the rover (a small vehicle that can move over rough ground, often used on the surface of other planets) is supposed to be filled to a pressure of 0.4 psi and 0.6 psi, respectively. Suppose the actual air pressure in each tire is a rv, X for the front tire and Y for the back tire, with joint pdf as $f(x,y) = ky,\ 0 < x < y < 1$. Then $\int\int f(x,y)\,dx\,dy = 1 \Rightarrow\ \int_0^1\int_0^y ky\,dx\,dy = 1 \Rightarrow\ k = 3$. The marginal density of X is

$$f_X(x) = \int f(x,y)\,dy = \int_x^1 3y\,dy = \frac{3}{2}(1-x^2),\ 0 < x < 1.$$

The marginal density of Y is

$$f_Y(y) = \int f(x,y)\,dx = \int_0^y 3y\,dx = 3y^2,\ 0 < y < 1.$$

Example...Contd..

Example

The conditional distribution of Y given $X = x$ is

$$f_{Y|X=x}(y) = \frac{f(x,y)}{f_X(x)} = \frac{2y}{1-x^2}, \ x < y < 1.$$

Independence

- (X, Y) is two-dimensional discrete random variable. Then X and Y are said to be independent if

$$p(x, y) = p_X(x)p_Y(y), \ \forall x, \ \forall y.$$

- (X, Y) is two-dimensional continuous random variable. Then X and Y are said to be independent if

$$f_{XY}(x, y) = f_X(x)f_Y(y).$$

- That is independence implies that the joint density is product of their marginal densities.

Example

A binary message is transmitted, which is either 0 or 1. Assume that the channel has an additive noise and it corrupts the transmission. Let X denote the transmitted message and Y denote the received message by the receiver. Consider the (X, Y) having density $p_{XY}(x, y)$ as

$X = x \downarrow Y = y \rightarrow$	0	1
0	$1/4$	$1/4$
1	$1/4$	$1/4$

Here marginal density of X is

x	0	1
$p_X(x)$	$1/2$	$1/2$

marginal density of Y is

y	0	1
$p_Y(y)$	$1/2$	$1/2$

example contd

Example

Here $p_{X,Y}(x,y) = p_X(x)p_Y(y),\ \forall x,\ \forall y.$ Hence X and Y are independent.

Covariance

- Let X and Y are random variable with means μ_X and μ_Y, respectively. Then covariance between X and Y is

$$Cov(X, Y) = E(X - \mu_X)(Y - \mu_Y) = E(XY) - \mu_X \mu_Y$$

- If X and Y are independent, then $Cov(X, Y) = 0$, i.e., $E(XY) = E(X)E(Y)$. But the converse may not be true.

Correlation

- Let X and Y are random variable with means μ_X and μ_Y and variances σ_X^2 and σ_Y^2, respectively. Then the correlation coefficient between X and Y is
$$\rho_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$
- $-1 \leq \rho_{XY} \leq 1$
- $|\rho_{XY}| = 1$ if and only if $Y = \alpha + \beta X$ for some real numbers α and $\beta \neq 0$.
- X and Y are uncorrelated if $Cov(X, Y) = 0 \Leftrightarrow \rho_{XY} = 0$
- If X and Y are independent, then $\rho_{XY} = 0$. But the converse may not be true.

Example

Consider the (X, Y) having density $p_{XY}(x, y)$ as

$X = x \downarrow Y = y \rightarrow$	-1	0	1
-1	0	1/4	0
0	1/4	0	1/4
1	0	1/4	0

Here marginal density of X is

x	-1	0	1
$p_X(x)$	1/4	1/2	1/4

marginal density of Y is

y	-1	0	1
$p_Y(y)$	1/4	1/2	1/4

Here $p_{X,Y}(-1,-1) \neq p_X(-1)p_Y(-1)$. Hence X and Y are not independent. But $Cov(X, Y) = 0$. Hence X and Y are uncorrelated.

Let X and Y are random variables and $Var(X)$ is finite , then

1. $E\left(E(X|Y)\right) = E(X)$.
2. $E\left(E(g(X)|Y)\right) = E(g(X))$.
3. $Var(X) = E\left(Var(X|Y)\right) + Var\left(E(X|Y)\right)$.

Summary

In this chapter we presented topics related to jointly distributed random variables. The concept of independence, covariance and correlation were introduced.

Introduction to Probability
Chapter 9
Bivariate Normal Distribution, Central Limit Theorem and Law of Large Numbers

Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

1. Linear combination of random variable
2. Sum of random variables
3. Bivariate Normal distibution
4. Central Limit Theorem
5. Law of Large Numbers

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.

Linear combination of random variables

Let $X_1, \ldots, X_n$ be random variables. Let X_i has mean μ_i and variance σ_i^2, $i = 1, 2, \ldots, n$. Let $a_0, a_1, \ldots, a_n$ are real valued constants. Then mean and variance of $Y = a_0 + \sum_{i=1}^{n} a_i X_i$, respectively, are

$$E\left(a_0 + \sum_{i=1}^{n} a_i X_i\right) = a_0 + \sum_{i=1}^{n} a_i \mu_i,$$

and if $X_1, \ldots, X_n$ be independent random variables, then further

$$Var\left(a_0 + \sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \sigma_i^2.$$

Example

Example

A patrol pump sells petrol, premium petrol and diesel. Their prices are Rs. 80, 90 and 60 respectively. Let $X_1,\ X_2$ and X_3 denote the amount of petrol, premium petrol and diesel purchased on a particular day. Let $X_1,\ X_2$ and X_3 are independent with $\mu_1 = 1000,\ \mu_2 = 300,\ \mu_3 = 500$, $\sigma_1 = 100,\ \sigma_2 = 80$ and $\sigma_3 = 50$. The revenue from sales is

$$Y = 80X_1 + 90X_2 + 60X_3.$$

Then

$$E(Y) = 80\mu_1 + 90\mu_2 + 60\mu_3 = 137000$$

and

$$V(Y) = (80)^2\sigma_1^2 + (90)^2\sigma_2^2 + (60)^2\sigma_3^2 = 124840000.$$

Sum of random variables

Let $X_1, \ldots, X_n$ be independent random variables. Then if $S_n = \sum_{i=1}^{n} X_i$, then the MGF of S_n is

$$\begin{aligned} M_{S_n}(t) &= E\left(e^{tS_n}\right) \\ &= \prod_{i=1}^{n} M_{X_i}(t) \end{aligned}$$

Sum of random variables

Let $X_1, \ldots, X_m$ be independent random variables. If

1. $X_i \sim Bin(n_i, p)$, then $\sum_{i=1}^m X_i \sim Bin\left(\sum_{i=1}^m n_i, p\right)$
2. $X_i \sim Poiss(\lambda_i)$, then $\sum_{i=1}^m X_i \sim Poiss\left(\sum_{i=1}^m \lambda_i\right)$
3. $X_i \sim Geo(p)$, then $\sum_{i=1}^m X_i \sim NB\left(m, p\right)$
4. $X_i \sim NB\left(n_i, p\right)$, then $\sum_{i=1}^m X_i \sim NB\left(\sum_{i=1}^m n_i, p\right)$
5. $X_i \sim Gamma(\alpha_i, \beta)$, then $\sum_{i=1}^m X_i \sim Gamma\left(\sum_{i=1}^m \alpha_i, \beta\right)$
6. $X_i \sim \chi^2_{r_i}$, then $\sum_{i=1}^m X_i \sim \chi^2_{\sum_{i=1}^m r_i}$

Sum of random variables

Let $X_1, \ldots, X_m$ be independent random variables. If

1. $X_i \sim N(\mu_i, \sigma_i^2)$, then $\sum_{i=1}^{m} X_i \sim N\left(\sum_{i=1}^{m} \mu_i, \sum_{i=1}^{m} \sigma_i^2\right)$
2. $X_i \sim N(\mu_i, \sigma_i^2)$, then

$$a_0 + \sum_{i=1}^{m} a_i X_i \sim N\left(a_0 + \sum_{i=1}^{m} a_i \mu_i, \sum_{i=1}^{m} a_i^2 \sigma_i^2\right) \quad \text{for } a_i \in \mathbb{R}.$$

Example contd

Example

Total revenue from sales is $Y = 80X_1 + 90X_2 + 60X_3$. Here $\mu_Y = E(Y) = 137000$ and $\sigma_Y^2 = V(Y) = 124840000$. If X_i's are normally distributed, the probability that revenue exceeds 100000 is

$$\begin{aligned} P(Y > 100000) &= P\left(Z > \frac{100000 - 137000}{11173.2}\right) \\ &= P(Z > -3.31) \\ &= \Phi(3.31) \\ &= 0.999533 \end{aligned}$$

Bivariate Normal Distribution (BVN)

$(X, Y) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ if the joint density is given by

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right)}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2\sigma_2^2(1-\rho^2)}\left(y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2},$$

$$-\infty < x < \infty, -\infty < y < \infty, -\infty < \mu_1 < \infty, -\infty < \mu_2 < \infty,$$

$$\sigma_1 \geq 0, \sigma_2 \geq 0, |\rho| < 1.$$

Note that $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$.
Hence $E(X) = \mu_1$, $Var(X) = \sigma_1^2$, $E(Y) = \mu_2$ and $Var(Y) = \sigma_2^2$.

Conditional density of Y given $X = x$ is

$$f_{Y|X=x}(y) = \frac{f_{XY}(x,y)}{f_X(x)}$$
$$= \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2\sigma_2^2(1-\rho^2)}\left(y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2}.$$

Hence $[Y|X = x] \sim N\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \sigma_2^2(1-\rho^2)\right)$.
$E(Y|X = x) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$, $Var(Y|X = x) = \sigma_2^2(1-\rho^2)$.
Similarly $[X|Y = y] \sim N\left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2), \sigma_1^2(1-\rho^2)\right)$
$E(X|Y = y) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)$, $Var(X|Y = y) = \sigma_1^2(1-\rho^2)$.
Also $Cov(X, Y) = \rho\sigma_1\sigma_2$. Hence $\rho = Corr(X, Y)$.

Bivariate Normal Distribution (BVN)

Example

The failure of tube can occur as the result of thermal wear of the internal components. Let X denote the modified life of tube and Y denote the modified thermal wear of the internal components. Let X and Y have a bivariate normal distribution with parameters
$\mu_X = 3,\ \mu_Y = 1,\ \sigma_X^2 = 16,\ \sigma_Y^2 = 25$ and $\rho = 3/5$. Determine the probabilities $P(-3 < X < 3)$ and $P(-3 < X < 3|Y = -4)$.

Solution: (1). $X \sim N(3, 16)$, therefore
$P(-3 < X < 3) = P\left(\frac{-3-3}{4} < Z < 0\right) = P(-1.5 < Z < 0) = 0.433.$
(2). $[X|Y = -4] \sim N\left(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y - \mu_Y), \sigma_X^2(1-\rho^2)\right) \equiv N(0.6, 10.24).$

$$\begin{aligned} P(-3 < X < 3|Y = -4) &= P\left(\frac{-3-0.6}{\sqrt{10.24}} < Z < \frac{3-0.6}{\sqrt{10.24}}\right) \\ &= P(-1.125 < Z < 0.75) \\ &= 0.64. \end{aligned}$$

Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables with mean μ and variance σ^2 then $Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}}$ has approximately $N(0,1)$ distribution for n large.

