MSDS 454

Applied Probability & Simulation Modeling

Module 1: Random Samples

Forsythe - Chapter 3: Basic Ideas in Probability

3.1 - Experiments, Outcomes, and Probability

3.1.1 - Outcomes and Probability

Experiments produce one outcome out of a set of outcomes.

Sample Space Ω - the set of all outcomes of an experiment.
(not necessarily finite)

Probability - relative frequency of an outcome.

$$P = \lim_{N \to \infty} \frac{\#A}{N}$$

where: $\#A$ = number of times outcome A occurs

N = number of experiments

- $0 \le P(A) \le 1$
- $\sum_{A_i \in \Omega} P(A_i) = 1$

Example **3.3**

A couple decides to have children until a girl and the a boy are born.

Sample space: $B^{*} G + B$ This has a lower bound of 2 (G,B) and no upper bound.

3.2 - Events

Event ε - a set of outcomes.

$P(\emptyset) = 0$ (The probability of the empty set is zero)

3.2.1 - Computing Event Probabilities by Counting Outcomes

Example **3.8**

A couple has children until the first girl is born or until they have three children. Calculate $P(B_i)$ and $P(C)$ where B_i is the number of male children and C is where the couple have more girls than boys.

$\Omega = \{G, BG, BBG, BBB\}$

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

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

Wrong ↑ The above outcomes do not have uniform probability. Adding fictitious outcomes to fill out the sample space provides the real probabilities

$$\Omega = \begin{Bmatrix} GGG, GGB, GBB, GBG \\ BBB, BBG, BGG, BGB \end{Bmatrix}$$

$B_1 = \{GGB, GBG, BGG\}$ $P(B_1) = \frac{3}{8}$

$C = \{GGG, GGB, GBG, BGG\}$ $P(C) = \frac{1}{2}$

Wrong ↑

Try again

Example 3.7: A couple decides to have three children. What are the probabilities that the couple have one boy $P(B_1)$ and that they have more girls than boys $P(C)$?

$$\Omega = \begin{Bmatrix} BBB, BBG, BGB, BGG \\ GGG, GGB, GBG, GBB \end{Bmatrix}$$

$B_1 = \{BGG, GBG, GGB\}$ $P(B_1) = \frac{3}{8}$

$C = \{GGG, GGB, GBG, BGG\}$ $P(C) = \frac{1}{2}$

~

Example 3.8: A couple decides to have children until the first girl is born or they have three children.
Calculate $P(B_1)$ and $P(C)$

$\Omega = \{G, BG, BBG, BBB\}$

B_1 only occurs in sequence BG, which occurs in previous Ω in $\{BGB, BGG\}$, thus $P(B_1) = \frac{2}{8} = \frac{1}{4}$

G>B only occurs in sequence G, which occurs in previous Ω $\{GGG, GBB, GGB, GBG\}$, thus $P(C) = \frac{4}{8} = \frac{1}{2}$

3.2.1 · Combinations and Permutations

Combinations are groupings in which order does not matter.

The number of possible combinations of n objects taken k at a time.

$${}_nC_k = \frac{n!}{k!(n-k)!}$$

Permutations are groupings in which order does matter.

The number of possible permutations of n objects taken k at a time

$${}_nP_k = \frac{n!}{(n-k)!}$$

Example **3.9**

What is the probability that one draws 2-8 of hearts in order when drawing seven cards from a fair deck?

$$\frac{1}{{}_{52}P_7} = \frac{1}{\frac{52!}{45!}} = \frac{45!}{52!} = \frac{(n-k)!}{n!}$$

Example **3.10**

What is the probability that one draws 2-8 of hearts in any order when drawing seven cards from a fair deck?

$$\frac{1}{{}_{52}C_7} = \frac{1}{\frac{52!}{7!45!}} = \frac{1}{\frac{52}{7}} = \frac{7}{52} = \frac{k}{n}$$

3.2.2 The Probability of Events

$P(A) + P(A^c) = 1$

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

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

3.3 Independence

Two events, A and B, are independent if $P(A \cap B) = P(A)P(B)$ $P(A|B) = P(A)$

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

Example **3.17**

Calculate the probability of rolling threes using two fair die. Show that the rolls are independent of one another.

