Probability and Statistics

Sanjit K. Kaul

April 26, 2022

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

What do you think this course is going to be about?

How many of you would rather not take the course?

Our brains are just not wired to do probability problems very well - Persi Diaconis

From the book Randomness by Deborah J. Bennett

Randomness

Uncertainty

Chance

Our Probabilistic World

- Chance, its quantifications, pervade our world

- Not just in a casino or when playing a game of cards or when solving a math puzzle

We All Must Grapple With Chance

- A fraction of natural births are identical twins

We All Must Grapple With Chance

- A fraction of natural births are identical twins

- Tests screening for disease result in false positives and false negatives
  - Your test is negative. Should the doctor put you on medication?

We All Must Grapple With Chance

- A fraction of natural births are identical twins
- Tests screening for disease result in false positives and false negatives
  - Your test is negative. Should the doctor put you on medication?
- Vaccine efficacy
  - You are safe, rather safer, probabilistically
  - Your odds of a severe infection are a fraction of that of one not vaccinated

We All Must Grapple With Chance

- The odds of becoming a lightning victim in the U.S. in any one year is 1 in 7,00,000 (NatGeo)
- The odds of being struck in your lifetime is 1 in 3,000 (NatGeo)
- Are commercial flights more dangerous than lightning?

We All Must Grapple With Chance

- The odds of even being in a plane crash are one in 11 million.

- You end up in a plane crash. What are the odds that you will survive?
  - 95%

Chance As a Decision Maker

Draw of lots to allocate flats

Delhi Development Authority has launched 1354 flats on offer on January 2, 2021. Around 33000 applications have been received till 16 February 2021. Out of these 33000 applications around 22,500 applicants have made payment... *On 10th March 2021, **the draw of lots** for the allotment of flats will be held.*

Chance As a Decision Maker

- Random Elimination using *"Akkad bakkad bumbe bo, assi nabbe pure so, sow mein nikala dhaaga, chor nikal ke bhaaga."*
  - *Random?*

Chance As a Decision Maker

- Assignment of grades via a draw of lots?

Did the Ancients Use Chance?

"Chance mechanisms have been used since antiquity: to divide property, delegate civic probabilities or privileges, settle disputes..."

- One die dated around 2750 BC from excavations of ancient Mesopotamia in Northern Iraq
- Dice found at the Mohenjo-Daro excavation in the Indus valley

From Randomness by Deborah J. Bennett

Probability

- It’s a number

- By convention in [0,1]

Probability

- A number that quantifies belief in the occurrence of something

- A number that, for our purposes, maybe arrived at as a result of experiment

Let's Toss a Coin (In class exercise)

https://www.google.com/search?q=toss+a+coin&oq=toss+a+coin

Measurements and Uncertainty

- You perform an experiment multiple times and obtain different results/measurements/outcomes

- What do you do?
  - Resign to such uncertainty?

- You capture the uncertainty using a probability model

Measurements and Uncertainty

- You could perform the experiment many times
- Associate with each of the different outcomes their frequency of outcome
  - Assign frequency of occurrence to the probability
- You have a probability model
  - Captures in a deterministic manner the uncertainty you saw in your experiments

Measurements and Uncertainty

- You could perform the experiment many times
- Associate with each of the different outcomes their frequency of outcome
  - Assign frequency of occurrence to the probability
- You have a probability model
  - Captures in a deterministic manner the uncertainty you saw in your experiments

- Often, in real world experiments, many repeats later you see that the uncertainty can be captured in such a manner

Probability Models

- Many commonly occurring probability models
  - They define the set of all possible outcomes
  - Associate probabilities with outcomes
- You can create your own
  - Must satisfy some commonsense rules
- Different models can be thought of as different randomizers/ methods of drawing lots

Why Probability Models?

- We don’t toss coins and roll dice often

- Maybe you don't play games that require tossing coins or rolling dice

Why Probability Models?

- We don’t toss coins and roll dice often
- Maybe you don't play games that require tossing coins or rolling dice
- For us coins and dice are models that we understand well
- Importantly, they often help model randomness in the real world

Why Probability Models?

- To exemplify, a fair coin can model the "natural process" that determines whether genetic sex of a newborn is male or female

Why Probability Models?

- To exemplify, a fair coin can model the "natural process" that determines whether genetic sex of a newborn is male or female

- A communications channel can be modelled as a coin too – a biased one however

Why Probability Models?

- To exemplify, a fair coin can model the "natural process" that determines whether genetic sex of a newborn is male or female
- A communications channel can be modelled as a coin too – a biased one however
- Of course, not everything is discrete
- Probability models may be more complex than a coin or dice

Why Think in Terms of Probability?

- Some observed phenomena are random (non-deterministic)
  - Thermal Noise
    - Impossible to predict the value of noise at a given instant
    - However, its average is 0 and deterministic
  - Gas contained in a closed vessel
    - Atoms have random trajectories and apply random pressures to the walls of the vessel.

Why Think in Terms of Probability?

- Phenomena may be deterministic but too complex to model deterministically
  - Coin tossing
  - Game of darts
    - Where will the dart hit?
  - Where will a missile hit?

Why Think in Terms of Probability?

- Phenomena may be deterministic but appear chaotic in nature
  - Very small changes in initial conditions, states can lead to very different observations (Butterfly Effect)
  - Your model is a differential equation.
  - However, predicting an output for a given input or state of the system is impossible
    - Small errors in input or state measurement can lead to very different measurements

Probability and Statistics

- Probability models help characterize experimental outcomes

- In the probability part of the course we will assume that a probability model is given/ known/ implicit
  - It *governs* your experiment's outcomes, how the lots are drawn

Probability and Statistics

- Probability models have parameters
  - Just as a line has the parameters slope and y-intercept
- You may not know the parameters in the real world
- You may have only observations/ experiment outcomes governed by the probability model

Probability and Statistics

- In the statistics part of the course, you will begin with measurements/outcomes of an experiment

- You will use the measurements to calculate statistics

Probability and Statistics

- Each parameter of a model *corresponds to* a *statistic* of a collection of numerical outcomes of the experiment

  - Parameter: Expected value, Statistic: *Mean* of outcomes (empirically calculated, comes with uncertainty).

Probability and Statistics

- Probability models will become hypotheses

- Given measurements, you will decide which hypothesis, from a set of hypotheses, it came from

Brief Detour

Textbooks, Grading and etc

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Textbooks

- Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers by Roy D. Yates (2nd Edition) [Indian Edition].
- Introduction to Probability and Statistics for Scientists and Engineers By Sheldon Ross (5th Edition)
- Most homework will be assigned from these books.
- Topics covered in both are very similar. However, organization of topics is different

A Few Dos and Don'ts

- Do not allow coursework to pile up
- Do not come to a class without revising what was taught in the previous class
- Do not read the book as if it were a novel
- Do not learn by rote, you will fail to learn with probability 1

A Few Dos and Don'ts

- The authors RY&DG make very important observations (read "Message to Students from the Authors" in the Preface)
  - Being able to use formulae is not enough
  - You must be able to solve *word* problems
- There is only one way of getting better at the probabilistic way of thinking and that is to keep working on word problems
- Work out problems on a sheet of paper, not in air, not in your head, not in your imagination or dreams
- Use proper notation (as in book by RY)

A Few Dos and Don'ts

- There is only one way of getting better at the probabilistic way of thinking and that is to keep working on word problems

- Work out problems on a sheet of paper, not in air, not in your head, not in your imagination or dreams

- Use proper **notation** (as in book by RY)

Gadgets Policy in Class

- No phones, laptops, and such like.
- Only pen/pencil and paper

On Copying

- **<u>Axiom 0</u>** If you copy from others you cannot learn the subject

- **If you want to learn** and are in a position where you think that copying is the only way out
  - Do not copy
  - Meet me/TA in person and we will work on your situation

On Copying

- **If you do not want** to learn, feel free to copy.

- **Penalties will apply**
  - **At least as harsh as the institute policy**
    - Copying and/or helping another student during an exam leads to an F as per institute policy.

On Attendance

- Must enter classroom within 10 minutes of start
  - This is non-negotiable
  - Important to maintain decorum for the 50 or so minutes we have
- If you leave, you will be marked absent
- If you leave after giving attendance, you will be awarded the F grade.
- Must submit in class exercise sheet (always one sheet, two sides) at end of lecture
  - Put your name and roll number on it
  - Will be graded with a non-zero probability

Grading (Most Likely Distribution)

- Exam based evaluations (85%)
  - Midsem (20%), Endsem (25%), Three quizzes (40%)

- Assignments (0 - 10%)
  - They are for you to learn well topics covered during lectures
  - While the percentage is just 10%, how sincerely you do your assignments will determine your score in the exams

- In class exercise submissions (0 - 5%)
- Coding component (0 – 15%)

More on Grading

- Show your steps, else you get a 0.
- Correct answers are important. The approach is more important

Keeping You on Track

- Assignments
- Exams
- Tutorials (please attend them)

Homework/Assignments

- Discussion is fine
  - Please give credit
- **Submitted solution must be yours**
- Will also feed into out of class interaction
  - Evidence of effort at your end will make me more likely to give you my time
- Solutions will be posted post deadline
- Please don't ignore assignments

Tutorials

- We will solve problems
- Your suggestions
  - Please suggest at least a few days in advance
- Some tutorial slots maybe used to conduct the quizzes

Notation

- Always use the notation we use in class
  - As in book by RY
  - It is not okay to write mathematics using arbitrary notation
    - Define any new notation that you may need to use. Else stick to RY

- You will not get marks for what is mathematically incorrect

Notation

- When you read the book, read the mathematics carefully

- The language of mathematics leaves nothing to chance
  - There should be no ambiguity in a mathematical statement
  - If you feel otherwise when you read, you may want to read more carefully

Plan

- Experiments, outcomes, sets, events, MECE, Bayes' rule, law of total probability, counting
  - Weeks 1 – 2
- Discrete RV
  - Weeks 2 – 3
- Continuous RV
  - Week 3 – 4
- Pairs/ Vectors of RV(s)
  - Weeks 4 – 6
- Chebyshev's Inequality, Sample mean, CLT, Weak Law, …
  - Weeks 7 – 8
- Estimation
  - Weeks 9 – 10
- Hypothesis Testing
  - Weeks 11 - 12

Post Conditions

- Students are able to derive first and second order statistics, conditional PDF(s) and PMF(s), of common random variables and functions of the same.
- Students are able to apply commonly found discrete and continuous state probabilistic models to various applications, as captured by word problems
- Students are able to derive properties of the sample mean and variance of samples from a normal population
- Students are able to derive estimates of the parameters of common probability distributions and test hypothesis from random samples
- Students are able to perform MATLAB (or equivalent) experiments to present data, generate data from distributions, calculate various statistics, and demonstrate basic hypothesis testing

You and Your Instructor

- You are free (in my class) to ask questions
  - Please be polite. Your politeness encourages me to pay attention
- Questions on what is being taught are welcome anytime during the class
  - You do not want me to go to slide 10 if you have questions regarding slide 9
- You will be graded based on what you do in exams and assignments. Not based on what you say or don't

You and Your Instructor

- Other things you may want to say are fine too
  - However, I would encourage you to keep general comments and observations, grievances, course feedback to office hours.

- Please remember that I can choose to not allow a certain interaction at any given time during the class or during office hours
  - We don’t have infinite time

You and Your Instructor

- You are not allowed to beg for marks
  - By doing so, you will be wasting my time. I will penalize you for the same.

- You are not entitled to marks. You must earn them.

You and Your TA(s)

- Please be polite (and expect the same in return)
- Respect the TA's time
  - Come prepared with questions you may have
  - Bring your attempt

Course Website

- On Classroom
  - Code: **idddndm**
  - Probability And Statistics Section B
- All messaging over Classroom
  - **In general, I won't respond to emails**
- Encourage you to share your questions with all
  - No question is a stupid question
  - It is stupid, however, to not ask a question
  - Encourage you to try working out the answer before you ask
  - ***Do not share entire solutions on Classroom***

Enjoy The Course

- Not a "math course"
- We will learn via fun toy problems
- We will learn to think probabilistically
- What we will learn has very broad applicability
- Don't fret, Have fun!