Example

Hard drives are packed 100 to a packet. Drives weights are independent random variable with mean of 0.5 kg and a standard deviation of 0.10 kg. 30 packets are loaded to a box. Suppose we want to find the probability that the drives on a box will exceed 1510 kg in weight. Neglecting both packet and box weight. Let X_i be the weight of ith hard drive $i = 1, 2, \ldots, 3000$. Then total weight is $X = X_1 + \cdots + X_{3000}$. $E(X) = 3000 \times 0.5 = 1500$ and $Var(X) = 3000 \times (0.10)^2 = 30$. Then using CLT, the required solution is

$$\begin{aligned} P(X > 1510) &= P\left(Z > \frac{1510 - 1500}{\sqrt{30}}\right) \\ &= P(Z > 1.83) \\ &= 1 - \Phi(1.83) \\ &= 1 - 0.96637 = 0.03363, \end{aligned}$$

here $Z = \frac{X-1500}{\sqrt{30}} \sim N(0, 1)$.

Law of large numbers

Let $X_1, \ldots, X_n$ be independent random variable with common mean μ and common variance σ^2. Let $S_n = \sum_{i=1}^n X_i$. then for any $\epsilon > 0$,

$$P(|\frac{S_n}{n} - \mu| < \epsilon) \to 1,$$

as $n \to \infty$.

Summary

In this chapter we presented the sum of random variables, bivariate normal distribution, central limit theorem and law of large numbers.

Random variable $\quad E \ (\Omega, \mathcal{F})$

A rule X which assigns a real value $X(\omega)$ to each $\omega \in \Omega$ (sample space) is called a r.v., i.e., X is a fn whose domain is sample space Ω of outcomes ω and whose range is (some subset of) the real numbers.

Example $\quad E$: Toss the coin two times

$$\Omega = \{\underbrace{HH}_{\omega_1}, \underbrace{HT}_{\omega_2}, \underbrace{TH}_{\omega_3}, \underbrace{TT}_{\omega_4}\}$$

$$\mathcal{F} = \mathbb{P}(\Omega)$$

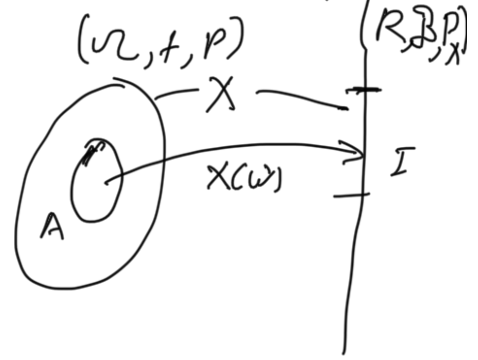
X : counting # of heads

Borel σ-field (containing subintervals of real line)

$$X(\omega_1) = 2 \ , \ X(\omega_2) = 1 = X(\omega_3) \ , \ X(\omega_4) = 0$$

$$A_1 = \{\omega_2, \omega_3\} \equiv I = \{1\}$$

$$A_2 = \{\omega_2, \omega_3, \omega_4\} \equiv I = [0, 1]$$



Event A

$$P(A) = \underline{P(\omega : X(\omega) \in I]} = P_X(I)$$

$$A \equiv I$$

equivalent event

X is r.v. if

$$\{\omega : X(\omega) \in I\} \in \mathcal{F} \ (\text{is an event})$$

$$\equiv \{\omega : X(\omega) \leq x\} \in \mathcal{F}$$

$\equiv \quad X^{-1}(-\infty, x] \in \mathcal{F}$ is an event

(CDF) Cummulative distribution function or distribution function : $(\Omega, \mathcal{F}, P)$

For event $[X \le x] \equiv \{\omega : X(\omega) \le x\}$, we define

$$F_X(x) = P(X \le x) = P(\underbrace{\{\omega : X(\omega) \le x\}}) \quad , x \in \mathbb{R}$$
$$= P_X((-\infty, x])$$

Example: $\Omega = \{HH, HT, TH, TT\}$

X : # of heads $\qquad \mathcal{F} = \mathbb{P}(\Omega)$

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

prob. mass function

pmf

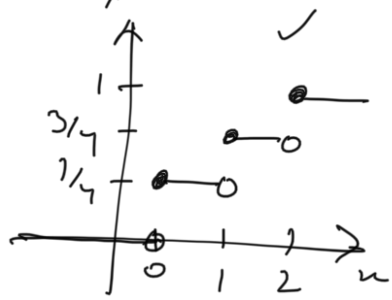
$\hookrightarrow p_X(x) = P(X = x)$

X	0	1	2
$p_X(x) = P(X=x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

CDF

$$F_X(x) = P(X \le x) = \sum_{x_i \le x} p(x_i) \qquad \text{discrete type r.v. } X$$

$$= \begin{cases} 0 & , x < 0 \\ \frac{1}{4} & , 0 \le x < 1 \\ 3/4 & , 1 \le x < 2 \\ 1 & ; x \ge 2 \end{cases}$$



Example A cathode ray tube is aged to failure

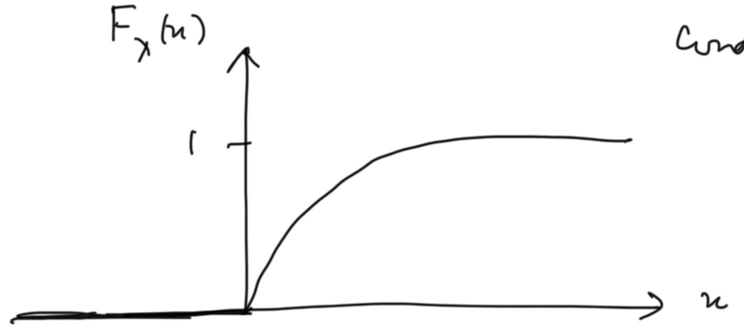
$\Omega = \{t \mid t \ge 0\} = [0, \infty)$

X : life of cathode ray tube till failure

CDF

$$F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases}, \quad \lambda > 0$$

continuous case ✓



pdf (prob. density fn)

$$\to f_X(x) = \frac{d}{dx} F_X(x)$$

<u>Properties of CDF</u> (i) $0 \leq F_X(x) \leq 1, \ \forall x$

(ii) $F_X(x)$ is non-decreasing in x, i.e.,

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

(iii) $F_X(x)$ is right continuous, i.e., $F_X(x+) = F_X(x)$

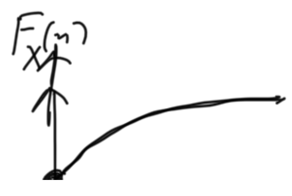
$$F_X(x+) = \lim_{\delta \to 0} F_X(x+\delta) = F_X(x)$$

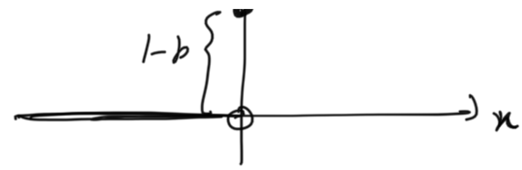
(iv) $\lim_{x \to +\infty} F_X(x) = 1, \quad \lim_{x \to -\infty} F_X(x) = 0$

$$P(X = a) = F_X(a) - F_X(a-)$$

— x —

<u>mixed type r.v.</u>





Example Let X be the lifetime of an instrument which may fail immediately on installation with prob. $(1-p)$ or it may live upto age x with prob. $p(1-e^{-\lambda x})$, $x>0$, $\lambda>0$

$$F_X(x) = P(X \le x)$$

$$= \begin{cases} 0 & , x<0 \\ (1-p) + p(1-e^{-\lambda x}) & , x \geq 0 \end{cases}$$

X mixed type r.v.

—x—

$(1,2,3,4)$ $\quad \dfrac{1+2+3+4}{4} = 1\times\frac{1}{4} + 2\times\frac{1}{4} + 3\times\frac{1}{4} + 4\times\frac{1}{4}$

$1\times p_1 + 2\times p_2 + 3\times p_3 + 4\times p_4$

$\sum_{i=1}^{4} p_i = 1$

Discrete distⁿ

Benoulli trail $\quad$ Ber(p)

$$1 \text{ trail} \begin{cases} \nearrow S & p \\ \searrow F & q \end{cases} \qquad p+q=1$$

X counts # of successes $\in \{0,1\}$

$P(X=0) = q \quad ; P(X=1) = p$

✓ $E(X) = \sum_{x=0}^{1} x p(x) = 0 \times q + 1 \times p = p$

$E(X^2) = \sum x^2 p(x) = 0 \times q + 1^2 \times p = p$

$V(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1-p) = pq$

Binomial distn : $X \sim Bin(n, p)$

n indep Bernoulli trail

X : # of successes in n trail $\in \{0, 1, 2, \dots, n\}$

$$P(X=x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & , x = 0, 1, \dots, n \\ 0 & \text{o.w.} \end{cases}$$

$E(X) = np$

$V(X) = npq$

Geometric distn : Let indep. Bernoulli trails are conducted till we observe a success.

X = # of trials to get a success

$$p(x) = P(X=x) = \begin{cases} q^{x-1} p & , x = 1, 2, \dots \\ 0 & \text{o.w.} \end{cases}$$

$q \quad q \; \dots \; q \quad p$

$F \quad F \; \dots \; F \quad S$

$X \sim Geo(p)$

memoryless property

$$P(X > m) = P(X > m+n \mid X > n)$$

Sol

$$P(X > m) = \sum^{\infty} P(X = x) = \sum^{\infty} p q^{x-1}$$

$x=m+1$ $\qquad\qquad\qquad$ $x=m+1$

$$= p\left[q^{m}+q^{m+1}+q^{m+2}+\cdots\right]$$

$$= pq^{m}\left[1+q+q^{2}+\cdots\right] = pq^{m}\times\frac{1}{1-q}$$

$$= \frac{\cancel{p}q^{m}}{\cancel{p}} = q^{m} \checkmark$$

$$P(X>m+n \mid X>n) = \frac{P(X>m+n,\, X>n)}{P(X>n)} = \frac{P(X>m+n)}{P(X>n)}$$

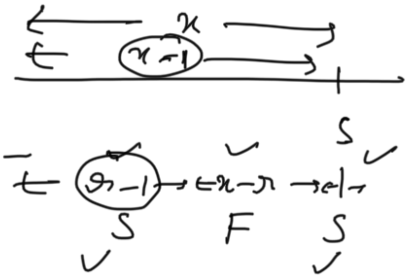
$$= \frac{q^{m+n}}{q^{n}} = q^{m}.$$

<u>Negative Binomial distn:</u> $X \sim NB(r,p)$

Let <u>indep Bernoulli trails are conducted till we have r successes.</u>

Let r.v. X : # of trails required to get <u>r successes</u>

$$P(X=x) = \begin{cases} \binom{x-1}{r-1} p^{r} q^{x-r}, & x=r, r+1, \ldots \\ 0 & \text{o.w.} \end{cases}$$

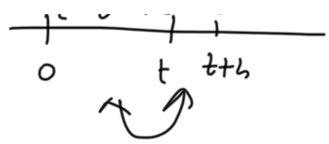


— x —

P.P. ✱ <u>h → small</u>

$$P_0(t+h) = P(N(0,t+h] = 0)$$

← 0 →



$$= P(\{N(0,t]=0\} \cap \{N(t,t+h]=0\})$$

$$= P(N(0,t]=0)\cdot P(N(t,t+h)=0)$$ | indep increments

$$= P_0(t)\cdot P(N(h)=0)$$ | stationary increments

$$= P_0(t)\cdot(1-\lambda h+o(h))$$

$$\Rightarrow \frac{P_0(t+h)-P_0(t)}{h} = -\lambda P_0(t) + \frac{o(h)\,P_0(t)}{h}$$