A = roll 3 on die 1

B = roll 3 on die 2 $P(A \cap B) = P(A)P(B)$ if A and B are independent

$P(A) = \frac{1}{6}$, $P(B) = \frac{1}{6}$ $\frac{1}{36} = \left(\frac{1}{6}\right)\left(\frac{1}{6}\right)$ > True!

$P(A \cap B) = \frac{1}{36}$

Example **3.18** · Cards (drawing)

What is the probability of drawing a red 10 from a fair deck?

$P(10 \cap R) = \frac{2}{52}$ (10 of hearts, 10 of diamonds)

Are drawing a 10 and a red card independent?

Yes. $P(10) = \frac{4}{52}$, $P(R) = \frac{1}{2}$

$P(10 \cap R) = \frac{2}{52} = \left(\frac{4}{52}\right)\left(\frac{1}{2}\right) = P(10)P(R)$

thus $P(10 \cap R) = P(10)P(R)$

With the 10 of hearts removed, what is new $P(10 \cap R)$? Independent?

$P(10 \cap R) = \frac{1}{51}$ (10 of diamonds)

$P(10) = \frac{3}{51}$, $P(R) = \frac{25}{51}$

$P(10 \cap R) = \frac{1}{51} \ne \left(\frac{3}{51}\right)\left(\frac{25}{51}\right) = P(10)P(R)$

Example **3.19** - Accidental DNA Matches

What is the probability of getting at least one accidental match in a DNA database of 20,000 when the probability of an accidental match during a single comparison is 1e-4.

At least one match includes all numbers greater than zero. So, the two events of interest are no matches and any matches.

$P(M) = 1e{-}4$

$P(NM) = 1 - 1e{-}4$

$P(NM \text{ in } 20{,}000) = (1 - 1e{-}4)^{20{,}000}$

$P(M \text{ in } 20{,}000) = 1 - (1 - 1e{-}4)^{20{,}000} \approx 0.86$

3.3.1 - Example: Airline Overbooking

Example 3.20 - Airline I

An airline sells seven tickets for a six-seater plane. If passengers have probability of p showing up, what is the probability that the plane is overbooked?

- Overbooked means that all 7 show up.

$P(\text{Overbooked}) = p^7$

Example 3.21 - Airline II

What if the airline sells eight tickets?

a) There are 16 variables, taken 8 at a time, $\Omega = {}_{16}C_8 = \frac{16!}{8!8!}$

b) Overbooking occurs when either 7 or 8 passengers show up

c) 7 show up $= (p)^7(1-p)$

8 show up $= (p)^8$

$P(\text{Overbooked}) = (p)^8 + 8(p^7(1-p))$

Example 3.22 - Airline III

What is the probability that the flight is fully booked?

- 16 variables taken two at a time
- Outcome is 6 arrive, 2 dont $= (p^6)(1-p)^2$
- There are ${}_8C_6$ possible versions: $\frac{8!}{6!\,2!}$

$P = \frac{8!}{6!\,2!}(p^6)(1-p)^2$

Example 3.23 - Airline IV

What is the probability of u passengers showing up to a s-seater plane when t tickets have been sold?

$P(u) = \frac{t!}{u!(t-u)!}(p)^u(1-p)^{t-u}$ = Binomial Distribution!

3.4 - Conditional Probability

$P(B|A) = \frac{P(A\cap B)}{P(A)}$ $\approx P(A)P(B)$ if A and B are independent.

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

$P(B|A) + P(B^c|A) = 1$

$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$

$P(A) = \sum_i P(A|B_i)P(B_i)$

3.4.2 - Detecting Rare Events is Hard

Example 3.31 - Rare Disease (False positives)

$P(D) = \frac{1}{100{,}000}$

$P(ND) = \frac{99{,}999}{100{,}000}$

$P(P|D) = 0.95$

$P(P|ND) = 1e-3$

What is $P(D|P)$?