Experiments, Models and Probabilities (Chapter 1 RY)

The fun begins!

Many slides courtesy RY (2nd Ed) instructor material

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Defining Our Universe of Interest

- Our universe of interest has to do with our experiment
- What was it for the coin tossing experiment we performed?
- Set theory provides the formalism we need to capture our universe of interest and its inhabitants

Set Theory

- A set is a collection of objects
- A = {Acushla, Sanjit}
  - A is the set of faculty members teaching probability and statistics in IIITD in 2022
- We will use **capital letters to denote sets**

Set Theory

- Various ways of defining a set

$$A = \{x | x \text{ is an Integer}\}$$

$$B = \{y | y\%2 = 0, \text{ y is an Integer}\}$$

$$B = \{y \text{ is an Integer} | y\%2 = 0\}$$

$$C = \{x^2 | x = 1, 2, 3\}$$

$$D = \{x^2 | x = 1, 2, 3, 4, \ldots\}$$

Set Theory

- Sets can be finite or infinite

- Sets can be finite (and hence also countable) and countably infinite
  - Set of days in a week
  - Set of students in MTH 201
  - Set of integers
  - Set of values that an image pixel may take

  - Set of rational numbers...? Yes/No? Why?

Set Theory

- Sets can be finite or infinite

- Sets can be a continuum
  - Set of all real numbers in the interval (0,1)
  - Set of temperatures
  - Set of heights of humans

Set Theory: Definitions and Notation

$A \subset B$ says that A is a subset of B.
If $A \subset B$, then by definition all members of A are also members of B.

$A \supset B$ implies that A is a superset of B.

Two sets A and B are equal, that is $A = B$, *if and only if* $A \subset B$ and $B \subset A$

Set Theory

- Universal Set S: It contains *everything*
  - All sets are a subset of S

- Null Set ϕ
  - It contains nothing
  - It is a subset of every set

For any set A, $\phi \subset A$

Set Theory

- To say that an element *a* belongs to (is a member of) the set A we write

$$a \in A$$

- We can also say

$$\{a\} \subset A$$

Set Operations

Union: $A \cup B$, $A \cup B \cup C$, $\cup_{i=1}^{\infty} A_i$

$\rightarrow x \in A \cup B$ iff $x \in A$ or $x \in B$

if $x \in A \cup B$, then $x \in A$ or $x \in B$

If $x \in A$ or $x \in B$, then $x \in A \cup B$

Set Operations

Union: $A \cup B$, $A \cup B \cup C$, $\cup_{i=1}^{\infty} A_i$

$\rightarrow x \in A \cup B$ iff $x \in A$ or $x \in B$

Intersection: $\cap_{i=1}^{\infty} A_i$

$\rightarrow x \in A \cap B$ iff $x \in A$ and $x \in B$

Complement of A is denoted as A^c

$\rightarrow x \in A^c$ iff $x \notin A$

Difference:

$\rightarrow x \in A - B$ iff $x \in A$ and $x \notin B$

Venn Diagrams

Set Theory

- **Mutually Exclusive Sets**

Sets A_i, $i = 1, 2, \ldots, n$ are mutually exclusive iff $A_i \cap A_j = \phi$ for $i \neq j$

Set Theory

- **Collectively Exhaustive Sets**

Sets A_i, $i = 1, 2, \ldots, n$ are collectively exhaustive iff $\cup_{i=1}^{n} A_i = S$

Theorem 1.1

De Morgan's law relates all three basic operations:

$$(A \cup B)^c = A^c \cap B^c.$$

Quiz 1.1 In-Class Submission

A pizza at Gerlanda's is either regular (R) or Tuscan (T). In addition, each slice may have mushrooms (M) or onions (O) as described by the Venn diagram at right. For the sets specified below, shade the corresponding region of the Venn diagram.

(1) R

(2) $M \cup O$

(3) $M \cap O$

(4) $R \cup M$

(5) $R \cap M$

(6) $T^c - M$

The Story That Must Not Be Forgotten

Examples Of Experiments

- Experiment consists of a repeatable procedure
  - Repeats may lead to different observations

- Example Procedures
  - Toss a coin
  - Toss a coin 10 times

- Example Observations
  - Number of heads
  - Sequence of heads and tails
  - Largest number of consecutive heads in the sequence

Outcome

- **Def 1.1:** Outcome
  - An outcome of an experiment is any observation of that experiment

- Outcomes are, by definition, mutually exclusive

Outcome

- Heads, Tails are the possible outcomes of a coin tossing experiment
  - They are mutually exclusive, cannot happen at the same time

Sample Space

- **Def 1.2:** The sample space S of an experiment is the
  - finest-grain,
  - mutually exclusive,
  - collectively exhaustive set of all possible outcomes

- Every element of the sample space is an outcome
  - Each of the outcomes are atomic in nature
    - Cannot be split into more than one outcome
  - The outcomes are also mutually exclusive
  - No outcomes are left out of the sample space

Sample Space

- By default, S is your universe of interest. It is a given.

- *Everything is conditioned on S (super set) having happened*
  - *What's the uncertainty about?*

$\{0\}, \{1\}, \ldots, \{10\}$

$\{0\} \subset S$

$S \supset \{0\}$

- In the ten coin toss experiment S = {0,1,2,…,10}

Sample Space

- You want to observe the gender sequence obtained after two people have walked past you
- We want to come up with the sample space

- Possible observations are …? Note that we are looking for outcomes

$S = \{MM, FF, FM, MF\}$

Sample Space

- Each one of the above is also an outcome
  - Note MM for your stated observation cannot be split into smaller observations
  - Also, the outcomes are mutually exclusive. Only one can happen at any given time

- S = {MM, FF, FM, MF}

Event and Event Space

- What if I am interested in whether a man was spotted among the first two people?

- No individual outcome is sufficient

Event and Event Space

- **Def 1.3 Event:** An Event is a set of outcomes

- For our example one event of interest, call it E, is
  - E = {MM, MF, FM}

- Note that we say that event E occurs when any of the outcomes MM, MF and FM occur.
  - Think in terms of the definition of union of sets

Event and Event Space

- **Def 1.4 Event Space:** A set of mutually exclusive and collectively exhaustive events is an event space

$S = \{1, 2, \ldots, 6\}$

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

$O = \{1, 3, 5\}$

$Z = \{E, O\}$. Given Def 1.4, Z is an event space.

$Z_1 = \{\{1\}, \{2\}, \ldots, \{6\}\}$

Example – Experiment Coin Flips (In Class Exercise Submission)

- Procedure: Flip three coins
- Observation: Sequence of heads (h)/tails (t) that is obtained
- Q1) Give an example outcome
- Q2) What is the Sample space of the experiment?
- How many elements does it contain?
- Let B_i be the event when the sequence contains *i* heads
  - Q3) What range of values can *i* take?
  - Q4) Are the B_i mutually exclusive?
- Q5) Is $B = \{B_0, B_1\}$ an event space?

Why Event Space?

- They maybe easier to handle

- For the above example, sample space contains ? outcomes, while the event space contains just ? events

Why Event Space?

- If the sequence length is increased to 50
  - Sample space has 2^{50} outcomes
  - Event Space has just 51 events, and is a much smaller set

- Clearly the event space does not contain all the information of the sample space
  - However, we may not always be interested in all the information

Partitioning An Event into Mutually Exclusive Events

- **Theorem 1.2**

For an event space $B = \{B_1, B_2, \ldots\}$ and any event A in the sample space, let $C_i = A \cap B_i$. For $i \neq j$, the events C_i and C_j are ME and $A = C_1 \cup C_2 \cup \ldots$.

- A very useful theorem!

Event and Event Space

- **Def 1.4 Event Space:** A set of mutually exclusive and collectively exhaustive events is an event space

$S = \{1, 2, \ldots, 6\}$

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

$O = \{1, 3, 5\}$

$Z = \{E, O\}$. Given Def 1.4, Z is an event space.

$Z_1 = \{\{1\}, \{2\}, \ldots, \{6\}\}$

Example – Experiment Coin Flips (In Class Exercise Submission)

- Procedure: Flip three coins
- Observation: Sequence of heads (h)/tails (t) that is obtained
- Q1) Give an example outcome
- Q2) What is the Sample space of the experiment?
- How many elements does it contain?
- Let B_i be the event when the sequence contains *i* heads
  - Q3) What range of values can *i* take?
  - Q4) Are the B_i mutually exclusive?
- Q5) Is $B = \{B_0, B_1\}$ an event space?

Why Event Space?

- They maybe easier to handle

- For the above example, sample space contains ? outcomes, while the event space contains just ? events

Why Event Space?

- If the sequence length is increased to 50
  - Sample space has 2^{50} outcomes
  - Event Space has just 51 events, and is a much smaller set

- Clearly the event space does not contain all the information of the sample space
  - However, we may not always be interested in all the information

Partitioning An Event into Mutually Exclusive Events

- **Theorem 1.2** $\bigcup_i B_i = S$ $\quad B_i \cap B_j = \phi \ \forall i, j \ i \neq j$

For an event space $B = \{B_1, B_2, \ldots\}$ and any event A in the sample space, let $C_i = A \cap B_i$. For $i \neq j$, the events C_i and C_j are ME and $A = C_1 \cup C_2 \cup \ldots$.

- A very useful theorem!

Experiment Tweet

- Procedure: You send a tweet over amateur radio using Morse code. Your tweet is 140 characters long.
- A character is received incorrectly with probability *p*.
- Your observation is the number of characters that were received correctly.
  - What is your sample space?

$S = \{0, 1, 2, \ldots, 140\}$

Experiment Tweet

- Let A_k be the event that at least k characters are received correctly
  - Do the A_k together form an event space?

- Can you think of another set of events that does?
- Express one in terms of the other…

Experiment Tweet

- Can you think of another set of events that does?

- Express one in terms of the other…

Quiz 1.2 In-class exercise

Monitor three consecutive phone calls going through a telephone switching office. Classify each one as a voice call (v) if someone is speaking, or a data call (d) if the call is carrying a modem or fax signal. Your observation is a sequence of three letters (each letter is either v or d). For example, two voice calls followed by one data call corresponds to vvd. Write the elements of the following sets:

(1) $A_1 = \{$first call is a voice call$\}$
(2) $B_1 = \{$first call is a data call$\}$
(3) $A_2 = \{$second call is a voice call$\}$
(4) $B_2 = \{$second call is a data call$\}$
(5) $A_3 = \{$all calls are the same$\}$
(6) $B_3 = \{$voice and data alternate$\}$
(7) $A_4 = \{$one or more voice calls$\}$
(8) $B_4 = \{$two or more data calls$\}$

For each pair of events A_1 and B_1, A_2 and B_2, and so on, identify whether the pair of events is either mutually exclusive or collectively exhaustive or both.

Problem 1.2.2

An integrated circuit factory has three machines X, Y, and Z. Test one integrated circuit produced by each machine. Either a circuit is acceptable (a) or it fails (f). An observation is a sequence of three test results corresponding to the circuits from machines X, Y, and Z, respectively. For example, aaf is the observation that the circuits from X and Y pass the test and the circuit from Z fails the test.

(a) What are the elements of the sample space of this experiment?

(b) What are the elements of the sets

$$Z_F = \{\text{circuit from } Z \text{ fails}\},$$
$$X_A = \{\text{circuit from } X \text{ is acceptable}\}.$$

(c) Are Z_F and X_A mutually exclusive?

(d) Are Z_F and X_A collectively exhaustive?

(e) What are the elements of the sets

$$C = \{\text{more than one circuit acceptable}\},$$
$$D = \{\text{at least two circuits fail}\}.$$

(f) Are C and D mutually exclusive?

(g) Are C and D collectively exhaustive?