$h \to 0$

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t)$$

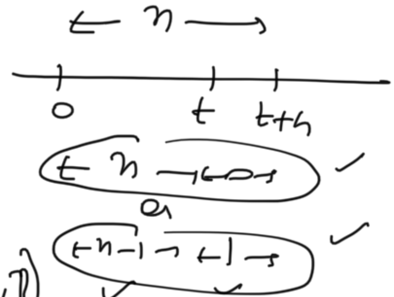
$$P_0(t) = Ce^{-\lambda t}$$

$$P_0(0) = 1 \Rightarrow C = 1$$

$$P_0(t) = e^{-\lambda t}$$

$h \to$ small

$$P_n(t+h) = P(N(0,t+h)=n)$$



$$= P\Big(\{\{N(0,t)=n\} \cap \{N(t,t+h)=0\}\}$$

$$\cup\ \{\{N(0,t)=n-1\} \cap \{N(t,t+h)=1\}\}\Big)$$

$$= P_n(t)\,P_0(h) + P_{n-1}(t)\,P_1(h)$$

$$= P_n(t)\,(1-\lambda h+o(h)) + P_{n-1}(t)\cdot(\lambda h+o(h))$$

$$\Rightarrow \frac{P_n(t+h)-P_n(t)}{h} = -\lambda(P_n(t)-P_{n-1}(t)) + o(h)$$

h $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad h$

$h \to 0$

$$\frac{dP_n(t)}{dt} = -\lambda\left(P_n(t) - P_{n-1}(t)\right)$$

know $\quad P_0(t) = e^{-\lambda t}$ ✓

Assume $\quad P_{n-1}(t) = \dfrac{e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!}$ ✓ $\qquad$ Use mathematical induction ✓

$\longrightarrow \underline{P_n(t)}$

$$\frac{dP_n(t)}{dt} = -\lambda P_n(t) + \lambda \frac{e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!}$$

$$e^{\lambda t}\frac{dP_n(t)}{dt} + \lambda e^{\lambda t} P_n(t) = \frac{\lambda^n t^{n-1}}{(n-1)!}$$

$$\frac{d}{dt}\left(e^{\lambda t} P_n(t)\right) = \frac{\lambda^n}{(n-1)!} t^{n-1}$$

$$e^{\lambda t} P_n(t) = \frac{\lambda^n}{n!} t^n$$

$$\Rightarrow \quad P_n(t) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \ldots$$

— x —

Stochastic Process : (S.P.)

→ A S.P. is a family of r.v.'s $\{X(t), t \in T\}$, defined on a given probability space, indexed by the parameter t, where $t \in T$.

→ The values assumed by the r.v. $X(t)$ are called states and set of all possible values form the state space (S) of the process

types (1) discrete state, discrete parameter/time SP

(2) " " , continuous " / " SP

(3) continuous state, " " / " SP

(4) " " , discrete " / " SP.

Example: Consider a queueing system with jobs arriving at random point in time, queueing for service and departing from the system after service completion.

(a) $X(t)$ # of jobs in the system at time t.

$$S = \{0, 1, 2, 3, ---\} \quad , T = \{t ; \ t \geq 0\}$$

$\{X(t)\}$ discrete state, continuous parameter SP.

(b) W_k time that the k^{th} customer has to wait in the system before receving service.

$$S = \{x, \; x \geq 0\}, \quad T = \{1, 2, 3, --\}$$

$\{W_k, k \in T\}$ continuous state, discrete parameter SP.

(c) $Y(t)$ cummulative service requirement (expulance) of all jobs in the system at time t.

$$S = [0, \infty), \quad T = [0, \infty)$$

$\{Y(t)\}$ cont. state, cont. parameter SP.

(d) Let N_k # of jobs in the system at the time of the departure of the k^{th} customer (after service completion).

$$S = \{0, 1, 2, ..\}, \quad T = \{1, 2, 3, --\}$$

$\{N_k, k \in T\}$ discrete state, discrete parameter S.P.

—x—

Discrete Time Markov Chain (DTMC):

discrete state, discrete parameter/time SP

SP $\{X_n, n = 0, 1, 2, --\}$ that takes on a finite or countable number of possible values.

$X_n = i \equiv$ process is in state i at time/step/transition n

$\{X_n\}$ DTMC if

$\{i, j, i_0, i_1, .. -\} \in S$ State space

$$P(X_{n+1}=j \mid X_0=i_0, X_1=i_1, \dots, X_n=i)$$
$$= P(X_{n+1}=j \mid X_n=i)$$
$$= p_{ij}^{(1)}(n) \rightarrow \text{transition probability}$$
$$= p_{ij}^{(1)} \rightarrow \text{stationary transition probability}$$
(Homogeneous M.C.)
$$= p_{ij}$$
$$\left(= P(X_{m+1}=j \mid X_m=i) = P(X_1=j \mid X_0=i\right)$$

$S = \{0, 1, 2, \dots\}$

$$p_{ij}^{(1)} = p_{ij} = P(\underset{\nearrow \text{final state}}{X_{n+1}=j} \mid \underset{\nearrow \text{initial state}}{X_n=i}), \quad i, j \in S$$

$$P = P^{(1)} = \begin{array}{c|c} & j \longrightarrow \\ i \downarrow & \begin{array}{cccc} 0 & 1 & 2 & \cdots \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \end{array} & \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \end{array}$$

Transition Probability Matrix (TPM)

$0 \le p_{ij} \le 1, \ \forall i, \forall j$

for fixed i

$$\sum_j p_{ij} = 1$$

$$p_{ij} = p_{ij}^{(1)}$$

Example (1) Consider a game of ladder climbing. There are 5 levels in the game, level 1 is lowest (bottom) and level 5 is the highest (top). A player start at the bottom. Each time, a fair coin is tossed.

If it turns up heads, the player moves up one rung.
If tails, the player moves down to the very bottom.
Once at the top level, the player moves to the very bottom if tails turns up and stays at the top if head turns up.

Let X_n be the level of the game in the n^{th} step / transition.

Find S, TPM

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

$\{X_n\}$ DTMC

$p_{ij} = P(X_{n+1} = j \mid X_n = i)$

tpm

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \end{array}$$

(2) Let $\{X_n\}_{n=0,1,2,\ldots}$ be a sequence of i.i.d. discrete r.v. with $P(X_1 = j) = \left(\frac{1}{2}\right)^{j+1}$ $\forall j = 0, 1, 2, \ldots$.

Determine whether each of the following chain is

Markovian or not. If so find its corresponding state space (S) and tpm.

(i) $\{S_n\}_{n=0,1,2,\dots}$ where $S_n = \sum_{i=1}^{n} X_i$

(ii) $\{M_n\}_{n=0,1,2,\dots}$, where $M_n = \max\{X_1, X_2, \dots, X_n\}$

Sol (i) $S_n = \sum_{i=1}^{n} X_i$

$\{S_n\}$ DTMC with $S = \{0, 1, 2, 3, \dots\}$

$p_{ij} = P(S_{n+1} = j \mid S_n = i)$

tpm

$$P = \begin{array}{c|cccccc} i \downarrow \backslash j \rightarrow & 0 & 1 & 2 & 3 & \cdots & \\ \hline 0 & 1/2 & 1/4 & 1/8 & 1/16 & \cdots & \\ 1 & 0 & 1/2 & 1/4 & 1/8 & \cdots & \\ 2 & 0 & 0 & 1/2 & 1/4 & \cdots & \\ 3 & 0 & 0 & 0 & 1/2 & \cdots & \\ \vdots & \cdots & & & & & \end{array}$$

(ii)

Ans $S = \{0, 1, 2, \dots\}$

$$P = \begin{bmatrix} 1/2 & 1/4 & 1/8 & 1/16 & \cdots \\ 0 & 3/4 & 1/8 & 1/16 & \cdots \\ 0 & 0 & 7/8 & 1/16 & \cdots \\ \cdots & & & & \end{bmatrix}$$

Example (transformation of a process into M.C.)

Suppose that whether or not it rains today depends on previous weather conditions through the last two days. Suppose that if it has rained for the past two days, then it will rain tommorow with prob. 0.7; if it has rained today but not yesterday, then it will rain tommorow with prob. 0.5; if it has rained yesterday but not today, then it will rain tommorow with prob. 0.4; if it has not rained in the past two days, then it will rain tommorow with prob. 0.2.

Let Y_n : weather condition on n^{th} day.

$$Y_n \mid Y_{n-1}, Y_{n-2}, \cancel{Y_{n-3}} \quad \text{not M.C}$$

Let X_n : State at any time is determined by the weather condition during both that day and the previous day

$$\underset{\text{tom, today}}{X_{n+1}} \mid \underset{\text{today, yest}}{X_n}, \cancel{X_{n-1}}$$

X_n DTMC

State X_n	Rained yesterday	Rained today
0	✓	✓
1	×	✓
2		

$p_{ij} = P(X_{n+1} = j \mid X_n = i)$

2	✓	×
3	×	×

$i, j \in S = \{0,1,2,3\}$

$X_{n+1} = j$

tprograph (X_n)

$$P = \begin{array}{cc} & \begin{array}{cccc} \checkmark\checkmark & \times\checkmark & \checkmark\times & \times\times \\ 0 & 1 & 2 & 3 \end{array} \\ \begin{array}{c} \checkmark\checkmark\ 0 \\ \times\checkmark\ 1 \\ \checkmark\times\ 2 \\ \times\times\ 3 \end{array} & \begin{bmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{bmatrix} \end{array}$$

yesterday today (rows); today tomorrow (columns)

$X_n = i$

—x—

n- step transition probability

$$p_{ij}^{(n)} = P(X_{m+n} = j \mid X_m = i) = P(X_n = j \mid X_0 = i)$$

$$0 \le p_{ij}^{(n)} \le 1 \quad , \forall i, \forall j \qquad (X_n) \text{ DTMC}$$

for fixed i $\qquad S = \{0, 1, 2, - -\}$ State space

$$\sum_j p_{ij}^{(n)} = 1$$

$$P^{(n)} = \left(\left(p_{ij}^{(n)} \right) \right) = \begin{array}{cc} & \begin{array}{cccc} 0 & 1 & 2 & \cdots \end{array} \ j \rightarrow \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \end{array} \ i \downarrow & \begin{bmatrix} p_{00}^{(n)} & p_{01}^{(n)} & p_{02}^{(n)} & \cdots \\ p_{10}^{(n)} & p_{11}^{(n)} & p_{12}^{(n)} & \cdots \\ p_{20}^{(n)} & p_{21}^{(n)} & p_{22}^{(n)} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \end{array}$$

Chapman Kolomgrov equations :

$$p_{ij}^{(m+n)} = \sum p^{(m)} p^{(n)} = \sum p^{(n)} p^{(m)}$$

$$p_{ij} = \sum_k p_{ik}\, p_{kj} = \sum_k p_{ik}\, p_{kj} \qquad i,j,k \in S$$

Sol: $(i,j)^{th}$ element of $P^{(m+n)}$

$$p_{ij}^{(m+n)} = P(X_{m+n}=j \mid X_0=i)$$

$$= \sum_k P(X_{m+n}=j, X_n=k \mid X_0=i)$$

thm of total prob.

$$P(AB|C) = \frac{P(ABC)}{P(C)} \times \frac{P(BC)}{P(BC)} = P(A|BC)P(B|C)$$

$$= \sum_k P(\underline{X_{m+n}=j} \mid \underline{X_n=k}, \underline{X_0=i}) \cdot P(X_n=k \mid X_0=i)$$

$$= \sum_k P(X_{m+n}=j \mid X_n=k)\, P(X_n=k \mid X_0=i)$$

$\{X_n\}$ DTMC

$$= \sum_k p_{kj}^{(m)}\, p_{ik}^{(n)} = \sum_k p_{ik}^{(n)}\, p_{kj}^{(m)}$$

$$p_{ij}^{(m+n)} = \sum_k p_{ik}^{(n)}\, p_{kj}^{(m)}$$

$$P^{(m+n)} = \begin{array}{c} P^{(n)} \\ \begin{array}{c} \\ i \end{array}\begin{pmatrix} - & - \\ p_{i0}^{(n)} & p_{i1}^{(n)} \cdots \end{pmatrix} \end{array} \begin{array}{c} P^{(m)} \\ j \\ \begin{pmatrix} \vdots & p_{0j}^{(m)} & \\ & p_{1j}^{(m)} & \\ \vdots & \vdots & \end{pmatrix} \end{array}$$

$$P^{(m+n)} = P^{(n)} P^{(m)}$$

$$P^{(1)} = P$$

$$P^{(2)} = P^{(1)} P^{(1)} = P.P = P^2$$

$$P^{(n)} = P^n$$

Example Suppose that the chance of rain tomorrow depends on previous weather conditions only through whether or not it is raining today and not on past weather condition. Suppose also that if it rains today, then it will rain tomorrow with probability α; and if it does not rain today, then it will rain tomorrow with prob. β.