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

$$= \frac{0.95\left(\frac{1}{100{,}000}\right)}{P(P)}$$

$$= \frac{0.95\left(\frac{1}{100{,}000}\right)}{\left(\frac{1}{100{,}000}\right)(0.95) + \left(\frac{99{,}999}{100{,}000}\right)(1e-3)}$$

3.4.3 - Conditional Probability and Various Forms of Independence

Events A and B are independent if $P(A\cap B) = P(A)P(B)$

Then, $P(A|B) = P(A)$

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

Pairwise Independence: Events $(A_1, A_2 \ldots A_i)$ are pairwise independent if each pair of events are independent

Example 3.33 - Pairwise Independence while drawing cards

When drawing three cards from a fair deck (with replacement):

A = C_1 and C_2 have same suit $\quad P(A) = \frac{1}{4}$

B = C_2 and C_3 have same suit $\quad P(B) = \frac{1}{4}$

C = C_1 and C_3 have same suit $\quad P(C) = \frac{1}{4}$

$P(A\cap B) = \frac{4}{64} = \frac{1}{16}$

$P(A\cap B\cap C) = \frac{1}{16}$, as it is implied in $P(A\cap B)$ $\neq \frac{1}{64}$ or $\left(\frac{1}{4}\right)^3$

3.4.3 - Conditional Probability and Various Forms of Independence

Conditional Independence: Events $A_1 \ldots A_i$ are conditionally independent if

$$P(A_1 \ldots A_i|B) = P(A_1|B) \ldots P(A_i|B)$$

Example 3.34 - Conditional Independence while drawing cards

Remove a red 10 and red 6 from a fair deck. Draw one card.

A: card is a 10

B: card is red

C: card is either a 10 or 6

Show that A and B are not independent, but are conditionally independent on C.

$P(A) = \frac{3}{50}$ $\quad P(A\cap B) = \frac{1}{50}$ $\quad P(A|B) = \frac{1}{24}$

$P(B) = \frac{24}{50}$ $\quad P(B|A) = \frac{1}{3}$

$P(C) = \frac{6}{50}$ $\quad P(A\cap C) = \frac{3}{50}$ $\quad P(A|C) = \frac{1}{2}$, $\quad P(B|C) = \frac{1}{3}$

$P(B\cap C) = \frac{2}{50}$

$P(A\cap B) = \frac{1}{50} \neq \frac{72}{250} = P(A)P(B)$, A and B are not independent

$P(A|C \ldots B|C) = \frac{1}{6} = \left(\frac{1}{2}\right)\left(\frac{1}{3}\right) = P(A|C)P(B|C)$, A and B are conditionally independent on C

3.4.4 - The Prosecutor's Fallacy

Suppose someone on trial can be either I or G with evidence ε

$P(\varepsilon|I) \neq P(I|\varepsilon)$

$P(\varepsilon|I)$ is the probability the evidence exists given the person is innocent.

$P(I|\varepsilon)$ is the probability the person is innocent given the evidence exists.

$P(I|\varepsilon) = \frac{P(\varepsilon|I)P(I)}{P(E)}$ $\quad P(E) = P(I|E)P(I) + P(I^c|\varepsilon)(1-P(I))$

$P(I|\varepsilon) = \frac{P(\varepsilon|I)P(I)}{P(\varepsilon|I)P(I) + P(\varepsilon|I^c)(1-P(I))}$ $= \frac{1}{1 + P(\varepsilon|I^c)(1-P(I))}$

If $P(I)$ is high or $P(\varepsilon|I)$ is larger than $P(\varepsilon|I^c)$, $P(I|\varepsilon)$ can be high while $P(\varepsilon|I)$ is low.

3.4.5 - The Monty Hall Problem

Module 1: Random Samples

Assignment 1: Forsythe Chapter 3 Exercises

Problem 1: 3.1

What is the sample space of rolling a 4 sided die?

$\Omega = \{1,2,3,4\}$

Problem 2: 3.2

King Lear decides to allocate three providences to his three daughters at random. What is the sample space?

$\Omega = \{GRC, GCR, CRG, CGR, RCG, RGC\}$

Problem 3: 3.5

What is the probability of rolling a 3 using a 4-sided die?

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

Problem 4: 3.6

What is the probability of drawing a King of hearts from a fair deck?

$P(K♡) = 1/52$

Problem 5: 3.8

At a university, $1/2$ drink and $1/3$ smoke

a) What is the largest possible fraction who do neither? $1/2$

b) If $1/3$ of students do neither, what fraction do both? $1/6$

Problem 6: 3.11

What is the probability of rolling an even number on a 20-sided die?

$P(E) = \frac{1}{2}$

Problem 7: 3.12

What is the probability of getting an even number when rolling a 5-sided die?

