Assignment IV: Probability and Stochastic Processes

106218 - Getrude Gichuhi

2022-08-25

Question 1.a

Clearly distinguish the following terms

i. Stochastic Process and Deterministic Data

A stochastic process is a collection of random variables that analyses and produces a random probability distribution or pattern within a measurable space or indexed by time. It has the capacity to handle uncertainties based on the inputs .

A deterministic Data is a collection of unique identifiers that match one user to one dataset and are used to predict the outcome with certainty. It is specified by a set of equations that describe exactly how the system will evolve over time

ii. Discrete and Continuous stochastic process.

Discrete stochastic process is essentially a random vector with components indexed by time and a time series is observed in an economical application as one realization of that random vector

A continuous stochastic process is a process for which the index variable takes a continuous set of values. A continuous stochastic process that is constructed from a discrete-time processes via a waiting time distribution is called continuous time random walks. It is like a drunk person is walking.

iii. sigma-algebra and Filtration

sigma-algebra is the smallest sigma algebra such that a random variable (Xt) is measurable for all time (tT)

Filtration are totally ordered collections of subsets that are used to model the information that is available at a given point and therefore play important roles in the formalization of random processes. they are part of the of the sigma-algebra

iv. Random Walk and Gambler's Ruin Problem

Random Walk is a stochastic process that consists of the sum of a sequence of changes in a random variable. There are no patterns to the changes in the random variable and these changes cannot be predicted

Gambler's Ruin Problem is the process where at any point in time n, A gambler X can have y wealth, where y also represents the state of the chains of gambles at time n. It is a sequence of wealth amounted by a gambler within a timeline.

v. Counting Process and Branching Process

Counting Process is a stochastic process that represents a total of events that occur at a certain time. For instance, the number of customers who enter a supermarket when a campaign was running.

Branching process is a stochastic process with properties that are distributing from each variable individually. A case study of the Offspring Distribution based on the generations Assuming that Steve is an individual in time 2020 represented as n = 0, and S(as steve). After one unit of time let's say 2031, Steve produces children and later on dies. Therefore Steve's children are considered as S1, (Say they are Sam, and Stone) If Sam (S1) happens not to produce a generation, then the gen ends there, however, if Stone (S2) continues the process and produces a generation who in turn produce more children the process continues.

vi. Birth and Death Process

They are both used as models for population growth and queue formation The Yule Furry Process (birth process) considers the number of species within a group and to conceive the creation (multiply) of new species by mutation as a random event (birth) whose probability of occurrence is proportional to the number of existing species.

Death Process the population either remains constant or it decreases, when it gets to zero it becomes extinct.

vii. Static Simulation Model and Dynamic Simulation Model

Static simulation model is one that represents the system at a particular point in time, referred to as Monte Carlo simulations.

Dynamic Simulation models represent systems as they evolve over time.

viii. Mathematical Model and Simulation Model

Mathematical Model is a description of a system using mathematical language and its concepts

Simulation model is when a system is observed or studied by running different models up until to the satisfaction of the analyst and models the system using random variables to their liking.

ix. Monte Carlo Simulation and Queuing System

A Monte Carlo simulation is used to model the probability of different outcomes in a process that cannot easily be predicted due to the intervention of random variables. It is a technique used to understand the impact of risk and uncertainty, such as investments, businesses, physics etc.

A Queuing system is a collection of mathematical models that inputs parameters of services and customers and provide quantitative parameters to describe the system performance.

x. Markov and Poisson processes

A Markov process is a random process in which the future is independent of the past, given the present.

A Poisson Process is a model of a series of discrete events where the average time or pausable time, which is known (could be 2 secs or 5 mins), but the exact timing of events is random.

xi. Queuing and Renewal processes

Queuing processes are a class of random process, describing the phenomena of waiting lines, constructed so that lengths and waiting times can be predicted.