X_n : weather condition on n^{th} day DTMC

$X_n \in \{0,1\}$ 0 : rain , 1 : not raining

$S = \{0,1\}$ $p_{ij} = P(X_{n+1}=j \mid X_n=i)$

$i,j \in S$

tpm

$$P = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{pmatrix} \end{array}$$

Let $\alpha = 0.7$ and $\beta = 0.3$. Calculate the prob. that it will rain 2 days from today given that it is raining today.

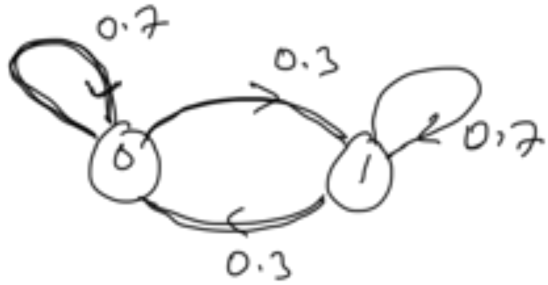
$$P = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix} \end{array} = \left(\left(p_{ij}^{(1)}\right)\right)$$

$$p_{00}^{(2)} = P(X_{n+2}=0 \mid X_n=0) = ?$$

$$P^{(2)} = P^2 = P.P = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix}\begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix}$$

$$= \begin{array}{c|cc} & 0 & 1 \\ 0 & \boxed{0.58} & 0.42 \\ 1 & 0.42 & 0.58 \end{array}$$

$$p_{00}^{(2)} = 0.58$$



$$p_{00}^{(2)} = 0.7 \times 0.7 + 0.3 \times 0.3 = 0.58$$

—x—

PMF of step X_n: $\{X_n\}$ DTMC

state space $S = \{0, 1, 2, --\}$

$$p_i^{(n)} = P(X_n = i) \qquad , i \in S$$

pmf of X_n is

$$\underset{\sim}{p}^{(n)} = \left(p_0^{(n)}, p_1^{(n)}, --, p_i^{(n)}, ---\right) \checkmark$$

$$\text{s.t. } \sum_i p_i^{(n)} = 1$$

initial state pmf of X_0 is

$$\underset{\sim}{p}^{(0)} = \left(p_0^{(0)}, p_1^{(0)}, ---\right) \qquad \text{s.t. } \sum_i p_i^{(0)} = 1$$

T.S. $\qquad \boxed{\underset{\sim}{p}^{(n)} = p^{(n-1)} P} \qquad P \leftarrow \text{tpm}$

$$\left(p_0^{(n)}, p_1^{(n)}, --, \boxed{p_i^{(n)}}, --\right) = \left(\underline{p_0^{(n-1)}, p_1^{(n-1)}, --p_i^{(n-1)}}, --\right)\begin{pmatrix} p_{00} p_{01} -- p_{0i} -- \\ p_{10} p_{11} -- p_{1i} -- \end{pmatrix}$$

T. S. $\boxed{p_i^{(n)} = \sum_k p_k^{(n-1)} p_{ki}}$

Sol $p_i^{(n)} = P(X_n = i) = \sum_k P(X_n = i, X_{n-1} = k)$

$= \sum_k P(X_n = i \mid X_{n-1} = k) \cdot P(X_{n-1} = k)$ Using thm of total prob.

$= \sum_k p_{ki}\, p_k^{(n-1)}$

$\Rightarrow \boxed{\underset{\sim}{p}^{(n)} = \underset{\sim}{p}^{(n-1)} P}$

$\underset{\sim}{p}^{(n)} = \underset{\sim}{p}^{(n-2)} P^2$

$\underset{\sim}{p}^{(n)} = \underset{\sim}{p}^{(0)} P^n$ $P \to$ tpm

$\cup A_i = \Omega$

$A_i \cap A_j = \phi$

$P(E) = \sum_i P(A_i \cap E)$

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

$\Leftrightarrow P(AB) = P(A|B)\, P(B)$

Example: Let (X_n) DTMC, with state space $S = \{1, 2, 3\}$, and

$$P = \begin{array}{c} \text{tpm} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{pmatrix} \end{array}$$

pmf of X_0 is $(0.7, \boxed{0.2}, 0.1) = \underset{\sim}{p}^{(0)}$

$P(X_0 = 2)$

(i) $P(X_0 = 2, X_2 = 3, X_1 = 3, X_3 = 2)$

$= P(X \ldots$

$= P(X_3=2, X_2=3, X_1=3, X_0=2)$

$= P(X_3=2 \mid X_2=3, X_1=3, X_0=2) \cdot P(X_2=3 \mid X_1=3, X_0=2)$

$\cdot P(X_1=3 \mid X_0=2) \cdot P(X_0=2)$

$P(ABCD) = P(A \mid BCD)\, P(B \mid CD)\, P(C \mid D)\, P(D)$

(X_n) DTMC

$= P(X_3=2 \mid X_2=3) \cdot P(X_2=3 \mid X_1=3) \cdot P(X_1=3 \mid X_0=2) \times 0.2$

$= p_{32}\; p_{33}\; p_{23} \times 0.2$

$= 0.4 \times 0.3 \times 0.2 \times 0.2 = 0.0048$

(ii) $P(X_2=3, X_1=3 \mid X_0=2)$

$P(AB \mid C) = \frac{P(ABC)}{P(C)} \times \frac{P(BC)}{P(BC)}$

$= P(A \mid BC) \cdot P(B \mid C)$

$= P(X_2=3 \mid X_1=3, X_0=2)$

$\cdot P(X_1=3 \mid X_0=2)$

(X_n) DTMC

$= P(X_2=3 \mid X_1=3) \cdot P(X_1=3 \mid X_0=2)$

$= p_{33}\; p_{23} = 0.3 \times 0.2 = 0.06$

(iii) $P(X_2=3)$

X_2 $\quad \underset{\sim}{p}^{(2)} = \left(p_1^{(2)}, p_2^{(2)}, p_3^{(2)}\right)$

$p_3^{(2)} \to P(X_2=3)$

$\underset{\sim}{p}^{(n)} = \underset{\sim}{p}^{(n-1)} P$ $\quad P \leftarrow$ tpm

X_1 $\quad \underset{\sim}{p}^{(1)} = \underset{\sim}{p}^{(0)} P = (0.7, 0.2, 0.1) \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{pmatrix}$

$$= (0.22, 0.43, 0.35)$$

X_2

$$\underset{\sim}{p}^{(2)} = \underset{\sim}{p}^{(1)} P = (0.22, 0.43, 0.35) \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{pmatrix}$$

$$= (0.385, 0.336, \boxed{0.279})$$

$$P(X_2 = 3) = p_3^{(2)} = 0.279$$

—x—

<u>Classification of states:</u> $\{X_n\}$ DTMC

$p_{ij}^{(n)} = P(X_n = j \mid X_0 = i)$ $\quad S = \{0, 1, 2, --\}$ $\quad i, j \in S$

Defn $i \to j$ state j is accessible from state i if $p_{ij}^{(n)} > 0$ for some n.

Defn $i \leftrightarrow j$ state i and j communicate with each other if $i \to j$ and $j \to i$

<u>Result</u> $i \leftrightarrow j$, $j \leftrightarrow k \Rightarrow i \leftrightarrow k$ $\quad , i, j, k \in S$

Sol
Given $\left[\; p_{ij}^{(n)} > 0 \quad p_{jk}^{(m)} > 0 \right.$ $\quad \exists\, n, m$ s.t. $i \to j$, $j \to k$

Now
$$p_{ik}^{(m+n)} = \sum_{l} p_{il}^{(n)} p_{lk}^{(m)} \geq p_{ij}^{(n)} p_{jk}^{(m)} > 0$$

$\therefore i \to k$

IIIly $k \to i$ $\quad \therefore i \leftrightarrow k$

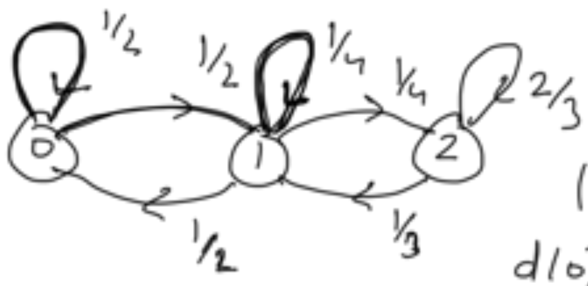
<u>Defn</u> $\{X_n\}$ DTMC is irreducible or indecomposable if every state communicates with every other state,

Otherwise it is reducible.

eg $\{X_n\}$ DTMC tpm

$$\begin{array}{c|ccc} & 0 & 1 & 2 \\ \hline 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 2 & 0 & \frac{1}{3} & \frac{2}{3} \end{array}$$

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



$\{X_n\}$ irreducible/countable

$d(0) = \gcd\{1, 2, 3, \ldots\} = 1$

$d(0) = d(1) = d(2) = 1$

All state has period 1 or aperiodic

$0 \leftrightarrow 1 \leftrightarrow 2$

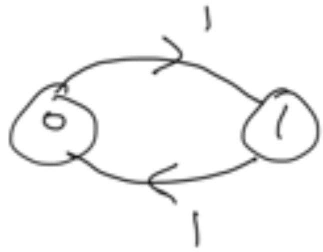
Def period of state i ; $d(i)$;

$d(i)$ is gcd of $I^+ = \{1, 2, 3, \ldots\}$ n s.t $p_{ii}^{(n)} > 0$

(If $p_{ii}^{(n)} = 0 \ \forall n \geq 1$, define $d(i) = 0$)

eg $\{X_n\}$ DTMC tpm

$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$



$d(0) = \gcd\{2, 4, 6, \ldots\} = 2 = d(1)$

eg

$$\begin{array}{c|cccccccc} & \cdots & -2 & -1 & 0 & 1 & 2 & \cdots \\ \hline \vdots & & & & & & & \\ -2 & \cdots & 0 & 1 & 0 & 0 & 0 & \cdots \\ -1 & \cdots & 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & \cdots & 0 & 0 & 0 & 1 & 0 & \cdots \end{array}$$

2

$d(i) = 0$

—x—

(X_n) DTMC, $\quad i \in S$

$$f_i^{(n)} \text{ or } f_{ii}^{(n)} = P(X_n = i,\ X_k \neq i, k=1,2,\dots,n-1 \mid X_0 = i)$$: prob of first visit to state i in n transitions / steps starting from state i

recurrence time prob

$$f_{ii}^{(0)} = 1$$

$$f_{ii} \text{ or } f_i = f_i^{(1)} + f_i^{(2)} + f_i^{(3)} + \cdots = \sum_{n=1}^{\infty} f_i^{(n)}$$

prob. of ever visiting state i

Def^n $f_i = 1$, return to state i is certain, starting from state i

i recurrent state

Def^n $f_i < 1$, return to state i is uncertain

i transient state.