$P(E) = \frac{2}{5}$

Problem 8: 3.15

When drawing a card from a deck:

a) What is the probability of drawing a King? $P(K) = \frac{1}{13}$

b) a heart? $P(H) = \frac{1}{4}$

red? $P(R) = \frac{1}{2}$

Problem 9: 3.16

When playing roulette: 38 slots, odds are red, even are black, two green 0's

a) What is the probability of green? $P(G) = \frac{1}{19}$

b) red and even? $P(R \cap E) = \frac{0}{38}$

red and $\%7 = 0$? $P(R \cap \%7 = 0) = 3/38$ (7, 21, 35)

Problem 10: 3.30

$P(A) = 0.5$

$P(B) = 0.2$

$P(A \cup B) = 0.65$ $\quad P(A \cap B) = 0.05$

Are A and B independent?

No.

Problem 11: 3.31

$P(A) = 0.5$

$P(B) = 0.5$

$P(A \cap B) = 0.25$

$P(A \cup B) = 0.75$

Problem 12: 3.44

Student takes a test where each question has N answers.

The student knows the answer to 70% of questions.

If they don't know, they guess randomly and uniformly.

K - the student knows the answer

R - the student gets the right answer

a) What is $P(K)$? $P(K) = 0.7$

b) What is $P(R|K)$? $P(R|K) = 1.0$

c) What is $P(K|R)$ as a function of N?

$P(K) = 0.7$

$P(R) = 0.7 + \frac{1}{N}$

$P(R|K) = 1$

$$P(K|R) = \frac{(1)(0.7)}{\left(0.7 + \frac{1}{N}\right)}$$

d) What value of N ensures $P(K|R) > 99\%$?

$$0.99 = \frac{0.7}{0.7 + \frac{1}{n}}$$

$(0.99)(0.7) + (0.99)\left(\frac{1}{n}\right) = 0.7$

$\frac{0.99}{n} = 0.7 - (0.99)(0.7)$

$0.99 = n(0.7 - (0.99)(0.7))$

$0.99 = 0.7n - 0.693n$

$0.99 = 0.007n$

$\frac{0.99}{0.007} = n$

142

Problem 13: 3.45

I = patient has illness

I^c = patient does not have illness

R = result is positive

R^c = result is negative

$P(I|R) = 0.5$

$P(R|I^c) = 0.1$

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

Problem 14

A: $X = 2$ $\quad P(A) = 1/6$ $\quad P(A|B) = 1/3$

B: $X \leq 3$ $\quad P(B) = 1/2$ $\quad P(B|A) = 1$

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

$$1 = \frac{(x)\left(\frac{1}{2}\right)}{\frac{1}{6}}$$

$\frac{1}{6} = \frac{1}{2}x$

$x = \frac{1}{3}$

$\boxed{1}$

$\boxed{2}$ ← intersection $= \frac{1}{6}$...?

$\boxed{3}$

4

5

6

Assignment 2 - Special Functions in Statistics

Topic 1: The Binomial Coefficient

The binomial coefficient is a counting technique for choosing groups of size k from a population size n. This is referred to as n choose k

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

1) Inviting Friends

3 friends: F_1, F_2, F_3, can only invite two

$${}_3C_2 = \frac{3!}{2!\,1!} = \frac{6}{2} = 3$$

$$\Omega = \{F_1F_2, F_1F_3, F_2F_3\}$$

2) Inviting Friends in R

choose(3,2) = 3

3) Maximization of the Binomial Coefficient

The binomial coefficient is maximized at $k = \frac{n}{2}$

4) Inviting Friends Again

$${}_5C_4 = \frac{5!}{4!\,1!} = \frac{120}{24} = 5$$

$${}_5C_3 = \frac{5!}{3!\,2!} = \frac{120}{12} = 10$$

5) Drawing Balls

What is the probability of drawing three red and three black balls from a bin with 10 red and 10 black balls?