A renewal Process is a stochastic model for events that occur randomly in time, known as renewals or arrivals

xii. Martingale and Brownian motion A brownian motion is a continuous time stochastic process when it has almost sure continuous paths and stationary independent increments such as X(s+t)-X(s) is Gaussian with mean 0 and variance t

A Martingale is a stochastic process M with respect to a filtration (Ft) if M is adapted to the biltration for any s<t then E(M(t)Fs) = M(s)

xiii. Hidden Markov model and semi-Markov process

A hidden Markov model(HMM) is a doubly stochastic process, whose discrete time finite state homogeneous Markov chain. The state sequence is not observable therefore hidden.

A semi - Markov Process is a renewal process defined by every given time.

Question 1.b

Collins bought a share of stock for 12 dollars, and it is believed that the stock price moves (day by day) as a simple random walk with p = 0.58. What is the probability that Collins' stock reaches the high value of 35dollars before the low value of 8 dollars?

$$Pa\frac{1-(q/p)^b}{1-(q/p)^{(a+b)}}$$

```
i <- 12
a <- 4
b <- 23
p <- .58
q <- .42

Prob <- (1-(q/p)^a) / (1-(q/p)^(a+b))
Prob

## [1] 0.7251491
```

Question 1.c

c) Explain clearly the difference between the following terms as used in Markov Chains

i. Communicating class and absorption state

In Communicating class, any two states i & j are said to communicate with each other if j is accessible from i and i is accessible from j.

In Absorption state is when the markov chains cannot jump or communicate from one state to the other and back but rather from let's say i to j, while j does not communicate back.

ii. Recurrence and nonrecurrence state A nonrecurrence state is when a communicating class has entered a Markov chain, it can never level the chain

A recurrence state is when a Markov chain returns to a state infinite times.

iii. Periodicity and aperiodic A state i has a period d if given the initial state at Xo = i, it is only possible to have Xn=i when n is a multiple of d. Therefore state is pediodic if d>1

Aperiodic is when each state is visited each at random rates.

iv. Ergodic chain and transient state Ergodic chain: Suppose Y is a Markov Chain with only finite states. If -Y is irreducible, and aperiodic, then it is ergodic. A transient sate is when a communicating class has entered a markov chain and can never leave. meaning it is periodic.

v. Reducible and irreducible

A markov chain is called reducible if there are two or more states communicate to each other.

A Markov chain is called irreducible if, given any two states, it is possible in constructing a sequence of jumps leading from one state to the other and back again.

Question 1.d

Consider an M/M/1 model at steady state, with μ as the service mechanism rate and λ as the arrival rate. Let Pn(t) = P [n customers in the system at time t] (Probability that there are n customers at time t). Derive Lq, Ls, Wq and Ws.

let

$$p = \frac{\lambda}{\mu}$$

$$Ws = waitingtimeinthesystem$$

$$Ws = \frac{1}{\mu - \lambda}$$

$$Lq = QueueLength$$

$$Lq = \frac{p^2}{1-p}$$

$$Wq = Waitingtimeinthequeue$$

$$Wq = \frac{Lq}{\lambda}$$

$$Ls = Numberinthesystem$$

$$Ls = (\lambda)Ws$$

Question 1.e

Clearly specify five components of a Hidden Markov Model Given an example of weather conditions(Hot, Wet, Cold) affecting the fabrics(Nylon, Wool, COtton) used. a) Q = Number of states = (Hot, Wet, COld)

b) A = Transition Probobaility matrix (using the current state of the weather and the previous state)

c) T = Sequence of the observations =(Nylon, Wool, Cotton)

d)

$$\theta = EmissionProbabilityMatrix$$

$$\pi = [\pi 1, \pi 2, \ldots, \pi N] = Set of prior probabilities(Hot, Cold, Wet)$$

e)Transitions and Emission Probabilities respectively.