Definitions of Probability

- Classical Definition

  Probability of an event A is given by

  $P[A] = \frac{N_A}{N}$

  - Basically, number of favorable outcomes divided by the total number of possible outcomes
    - Probability that roll of a die leads to an even number…?
  - How about an improvement to the above?
    - [redacted]

- Relative Frequency Definition

$$P[A] = \lim_{n \to \infty} \frac{n_A}{n}$$

  - You perform *n* experiments and observe the event A n_A times to come up with the probability above.

Definitions of Probability

- Axiomatic Definition
  - A few axioms and a few definitions is all you need ☺
  - In this course we will use the Axiomatic definition
  - Once in a while we will go back to the Relative Frequency Definition

Axioms of Probability

- We started with an Experiment
- Experiments involved a procedure, observations and a model
- We started with a model by mapping observations to outcomes, sample space, events, and event space
- Now we want to associate a probability (a number, a measure) with each event in S

Let $P[A]$ denote the probability of event A

Three Axioms of Probability

A1: For any event A, $P[A] \geq 0$

A2: $P[S] = 1$

A3: For any countable collection $A_1, A_2, \ldots$ of ME events $P[A_1 \cup A_2 \cup \ldots] = P[A_1] + P[A_2] + \ldots$

- We assume that the above are true.
- All of the rest follows from these axioms
- This is also called the Axiomatic approach and is a fairly recent approach (starting early 1900(s))

Use A1, A2, A3 to show that

$$P[A_1 \cup A_2 \cup \dots \cup A_n] = P[A_1] + P[A_2] + \dots + P[A_n]$$

when $A_1, A_2, \dots, A_n$ are mutually exclusive.

$B_i = A_i, \quad i = 1, 2, \dots, n$

$B_j = \Phi, \quad j = n+1, \dots$

$$\underbrace{P[A_1 \cup A_2 \cup \dots \cup A_n]} = P\left[\bigcup_{j=1}^{\infty} B_j\right]$$

$$= \underbrace{P[B_1] + P[B_2] + \dots + P[B_n] + P[\Phi] + P[\Phi] + \dots}_{\downarrow} = P[A_1] + \dots + P[A_n] + P[\Phi] + P[\Phi] + \dots$$

$$\Phi = \underbrace{(\Phi \cap B_1)}_{\triangleq C_1} \cup \underbrace{(\Phi \cap B_2)}_{\triangleq C_2} \dots$$

let $\Phi \cap B_i = C_i$

$$\Phi = C_1 \cup C_2 \cup C_3 \dots \qquad C_i \cap C_j = \Phi \quad i \neq j$$

$$P[\Phi] = P[(\Phi \cap B_1) \cup (\Phi \cap B_2) \cup \dots]$$

$$= \sum_{i=1}^{\infty} P[\Phi \cap B_i] = \sum_{i=1}^{\infty} P[\Phi]$$

Theorem 1.4

If the events A_i, $i = 1, 2, \ldots, m$ are ME, then $P[A_1 \cup A_2 \cup \ldots \cup A_m] = \sum_{i=1}^{m} P[A_i]$

We will use just the three axioms!

Let $B_1, B_2, \ldots$ be mutually exclusive sets. Axiom $A3$ applies to the sets B_i.

Let $B_i = A_i$ for $1 \leq i \leq m$ and for $i > m$, let $B_i = \phi$

Proof of Theorem 1.4

Therefore, we have $\cup_{i=1}^{m} A_i = \cup_{i=1}^{\infty} B_i$

Also $P[\cup_{i=1}^{m} A_i] = P[\cup_{i=1}^{\infty} B_i]$

$$= \sum_{i=1}^{m} P[A_i] + \sum_{i=m+1}^{\infty} P[\phi]$$

Show that $P[\phi] = 0$ and we are done!

Using Theorem 1.2 from basic set theory, write $\phi = (\phi \cap B_1) \cup (\phi \cap B_2) \cup \ldots$ and... we are done (How?)

Theorem 1.5

The probability of an event $B = \{s_1, s_2, \ldots, s_m\}$ is the sum of the probabilities of the outcomes contained in the event:

$$P[B] = \sum_{i=1}^{m} P[\{s_i\}].$$

Theorem 1.6

For an experiment with $S = \{s_1, s_2, \ldots, s_n\}$ in which each outcome s_i is equally likely $P[s_i] = 1/n,\ 1 \leq i \leq n$

Proof: Simple. How?

Example – In Class Exercise

Score T is an integer between 0 and 100 corresponding to the outcomes $s_0, \ldots, s_{100}$. A score of 90 to 100 is an A and below 60 is a failing grade of F. Given that all scores between 51 and 100 are equally likely and a score of 50 or less never occurs,
find $P[\{s_{79}\}]$, $P[T \geq 90]$, $P[student\ passes]$.

Theorem 1.8

For any event A, and event space $\{B_1, B_2, \ldots, B_m\}$,

- Think Event Space…
- How are the B(s) related?
- Can I express A in terms of the event space?
- How about the P[A]?

$$P[A] = \sum_{i=1}^{m} P[A \cap B_i]$$

Quiz 1.4

Monitor a phone call. Classify the call as a voice call (V) if someone is speaking, or a data call (D) if the call is carrying a modem or fax signal. Classify the call as long (L) if the call lasts for more than three minutes; otherwise classify the call as brief (B). Based on data collected by the telephone company, we use the following probability model: $P[V] = 0.7$, $P[L] = 0.6$, $P[VL] = 0.35$. Find the following probabilities:

(1) $P[DL]$

(2) $P[D \cup L]$

(3) $P[VB]$

(4) $P[V \cup L]$

(5) $P[V \cup D]$

(6) $P[LB]$

$L = (L \cap D) \cup (L \cap V)$

$\because \{D, V\}$ is an event space

$P[L] = P[(L \cap D) \cup (L \cap V)]$

$= P[L \cap D] + P[L \cap V]$

$= P[LD] + P[LV]$

$P[H_0] + P[H_1] + P[H_2] = 1$

$P[H_0] = P[H_1] = P[H_2] = 1/3$

$P[F] + P[V] = 1$

$P[F] = 5/12$

$P[V] = 7/12$

Problem 1.4.6

Suppose a cellular telephone is equally likely to make zero handoffs (H_0), one handoff (H_1), or more than one handoff (H_2). Also, a caller is either on foot (F) with probability $5/12$ or in a vehicle (V).

(a) Given the preceding information, find three ways to fill in the following probability table:

	H_0	H_1	H_2
F			
V			

(b) Suppose we also learn that $1/4$ of all callers are on foot making calls with no handoffs and that $1/6$ of all callers are vehicle users making calls with a single handoff. Given these additional facts, find all possible ways to fill in the table of probabilities.

$P[FH_0] = 1/4$

$P[VH_1] = 1/6$

$$P[H_0] + P[H_1] + P[H_2] = 1$$

$$P[H_0] = P[H_1] = P[H_2] = 1/3$$

$$P[F] + P[V] = 1$$

$$P[F] = 5/12$$

$$P[V] = 7/12$$

	H_0	H_1	H_2
F	$P[FH_0]$	$P[FH_1]$	$P[FH_2]$
V	$P[VH_0]$	$P[VH_1]$	$P[VH_2]$

$P[F]P[H_0]$ $\quad P[F]P[H_1]$ $\quad P[F]P[H_2]$

$P[V]P[H_0]$ $\quad P[V]P[H_1]$ $\quad P[V]P[H_2]$

$$\longrightarrow \quad 0 \quad 1/3 \quad 5/12 - 1/3 \rightarrow 5/12$$

$$1/3 \quad 0 \quad 7/12 - 1/3 \rightarrow 7/12$$

$$\downarrow$$

Conditional Probability

- What is the probability that it is raining right now?
- The number you state is an expression of your belief
- What if I tell you that it is very cloudy outside?
  - Does this change your belief about rain?
- If A is the event {rain} and B is the event {cloudy sky}, then P[A] is your belief in absence of any other information
- P[A|B] is your belief after you are told that B has occurred

$P\left[\frac{A}{B}\right]$

Conditional Probability

- P[A] is called the a priori probability of event A
- It is the prior probability of A, prior to you knowing other facts
- P[A|B] is the probability of A given that B has occurred.
  - It is the conditional probability of A given B
- What is P[A|A]?
- Roll of a die: What is the probability of the outcome {2}?
- Let E be the event that an even number was obtained. What is P[{2}|E]?

Conditional Probability

- **Definition 1.6**

$$P[A|B] = \frac{P[AB]}{P[B]}$$

If A and B are mutually exclusive, then P[AB] = ?

P[A|B] = ?

Theorem 1.9

A conditional probability measure $P[A|B]$ has the following properties that correspond to the axioms of probability.

Axiom 1: $P[A|B] \geq 0$.

Axiom 2: $P[B|B] = 1$.

Axiom 3: If $A = A_1 \cup A_2 \cup \cdots$ with $A_i \cap A_j = \phi$ for $i \neq j$, then

$$P\left[A|B\right] = P\left[A_1|B\right] + P\left[A_2|B\right] + \cdots$$

Law of Total Probability (Theorem 1.10)

For an event space $\{B_1, \ldots, B_m\}$ with $P[B_i] > 0$ for all i, $P[A] = \sum_{i=1}^{i=m} P[A|B_i]P[B_i]$

Why this?

Proof…?

Q is the event that a resistor is within 50Ω of the nominal value.

$$Q = (Q \cap B_1) \cup (Q \cap B_2) \cup (Q \cap B_3),$$

where $\{B_1, B_2, B_3\}$ is an event space

Example 1.19 Problem

A company has three machines B_1, B_2, and B_3 for making 1 kΩ resistors. It has been observed that 80% of resistors produced by B_1 are within 50 Ω of the nominal value. Machine B_2 produces 90% of resistors within 50 Ω of the nominal value. The percentage for machine B_3 is 60%. Each hour, machine B_1 produces 3000 resistors, B_2 produces 4000 resistors, and B_3 produces 3000 resistors. All of the resistors are mixed together at random in one bin and packed for shipment. What is the probability that the company ships a resistor that is within 50 Ω of the nominal value?

$$P(Q) = P[QB_1] + P[QB_2] + P[QB_3]$$

$$= P[Q|B_1]\,P[B_1] + \cdots$$

$P[Q|B_1] = 0.8$

$P[Q|B_2] = 0.9$

$P[Q|B_3] = 0.6$

Bayes' Theorem – The Most Important Theorem Ever!

$$P[B|A] = \frac{P[A|B]P[B]}{P[A]}$$

Proof follows from the definition of conditional probability

Quiz 1.5

Monitor three consecutive phone calls going through a telephone switching office. Classify each one as a voice call (v) if someone is speaking, or a data call (d) if the call is carrying a modem or fax signal. Your observation is a sequence of three letters (each one is either v or d). For example, three voice calls corresponds to vvv. The outcomes vvv and ddd have probability 0.2 whereas each of the other outcomes vvd, vdv, vdd, dvv, dvd, and ddv has probability 0.1. Count the number of voice calls N_V in the three calls you have observed. Consider the four events $N_V = 0$, $N_V = 1$, $N_V = 2$, $N_V = 3$. Describe in words and also calculate the following probabilities:

(1) $P[N_V = 2]$

(2) $P[N_V \geq 1]$

(3) $P[\{vvd\}|N_V = 2]$

(4) $P[\{ddv\}|N_V = 2]$

(5) $P[N_V = 2|N_V \geq 1]$

(6) $P[N_V \geq 1|N_V = 2]$

Problem 1.5.6

Deer ticks can carry both Lyme disease and human granulocytic ehrlichiosis (HGE). In a study of ticks in the Midwest, it was found that 16% carried Lyme disease, 10% had HGE, and that 10% of the ticks that had either Lyme disease or HGE carried both diseases.

(a) What is the probability $P[LH]$ that a tick carries both Lyme disease (L) and HGE (H)?

(b) What is the conditional probability that a tick has HGE given that it has Lyme disease?