—x—

Let $\quad I_n = \begin{cases} 1 & \text{if } X_n = i \\ 0 & \text{if } X_n \neq i \end{cases}$

$\sum^{\infty} I_n$: # of time periods ...

$\sum_{n=1} I_n$: # of time periods, the process is in state i

$$E\left(\sum_{n=1}^{\infty} I_n \mid X_0=i\right) = \sum_{n=1}^{\infty} E(I_n \mid X_0=i)$$

$$= \sum_{n=1}^{\infty}\left[1.P(X_n=i \mid X_0=i) + 0\,P(X_n\neq i \mid X_0=i)\right]$$

$$= \sum_{n=1}^{\infty} P(X_n=i \mid X_0=i)$$

$$= \sum_{n=1}^{\infty} p_{ii}^{(n)}$$

i recurrent $\Leftrightarrow \sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty \Leftrightarrow f_i = 1$

i transient $\Leftrightarrow \sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty \Leftrightarrow f_i < 1$

<u>Defn</u> Let i recurrent

m_{ii} or $m_i = \sum_{n=1}^{\infty} n f_i^{(n)}$: mean recurrence time (expected time the process return to state i)

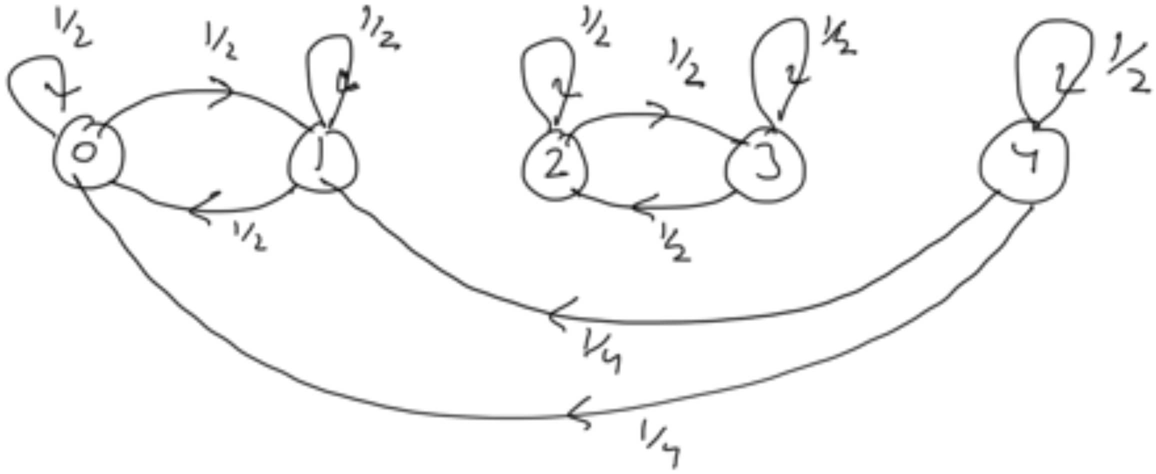
→ If $m_i = \infty$ $\quad$ i null recurrent

→ If $m_i < \infty$ $\quad$ i non-null recurrent/positive recurrent

<u>Example</u>: $\{X_n\}$ DTMC $\quad S=\{0,1,2,3,4\}$ tpm

$$P = \begin{array}{c|ccccc} & 0 & 1 & 2 & 3 & 4 \\ \hline 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 1 & 1/2 & 1/2 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1/2 & 1/2 & 0 \end{array}$$

$$\begin{array}{c} 3 \\ 4 \end{array}\begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$



Classes $\{0,1\}$ $\{2,3\}$ $\{4\}$ $\qquad 0 \leftrightarrow 1 \quad 2 \leftrightarrow 3 \quad 4$

recurrent aperiodic, recurrent aperiodic, transient — Reducible M.C

$$f_0 = f_0^{(1)} + f_0^{(2)} + f_0^{(3)} + \cdots$$

$$= \frac{1}{2} + \frac{1}{2}\times\frac{1}{2} + \frac{1}{2}\times\frac{1}{2}\times\frac{1}{2} + \cdots$$

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

State 0 recurrent

—x—

P1 $\quad \underbrace{i \leftrightarrow j, \; i \text{ recurrent}}$ $\Rightarrow j$ recurrent

Given $\exists\, n, m$ $\quad p_{ij}^{(n)} > 0,\; p_{ji}^{(m)} > 0;\; \sum_{\nu=1}^{\infty} p_{ii}^{(\nu)} = \infty$

$$p_{jj}^{(m+n+v)} \geq p_{ji}^{(n)} p_{ii}^{(v)} p_{ij}^{(m)} \quad \Big] \text{using C.K. equation}$$

$$\sum_v p_{jj}^{(m+n+v)} \geq p_{ji}^{(n)} p_{ij}^{(m)} \sum_v p_{ii}^{(v)}$$

∞

$$\Rightarrow \sum_n p_{jj}^{(n)} = \infty$$

, i.e. j recurrent state

P2 $\quad$ i transient, $i \leftrightarrow j \Rightarrow j$ transient

Sol $\quad$ On contrary suppose j recurrent, since $i \leftrightarrow j$

$\Rightarrow i$ recurrent (Using P1)

a contradiction

P3 $\quad$ In a finite state M.C. all states can not be transient.

P4 $\quad$ In a finite state, irreducible M.C. all states are recurrent.

Sol $\quad$ using P3, P1

P5 $\quad$ In irreducible M.C, all states are recurrent or transient.

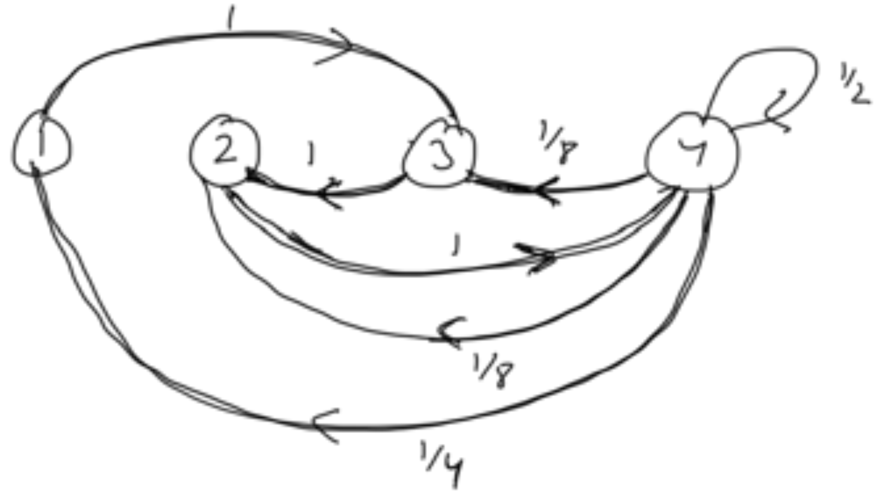
Example: $\{X_n\}$ DTMC $\quad S = \{1,2,3,4\}$

tpm

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \left[\begin{array}{cccc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ \end{array}\right. \end{array}$$

Classify the states

$$4\begin{bmatrix} \frac{1}{4} & \frac{1}{8} & \frac{1}{8} & \frac{1}{2} \end{bmatrix}$$



$1 \leftrightarrow 2 \leftrightarrow 3 \leftrightarrow 4$ Irreducible M.C.

Class $\{1,2,3,4\}$

Finite state, irreducible M.C all states are positive recurrent. (using P4)

mean recurrence time for state 4

$$m_4 = \sum_{n=1}^{\infty} n f_4^{(n)}$$

$$= 1\times\frac{1}{2} + 2\times\frac{1}{8} + 3\times\frac{1}{8} + 4\times\frac{1}{4} + 0 + 0 \ldots$$

$$= \frac{17}{8} < \infty$$

State 4 positive recurrent / non-null recurrent.

$f_4^{(1)} = \frac{1}{2}$, $f_4^{(2)} = \frac{1}{8}\times 1$,

$f_4^{(3)} = \frac{1}{8}\times 1\times 1$

$f_4^{(4)} = \frac{1}{4}\times 1\times 1\times 1$

$f_4^{(5)} = 0$

—x—

Gambler's Ruin problem

$i = 0, 1, 2, \ldots, N$

initial capital Rs i Aim Rs N

Z_i : i^{th} bet / step / transition / time

$Z_1, Z_2, \ldots$ are independent

$P(Z_i = 1) = p$, $P(Z_i = -1) = 1-p = q$

$X_n = Z_1 + Z_2 + \dots + Z_n + i$

: fortune of the gambler after n steps

$p_{ij} = P(X_{n+1} = j \mid X_n = i)$; $i, j \in S = \{1, 2, \dots, N\}$

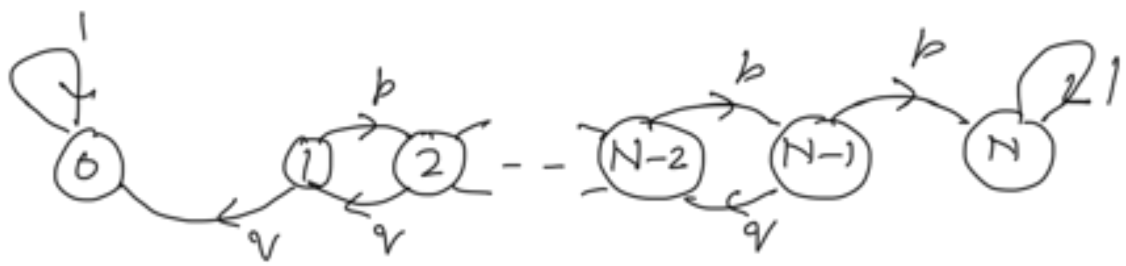
$\{X_n\}_{n=1,2,\dots}$ DTMC

$p_{0,0} = 1 = p_{N,N}$

$p_{i,i+1} = p$, $p_{i,i-1} = q$, $i = 1, \dots, N-1$

tpm

$$P = \begin{array}{c|c} & j \to \\ & \begin{array}{cccccc} 0 & 1 & 2 & \cdots & & N \end{array} \\ \begin{array}{c} i \downarrow \\ 0 \\ 1 \\ 2 \\ \vdots \\ N \end{array} & \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ q & 0 & p & \cdots & 0 \\ 0 & q & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \end{array}$$



Classes $\{0\}$ $\{1, 2, \dots, N-1\}$ $\{N\}$

recurrent /absorbing ; $\underbrace{}_{\text{transient}}$; recurrent/ absorbing

T_0 = time he broke

$= \inf\{n : X_n = 0\}$

T_N = time he has Rs N

$$= \inf \{ n : X_n = N \}$$

$P_i = P(T_N < T_0)$: prob. that starting with Rs i, the gamblers fortune will reach N before reaching 0 ?

$$P_i = P(T_N < T_0 | Z_1 = 1) P(Z_1 = 1) + P(T_N < T_0 | Z_1 = -1) P(Z_1 = -1)$$

$$P_i = P_{i+1} p + P_{i-1} q$$

$$p P_i + q P_i = P_{i+1} p + P_{i-1} q$$

$$\Rightarrow (P_{i+1} - P_i) p = q (P_i - P_{i-1})$$

$$\Rightarrow P_{i+1} - P_i = \frac{q}{p} (P_i - P_{i-1})$$

$P_0 = 0, P_N = 1$

$i = 1 \qquad P_2 - P_1 = \frac{q}{p} (P_1 - P_0) = \frac{q}{p} P_1$

$i = 2 \qquad P_3 - P_2 = \frac{q}{p} (P_2 - P_1) = \left(\frac{q}{p}\right)^2 P_1$