$\left(\frac{1}{2}\right)^6 = 1/256$ or... $P(3\text{ black}, 3\text{ red}) = \dfrac{\#\text{ ways to choose 3 black} \times \#\text{ ways to choose 3 red}}{\#\text{ ways to choose 6 balls}}$

$$= \frac{(20)(20)}{{}_{20}C_6} = \frac{400}{\frac{20!}{6!\,14!}} = \frac{400}{38760} = 0.0103199$$

$$1/256 = 0.003906$$

$${}_6P_3 = \frac{6!}{3!} = \frac{600}{6} = 100 \quad \frac{10{,}000}{38760} = 0.25799$$

$$\frac{144400}{38760} = 0.371517$$

Topic 2: The Factorial Function

The factorial function is defined as $n! = n \cdot (n-1) \cdot \ldots \cdot 1$

Problem 7: Computing the Binomial Function

$${}_{10}C_3 = \frac{10!}{3!\,7!} = 120$$

Problem 8: The Gamma Function

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\,dx$$

$$\Gamma(4.1) = 6.81263$$

Problem 9: The Gamma Function - Special Case #1

$$\Gamma(\alpha > 1) = (\alpha-1)\cdot\Gamma(\alpha-1)$$

$$\Gamma(2.62) = 1.451396 = (1.62)(\Gamma(1.62)) = (\alpha-1)\cdot\Gamma(\alpha-1)$$

Problem 10: The Gamma Function - Special Case #2

$$\Gamma(\alpha \geq 2) = (\alpha-1)!$$

$$\Gamma(100) = 9.33e^{155} = 99! = (\alpha-1)!$$

Problem 11: The Beta Function

$$\text{Beta}(\alpha,\beta) = \frac{\Gamma(\alpha)\cdot\Gamma(\beta)}{\Gamma(\alpha+\beta)} \quad \text{for } \alpha > 0 \text{ and } \beta > 0$$

$$\text{beta}(3,2) = 0.08333 = \frac{\Gamma(3)\Gamma(2)}{\Gamma(5)} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

MSDS 454

Applied Probability & Simulation Modeling

Module 2: Markov Chains and Hidden Markov Models

Forsythe: Chapter 14 - Markov Chains (p. 331-344)

14.1 - Markov Chains

A markov chain is a sequence of random variables X_n for which the conditional probability at time t is solely dependent on $t-1$.

$P(X_n = j \mid t_0 \dots t_{n-1}) = P(X_n = j \mid x_{n-1})$

Transition Probabilities: $P(X_n = j \mid x_{n-1} = i)$ at each state.

*This text uses discrete-time, time-homogenous Markov Chains in finite state space

Markov Chains can be represented by a biased random walk over a finite directed graph $G(V,A)$ where V is a set of nodes and A is a set of arcs with probabilities $P(X_n = j \mid x_{n-1} = i)$. The sum of outgoing probabilities from each node V_i sum to 1.

Absorbing State: state at which the markov chain ends. This node has incoming but no outgoing nodes.

Recurrent: markov chains without an absorbing state exist in infinite sequence.

14.1.1 - Transition Probability Matrices

Matrix P with $p_{ij} = P(X_n = j \mid X_{n-1} = i)$ where $\sum_j P_{ij} = 1$

$$P_{ij} = \begin{array}{c c} & \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{array}{c} S_1 \\ S_2 \\ S_3 \end{array} & \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} \end{array}$$

← Stochastic Matrix ($P_{ij} > 0$)

π: k-dimensional row vector with the probability distribution for the first state.

$$\pi = P(X_1 = j) = \sum_i P(X_1 = j, X_0 = i) \leftarrow \text{each } (i,j) \text{ combination}$$
$$= \sum_i P(X_1 = j \mid X_0 = i) P(X_0 = i) \leftarrow$$
$$= \sum_i P_{ij} \pi_i$$

Example 14.3: Virus Mutation

A virus mutates between three states. At each point, the probability of staying at the same state is α. The probability of switching to each of the other states is $\frac{1-\alpha}{2}$.

With $\alpha = 0.8$, the transition probability matrix $P_{ij} = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$

14.1.2 - Stationary Distributions

Irreducibility: markov chains without absorbing states are irreducible.

Stationary Distribution: S = probability distribution that an irreducible markov chain takes over ~~time~~

$$s = \lim_{n \to \infty} \pi P^{(T)}$$

S is an eigenvector of P^T

14.2 - Estimating Properties of Markov Chains

14.2.1 - Simulation

The Weak Law of Large Numbers (Bernoulli's Theorem)

States that for independent and identically distributed random variables, the sample mean $\bar{x}$ approaches the population mean x as sample size n increases. $X_i \sim \text{iid}$, $\lim_{n \to \infty} P(|\bar{X}_n - \mu| \geq \varepsilon) = 0$

For simulation, the outcomes assume the true distribution of outcomes as the number of experiments increases.