Bayes' and Law of Total Probability

For an event space $\{B_1, \ldots, B_m\}$ with $P[B_i] > 0$ for all i,

$$P[B_i|A] = \frac{P[AB_i]}{P[A]} = \frac{P[A|B_i]P[B_i]}{P[A]}$$

$$= \frac{P[A|B_i]P[B_i]}{\sum_{i=1}^{m} P[A|B_i]P[B_i]}$$

We know the a priori probabilities of the B_i(s) and also the probability of an observable event A given B_i. Having seen A (the effect) we want to know the chance that a certain B_i is the cause (that caused A)

Bayes' and Law of Total Probability

For an event space $\{B_1, \ldots, B_m\}$ with $P[B_i] > 0$ for all i,

$$P[B_i|A] = \frac{P[AB_i]}{P[A]} = \frac{P[A|B_i]P[B_i]}{P[A]}$$

$$= \frac{P[A|B_i]P[B_i]}{\sum_{i=1}^{m} P[A|B_i]P[B_i]}$$

We know the a priori probabilities of the B_i(s) and also the probability of an observable event A given B_i. Having seen A (the effect) we want to know the chance that a certain B_i is the cause (that caused A)

Bayes' and Law of Total Probability

- Let A be the event {A world class athlete is seen}
- Consider the event space (World) contains countries of origins. It is W = {China, Jamaica, India, Malaysia, Pakistan, Sri Lanka, UK, US,…}. Let B_1 = China and so on…
- $P[B_i]$ is the probability of the event that the country of origin of a human is B_i
- A priori we could calculate $P[B_i]$ as the ratio of the population of B_i and the population of the world.
- We could also arrive at $P[A|B_i]$, say based on past statistics
- Now say I tell you that I saw an athlete, that is observed A, and I want to know what is probability that the athlete is from India
  - I want to know P[India|A] or $P[B_3|A]$

INNOCENT OR GUILTY

From the book "Randomness"

85% of cabs in your city are red.
The rest are blue.

$P[R] = 0.85$
$P[B] = 0.15$

$R \triangleq$ Event that a randomly chosen cab is red.

A cab is involved in a hit and run incident.

A witness identifies the cab as blue.

$P[\text{Witness identifies blue} \mid B] = 0.8$

$P[\text{Witness identifies red} \mid R] = 0.8$

A judge declares that a blue cab is guilty.

Did the judge get this right?

What may make this problem interesting?

$$P[B \mid \underbrace{\text{Witness identified blue}}_{W_B}] = \frac{P[B, W_B]}{P[W_B]} = \frac{P[W_B \mid B]\, P[B]}{P[(W_B \cap B) \cup (W_B \cap R)]}$$

$\{B, R\}$ is an Event Space.

$$\frac{P[W_B \mid B]\, P[B]}{P[W_B \mid B]\, P[B] + P[W_B \mid R]\, P[R]} = \frac{(0.8)(0.15)}{(0.8)(0.15) + (0.2)(0.85)} = \frac{12}{29}$$

$$P[R \mid \text{Witness identified blue}] = \frac{17}{29}$$

Total identified as Blue

Explain your calculated probabilities

DIAGNOSTIC TESTS

A disease has a prevalance rate of 1/1000.

A person decides to get tested. What is the probability that the person is infected before the test result is out?

The test has a false +ve rate of 1/1000 & a false −ve rate of 0.

⇓ The corresponding probabilities?

$$P[\text{Test is +ve} \mid \text{Person is not infected}] = 10^{-3}$$

$$P[\text{Test is −ve} \mid \text{Person is infected}] = 0$$

The test result is +ve. Is the person infected?

$$P[\text{Person is infected} \mid \text{Test result is +ve}]$$

$$= \frac{P[\text{Test result is +ve} \mid \text{Person is infected}]\, P[\text{Person is infected}]}{P[\text{Test result is +ve} \mid \text{Person is inf}]\, P[\text{Per is inf}] + P[\text{Test result is +ve} \mid \text{Person is not inf}]\, P[\text{Person is not inf}]}$$

SCREAM DETECTION APP

$P[\text{Voice sample is scream}] = p.$

App has false +ve rate > 0 → FPR

a false −ve rate > 0. → FNR

$P[\text{App detects scream} \mid \text{Audio isn't scream}]$

TDR : $P[\text{App detects scream} \mid \text{Sample is a scream}]$

$P[\text{Scream happened} \mid \text{App detected scream}]$

$=$

$$= \frac{(TDR)\,p}{(TDR)\,p + FPR(1-p)} = \frac{p}{p + \frac{FPR}{TDR}(1-p)}$$

Independence

- **Two Events** A and B are independent *iff* P[AB] = P[A]P[B]
- The occurrence of event B does not change your belief about whether event A has occurred or not
  - Implies P[A|B] = P[?]
- **Definition 1.9**
  - If n>3, the sets $A_1, A_2, \ldots, A_n$ are independent *if*
    - *Every set of n-1 sets taken from* $A_1, A_2, \ldots, A_n$ is independent,
    - $P[A_1 \cap A_2 \cap \ldots \cap A_n] = P[A_1]\, P[A_2] \ldots P[A_n]$

Quiz 1.6

Monitor two consecutive phone calls going through a telephone switching office. Classify each one as a voice call (v) if someone is speaking, or a data call (d) if the call is carrying a modem or fax signal. Your observation is a sequence of two letters (either v or d). For example, two voice calls corresponds to vv. The two calls are independent and the probability that any one of them is a voice call is 0.8. Denote the identity of call i by C_i. If call i is a voice call, then $C_i = v$; otherwise, $C_i = d$. Count the number of voice calls in the two calls you have observed. N_V is the number of voice calls. Consider the three events $N_V = 0$, $N_V = 1$, $N_V = 2$. Determine whether the following pairs of events are independent:

(1) $\{N_V = 2\}$ and $\{N_V \geq 1\}$

(2) $\{N_V \geq 1\}$ and $\{C_1 = v\}$

(3) $\{C_2 = v\}$ and $\{C_1 = d\}$

(4) $\{C_2 = v\}$ and $\{N_V$ is even$\}$

Problem 1.6.7

For independent events A and B, prove that

(a) A and B^c are independent.

(b) A^c and B are independent.

(c) A^c and B^c are independent.

Sequential Diagrams

Example 1.25 Problem

Suppose traffic engineers have coordinated the timing of two traffic lights to encourage a run of green lights. In particular, the timing was designed so that with probability 0.8 a driver will find the second light to have the same color as the first. Assuming the first light is equally likely to be red or green, what is the probability $P[G_2]$ that the second light is green? Also, what is $P[W]$, the probability that you wait for at least one light? Lastly, what is $P[G_1|R_2]$, the conditional probability of a green first light given a red second light?

Sequential Diagrams

Example 1.25 Problem

Suppose traffic engineers have coordinated the timing of two traffic lights to encourage a run of green lights. In particular, the timing was designed so that with probability 0.8 a driver will find the second light to have the same color as the first. Assuming the first light is equally likely to be red or green, what is the probability $P[G_2]$ that the second light is green? Also, what is $P[W]$, the probability that you wait for at least one light? Lastly, what is $P[G_1|R_2]$, the conditional probability of a green first light given a red second light?

Quiz 1.7

In a cellular phone system, a mobile phone must be paged to receive a phone call. However, paging attempts don't always succeed because the mobile phone may not receive the paging signal clearly. Consequently, the system will page a phone up to three times before giving up. If a single paging attempt succeeds with probability 0.8, sketch a probability tree for this experiment and find the probability $P[F]$ that the phone is found.

Problem 1.7.6

A machine produces photo detectors in pairs. Tests show that the first photo detector is acceptable with probability $3/5$. When the first photo detector is acceptable, the second photo detector is acceptable with probability $4/5$. If the first photo detector is defective, the second photo detector is acceptable with probability $2/5$.

(a) What is the probability that exactly one photo detector of a pair is acceptable?

(b) What is the probability that both photo detectors in a pair are defective?

Problem 1.7.9

The quality of each pair of photodiodes produced by the machine in Problem 1.7.6 is independent of the quality of every other pair of diodes.

(a) What is the probability of finding no good diodes in a collection of n pairs produced by the machine?

(b) How many pairs of diodes must the machine produce to reach a probability of 0.99 that there will be at least one acceptable diode?

$$P[A_1^c \cap A_2^c]$$

$$P[(A_{1,1}^c \cap A_{2,1}^c) \cap (A_{1,2}^c \cap A_{2,2}^c) \cdots \cap (A_{1,n}^c \cap A_{2,n}^c)]$$

$$= (P[(A_{1,1}^c \cap A_{2,1}^c)]) \cdots (P[A_{1,n}^c \cap A_{2,n}^c])$$

Section 1.8

Counting Methods

Definition 1.10 Fundamental Principle of Counting

If subexperiment A has n possible outcomes, and subexperiment B has k possible outcomes, then there are nk possible outcomes when you perform both subexperiments.

Example 1.29 Problem

Shuffle a deck and observe each card starting from the top. The outcome of the experiment is an ordered sequence of the 52 cards of the deck. How many possible outcomes are there?

Example 1.30 Problem

Shuffle the deck and choose three cards in order. How many outcomes are there?

Theorem 1.12

The number of k-permutations of n distinguishable objects is

$$(n)_k = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

Theorem 1.13

The number of ways to choose k objects out of n distinguishable objects is

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

Definition 1.11 n choose k

For an integer $n \geq 0$*, we define*

$$\binom{n}{k} = \begin{cases} \dfrac{n!}{k!(n-k)!} & k = 0, 1, \ldots, n, \\ 0 & \textit{otherwise.} \end{cases}$$

Example 1.33 Problem

A laptop computer has PCMCIA expansion card slots A and B. Each slot can be filled with either a modem card (m), a SCSI interface (i), or a GPS card (g). From the set $\{m, i, g\}$ of possible cards, what is the set of possible ways to fill the two slots when we sample with replacement? In other words, how many ways can we fill the two card slots when we allow both slots to hold the same type of card?

Theorem 1.14

Given m distinguishable objects, there are m^n ways to choose with replacement an ordered sample of n objects.

Theorem 1.15

For n repetitions of a subexperiment with sample space $S = \{s_0, \ldots, s_{m-1}\}$, there are m^n possible observation sequences.

Example 1.38 Problem

For five subexperiments with sample space $S = \{0, 1\}$, how many observation sequences are there in which 0 appears $n_0 = 2$ times and 1 appears $n_1 = 3$ times?

Theorem 1.16

The number of observation sequences for n subexperiments with sample space $S = \{0, 1\}$ with 0 appearing n_0 times and 1 appearing $n_1 = n - n_0$ times is $\binom{n}{n_1}$.

Theorem 1.17

For n repetitions of a subexperiment with sample space $S = \{s_0, \ldots, s_{m-1}\}$, the number of length $n = n_0 + \cdots + n_{m-1}$ observation sequences with s_i appearing n_i times is

$$\binom{n}{n_0, \ldots, n_{m-1}} = \frac{n!}{n_0! n_1! \cdots n_{m-1}!}.$$

$${}^{n}C_{n_0} \quad {}^{n-n_0}C_{n_1} \quad {}^{n-n_1-n_0}C_{n_2} \quad - \quad - \quad - \quad - \quad {}^{n-(n_1+n_2+\cdots+n_{m-2})}C_{n_{m-1}}$$

Definition 1.12 ***Multinomial Coefficient***

For an integer $n \geq 0$*, we define*

$$\binom{n}{n_0, \ldots, n_{m-1}} = \begin{cases} \dfrac{n!}{n_0! n_1! \cdots n_{m-1}!} & n_0 + \cdots + n_{m-1} = n; \\ & n_i \in \{0, 1, \ldots, n\}, i = 0, 1, \ldots, m-1, \\ 0 & \textit{otherwise.} \end{cases}$$

Quiz 1.8

Consider a binary code with 4 bits (0 or 1) in each code word. An example of a code word is 0110.

(1) How many different code words are there?

(2) How many code words have exactly two zeroes?

(3) How many code words begin with a zero?

(4) In a constant-ratio binary code, each code word has N bits. In every word, M of the N bits are 1 and the other $N - M$ bits are 0. How many different code words are in the code with $N = 8$ and $M = 3$?