$\vdots$

$$P_i - P_{i-1} = \left(\frac{q}{p}\right)^{i-1} P_1$$

$$P_i - P_1 = P_1 \left[\frac{q}{p} + \left(\frac{q}{p}\right)^2 + \cdots + \left(\frac{q}{p}\right)^{i-1} \right]$$

$$P_i = P_1 \left[1 + \left(\frac{q}{p}\right) + \left(\frac{q}{p}\right)^2 + \cdots + \left(\frac{q}{p}\right)^{i-1} \right]$$

$$= \left[1 - \left(\frac{q}{p}\right)^i \right.$$

$$P_i = \begin{cases} \dfrac{1-\left(\frac{q}{p}\right)^i}{1-\frac{q}{p}} P_1 & , \ \frac{q}{p} \neq 1 \\ i P_1 & , \ \frac{q}{p} = 1 \end{cases} \quad \text{--- (1)}$$

$$P_N = 1$$

If $i = N$ in (1), we have
$$P_1 = \begin{cases} \dfrac{1-\frac{q}{p}}{1-\left(\frac{q}{p}\right)^N} & , \ \frac{q}{p} \neq 1 \\ \dfrac{1}{N} & , \ \frac{q}{p} = 1 \end{cases} \quad \text{--- (2)}$$

using (2) in (1)

$$P_i = \begin{cases} \dfrac{1-\left(\frac{q}{p}\right)^i}{1-\left(\frac{q}{p}\right)^N} & , \ \frac{q}{p} \neq 1 \Leftrightarrow p \neq \frac{1}{2} \\ \dfrac{i}{N} & , \ \frac{q}{p} = 1 \Leftrightarrow p = \frac{1}{2} \end{cases}$$

If $N \to \infty$, then
$$P_i = \begin{cases} 1-\left(\frac{q}{p}\right)^i & , \ \frac{q}{p} < 1 \Leftrightarrow p > \frac{1}{2} \\ 0 & , \ \frac{q}{p} \geq 1 \Leftrightarrow p \leq \frac{1}{2} \end{cases}$$

—x—

Example (assign 1) A rat is put into the linear maze as shown below

0	1	2	3	4	5
Shock					Food

each step $1/4 \leftarrow \quad \rightarrow 3/4$

Sol P(rat finds the food before getting shocked)

Gamblers ruin problem

$i=2\ ,\ N=5\ ,\ p=\frac{3}{4} \qquad \frac{q}{p}=\frac{1}{4}\times\frac{4}{3}=\frac{1}{3}\neq 1$

$$P_5 = \frac{1-\left(\frac{q}{p}\right)^{i}}{1-\left(\frac{q}{p}\right)^{N}} = \frac{1-\left(\frac{1}{3}\right)^{2}}{1-\left(\frac{1}{3}\right)^{5}} = 0.892$$

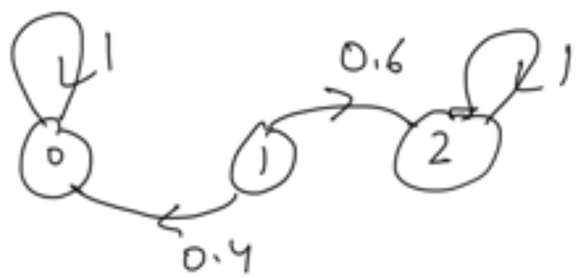
$1-P_5 = P$(rat get shocked before getting food)

Example (o2amp) $\{X_n\}$ DTMC tpm

$$\begin{array}{c|ccc} & 0 & 1 & 2 \\ \hline 0 & 1 & 0 & 0 \\ 1 & 0.4 & 0 & 0.6 \\ 2 & 0 & 0 & 1 \end{array}$$

Starting with 1, determine the prob. that the M.C ends in state 0.

Sol.



absorbing/recurrent

$[0]\ \ [1]\ \ [2]$

transient

$p=0.6,\ q=0.4$

$\frac{q}{p}=\frac{4}{6}=\frac{2}{3}\neq 1$

$$\text{Reqd. prob} = 1-\frac{1-\left(\frac{2}{3}\right)^{1}}{1-\left(\frac{2}{3}\right)^{2}} = 1-\frac{\frac{1}{3}}{5/9}$$

$$= 0.4$$

—x—

Q. $x,y \in S \qquad x \leftrightarrow y$, then $d(x)=d(y)$

Sol. $d(x) = \gcd\{n \geq 1 : p_{xx}^{(n)} > 0\}$

Since $x \leftrightarrow y$

$p_{xy}^{(m)} > 0, \quad p_{y,x}^{(n)} > 0$ for some m, n

$$p_{y,y}^{(n+m)} \geq p_{y,x}^{(n)}\, p_{x,y}^{(m)} > 0$$

$$p_{yy}^{(n+s+m)} \geq p_{y,x}^{(n)}\, p_{xx}^{(s)}\, p_{xy}^{(m)} > 0$$

$d(y)$ divides both $n+m$ and $n+s+m$

$\therefore$ $d(y)$ divides every s with $p_{xx}^{(s)} > 0$

$\Rightarrow$ $d(y)$ divides gcd of such s

$\Rightarrow$ $d(y)$ divides $d(x)$

Repeat by changing the roles

$x \leftrightarrow y \Rightarrow d(y)$ divides $d(x)$

period $d(x)$ and $d(y)$ divides each other $\Rightarrow$ they must be equal.

— x —

<u>Limiting Probabilities</u>

<u>long run behaviour of M.C.s:</u>

finite state space $\to$ $S = \{0, 1, \ldots, N\}$

<u>Regular TPM</u> Given tpm P is regular if P^k has all its elt > 0.

such tpm or corresponding MC is called regular

The most important fact concerning a regular M.C is the existance of a limiting prob. distn $\underline{\Pi} = (\Pi_0, \ldots, \Pi_N)$ when $\Pi_j > 0 \; \forall j = 0, 1, \ldots, N$, $\sum_j \Pi_j = 1$ and this distn is indep. of initial state.

regular tpm $P = ((p_{ij}))$, we have the convergence

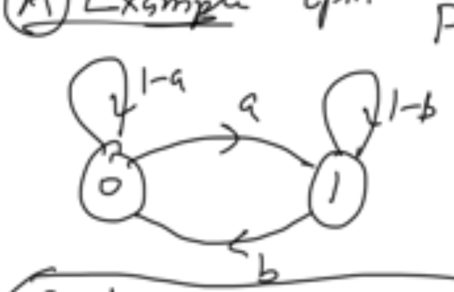
$$\lim_{n\to\infty} p_{ij}^{(n)} = \Pi_j > 0 \quad \forall j = 0, 1, \ldots, N$$

i.e., $$\lim_{n\to\infty} P(X_n = j \mid X_0 = i) = \Pi_j > 0 \quad \forall j = 0, 1, \ldots, N$$

This convergence means that, in long run $(n \to \infty)$, the prob. of finding the M.C in state j is $\simeq \Pi_j$; no matter in which state the chain begins at time 0.

(A) Example tpm

$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1-a & a \\ 1 & b & 1-b \end{array}, \quad 0 \le a, b \le 1$$



here $|1-a-b| < 1$

irreducible, aperiodic, finite state M.C. ~~a = b = 0 or a = b = 1~~ we see separately

P regular tpm

$$P^n ?$$

$p_{00} = 1-a, \; p_{01} = a, \; p_{10} = b, \; p_{11} = 1-b$

$$p_{01}^{(n-1)} = 1 - p_{00}^{(n-1)}$$

$$p_{00}^{(n)} = p_{00}^{(n-1)} p_{00} + p_{01}^{(n-1)} p_{10} = (1-a) p_{00}^{(n-1)} + b(1 - p_{00}^{(n-1)})$$

$$= b + (1 - a - b) p_{00}^{(n-1)}$$

$$= b + (1-a-b)\,p_{00}^{(n-1)} \quad , n > 1$$

$$= b + (1-a-b)\left[b + (1-a-b)\,p_{00}^{(n-2)}\right]$$

$$= b + b(1-a-b) + (1-a-b)^2\,p_{00}^{(n-2)}$$

$$\vdots$$

$$= b + b(1-a-b) + b(1-a-b)^2 + \cdots + b(1-a-b)^{n-2} + p_{00}(1-a-b)^{n-1}$$

$$p_{00} \rightarrow (1-a)$$

$$= b\sum_{k=0}^{n-2}(1-a-b)^k + (1-a)(1-a-b)^{n-1}$$

$$\sum_{k=0}^{n-2}(1-a-b)^k \rightarrow \frac{1-(1-a-b)^{n-1}}{1-(1-a-b)}$$

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

$\therefore$ Similarly

$$\begin{array}{cc} & \begin{array}{cc} \quad 0 \quad & \quad 1 \quad \end{array} \\ P^n = \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} \dfrac{b + a(1-a-b)^n}{a+b} & \dfrac{a - a(1-a-b)^n}{a+b} \\[2ex] \dfrac{b - b(1-a-b)^n}{a+b} & \dfrac{a + b(1-a-b)^n}{a+b} \end{bmatrix} \end{array} \quad (2)$$

Note that $|1-a-b| < 1$, $0 < a, b < 1$

thus $|1-a-b|^n \rightarrow 0$ as $n \rightarrow \infty$

$$\begin{array}{cc} & \begin{array}{cc} \quad 0 \quad & \quad 1 \quad \end{array} \\ \lim_{n\rightarrow\infty} P^n = \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\[2ex] b & \end{bmatrix} \end{array}$$

$$\begin{matrix} 1 \left[\frac{-}{a+b} \right. & \left. \frac{a}{a+b} \right] \end{matrix}$$

Let

$$P(X_0=0)=\frac{1}{3}, \quad P(X_0=1)=\frac{2}{3}$$

initial state pmf $\quad \underset{\sim}{p}^{(0)}=\left(\frac{1}{3}, \frac{2}{3}\right)$

$$\underset{\sim}{p}^{(n)}=\underset{\sim}{p}^{(n-1)}P=\underset{\sim}{p}^{(0)}P^n$$

$$=\left(\frac{1}{3}, \frac{2}{3}\right)P^n$$

Take P^n from (*)

$$\lim_{n\to\infty}\underset{\sim}{p}^{(n)}=\left(\frac{b}{a+b}, \frac{a}{a+b}\right)$$

Rk: (1) Finite state aperiodic irreducible M.C. is regular and recurrent.