14.2.2 - Simulation Results as Random Variables

The outcome of a simulation is a random variable that is normally distributed. The standard deviation of simulation outcomes behaves according to $C/\sqrt{N}$ is some constant and N is the number of simulations.

In effect, this results in doubling accuracy requiring four times as many simulations.

Canvas: Classifying Markov Chains

1) Classification of States for Markov Chains

$$P_{ij} = t \left\{ \begin{array}{c c} & \overbrace{\begin{array}{ccc} S_1 & S_2 & S_3 \end{array}}^{t+1} \\ \begin{array}{c} S_1 \\ S_2 \\ S_3 \end{array} & \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} \end{array} \right.$$

$P_{11} = P(X_{t+1} = S_1 \mid X_t = S_1)$

$P_{ij} = \sum_j P(X_{t+1} = S_j \mid X_t = S_i)$

$P_{i1} = \sum_i P(X_{t+1} = S_1 \mid X_t = S_i)$

Accessibility: A state S_j is accessible from another state S_i if a starting state S_i has a positive probability to reach S_j at some point.

$S_i \; (S_j \leftarrow S_i)$ if $(P_{ij}^n > 0 \; \exists n > 0)$

Example 1) The number of steps from state to state

It is easy to conceptualize this in 2x2 transition matrix like the one showed below

$P_{ij} = \begin{bmatrix} 0.9 & 0.1 \\ 0.8 & 0.1 \end{bmatrix}$ $S_1 \to S_2$ and $S_1 \leftarrow S_2$ in this matrix.

Communicate: states accessible to one another are said to communicate

Example 2) Do the states communicate across the matrix?

$P_{ij} = \begin{bmatrix} 0.5 & 0.3 & 0 \\ 0 & 0.8 & 0.2 \\ 0 & 0 & 1 \end{bmatrix}$

$S_1 \to S_2$ but $S_1 \nleftarrow S_2$

$S_1 \nrightarrow S_3$ and $S_1 \nleftarrow S_3$

$S_2 \to S_3$ but $S_2 \nleftarrow S_3$

Absorbing State: For S_i where $P_{ii} = 1$, states are called absorbing states.

Also referred to as 'sinks.'

Communicating Class: A set of mutually accessible states.

Example 3) Block Diagonal Matrix

$P_{ij} = \begin{bmatrix} 0.6 & 0.4 & 0 & 0 \\ 0.7 & 0.3 & 0 & 0 \\ 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0.8 & 0.2 \end{bmatrix}$

The communicating classes are $\{S_1, S_2\}$ and $\{S_3, S_4\}$. The two classes are not accessible from one another.

Irreducibility: Discrete time markov chains are said to be irreducible if they contain a single communicating class.

2) Hitting Times

Hitting Time (First Passage Time): the number of steps it takes for a markov chain to reach a specified state for the first time. This is a random variable. T_{ij}. The mean/expected hitting time is usually of interest.

3) Recurrent vs. Transient States

Recurrent State: a state likely to be repeatedly taken by the markov chain.

T_{ii} where $P(T_{ii} < \infty) = 1$

Transient State: a state not likely to be taken by the markov chain again.

$P(T_{ii} < \infty) = 0$ or $P(T_{ii} = \infty) = 1$

Absorbing State: $S_i \sim P_{ii} = 1$ (will be on the main diagonal)

Example 1) Recurrent States

$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix}$ Both S_1 and S_2 in this markov chain.

Example 2) Recurrent and Transient states within a matrix

$$P = \begin{array}{c c} & \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{array}{c} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} & \begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \end{array}$$

$\{S_1, S_2\}$ are transient states

$\{S_3, S_4, S_5\}$ are recurrent states.