Problem 1.8.7

An instant lottery ticket consists of a collection of boxes covered with gray wax. For a subset of the boxes, the gray wax hides a special mark. If a player scratches off the correct number of the marked boxes (and no boxes without the mark), then that ticket is a winner. Design an instant lottery game in which a player scratches five boxes and the probability that a ticket is a winner is approximately 0.01.

Section 1.9

Independent Trials

Example 1.39 Problem

What is the probability $P[S_{2,3}]$ of two failures and three successes in five independent trials with success probability p.

$^5C_3$ $^3C_3$ mutually exclusive ways.

Any way occurs with prob $(1-p)(1-p)ppp$

$= (1-p)^2 p^3$.

$\therefore P[S_{2,3}] = {}^5C_3 (1-p)^2 p^3$.

Theorem 1.18

The probability of n_0 failures and n_1 successes in $n = n_0 + n_1$ independent trials is

$$P\left[S_{n_0,n_1}\right] = \binom{n}{n_1}(1-p)^{n-n_1}p^{n_1} = \binom{n}{n_0}(1-p)^{n_0}p^{n-n_0}.$$

Example 1.41 Problem

To communicate one bit of information reliably, cellular phones transmit the same binary symbol five times. Thus the information "zero" is transmitted as 00000 and "one" is 11111. The receiver detects the correct information if three or more binary symbols are received correctly. What is the information error probability $P[E]$, if the binary symbol error probability is $q = 0.1$?

$$P[E] = P[E \mid \text{Info } 0]\, P[\text{Info } 0] + P[E \mid \text{Info } 1]\, P[\text{Info } 1]$$

$$= 1 - \Big(P[\text{No bit flipped}] + P[\text{Exactly 1 flipped}] + P[\text{Exactly 2 flipped}] \Big)$$

Theorem 1.19

A subexperiment has sample space $S = \{s_0, \ldots, s_{m-1}\}$ with $P[s_i] = p_i$. For $n = n_0 + \cdots + n_{m-1}$ independent trials, the probability of n_i occurences of s_i, $i = 0, 1, \ldots, m-1$, is

$$P\left[S_{n_0,\ldots,n_{m-1}}\right] = \binom{n}{n_0, \ldots, n_{m-1}} p_0^{n_0} \cdots p_{m-1}^{n_{m-1}}.$$

Quiz 1.9

Data packets containing 100 bits are transmitted over a communication link. A transmitted bit is received in error (either a 0 sent is mistaken for a 1, or a 1 sent is mistaken for a 0) with probability $\epsilon = 0.01$, independent of the correctness of any other bit. The packet has been coded in such a way that if three or fewer bits are received in error, then those bits can be corrected. If more than three bits are received in error, then the packet is decoded with errors.

(1) Let $S_{k,100-k}$ denote the event that a received packet has k bits in error and $100 - k$ correctly decoded bits. What is $P[S_{k,100-k}]$ for $k = 0, 1, 2, 3$?

(2) Let C denote the event that a packet is decoded correctly. What is $P[C]$?

$P[\text{Group 1 kicker kicks a field goal}] = 1/2$

$P[\text{Group 2 } \underline{\qquad\qquad\qquad}] = \frac{1}{3}$

kicker from G_1

kicker from G_2

$$K = (K \cap G_1) \cup (K \cap G_2)$$

$$P[K] = P[K \cap G_1] + P[K \cap G_2]$$

$$= \underbrace{P[K|G_1]}\, P[G_1] + P[K|G_2]\, P[G_2]$$

$$= (1/2)\, P[G_1] + \frac{1}{3}\, P[G_2]$$

$$= \left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)\left(2/3\right) =$$

Problem 1.9.5

There is a collection of field goal kickers, which can be divided into two groups 1 and 2. Group i has $3i$ kickers. On any kick, a kicker from group i will kick a field goal with probability $1/(i+1)$, independent of the outcome of any other kicks by that kicker or any other kicker.

(a) A kicker is selected at random from among all the kickers and attempts one field goal. Let K be the event that a field goal is kicked. Find $P[K]$.

(b) Two kickers are selected at random. For $j = 1, 2$, let K_j be the event that kicker j kicks a field goal. Find $P[K_1 \cap K_2]$. Are K_1 and K_2 independent events?

(c) A kicker is selected at random and attempts 10 field goals. Let M be the number of misses. Find $P[M = 5]$.

$P[\text{Group 1 kicker kicks a field goal}] = 1/2$

$P[\text{Group 2 kicker from } G_1] = \frac{1}{3}$

$$K = (K \cap G_1) \cup (K \cap G_2)$$

kicker from G_1 — kicker from G_2

$$P[K] = P[K \cap G_1] + P[K \cap G_2]$$

$$= P[K|G_1]\,P[G_1] + P[K|G_2]\,P[G_2]$$

$$= (1/2)\,P[G_1] + \tfrac{1}{3}\,P[G_2]$$

$$= \left(\tfrac{1}{2}\right)\left(\tfrac{1}{3}\right) + \left(\tfrac{1}{3}\right)\left(2/3\right) =$$

K_i is the event that kicker i kicks a field goal.

$G_1^{(i)}$ is the event that the first kicker is from group i

$$K_1 = (K_1 \cap G_1^{(1)}) \cup (K_1 \cap G_1^{(2)})$$

$$P[K_1] = P[K_1 \cap G_1^{(1)}] + P[K_1 \cap G_1^{(2)}]$$

$$= P[K_1|G_1^{(1)}]\,P[G_1^{(1)}] + P[K_1 \quad |G_1^{(2)}]\,P[G_1^{(2)}]$$

$$= (0.5)\left(3/9\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{6}{9}\right)$$

$$P[K_2] = P[K_2 \cap (G_1^{(1)} \cap G_2^{(1)})] + P[K_2 \cap (G_1^{(1)} \cap G_2^{(2)})]$$

$$+ P[K_2 \cap (G_1^{(2)} \cap G_2^{(1)})] + P[K_2 \cap (G_1^{(2)} \cap G_2^{(2)})]$$

$$P[K_2 | G_1^{(1)} \cap G_2^{(1)}]\;P[G_1^{(1)} \cap G_2^{(1)}]$$

$$= \left(\tfrac{1}{2}\right) P[G_1^{(1)} \cap G_2^{(1)}]$$

$$P[G_1^{(1)} \cap G_2^{(1)}] = P[G_2^{(1)}|G_1^{(1)}]\,P[G_1^{(1)}]$$

$$= \left(\tfrac{2}{2+6}\right)\left(\tfrac{3}{9}\right) = \left(\tfrac{1}{4}\right)\left(\tfrac{1}{3}\right)$$

$$P[G_1^{(1)}] = \tfrac{1}{3}$$

$$P[G_2^{(1)}] = ?$$

$$G_2^{(1)} = (G_2^{(1)} \cap G_1^{(1)}) \cup (G_2^{(1)} \cap G_1^{(2)})$$

$$P[G_2^{(1)}] = P[G_2^{(1)}|G_1^{(1)}]\,P[G_1^{(1)}] + P[G_2^{(1)}|G_1^{(2)}]\,P[G_1^{(2)}]$$

$$= \left(\tfrac{1}{4}\right)\left(\tfrac{1}{3}\right) + \left(\tfrac{2}{8}\right)\left(\tfrac{2}{3}\right)$$

$$= \tfrac{1}{12} + \tfrac{1}{4} = \tfrac{4}{12} = \left(\tfrac{1}{3}\right)$$

Problem 1.9.5

There is a collection of field goal kickers, which can be divided into two groups 1 and 2. Group i has $3i$ kickers. On any kick, a kicker from group i will kick a field goal with probability $1/(i+1)$, independent of the outcome of any other kicks by that kicker or any other kicker.

(a) A kicker is selected at random from among all the kickers and attempts one field goal. Let K be the event that a field goal is kicked. Find $P[K]$.

(b) Two kickers are selected at random. For $j = 1, 2$, let K_j be the event that kicker j kicks a field goal. Find $P[K_1 \cap K_2]$. Are K_1 and K_2 independent events?

(c) A kicker is selected at random and attempts 10 field goals. Let M be the number of misses. Find $P[M = 5]$.

Discrete Random Variables

Chapter 2 in the book by RY

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Random Variables

- Our Experiment contained
  - Procedure
  - Observation
  - Model

- Sample space
  - The set of all possible outcomes

- Examples
  - {Excellent, Good, Fair, Poor}
  - {Heads, Tails}
  - {A+, A, A-, B, B-, C, C-, D, F}

Random Variables

- Our Experiment contained
  - Procedure
  - Observation
  - Model

- Under the model we assigned probabilities to outcomes in the sample space S

Random Variables

- Our Experiment contained
  - Procedure
  - Observation
  - Model

- Under the model we assigned probabilities to outcomes in the sample space S

- Now we will assign numbers to outcomes in S
  - How should we?

Random Variables

- Now we will assign numbers to outcomes in S
  - How should we?
- We have a sample space
- In addition, define a range space
  - Range space is a set of numbers
- An experiment must result in exactly one number in the range space

Random Variable

- An experiment must result in exactly one number in the range space

Random Variable

- An experiment must result in exactly one number in the range space

- A Random Variable is **a function** and maps outcomes in the sample space to numbers in a range space

Random Variable

Random Variable

- Coin Tossing Experiment
  - Procedure: Toss a coin once
  - Observation: Heads (1) or Tails (0)
    - Absence/Presence; Pass/Fail; Success/Failure; Accept/Reject; Rich/Poor; Male/Female; Infected/Healthy
    - You are happy to make Boolean observations! That is good enough for your experiment
  - Model: P[0] = 0.5, P[1] = 0.5
  - Our Random Variable, say X(.), takes values of 0 and 1

Random Variable

- **Some Notation**

Random variable X has a range $S_X = \{0, 1\}$.
Range of X is the set of possible values of X
Random variable Y has a range S_Y...

Random Variable

- The Sample and Range space may be the same
  - You are asked to count the number of buses that pass in 10 minutes.
  - The number of buses belongs to the set of non-negative integers (Range space and Sample space)

Random Variable

- Random variable maybe a function of the observations in the sample space
  - Toss a coin five times
  - Number of heads belongs to the set {0,1,2,3,4,5} and is a random variable.

- The range space is … $S_X = \{0, 1, 2, 3, 4, 5\}$

- When an experiment is performed the RV … X takes a value in ~~the~~ … S_X

Experiment and the Random Variable

- The RV may be a function of another RV

  - Suppose you earn Rs 100 if the number of heads is even and have to pay Rs 100 if the number of heads is odd

  - Let RV G denote your gain at the end of the experiment of 5 coin tosses

$$G = f(X) = \begin{cases} 100 & \text{if } X = 0, 2, 4, \\ -100 & \text{if } X = 1, 3, 5. \end{cases}$$

S_G S_X

Def 2.1: Random Variable

- **Random Variable**: A RV **consists of** an experiment with a probability measure P[.] defined on a sample space S and a **function** that assigns a real number to each outcome in the sample space of the experiment

  - There is an underlying experiment

  - The function defined is just a mapping from outcomes to numbers and is in no way related to a specific experiment

Def 2.1: Random Variable

- Notation
  - RV is denoted by a capital letter. The value it assumes is denoted by the corresponding small letter
  - X = x, X is the RV and x is the value
  - Writing X = z is correct too. Here z is the value
  - X(.) denotes the function corresponding to the RV X

Never Forget the Experiment!

The event

$$\{X = x\} = \boxed{\{s \in S \mid X(s) = x\}}$$

What is the set of outcomes corresponding to the event X=x?