(2) If a tpm P of N states is regular, the P^{N^2} will have no zero elts. Equivalently if P^{N^2} is not strictly +ve, then M.C is not regular.

eg

$$P=\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \quad \text{regular tpm}$$

$$\overset{P}{\begin{pmatrix} + & + & 0 \\ + & 0 & + \\ + & 0 & + \end{pmatrix}}\overset{P}{\begin{pmatrix} + & + & 0 \\ + & 0 & + \\ + & 0 & + \end{pmatrix}}=\overset{P^2}{\begin{pmatrix} + & + & + \\ + & + & + \\ + & + & + \end{pmatrix}}$$

Thm: Let P be a regular tpm on states $0, 1, \to N$. $S = \{0, 1, \to N\}$

Then the limiting distn $\underset{\sim}{\Pi} = (\Pi_0, \Pi_1, -, \Pi_N)$ is the unique non-negative soln of equations

$$\Pi_j = \sum_{k=0}^{N} \pi_k \, p_{kj} \quad , j = 0, 1, \to N$$
$$\sum_{k=0}^{N} \Pi_k = 1$$

$$\longleftrightarrow \boxed{\begin{array}{l} \underset{\sim}{\Pi} = \underset{\sim}{\Pi} P \\ \sum_{k=0}^{N} \Pi_k = 1 \end{array}}$$

Sol: $\because$ MC regular, we have limiting prob.

$$\lim_{n \to \infty} p_{ij}^{(n)} = \Pi_j \quad \text{for which } \sum_{k=0}^{N} \Pi_k = 1$$

$$P^n = P^{n-1} P$$

$$p_{ij}^{(n)} = \sum_{k=0}^{N} p_{ik}^{(n-1)} p_{kj} \quad , j = 0, 1, \to N$$

as $n \to \infty$, $p_{ij}^{(n)} \to \Pi_j$, $p_{ik}^{(n-1)} \to \Pi_k$

$n \to \infty$

$$\Pi_j = \sum_{k=0}^{N} \Pi_k \, p_{kj} \quad \text{as claimed}$$

T.S. soln is unique. suppose that $x_0, x_1, \to x_N$ solve

$$x_j = \sum_{k=0}^{N} x_k \, p_{kj} \quad (1) \quad \text{for } j = 0, 1, \to N, \quad \sum_{k=0}^{N} x_k = 1$$

We wish to show $x_j = \Pi_j$

$$\text{Now } \underbrace{\sum_{j=0}^{N} x_j \, p_{jl}}_{1} = \sum_{j=0}^{N} \sum_{k=0}^{N} x_k \, p_{kj} \, p_{jl} = \sum_{k=0}^{N} x_k \underbrace{\sum_{j=0}^{N} p_{kj} \, p_{jl}}_{}$$

$p_{k\ell}^{(2)}$

using ① $\downarrow$

$$x_\ell = \sum_{k=0}^{N} x_k \, p_{k\ell}^{(2)}$$

repeating the argument

$$x_\ell = \sum_{k=0}^{N} x_k \, p_{k\ell}^{(n)} \quad , \ell = 0, 1, \ldots, N$$

$$n \to \infty \; , \quad p_{k\ell}^{(n)} \to \pi_\ell$$

$$\therefore \quad x_\ell = \sum_{k=0}^{N} x_k \, \pi_\ell = \pi_\ell \quad , \ell = 0, 1, \ldots, N$$

$\therefore$ $x_\ell = \pi_\ell$ as claimed.

—x—

Example An No claim discount (NCD) system has the discount classes E_0 (no discount), E_1 (20% discount) and E_2 (40% discount). Movement in the system is determined by the rule whereby one step back one discount level (or stays in E_0) with one claims in a year, and return to a level of no discount if more than one claim is made. A claim-free year results in a step up to a higher discount level (or one remains in class E_2 if already there). NCD system

NCD class	E_0	E_1	E_2
% discount	0	20	40

% discount	0	20	40
annual premium	100	80	60

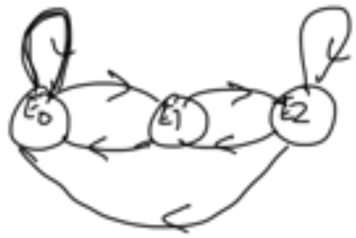
If we suppose that for anyone in this scheme the probabilities of one claim in a year is 0.2 while the prob. of two or more claims is 0.1.

(i) Find tpm

$$P = \begin{array}{c|ccc} & E_0 & E_1 & E_2 \\ \hline E_0 & 0.3 & 0.7 & 0 \\ E_1 & 0.3 & 0 & 0.7 \\ E_2 & 0.1 & 0.2 & 0.7 \end{array} \quad \text{regular}$$

(ii) In long run, what proportion of time in the process in each of three discount classes.

$$\overset{P}{\begin{pmatrix} + & + & 0 \\ + & 0 & + \\ + & + & + \end{pmatrix}} \overset{P}{\begin{pmatrix} + & + & 0 \\ + & 0 & + \\ + & + & + \end{pmatrix}} = \underset{\text{regular}}{\begin{pmatrix} + & + & + \\ + & + & + \\ + & + & + \end{pmatrix}} P^2$$



irreducible, aperiodic, finite State M.C.

$\Rightarrow$ regular

$$\underset{\sim}{\Pi} = (\Pi_0, \Pi_1, \Pi_2)$$

$$\underset{\sim}{\Pi} = \underset{\sim}{\Pi} P$$

$$\Pi_0 + \Pi_1 + \Pi_2 = 1$$

$$(\Pi_0, \Pi_1, \Pi_2) = (\Pi_0, \Pi_1, \Pi_2) \begin{pmatrix} 0.3 & 0.7 & 0 \\ \ldots \end{pmatrix}$$

$$\begin{pmatrix} 0.3 & 0 & 0.7 \\ 0.1 & 0.2 & 0.7 \end{pmatrix}$$

$$\pi_0 + \pi_1 + \pi_2 = 1$$

$$\left.\begin{aligned} \pi_0 &= 0.3\,\pi_0 + 0.3\,\pi_1 + 0.1\,\pi_2 \\ \pi_1 &= 0.7\,\pi_0 \qquad\qquad + 0.2\,\pi_2 \\ \pi_0 &+ \pi_1 + \pi_2 = 1 \end{aligned}\right\}$$

$\pi_0 = 0.1860$, $\pi_1 = 0.2442$, $\pi_2 = 0.5698$

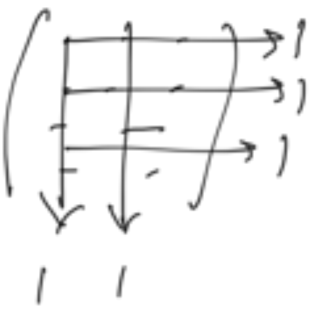
(iii) Find the average annual premium paid

$= 100 \times 0.1860 + 80 \times 0.2442 + 60 \times 0.5698$

$= 72.324$

—x—

Doubly Stochastic Matrices : tpm doubly stochastic $P = ((p_{ij}))$



Columns sum to one as well as rows

$$\underline{\sum_k p_{ik} = \sum_k p_{kj} = 1 \ \forall i, j}$$

Props :- $S = \{0, 1, \ldots, N-1\}$

P regular tpm and doubly stochastic, then the unique limiting dist is the uniform dist $\underline{\pi} = \left(\frac{1}{N}, \frac{1}{N}, \ldots, \frac{1}{N}\right)$

Sol ~~(one sol)~~ from last theorem

$$\pi_j = \sum_k \pi_k \, p_{kj} \checkmark$$

$$\sum \pi_k = 1$$

only need to check that is a sol

$$\frac{1}{N} = \sum_k \frac{1}{N} p_{kj} = \frac{1}{N} \underbrace{\sum_k p_{kj}}_{1} = \frac{1}{N}$$

$\because$ doubly stochastic tpm

— x —

Example: Let Y_n be the sum of n indep rolls of a fair die and consider the problem of determining with what prob. Y_n is a multiple of 7 in the long run.

Let X_n be the remainder when Y_n is divided by 7

$X_n \in \{0, 1, \dots, 6\}$ $\quad X_n$ is a M.C. $S = \{0, 1, \dots, 6\}$

tpm

$$P = \begin{array}{c|ccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 2 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 1/6 \\ 3 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 \\ 4 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 \\ 5 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 1/6 \\ 6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 \end{array}$$

regular and doubly stochastic

$$\Pi = \left(\frac{1}{-}, \frac{1}{-}, \dots, \frac{1}{-}\right)$$

$\left(\frac{1}{7}, \frac{1}{7}, \ldots, \frac{1}{7}\right)$

Y_n is a multiple of 7 iff $X_n = 0$

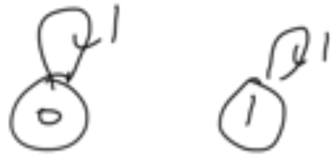
$\therefore$ limiting prob. that Y_n is a multiple of 7 is $\frac{1}{7}$.

— x —

(A) Example (i) $a = b = 0$

$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 0 & 1 \end{array} = I$$

$P^n = I$

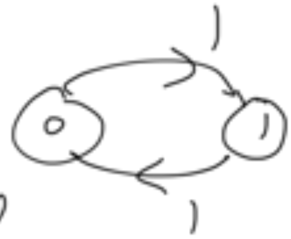


(ii) $a = b = 1$

$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}, \quad P^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad P^3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \ldots$$

periodic M.C. (with period 2)

$$P^n = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \text{if } n \text{ even} \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & \text{if } n \text{ odd} \end{cases}$$



$d(0) = \gcd(2, 4, 6, \ldots) = 2 = d(1)$

Here limiting prob. DNE.

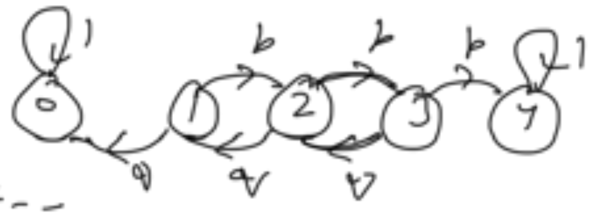
stationary prob. distn $\underline{\pi} = (\pi_0, \pi_1)$

$$\left.\begin{array}{l} \underline{\pi} = \underline{\pi} P \\ \pi_0 + \pi_1 = 1 \end{array}\right] \Rightarrow \left.\begin{array}{l} \pi_0 = \pi_1 \\ \pi_0 + \pi_1 = 1 \end{array}\right] \Rightarrow \pi_0 = \pi_1 = \frac{1}{2}$$

$$\underset{\sim}{\Pi} = \left(\frac{1}{2}, \frac{1}{2}\right)$$

— x —

Example



$$f_{2,1} = q + pq^2 + p^2q^3 + p^3q^4 + \cdots$$

$$= q + pq(q + pq^2 + \cdots) = q + pq\, f_{2,1}$$

$$\Rightarrow f_{2,1} = \frac{q}{1-pq}$$

If $\underline{p = 0.4}$

$$f_{2,1} = \frac{0.6}{1 - 0.6 \times 0.4} = \frac{0.6}{1 - 0.24} = \frac{0.6}{0.76} = \underline{0.78}$$

Gambler's ruin problem

$$f_{2,1} = P(\text{Start } 1 \text{ goes broke before reaching } 3)$$

$$= 1 - \frac{1 - \left(\frac{0.6}{0.4}\right)^1}{1 - \left(\frac{0.6}{0.4}\right)^3} = 0.78$$

— x —

Mean time spent in transient states:

finite state M.C

$T = \{1, 2, \ldots, t\}$ set of transient states

$$P_T = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1t} \\ \cdots & \cdots & & \cdots \\ \cdots & \cdots & & \\ p_{t1} & p_{t2} & \cdots & p_{tt} \end{bmatrix}$$ sum of rows < 1

$i, j \in T$

s_{ij} : expected number of time periods that the MC is in state j, given that it starts in state i.