Canvas: Classifying Markov Chains

3. Recurrent vs. Transient states

Example 3) Recurrent, Transient, and Absorbing states within a matrix

$$P: \begin{array}{c|cccccc} & S_1 & S_2 & S_3 & S_4 & S_5 & S_6 \\ \hline S_1 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ S_2 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ S_3 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 \\ S_4 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 \\ S_5 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 \\ S_6 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}$$

$\{S_1, S_2\}$ are transient states

$\{S_3, S_4, S_5\}$ are recurrent states

$\{S_6\}$ is an absorbing state

4. Stationary Distribution

Stationary Distribution: the distribution a markov chain takes as steps go to infinity. $\lim_{n \to \infty} P^n = Z$

Canvas: Markov Properties

Stochastic Process: a random variable indexed over time.

$X_t(\omega) = E, \omega \in \Omega$

where: E = state space

Ω = sample space

Examples of Markov Processes

1) Poisson: models frequency of events
2) Geometric Brownian Motion: movement of stock prices
3) Brownian Motion: movements of particles in physics

Markov Property: $P(X(\omega) = X_t \mid X_0, X_1, \dots, X_{t-1}) = P(X(\omega) = X_t \mid X_{t-1})$

$P(X_t \mid X_{t-1})$ is always the same. It is time homogenous

Module 3: Markov Chain Modeling and Simulation

Assignment 5: Modeling an Epidemic

An individual can take one of four states: susceptile, contagious, immune, and dead.

In a community of 100 humans, everyone starts as susceptile at t_0.

At t_1, the first person gets infected.

Assuming a contagion rate of three, meaning a contagious person infects three people per t.

If fatality rates is 2%, and recovery occurs after t_{t+1}.

P:

	S_1	S_2	S_3	S_4
S_1	0.97	0.03	0	0
S_2	0	0	0.98	0.02
S_3	0	0	1	0
S_4	0	0	0	1

Module 5: Random Number Generation and Simulation

Canvas 1) Linear Congruential RNG

The linear congruental method is an algorithm that produces pseudo-random numbers that are uniformly distributed integers between zero and an upper limit. The equation the method uses is:

$$X_{n+1} = AX_n + c \ \% \ m$$

where: A = multiplier

c = increment

m = modules

* X_0 must be specified

The produced sequence can have a periodicity of m under certain conditions

Canvas 2) Box-Mueller RNG

The Box-Mueller Method is an algorithm that produces pairs of indepependent, normally distributed random variables from pairs of a uniform distribution between 0 and 1.

Given U_1 and U_2 where $0 \le U_1, U_2 \le 1$ and are uniformly distributed.

$z_1 = R\cos(\theta) = \sqrt{-2\ln U_1}\cos(2\pi U_2)$

$z_2 = R\sin(\theta) = \sqrt{-2\ln U_1}\sin(2\pi U_2)$

Canvas 3) Quantile / Inverse Transform Method

Algorithm for generating sequences of random numbers that follow a specified statistical distribution given an intial value selected from a uniform distribution between 0 and 1, which is then transformed using the inverse of the desired cumulative distribution function.

Module 4: Hidden Markov Models

Canvas: Biological Sequence Analysis

Chapter 3: Markov Chains and Hidden Markov Models

3.1 - Markov Chains

Start and end states can be added as quiet states to markov chains to allow for measurement of the distribution of chain lengths.

Beginning sequence: $x_0 = B$, $P(x_1 = x_i) = P(x_1 = x_i \mid x_0 = B)$

Thus, the model defines a probability distribution over all possible lengths.

3.2 - Hidden Markov Models

Hidden Markov Model: an observable process Y that is dependent on a markov process X in a known way where X is unobservable. Thus, the state and state sequence of X are unknown.

Canvas: The Viterbi Algorithm

State sequence path: π

Transition probabilities: $a_{kl} = P(\pi_i = l \mid \pi_{i-1} = k)$

Emission probabilities: $e_k(b) = P(x_i = b \mid \pi_i = k)$

Joint probabilities: $P(X, \pi) = \underbrace{a_{0\pi_1}}_{\text{transition probability}} \prod_{i=1}^{L} \underbrace{e_{\pi_i}(x_i)}_{\text{emission probability}} \underbrace{a_{\pi_i \pi_{i+1}}}_{\text{transition probability}}$ where $\pi_{L+1} = 0$

Most probable path: $\pi^* = \underset{\pi}{\operatorname{argmax}} P(x, \pi)$

Suppose $P(\pi^*)$ for chain ending in state k at observation i $v_k(i)$ is known for all states k. Then $v_l(i+1) = \underbrace{e_l(x_{i+1})}_{\text{emission probability}} \max(\underbrace{v_k(i)}_{P(\pi^*)} \underbrace{a_{kl}}_{\text{transition probability}})$

The Forward Algorithm

Similar to the Viterbi Algorithm, but with summation instead of maximization.