$\{X = x\} = \{\text{goats}, \text{sheep}\}$

Discrete Random Variable

- X is a **discrete** random variable if its range S_X is a countable set

$$S_X = \{x_1, x_2, \ldots\}$$

- $\{X = x_i\}$ and $\{X = x_j\}$, where $x_i \neq x_j$ are disjoint!
- Note that there is no constraint on the values x_1, x_2,…

Discrete Random Variable

- The range must be a countable set
  - Finite, Countably infinite

- Range cannot be the interval (0,1)
- It can be {0.01, 0.05, 0.08, 0.1,…}
- It can be {0.001, 0.5, 0.08, 0.9999}
- It can be the set of integers

- X(.) maps outcomes in sample space S to values in the range S_X

Quiz 2.1

A student takes two courses. In each course, the student will earn a B with probability 0.6 or a C with probability 0.4, independent of the other course. To calculate a grade point average (GPA), a B is worth 3 points and a C is worth 2 points. The student's GPA is the sum of the GPA for each course divided by 2. Make a table of the sample space of the experiment and the corresponding values of the student's GPA, G.

$S = \{BB, BC, CB, CC\}$

$S_X = \{2, 2.5, 3\}$

$S = (0, 1)$

Probability Mass Function

- **<u>Def 2.4</u>** The probability mass function (PMF) of a discrete random variable X is

$$P_X(x) = P[X = x]$$

Probability of the event X=x

Notation!

Probability Mass Function

- **<u>Def 2.4</u>** The probability mass function (PMF) of a discrete random variable X is

$$P_X(x) = P[X = x]$$

- Note that the PMF is defined for all *x*, not just the *x* that have a mapping to one or more outcomes
  - **The domain of P_X is the set of real numbers**

Probability Mass Function

- There is nothing sacrosanct about x. Just a convention to use a capital letter for a RV and corresponding small for the value

$$P_X(u) = P[X = u]$$

Fickle Coin

- You toss a fickle coin once. The sample space S={heads, tails, standing}.

- The outcomes heads, tails and standing occur with probabilities 0.7, 0.2 and 0.1 respectively.

- In our model, if tails is observed, then the RV X = 0. For the outcome heads, X = 1. For the outcome standing X=2.

Fickle Coin

- P[X=0] = 0.7, P[X=1] = 0.2, P[X=2] = 0.1

- The PMF is

$$P_X(x) = \begin{cases} 0.2 & x = 0, \\ 0.7 & x = 1, \\ 0.1 & x = 2, \\ 0 & otherwise. \end{cases}$$

PMF: Tossing a Biased Coin…

PMF

- Is $P_X(\pi)$ defined for the experiment? What is its value?

- **The PMF contains all the information about the RV X**

$X \sim$ Binomial(n,p) RV

- Here the number of trials is fixed to n

- The Binomial RV counts the number of successful trials amongst the total of n trials

$$S_X = \{0, 1, 2, \ldots, n\}$$

- PMF? Simply calculate probabilities for all values in the range space...

$$P[X = x] = \begin{cases} {}^{n}C_{k}(1-p)^{n-k}p^{k} & k = 0, 1, 2, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

$P[X = k] = P[$k successes in n indep and identical Bern trials$]$

$\{X=x\}$ = You obtain the kth 1 at the xth trial.

$x \geq k$

$\underline{\quad}_{1} \quad \underline{\quad}_{2} \quad \underline{\quad}_{3} \quad \cdots \quad \underline{\quad}_{x}$

We have k 1(s) &
$x-k$ 0(s) &
the xth outcome must be 1.

Any such sequence has a probability $p^k(1-p)^{x-k}$.

How many such sequences do I have?

$${}^{x-1}C_{k-1}$$

$$P[X=x] = {}^{x-1}C_{k-1}\, p^k (1-p)^{x-k}$$

Binomial(n,p) RV

- You are not just interested in the first bit error
- You are interested in the number of bit errors in a packet of size n bits
- What is the probability that x out of n bits are in error, when each bit is in error with probability p, independent of other bits?
- **Def 2.7:** X is a binomial(n,p) random variable if the PMF of X has the form

$$P_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x \in \{0, 1, \ldots\}, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $n \geq 1$ and $0 < p < 1$.

Pascal(k,p) RV

- Example 2.15 and Definition 2.8
- Straightforward extension of Binomial
- Read from book!

Discrete Uniform(k,l) RV

- A fair coin has an equal probability of landing heads up and tails up
- The roll of a fair die has an equal probability of giving each of 1,2,3,4,5 and 6
- If 10 outcomes are possible and you have no reason to believe that any one outcome is more likely than another, you model the outcomes to be equiprobable

Discrete uniform*(k,l)* RV

- **Def 2.9** X is a discrete uniform*(k,l)* RV if the PMF of X has the form

$$P_X(x) = \begin{cases} 1/(l-k+1) & x = k, k+1, \ldots, l, \\ 0 & \text{otherwise.} \end{cases}$$

- Note that bit *k* and *l*, *k* < *l*, are integers by the definition of a **discrete** uniform RV.

Discrete uniform*(k,l)* RV

- **Def 2.9** X is a discrete uniform*(k,l)* RV if the PMF of X has the form

$$P_X(x) = \begin{cases} 1/(l-k+1) & x = k, k+1, \ldots, l, \\ 0 & \text{otherwise.} \end{cases}$$

- Tossing a fair coin can be modeled using a discrete uniform*(100,101)* RV
  - Of course, typically we use discrete uniform*(0,1)*!
- For an experiment that requires rolling a fair die and noting the number
  - Use the discrete uniform(1,6) RV

The Binomial Picture

Poisson (α) Random Variable

- You stand at a railway ticket counter and note the number of customers that arrive at the counter every minute

- Turns out that in many such real world situations
  - Customers arriving at a restaurant
  - Packets arriving at a router
  - Students arriving to a class?

the **number of arrivals in a fixed time interval** can be modeled as a Poisson RV

Poisson (α) Random Variable

- The observation is a count and hence the range S_X is the set of non-negative integers
  - S_X is countably infinite
  - Poisson is indeed a discrete RV

Poisson (α) Random Variable

- **Def 2.10** X is a Poisson(α) RV if the PMF of X has the form

$$P_X(x) = \begin{cases} \alpha^x e^{-\alpha}/x! & x = 0, 1, 2, \ldots, \\ 0 & otherwise. \end{cases}$$

where the parameter $\alpha > 0$.

- If the average rate of arrivals is λ/second and time interval is of T seconds, α = λT.

Poisson (α) Random Variable

- **Def 2.10** X is a Poisson(α) RV if the PMF of X has the form

$$P_X(x) = \begin{cases} \alpha^x e^{-\alpha}/x! & x = 0, 1, 2, \ldots, \\ 0 & otherwise. \end{cases}$$

where the parameter $\alpha > 0$.

- If the average rate of arrivals is λ/second and time interval is of T seconds, α = λT.

- What explains the crazy looking expression?

Word Problem (Example 2.20)

- The number of database queries processed by a computer in any 10-second interval is a Poisson RV K with α=5 queries.

  - What is the probability that there will be no queries processed in a 10-second interval?

$$P_X(x) = \begin{cases} \alpha^x e^{-\alpha}/x! & x = 0, 1, 2, \ldots, \\ 0 & otherwise. \end{cases}$$

where the parameter $\alpha > 0$.

Word Problem (Example 2.20)

- The number of database queries processed by a computer in any 10-second interval is a Poisson RV K with α=5 queries.
  - What is the probability that at least two queries will be processed in a 2-second interval?

$$P_X(x) = \begin{cases} \alpha^x e^{-\alpha}/x! & x = 0, 1, 2, \ldots, \\ 0 & otherwise. \end{cases}$$

where the parameter $\alpha > 0$.

Quiz 2.3

Each time a modem transmits one bit, the receiving modem analyzes the signal that arrives and decides whether the transmitted bit is 0 or 1. It makes an error with probability p, independent of whether any other bit is received correctly.

(1) If the transmission continues until the receiving modem makes its first error, what is the PMF of X, the number of bits transmitted?

Quiz 2.3

Each time a modem transmits one bit, the receiving modem analyzes the signal that arrives and decides whether the transmitted bit is 0 or 1. It makes an error with probability p, independent of whether any other bit is received correctly.

(1) If the transmission continues until the receiving modem makes its first error, what is the PMF of X, the number of bits transmitted?

(2) If $p = 0.1$, what is the probability that $X = 10$? What is the probability that $X \geq 10$?

(3) If the modem transmits 100 bits, what is the PMF of Y, the number of errors?

(4) If $p = 0.01$ and the modem transmits 100 bits, what is the probability of $Y = 2$ errors at the receiver? What is the probability that $Y \leq 2$?

(5) If the transmission continues until the receiving modem makes three errors, what is the PMF of Z, the number of bits transmitted?

(6) If $p = 0.25$, what is the probability of $Z = 12$ bits transmitted?

Problem 2.3.12

Suppose each day (starting on day 1) you buy one lottery ticket with probability $1/2$; otherwise, you buy no tickets. A ticket is a winner with probability p independent of the outcome of all other tickets. Let N_i be the event that on day i you do *not* buy a ticket. Let W_i be the event that on day i, you buy a winning ticket. Let L_i be the event that on day i you buy a losing ticket.

(a) What are $P[W_{33}]$, $P[L_{87}]$, and $P[N_{99}]$?

$P[W_{33}] = \left(\frac{1}{2}\right)p$

$P[L_{87}] = \left(\frac{1}{2}\right)(1-p)$

$P[N_{99}] = 1/2$

Suppose each day (starting on day 1) you buy one lottery ticket with probability $1/2$; otherwise, you buy no tickets. A ticket is a winner with probability p independent of the outcome of all other tickets. Let N_i be the event that on day i you do *not* buy a ticket. Let W_i be the event that on day i, you buy a winning ticket. Let L_i be the event that on day i you buy a losing ticket.

(a) What are $P[W_{33}]$, $P[L_{87}]$, and $P[N_{99}]$?

(b) Let K be the number of the day on which you buy your first lottery ticket. Find the PMF $P_K(k)$.

$$P_K(k) = \begin{cases} \left(\frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right) & k = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

$$\sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k$$

(c) Find the PMF of R, the number of losing lottery tickets you have purchased in m days.

(d) Let D be the number of the day on which you buy your jth losing ticket. What is $P_D(d)$? Hint: If you buy your jth losing ticket on day d, how many losers did you have after $d-1$ days?

$$P[L_i] = \left(\frac{1}{2}\right)(1-p), \quad i = 1, 2, \ldots, m.$$

$$P_R(r) = P[R = r] = \begin{cases} {}_mC_r \left(P[L_i]\right)^r \left(1 - P[L_i]\right)^{m-r} & r = 0, 1, \ldots, m \\ 0 & \text{otherwise} \end{cases}$$

D is Pascal$(P[L_i], j)$

A student attends a lecture with probability 0.3. Every lecture attendance is taken with probability 0.7. A student's attendance is recorded when the student attends the lecture and the attendance is taken. Suppose there are a total of 10 lectures.

(a) What is the probability that the student attends 7 out of 10 lectures?

$$^{10}C_7 (0.3)^7 (0.7)^3$$

A student attends a lecture with probability 0.3. Every lecture attendance is taken with probability 0.7. A student's attendance is recorded when the student attends the lecture and the attendance is taken. Suppose there are a total of 10 lectures.

(b) What is the probability that the student attends 7 out of 10 lectures and has an attendance of 4 out of 10?

P[Student attends 7 out of 10 & has an attendance of 4 out of 10]

= P[Student attends 4 when attendance is taken, Student attends 3 when attendance isn't taken, Student doesn't attend 3]

$= {}^{10}C_4 \, {}^{6}C_3 \, {}^{3}C_3 \, ((0.3)(0.7))^4 ((0.3)(0.3))^3 (0.7)^3$

A student attends a lecture with probability 0.3. Every lecture attendance is taken with probability 0.7. A student's attendance is recorded when the student attends the lecture and the attendance is taken. Suppose there are a total of 10 lectures.

(c) Suppose you are told that the student attended the 10th lecture and attendance was taken in the lecture. What is your updated belief about the occurrence of the event in part (b)?

P[Student attends 7 out of 10 & has an attendance of 6 out of 6 | Student attended 10th lecture in which attendance was taken]

A student attends a lecture with probability 0.3. Every lecture attendance is taken with probability 0.7. A student's attendance is recorded when the student attends the lecture and the attendance is taken. Suppose there are a total of 10 lectures.