Question 1.f

Use Chapman Kolmogorov postulates to derive the Poisson Process. Also derive the mean and variance of the Poisson process.

$$\begin{pmatrix} p_{00}{}'(t) & p_{01}{}'(t) & p_{02}{}'(t) & \cdots \\ p_{10}{}'(t) & p_{11}{}'(t) & p_{12}{}'(t) & \cdots \\ p_{20}{}'(t) & p_{21}{}'(t) & p_{22}{}'(t) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} p_{00}(t) & p_{01}(t) & p_{02}(t) & \cdots \\ p_{10}(t) & p_{11}(t) & p_{12}(t) & \cdots \\ p_{20}(t) & p_{21}(t) & p_{22}(t) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$= \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Question 1.g

A certain stock price has been observed to follow a pattern. If the stock price goes up one day, there's a 25% chance of it rising tomorrow, a 35% chance of it falling, and a 40% chance of it remaining the same. If the stock price falls one day, there's a 25% chance of it rising tomorrow, a 50% chance of it falling, and a 25% chance of it remaining the same.

Finally, if the price is stable on one day, then it has a 50-50 change of rising or falling the next day.

i. Generate the transition matrix

```
library("ChannelAttribution")

## Warning: package 'ChannelAttribution' was built under R version 4.1.3

## ChannelAttribution 2.0.5

## Looking for attribution at path level? Try ChannelAttributionPro! Visit
https://channelattribution.io for more information.

library("ggplot2")

## Warning: package 'ggplot2' was built under R version 4.1.3

library("reshape")

## Warning: package 'reshape' was built under R version 4.1.3

library("dplyr")

## Warning: package 'dplyr' was built under R version 4.1.3

##
## Attaching package: 'dplyr'

## The following object is masked from 'package:reshape':
##
##     rename

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

library("rlang")

## Warning: package 'rlang' was built under R version 4.1.3

library("igraph")

## Warning: package 'igraph' was built under R version 4.1.3

##
## Attaching package: 'igraph'

## The following object is masked from 'package:rlang':
##
##     is_named
```

```
## The following objects are masked from 'package:dplyr':
##
##     as_data_frame, groups, union

## The following objects are masked from 'package:stats':
##
##     decompose, spectrum

## The following object is masked from 'package:base':
##
##     union

library("plyr")

## Warning: package 'plyr' was built under R version 4.1.3

## ----------------------------------------------------------------------------
----

## You have loaded plyr after dplyr - this is likely to cause problems.
## If you need functions from both plyr and dplyr, please load plyr first,
then dplyr:
## library(plyr); library(dplyr)

## ----------------------------------------------------------------------------
----

##
## Attaching package: 'plyr'

## The following objects are masked from 'package:dplyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize

## The following objects are masked from 'package:reshape':
##
##     rename, round_any

library("reshape2")

## Warning: package 'reshape2' was built under R version 4.1.3

##
## Attaching package: 'reshape2'

## The following objects are masked from 'package:reshape':
##
##     colsplit, melt, recast

library("markovchain")

## Warning: package 'markovchain' was built under R version 4.1.3
```

```
## Package:  markovchain
## Version:  0.9.0
## Date:     2022-07-01
## BugReport: https://github.com/spedygiorgio/markovchain/issues
```

```
library("plotly")
```

```
## Warning: package 'plotly' was built under R version 4.1.3

##
## Attaching package: 'plotly'

## The following objects are masked from 'package:plyr':
##
##     arrange, mutate, rename, summarise

## The following object is masked from 'package:igraph':
##
##     groups

## The following object is masked from 'package:reshape':
##
##     rename

## The following object is masked from 'package:ggplot2':
##
##     last_plot

## The following object is masked from 'package:stats':
##
##     filter

## The following object is masked from 'package:graphics':
##
##     layout
```

```
library("diagram")
```

```
## Warning: package 'diagram' was built under R version 4.1.1

## Loading required package: shape

## Warning: package 'shape' was built under R version 4.1.1
```

```
#Up (shows the changes if it goes up one day)
#Down(Shows the changes if it goes down one day)
#Stable(Shows the changes if it remains the same)
#R,F,S (Represents, rise, fall, and same)

Stock <- matrix(c(0.25, 0.35, 0.40, 0.25, 0.50, 0.25, 0.50, 0.50, 0), nrow =
3, byrow =TRUE)
rownames(Stock) <- c("Up","Down", "Stable")
```

```
colnames(Stock) <- c("R", "F", "S")
Stock

##           R    F    S
## Up     0.25 0.35 0.40
## Down   0.25 0.50 0.25
## Stable 0.50 0.50 0.00
```