$v_k(i) \approx f_k(i) = P(x_1 \dots x_i, \pi_i = k)$

The recursive equation is $f_l(i+1) = e_l(x_{i+1}) \sum_k f_k(i) a_{kl}$

The Backward Algorithm and Posterior State Probabilities

Backward recursive equivalent of the forward algorithm.

$b_k(i) = P(x_{i+1} \dots x_L \mid \pi_i = k)$

Module 6: Discrete Event Simulation

Topic: Queueing Theory

Arrival Process: defines the arrival of customers into the queueing system, often modeled as a Poisson process, assuming arrivals are random and independent.

Service Time: how long it takes to service a customer, modeled as a random variable, often assuming an exponential distribution.

Queue Length: the number of customers in a queue, can serve as a performance measure of a queuing system.

Queueing Models

Three Character Kendall's Notation: A/S/c

where: A = arrival

S = service time

c = servers

Module 5: Random Number Generation and Simulation

Canvas 1) Linear Congruential RNG

The linear congruential method is an algorithm that produces pseudo-random numbers that are uniformly distributed integers between zero and an upper limit. The equation the method uses is:

$$X_{n+1} = AX_n + c \,\%\, m$$

where: A = multiplier

c = increment

m = modulus

- X_0 must be specified

The produced sequence can have a periodicity of m under certain conditions

Canvas 2) Box-Mueller RNG

The Box-Mueller Method is an algorithm that produces pairs of independent, normally distributed random variables from pairs of a uniform distribution between 0 and 1.

Given U_1 and U_2 where $0 \le U_1, U_2 \le 1$ and are uniformly distributed.

$Z_1 = R\cos(\theta) = \sqrt{-2\ln U_1}\cos(2\pi U_2)$

$Z_2 = R\sin(\theta) = \sqrt{-2\ln U_1}\sin(2\pi U_2)$

Canvas 3) Quantile / Inverse Transform Method

Algorithm for generating sequences of random numbers that follow a specified statistical distribution given an initial value selected from a uniform distribution between 0 and 1, which is then transformed using the inverse of the desired cumulative distribution function.

Module 6: Discrete Event Simulation

Topic: Queueing Theory

Arrival Process: defines the arrival of customers into the queueing system, often modeled as a Poisson process, assuming arrivals are random and independent.

Service Time: how long it takes to service a customer, modeled as a random variable, often assuming an exponential distribution.

Queue Length: the number of customers in a queue, can serve as a performance measure of a queueing system.

Queueing Models

Three Character Kendall's Notation: $A/S/c$

where: A = arrival

S = service time

c = servers

Module 7: Continuous Random Variable Generation

Reading 1: Essentials of Monte Carlo Simulation

Chapter 4: Generating Continuous Random Variates

Probability Density Function: determines the relative likelihood that a continuous random variable will take a given value within its domain.

Continuous Uniform Distribution

Probability Density Function: $f(x) = \frac{1}{b-a}$ for $a \le x \le b$

Cumulative Probability Density Function: $F(x) = \frac{x-a}{b-a}$ for $a \le x \le b$

Expected Value: $E[x] = \frac{b+a}{2}$

Variance: $V(x) = \frac{(b-a)^2}{12}$

Routine:
1) Generate a random uniform $u \sim U(0,1)$
2) $x = a + u(b-a)$
3) return(x)

Exponential Distribution

The exponential distribution describes the time inbetween events in a Poisson process.

Probability Density Function: $f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$

Cumulative Probability Density Function: $F(x;\lambda) = \begin{cases} 1 - e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$

Expected Value: $E[x] = 1/\lambda$

Variance: $V(x) = 1/\lambda^2$

Routine: Using the inverse transform method,
1) Generate a continuous uniform variable $u \sim U(0,1)$
2) Define PDF of exponential as $F(x) = 1 - e^{-\lambda x}$, set equal to $F(x) = u$
3) Solving produces $x = -\frac{1}{\lambda}\ln(1-u)$

Erlang Distribution

Gamma Distribution