(d) Suppose you are told that the student attended the 10th lecture. What is your updated belief about the occurrence of the event in part (b)?

P[Student attends 7 out of 10 & has an attendance of 4 out of 6 | Student attended 10th lecture] =

$$\frac{1}{P[\text{Student attends } 10^{th}]}\Big\{P[\text{Student attends 7 out of 10 \& has an attendance of 4 out of 10 \& Student attended } 10^{th} \text{ lecture in which attendance was taken}]$$

$$+ P[\ldots\ldots 10^{th} \text{ lecture in which attendance was not taken}]\Big\}$$

Cumulative Distribution Function (CDF)

- **Def 2.11** The cumulative distribution function of a RV *X* is

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

- **It is defined for all real *x***
- **Theorem 2.2 (part)**

For any DRV X with range $S_X = \{x_1, x_2, \ldots\}$ satisfying $x_1 \leq x_2 \leq x_3 \ldots$,

$$F_X(-\infty) = 0$$

$$F_X(\infty) = 1$$

For all $x' \geq x$, $\underbrace{F_X(x')}_{P[X \leq x']} \geq \underbrace{F_X(x)}_{P[X \leq x]}$

$\{X \leq x\} = X \in (-\infty, x]$

$\{X \leq x'\} = X \in (-\infty, x']$

PMF -> CDF

CDF

Starts at 0 and goes all the way to 1
Is non-decreasing
Jumps at points in the range of the RV
Jump at a point x equal to P[X=x]

Accumulate the probabilities in a PMF to get a CDF

CDF

$P[X \in (x_i - \epsilon, x_i]]$

$= P[X \le x_i] - P[X \le x_i - \epsilon]$

For $x_i \in S_X$ and ϵ, an arbitrary small +ve no.,

$$F_X(x_i) - F_X(x_i - \epsilon) = P_X(x_i)$$

$F_X(x) = F_X(x_i)$ for all x such that $x_i \le x < x_{i+1}$.

CDF: Theorem 2.3

$$F_X(b) - F_X(a) = P[a < X \leq b]$$

What if we want

$$P[a \leq X \leq b]$$

CDF

Let $S_X = \{x_{min}, \ldots, x_{max}\}$.

$F_X(x_{max}) = ?$

$F_X(x_{min}) = ?$

For $x > x_{max}$, $F_X(x) = ?$

Example 2.24 Problem

In Example 2.11, let the probability that a circuit is rejected equal $p = 1/4$. The PMF of Y, the number of tests up to and including the first reject, is the geometric $(1/4)$ random variable with PMF

$$P_Y(y) = \begin{cases} (1/4)(3/4)^{y-1} & y = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases} \tag{2.40}$$

What is the CDF of Y?

$$P[Y \le y] = \sum_{k=1}^{y} \underbrace{P[Y = k]}_{(1/4)(3/4)^{k-1}}$$

Average of a Discrete Random Variable

- Mode (Def 2.12 – read from book)

- Median (Def 2.13 – read from book)

- Expected Value
  - When someone says average of a random variable, they usually mean the expected value of the RV!

Average of a Discrete Random Variable

- **Def 2.14** The expected value of X is

$$E[X] = \mu_X = \sum_{x \in S_X} x P_X(x).$$

Average of a Discrete Random Variable

- **Def 2.14** The expected value of X is

$$E[X] = \mu_X = \sum_{x \in S_X} x P_X(x).$$

$$P_X(x) = \begin{cases} 9/10 & x = 0 \\ 1/10 & x = 10 \end{cases}$$

$$P_Y(x) = \begin{cases} \frac{999}{1000} & x = 0 \\ \frac{1}{1000} & x = 10^6 \end{cases}$$

$E[X] =$

$E[Y] =$

Average of a Discrete Random Variable

- **Def 2.14** The expected value of X is

$$E[X] = \mu_X = \sum_{x \in S_X} x P_X(x).$$

Is E[X] a random variable?

Average of a Discrete Random Variable

- **Def 2.14** The expected value of X is

$$E[X] = \mu_X = \sum_{x \in S_X} x P_X(x).$$

$x(1), x(2), \ldots, x(n)$

Thinking like a frequentist...

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n} x(i)}_{\phi} = \underbrace{\frac{1}{n}\sum_{x \in S_X} (n(x))\, x}$$

$$\text{if } n \to \infty \quad \frac{1}{n}\sum_{x \in S_X} n(x)\, x$$

Expected Value

- Consider n trials of an experiment.
- We obtain x(1), x(2), …, x(n)

- The empirical average is

$$m_n = \frac{1}{n} \sum_{i=1}^{n} x(i)$$

Expected Value

- For a range $S_{X,}$ write as a sum over the range space

- Say N_x is the number of times the value x in S_X occurs during the n trials

Expected Value

$$m_n = \frac{1}{n} \sum_{x \in S_X} N_x x$$

As $n \to \infty$, $N_x/n \to P_X(x)$,
and we say that $m_n \to E[X]$.

We used the relative frequency definition of probability to link empirical averages obtained from multiple experiment trials with the expected value of the RV X.

Some Examples of E[X]

- **Def 2.5:** X is a Bernoulli(p) RV if the PMF of X has the form

$$P_X(x) = \begin{cases} 1-p & x = 0, \\ p & x = 1, \\ 0 & \text{otherwise.} \end{cases}$$

where $0 < p < 1$.

- E[X] = ?
  - p

Some Examples of E[X]

- Geometric(p)

$$P_X(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{aligned} E[X] &= \sum_{x=1}^{\infty} x P_X(x) \\ &= \sum_{x=1}^{\infty} x p (1-p)^{x-1} \\ &= p \sum_{x=1}^{\infty} x q^{x-1} \\ &= \frac{1}{p}. \end{aligned}$$

We define q=1-p

E[X] Examples

- When X is Geometric(p)
  - E[X] = 1/p
- A 0.1 probability of a faulty device =>
  - On an average the 10th device will be the first to be found faulty
  - This does not mean that every 10th device will be faulty. In fact P[X=10] = 0.03!
- Please see Theorem 2.6 and 2.7 for E[.] of other RVs

$\underline{X_n} \sim \text{Binomial}(n, p)$

$\rightarrow Y_i \sim \text{Bern}(p)$

$\frac{0}{1} \quad \frac{1}{2} \quad \frac{0}{3} \quad \frac{1}{4} \qquad n = 4$

$Y_1 = y_1 = 0$
$Y_2 = y_2 = 1$
$Y_3 = y_3 = 0$
$Y_4 = y_4 = 1$

Problem 2.5.10

Let binomial random variable X_n denote the number of successes in n Bernoulli trials with success probability p. Prove that $E[X_n] = np$. Hint: Use the fact that $\sum_{x=0}^{n-1} P_{X_{n-1}}(x) = 1$.

$$X_n = \sum_{i=1}^{n} Y_i$$

Given the properties of the expectation operator $E[\cdot]$

$$E[X_n] = E\left[\sum_{i=1}^{n} Y_i\right]$$

$$= \sum_{i=1}^{n} E[Y_i] = \sum_{i=1}^{n} p = np.$$

$$X \sim \text{Pascal}(k, p)$$

$k=3$:

$N_1 \sim$
$N_2 \sim$ Geom(p)
$N_3 \sim$

$$X(k,p) = \sum_{i=1}^{k} N_i \quad , \quad N_i \sim \text{Geom}(p)$$

$$E[X] = \sum_{i=1}^{k} E[N_i] = \sum_{i=1}^{k} \frac{1}{p} = \frac{k}{p}$$

Your experiment involves observing arrival over time T sec

$$T = n\Delta t$$

$$P[X_n = x] = {}^{n}C_x \, p^x (1-p)^{n-x}$$

Let $\boxed{np = \lambda T}$

$$P(X_n = x) = {}^{n}C_x \left(\frac{\lambda T}{n}\right)^x \left(1 - \frac{\lambda T}{n}\right)^{n-x}$$

$$\lim_{n\to\infty} P[X_n = x] = \lim_{n\to\infty} \frac{\lfloor n}{\underbrace{\lfloor x}_{1} \lfloor n-x} \frac{(\lambda T/n)^x}{(1-\lambda T/n)^x} \left(1 - \frac{\lambda T}{n}\right)^n$$

$$= \lim_{n\to\infty} \frac{(\lambda T)^x}{\lfloor x} \frac{(n(n-1)(n-2)\cdots 1)}{n^x \lfloor n-x} \underbrace{\boxed{\frac{1}{\left(1-\frac{\lambda T}{n}\right)^x}}}_{\to 1} \left(1 - \frac{\lambda T}{n}\right)^n$$

$Y \sim \text{Poisson}(\lambda T)$

$$P[Y = x] = \frac{e^{-\lambda T} (\lambda T)^x}{\lfloor x}$$

$$= \lim_{n\to\infty} \frac{(\lambda T)^x}{\lfloor x} \frac{n(n-1)\cdots(n-x+1)}{n^x} \left(1 - \frac{\lambda T}{n}\right)^n$$

Theorem 2.8

Perform n Bernoulli trials. In each trial, let the probability of success be α/n, where $\alpha > 0$ is a constant and $n > \alpha$. Let the random variable K_n be the number of successes in the n trials. As $n \to \infty$, $P_{K_n}(k)$ converges to the PMF of a Poisson (α) random variable.

The Properties of the E[.] Operator

- E[.] is a linear operator

$$E[cX+bY] = E[cX] + E[bY] = cE[X] + bE[Y]$$

- Every linear operator has two properties:

$$\hat{A}(f+g) = \hat{A}f + \hat{A}g \qquad (1)$$
$$\hat{A}(cf) = c\hat{A}f \qquad (2)$$

where $\hat{A}$ is the operator, c is a scalar, f and g are functions.

$$E[cX] = cE[X]$$

where c is a scalar

$$E[X+Y] = E[X] + E[Y]$$

$$E\left[\sum_{i=1}^{\infty} X_i\right] = \sum_{i=1}^{\infty} E[X_i]$$

See http://vergil.chemistry.gatech.edu/notes/quantrev/node14.html

The Properties of the E[.] Operator

- Scalar Multiplication of a Random Variable
  - $E[aX] = a\,E[X]$
  - $E[aX^2] = a\,E[X^2]$ (Note that X^2 is a RV)

- $E[X + X^2] = E[X] + E[X^2]$

- $E[(X + c)^2] = E[X^2 + c^2 + 2cX] = E[X^2] + E[c^2] + 2E[cX]$
  $= E[X^2] + c^2 + 2cE[X]$

Functions of a Random Variable

- You are measuring the power consumed by your appliance

- The power observed is a random variable, say W and is in watts
  - The randomness in the observed power can be due to measurement errors introduced by your equipment

- The power in dB, Y = 10 * log10(W) is also a random variable and is a function of the RV W.
  - We say that Y is derived from W

Derived Random Variable

- **Def 2.15** Each sample value *y* of a derived random variable Y is a **mathematical function** g(x) of a sample value *x* of another random variable X.
  - The notation Y=g(X) is used to describe the relationship of the two random variables.