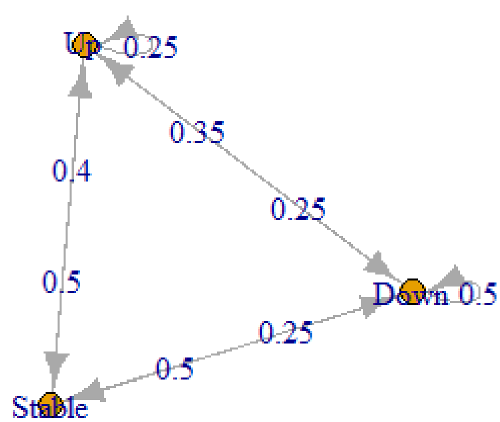
ii. Draw the Markov chain using R

```
print(Stock)

##           R    F    S
## Up     0.25 0.35 0.40
## Down   0.25 0.50 0.25
## Stable 0.50 0.50 0.00

stock2 <- new('markovchain',
              transitionMatrix=Stock,
              name="RFS"
              )


plot(stock2, relsize = 0.75)
```



iii. Determine if the chain is Ergodic Yes it is Ergodic, since all the states are communicating to each other and they are aperiodic

iv. Find the limiting distribution of the transition matrix

```
apply(steadyStates(stock2),1,sum)

## [1] 1

steadyStates(stock2)

##             Up      Down    Stable
## [1,] 0.3092784 0.4536082 0.2371134
```

Question 1.h

A telephone attendant receives 110 calls during the busy hour. Each call takes, on average, 2.1 minutes to process.

a) What percentage of the attendant's time is devoted to answering calls?

$$\lambda = 110$$

$$\mu = 2.1$$

```
mu=2.1
Service_rate = (1/mu)*60

Service_rate

## [1] 28.57143
```

Around 29 calls per hour

$$T = \frac{1}{\mu - \lambda} = \frac{1}{29 - 110} * 60 = 0.7407 Sec$$

Around 0.62% of the hour is spent to make a call

b) How long must people wait, on average, before their call is processed?

$$W = T - \frac{1}{\mu} = \left(0.0123 - \frac{1}{2.1}\right) = -0.463$$

Around 27 secs.

Question 1.i

Jobs arrive to a computer system (consisting of a CPU and an I/O device) according to a Poisson process with rate 8 jobs per minute. Once in the system, a job requires on average 30 seconds of CPU time and 9 minutes of I/O time, in which the CPU and I/O time required by the jobs are exponentially distributed.

a)What is the probability that a job will have to wait before being processed by the devices? (Hint: replace the CPU and I/O subsystem as equivalent to single server)

$$\lambda = 8, \mu = 9.5$$

to get the probability

$$p = \frac{\lambda}{\mu} = \frac{8}{9.5} = 0.842$$

b) What proportion of time is the system busy?

$$T = Delay + Servicetime = \frac{1}{\mu} = \frac{1}{9.5} = 0.105 mins$$

c)On average, how many jobs are waiting in line to be processed?

$$Lq = \frac{p2}{1 - p} = \frac{(0.842)^2}{(1 - 0.842)} = 4.487$$

Around 4 jobs are waiting

d) On average, how long will a job spend in the system?

$$Wq = \frac{Lq}{\lambda} = \frac{4}{9.5} = 0.421 mins$$

e) What is the probability that exactly 10 jobs arrive to the system in one minute?

$$P_0 = 1 - (\frac{\lambda}{\mu})^n = 1 - 0.8421 = 0.1579$$

$$P_n = P_0 - (\frac{\lambda}{\mu})^n$$

$$P_{10} = P_0 (\frac{\lambda}{\mu})^{10} = 0.1579 \times 0.8421^{10} = 0.02831$$