Let $\delta_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{otherwise} \end{cases}$

$I_{n,j} = \begin{cases} 1 & \text{if } X_n = j \\ 0 & \text{o.w.} \end{cases}$

$$s_{ij} = \delta_{ij} + E\left(\sum_{n=1}^{\infty} I_{n,j} \mid X_0 = i\right)$$

$$= \delta_{ij} + \sum_{n=1}^{\infty} E(I_{n,j} \mid X_0 = i)$$

$$= \delta_{ij} + \sum_{n=1}^{\infty} P(X_n = j \mid X_0 = i)$$

$$= \delta_{ij} + \sum_{n=1}^{\infty} p_{ij}^{(n)} \qquad (*)$$

$$\sum_k p_{ik} p_{kj}^{(n-1)}$$ Using Ck eq's

$$= \delta_{ij} + \sum_k p_{ik} \sum_{n=1}^{\infty} p_{kj}^{(n-1)}$$

$$= \delta_{ij} + \sum_k p_{ik} \left[\delta_{kj} + \sum_{n=2}^{\infty} p_{kj}^{(n-1)} \right]$$

$$= \delta_{ij} + \sum_k p_{ik} \underbrace{\left(\delta_{kj} + \sum_{n=1}^{\infty} p_{kj}^{(n)} \right)}_{s_{kj} \text{ from } (*)}$$

$$= \delta_{ij} + \sum_{k} p_{ik} s_{kj}$$

$$= \delta_{ij} + \sum_{k=1}^{t} p_{ik} s_{kj} ,$$

since it is impossible to go from a recurrent to a transient state $\Rightarrow s_{kj} = 0$, where k is a recurrent state

$$\therefore \boxed{s_{ij} = \delta_{ij} + \sum_{k=1}^{t} p_{ik} s_{kj}} \quad ; i, j \in T$$

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1t} \\ - & - & & \text{---} \\ s_{t1} & s_{t2} & \cdots & s_{tt} \end{bmatrix} \quad ; \quad I = \begin{bmatrix} 1 & 0 & \text{---} & 0 \\ 0 & 1 & \cdots & 0 \\ & \text{---} & & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$S = I + P_T S$$

$$(I - P_T) S = I$$

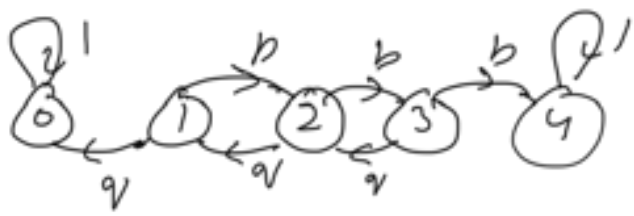
$$\Rightarrow \quad S = (I - P_T)^{-1}$$

<u>Example</u> : Consider the gamblers ruin problem with $p = 0.4$ and $N = 4$. Starting with 2 units, determine

a) the expected amt of time the gambler has 3 units.

b) " " " " " " " " 1 unit.

Sol



Classes $\{0\}$ $\{1, 2, 3\}$ $\{4\}$

absorbing transient absorbing

$T = \{1,2,3\}$

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 \\ 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

$$P_T = \begin{bmatrix} 0 & 0.4 & 0 \\ 0.6 & 0 & 0.4 \\ 0 & 0.6 & 0 \end{bmatrix}$$

$$I - P_T = \begin{bmatrix} 1 & -0.4 & 0 \\ -0.6 & 1 & -0.4 \\ 0 & -0.6 & 1 \end{bmatrix}$$

$$S = (I - P_T)^{-1} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1.46 & 0.76 & 0.31 \\ 1.15 & 1.92 & 0.76 \\ 0.69 & 1.15 & 1.46 \end{bmatrix} \end{array}$$

$$s_{2,3} = 0.76 \quad , \quad s_{2,1} = 1.15$$

(4) example (contd) Starting with state 2, what is the prob. that the gambler ever has a fortune of 1?

$$f_{2,1} = 0.78$$

—x—

$f_{i,j}$ = prob. that the M.C ever makes a transition into state j given that it starts in state i

$$s_{ij} = E(\text{time in } j \mid \text{start in } i)$$

$$= \underline{E(\text{time in } j \mid \text{start in } i, \text{ ever transit to } j)} \cdot f_{ij}$$

$+ E(\text{time in } j \mid \text{start } i, \text{ never transit to } j) \cdot (1 - f_{ij})$

$$= \underline{(\delta_{ij} + s_{jj}) f_{ij}} + \delta_{ij}(1 - f_{ij})$$

$$= \delta_{ij} + f_{ij} s_{jj}$$

$$\Rightarrow \boxed{f_{ij} = \frac{s_{ij} - \delta_{ij}}{s_{jj}}}$$

example (4)

$$f_{2,1} = \frac{s_{2,1} - \delta_{2,1}}{s_{11}} = \frac{1.15 - 0}{1.46} = 0.78$$

—x—

<u>Branching Process</u>:

each individual by the end of its lifetime produced j new offspring with prob $p_j < 1 \ \forall j \geq 0$

(let independently of number produced by other individuals)

X_0 size of zeroth generation

X_1 : all offsprings of zeroth generation or first generation

X_n : size of n^{th} generation.

$X_n \in \{0, 1, 2, --\} = S \to$ State space

(X_n) DTMC

$f_{00} = 1$ 0 recurrent

the popln will either die out or its size will

Converge to ∞.

mean # of offsprings of a single individual $\mu = \sum_{j=0}^{\infty} j p_j$

var. of " " " " " " $\sigma^2 = \sum_{j=0}^{\infty} (j-\mu)^2 p_j$

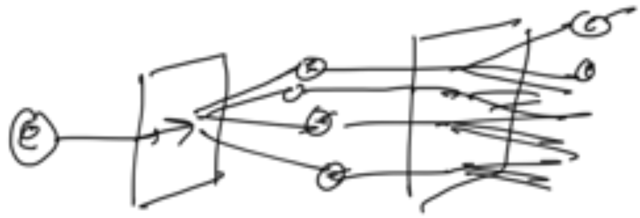
Let $X_0 = 1$

$$X_n = Z_1 + Z_2 + \cdots + Z_{X_{n-1}} = \sum_{i=1}^{X_{n-1}} Z_i$$

where Z_i # of offspring of ith individual of the $(n-1)^{st}$ generation

$E(Z_i) = \mu$, $V(Z_i) = \sigma^2$

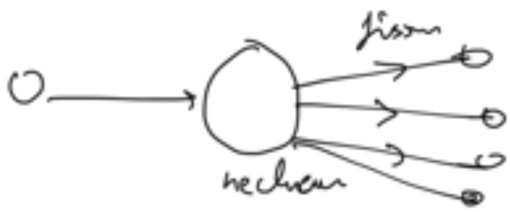
Application (i) electron multiplier



X_n # of electron emitted from n^{th} plate due to electron emanating from $(n-1)^{th}$ plate

(X_n) Branching process

(ii) Neutron chain reaction



X_n

(iii) Survival of family name

$$E(X_n) = E\big(\underline{E(X_n \mid X_{n-1})}\big)$$

$$= E\left(E\left(\sum_{i=1}^{X_{n-1}} Z_i \,\middle|\, X_{n-1}\right)\right)$$

$$X_{n-1}\mu$$

$$= E(X_{n-1}\mu)$$

$$= \mu E(X_{n-1})$$

$E(X_0)=1$, $E(X_1)=\mu$, $E(X_2)=\mu^2$, ---, $\boxed{E(X_n)=\mu^n}$ ✓

$$E\left(\sum_{i=1}^{X_{n-1}} Z_i \,\middle|\, X_{n-1}=x\right)$$

$$= E\left(\sum_{i=1}^{x} Z_i\right)$$

$$= \sum_{i=1}^{x} \underbrace{E(Z_i)}_{\mu} = x\mu$$

$$V(X_n) = E\Big[\underbrace{V(X_n|X_{n-1})}_{\sigma^2 X_{n-1}}\Big] + V\Big[\underbrace{E(X_n|X_{n-1})}_{\mu X_{n-1}}\Big]$$

$$= \sigma^2 E(X_{n-1}) + \mu^2 V(X_{n-1})$$

$$= \sigma^2\mu^{n-1} + \mu^2 V(X_{n-1}) \quad \text{——} *$$

$$= \sigma^2\mu^{n-1} + \mu^2\left[\sigma^2\mu^{n-2} + \mu^2 V(X_{n-2})\right]$$

using *

$$V\left(\sum_{i=1}^{X_{n-1}} Z_i \,\middle|\, X_{n-1}=x\right)$$

$$= V\left(\sum_{i=1}^{x} Z_i\right)$$

Z_i indp

$$= \sum_{i=1}^{x} \underbrace{V(Z_i)}_{\sigma^2} = x\sigma^2$$

$$= \sigma^2\left[\mu^{n-1}+\mu^n\right] + \mu^4 V(X_{n-2})$$

$$= \sigma^2(\mu^{n-1}+\mu^n) + \mu^4\left(\sigma^2\mu^{n-3} + \mu^2 V(X_{n-3})\right)$$

$$= \sigma^2(\mu^{n-1}+\mu^n+\mu^{n+1}) + \mu^6 V(X_{n-3})$$

$$= \sigma (\mu + \mu + \cdots + \mu) \quad + \mu \underbrace{(V(X_0))}_{\to 0}$$

$$= \sigma^2 \mu^{n-1} (1 + \mu + \cdots + \mu^{n-1})$$

$$\therefore V(X_n) = \begin{cases} \sigma^2 \mu^{n-1} \left(\dfrac{1-\mu^n}{1-\mu} \right) & \text{if } \mu \neq 1 \\ n\sigma^2 & \text{if } \mu = 1 \end{cases}$$

$\longrightarrow$ $$u_{n+1} = P(X_{n+1} = 0) = \sum_j P(X_{n+1} = 0 \mid X_1 = j)\, p_j$$

$$= \sum_j \left(P(X_n = 0) \right)^j p_j$$

$$= \sum_j u_n^j\, p_j$$

$$\therefore \boxed{u_{n+1} = \sum_j u_n^j\, p_j}$$

$\longrightarrow$ <u>Prob of ultimate extinction</u> (Π_0)

Π_0 : prob. that popl^n will eventually die out (under the assumption that $X_0 = 1$)

$$\Pi_0 = \lim_{n \to \infty} P(X_n = 0 \mid X_0 = 1)$$

$\rightarrow$ $\Pi_0 = 1$ if $\mu < 1$

$$\mu^n = E(X_n) = \sum_{j=\cancel{0}1}^{\infty} j\, P(X_n = j)$$

$$\geq \sum_{j=1}^{\infty} 1 \cdot P(X_n = j)$$

$$= P(X_n \geq 1)$$

Since $\mu^n \to 0$ if $\mu < 1$ as $n \to \infty$

$$P(X_n \geq 1) \to 0$$

, i.e., $P(X_n = 0) \to 1$, i.e., $\Pi_0 = 1$

→ It can shown that $\Pi_0 = 1$ even when $\mu = 1$

→ When $\mu > 1$ it turns out $\Pi_0 < 1$

$$\Pi_0 = P(\text{popl}^n \text{ dies out})$$

$$= \sum_{j=0}^{\infty} P(\text{popl}^n \text{ dies out} \mid X_1 = j)\, p_j$$

$$\leq \sum_{j=0}^{\infty} \Pi_0^j\, p_j$$

$$\boxed{\Pi_0 = \sum_{j=0}^{\infty} \Pi_0^j\, p_j} \longrightarrow (✡)$$

When $\mu > 1$, it can be shown that Π_0 is the smallest +ve number satisfying equation (✡).

Example (i) $X_0 = 1$, $p_0 = \frac{1}{2}$, $p_1 = \frac{1}{4}$, $p_2 = \frac{1}{4}$

$$\Pi_0 = ?$$

Sol $\mu = 0 \times \frac{1}{2} + 1 \times \frac{1}{4} + 2 \times \frac{1}{4} = \frac{3}{4} \leq 1$

$$\therefore \Pi_0 = 1$$

(ii) $X_0 = 1$; $p_0 = \frac{1}{4}$, $p_1 = \frac{1}{4}$, $p_2 = \frac{1}{2}$

$$\Pi_0 = ?$$

Sol

$$\mu = 0 \times \frac{1}{4} + 1 \times \frac{1}{4} + 2 \times \frac{1}{2} = \frac{5}{4} > 1$$

$$\pi_0 = \sum_{j=0}^{\infty} \pi_0^j \, p_j \qquad \Big| \text{ using } \star$$

$$\Rightarrow \pi_0 = \frac{1}{4} + \frac{\pi_0}{4} + \frac{\pi_0^2}{2}$$

$$\Rightarrow 2\pi_0^2 - 3\pi_0 + 1 = 0$$

$$\Rightarrow \pi_0 = 1, \frac{1}{2}$$

$$\therefore \pi_0 = \frac{1}{2}$$