<u>Mathematical function:</u>
One-to-one or a many-to-one relation

g(.) maps the range of X into the range of Y

Given PMF of X

- Calculate PMF of Y
  - Think the event {Y = y} in terms of the range space of X

Given: $P_X(x), x \in R$

Calculate $\underbrace{P_Y(y)}, y \in R$

$$P[Y=y] = \sum_{x \in S_X : g(x) = y} P[X=x]$$

$$\{Y=y\} = \underbrace{\{x : x \in S_X, g(x) = y\}}$$

$P_X(x)$ -> $P_Y(y)$

- **Theorem 2.9** For a discrete random variable X, the PMF of Y=g(X) is

$$P_Y(y) = \sum_{x:g(x)=y} P_X(x)$$

Given PMF of X

- Calculate E[Y]

$$E[Y] = \sum_{y \in S_Y} y \, P[Y=y]$$

$$= \sum_{y \in S_Y} y \sum_{x : g(x) = y} P[X=x]$$

$$= \sum_{y \in S_Y} \sum_{x : g(x) = y} g(x) P[X=x]$$

$$= \sum_{x \in S_X} g(x) P[X=x] = E[g(X)]$$

Derived Random Variable

- **Theorem 2.10** Given a random variable X with PMF $P_X(x)$ and the derived random variable Y=g(X), the expected value of Y is

$$E[Y] = \mu_Y = \sum_{x \in S_X} g(x) P_X(x).$$

More on the E[.] Operator

- Consider the RV $Y = g(X) = X - E[X] = X - \mu_X$
- The expected value of X is subtracted from each number in S_X. The resulting range is the range S_Y of the random variable Y
- What is E[Y]?
- Using First principles

$Y = g(X)$

$Y = X + E[X]$

$E[Y] = E[X + E[X]]$

$= E[X] + E[E[X]]$

$= E[X] + E[X]$

$= 2E[X]$

More on the E[.] Operator

- Consider the RV $Y = g(X) = X - E[X] = X - \mu_X$
- The expected value of X is subtracted from each number in S_X. The resulting range is the range S_Y of the random variable Y
- What is E[Y]?

- Using properties of E[]

The E[.] Operator

- **Theorem 2.12** For any random variable X,

$$E[aX + b] = aE[X] + b.$$

- Proof is similar to the previous theorem

Variance

- **Def 2.16** Variance of X is

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

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

Think of the relative frequency interpretation of variance (which is also the empirical estimate)

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

$$= E[X^2 + \mu_X^2 - 2\mu_X X]$$

$$= E[X^2] + E[\mu_X^2] - E[2\mu_X X]$$

- Standard deviation is just the square-root of Variance
  - Values of the RV within the SD of the Expected Value are referred to as its *typical values*

$$E[X^2] + \mu_X^2 - 2\mu_X E[X] = E[X^2] - (E[X])^2$$

Properties of Variance

- **Theorem 2.13**

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

$$Var[aX + b] = ?$$

$$E[(aX+b)^2] - (E[aX+b])^2$$

$$= a^2 Var[X]$$

- **Theorem 2.14**

$$Var[aX + b] = a^2 Var[X]$$

  - Offsetting a RV by a constant does not change its variance!

•Show both results using the properties of the E[.] operator

•Also, show using first principles, that is by expanding E[g(x)]

Moments

- For random variable X
  - The nth moment is $E[X^n]$
  - The nth central moment is $E[(X - E[X])^n]$

VVS Word Problems

- X is a discrete uniform RV with range S_X = {1,2,3,4,5,…,100}. Find E[X] and E[X^2].

- Which are valid CDFs?

VVS Word Problems

- $Y = X + 10$
- Express Var[Y] in terms of Var[X]

- $Y = 10X + 10$
- Express Var[Y] in terms of Var[X]

Quiz 2.6

Monitor three phone calls and observe whether each one is a voice call or a data call. The random variable N is the number of voice calls. Assume N has PMF

$E[N] = 0.3(1+2+3)$

$= 1.8$

$$P_N(n) = \begin{cases} 0.1 & n = 0, \\ 0.3 & n = 1, 2, 3, \\ 0 & \text{otherwise.} \end{cases} \tag{2.75}$$

Voice calls cost 25 cents each and data calls cost 40 cents each. T cents is the cost of the three telephone calls monitored in the experiment.

$T = 25N + 40(3-N)$

(1) Express T as a function of N.

$E[T] = 25E[N] + 120 - 40E[N]$

$S_T = \{$

(2) Find $P_T(t)$ and $E[T]$.

$\hookrightarrow P[T=t]$

Conditional PMF of a Discrete RV

- The PMF of a RV X gives the probability of the event {X=x}
- PMF $P_X(x) = P[X=x]$
- The PMF of a RV conditioned on an event B is $P[X=x|B]$

$$\sum_{x \in B} P[X=x] = P[B]$$

- **Def 2.19** Given the event B, with $P[B] > 0$, the conditional probability mass function of X is $P_{X|B}(x)$ $= P[X=x|B]$.

$$= \frac{P[X=x, B]}{P[B]} = \begin{cases} \dfrac{P[X=x]}{P[B]} & x \in B \\ 0 & \text{otherwise} \end{cases}$$

Theorem 2.16

- A RV X resulting from an experiment with event space $B_1, B_2, \ldots, B_m$ has PMF _____?
- What properties do $B_1, B_2, \ldots, B_m$ satisfy?

$$P_X(x) = \sum_{i=1}^{m} P_{X|B_i}(x) P[B_i].$$

VS Word Problem (RY)

Example 2.38 Problem

Let X denote the number of additional years that a randomly chosen 70 year old person will live. If the person has high blood pressure, denoted as event H, then X is a geometric $(p = 0.1)$ random variable. Otherwise, if the person's blood pressure is regular, event R, then X has a geometric $(p = 0.05)$ PMF with parameter. Find the conditional PMFs $P_{X|H}(x)$ and $P_{X|R}(x)$.

$$P[B|A] = \frac{P[A|B]\,P[B]}{P[A]}$$

$$= \frac{P[A|B]\,P[B]}{P[A|B]\,P[B] + P[A|B^c]\,P[B^c]}$$

$$= \frac{P[A|B]\,P[B]}{P[A|B]\,P[B] + P[A|B^c]\,(1 - P[B])}$$

Experiment outcomes $S \xrightarrow{X(\ )} S_X$, where X is the RV.

$$P_X(x) = P[X = x], \quad \forall x \in \mathbb{R}$$

Let's say event A occurred.

$$P_{X|A}(x) = P[X = x \mid A], \quad \forall x \in \mathbb{R}$$

$$P[X = x \mid A] = \frac{P[X = x, A]}{P[A]}$$

VS Word Problem (RY)

Example 2.38 Problem

Let X denote the number of additional years that a randomly chosen 70 year old person will live. If the person has high blood pressure, denoted as event H, then X is a geometric $(p = 0.1)$ random variable. Otherwise, if the person's blood pressure is regular, event R, then X has a geometric $(p = 0.05)$ PMF with parameter. Find the conditional PMFs $P_{X|H}(x)$ and $P_{X|R}(x)$. If 40 percent of all seventy year olds have high blood pressure, $\rightarrow P[H] = 0.4$ what is the PMF of X?

$$P_{X|H}(x) = P[X = x \mid H]$$

$$P_{X|R}(x) = \begin{cases} (0.95)^{x-1}(0.05) & x = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} (0.9)^{x-1}(0.1) & x = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Theorem 2.17

- Consider an Event B. Let B be a subset of the range S_X of RV X.
- What is $P_{X|B}(x)$?

- Start with the definition of conditional probability…

Theorem 2.17

- We have

$$P_{X|B}(x) = \begin{cases} \frac{P_X(x)}{P[B]} & x \in B \\ 0 & \text{otherwise.} \end{cases}$$

Note that this is > $P_X(x)$. Knowledge that B occurred and that {X=x} is in B increases our belief that {X=x} occurred

VVS Word Problem

Example 2.40 Problem

$X \sim P_X(x)$

Suppose X, the time in integer minutes you must wait for a bus, has the uniform PMF

$$P_X(x) = \begin{cases} 1/20 & x = 1, 2, \ldots, 20, \\ 0 & \text{otherwise.} \end{cases} \tag{2.119}$$

Suppose the bus has not arrived by the eighth minute, what is the conditional PMF of your waiting time X?

Event $\{X > 8\} \triangleq A$

$$P_{X|A}(x) = \begin{cases} \dfrac{P(X = x)}{P(A)} & x \in \{X > 8\} = \{9, 10, 11, \ldots, 20\} \\ 0 & \text{otherwise} \end{cases}$$

No Surprises Here!

Theorem 2.18

(a) For any $x \in B$, $P_{X|B}(x) \geq 0$.

(b) $\sum_{x \in B} P_{X|B}(x) = 1$.

(c) For any event $C \subset B$, $P[C|B]$, the conditional probability that X is in the set C, is

$$P[C|B] = \sum_{x \in C} P_{X|B}(x).$$

Conditional Expected Value

- Just replace the PMF $P_X(x)$ in the formula for expectation by the conditional PMF $P_{X|B}(x)$
- The same holds for the conditional expected value of a function of a RV too

Definition 2.20 ***Conditional Expected Value***

The conditional expected value *of random variable* X *given condition* B *is*

$$E[X|B] = \mu_{X|B} = \sum_{x \in B} x P_{X|B}(x) .$$

Expectation of X in Terms of Conditional Expectations Over an Event Space

Theorem 2.19

For a random variable X resulting from an experiment with event space $B_1, \ldots, B_m$,

$$E[X] = \sum_{i=1}^{m} E[X|B_i]\, P[B_i].$$

$$\cancel{E[X] = \sum_{i=1}^{m} E[X|B_i]}.$$

$$E[X] = \sum_{x \in S_X} x \underbrace{P[X=x]}_{\sum_{i=1}^{m} P[X=x|B_i]\, P[B_i]}$$

$$= \sum_{x \in S_X} x \sum_{i=1}^{m} P[X=x|B_i]\, P[B_i]$$

$$= \sum_{i=1}^{m} \underbrace{\left[\sum_{x \in S_X} x\, P[X=x|B_i]\right]}_{\sum_{x \in B_i} x\, P[X=x|B_i]} P[B_i]$$

$\because P[X=x|B_i] = 0, \ \forall x \notin B_i$

Conditional Expected Value of Y=g(X)

Theorem 2.20

The conditional expected value of $Y = g(X)$ given condition B is

$$E[Y|B] = E[g(X)|B] = \sum_{x \in B} g(x) P_{X|B}(x).$$

- Proof: Do as we did for E[Y] in Theorem 2.10.

Word Problem

Quiz 2.9

On the Internet, data is transmitted in packets. In a simple model for World Wide Web traffic, the number of packets N needed to transmit a Web page depends on whether the page has graphic images. If the page has images (event I), then N is uniformly distributed between 1 and 50 packets. If the page is just text (event T), then N is uniform between 1 and 5 packets. Assuming a page has images with probability $1/4$, find the

(1) conditional PMF $P_{N|I}(n)$ $\sim \text{Unif}(1,50)$

(2) conditional PMF $P_{N|T}(n)$ $\sim \text{Unif}(1,5)$

(3) PMF $P_N(n)$ $= P_{N|I}(n)\,P(I) + P_{N|T}(n)\,P[T]$

(4) conditional PMF $P_{N|N\leq 10}(n)$ $\rightarrow \dfrac{P[N=n, N\leq 10]}{P[N\leq 10]} = \begin{cases} \dfrac{P[N=n]}{P[N\leq 10]} & n\leq 10 \\ 0 & \text{otherwise} \end{cases}$

(5) conditional expected value $E[N|N \leq 10]$

(6) conditional variance $\text{Var}[N|N \leq 10]$

Problem 2.7.7

A particular circuit works if all 10 of its component devices work. Each circuit is tested before leaving the factory. Each working circuit can be sold for k dollars, but each nonworking circuit is worthless and must be thrown away. Each circuit can be built with either ordinary devices or ultrareliable devices. An ordinary device has a failure probability of $q = 0.1$ while an ultrareliable device has a failure probability of $q/2$, independent of any other device. However, each ordinary device costs \$1 while an ultrareliable device costs \$3. Should you build your circuit with ordinary devices or ultrareliable devices in order to maximize your expected profit $E[R]$? Keep in mind that your answer will depend on k.

$$E[W_n] = E\left[\frac{n}{K_n+1}\right]$$

$$= E[g(K_n)]$$

$$\neq \frac{n}{E[K_n]+1}$$

Problem 2.7.8

In the New Jersey state lottery, each \$1 ticket has six randomly marked numbers out of $1, \ldots, 46$. A ticket is a winner if the six marked numbers match six numbers drawn at random at the end of a week. For each ticket sold, 50 cents is added to the pot for the winners. If there are k winning tickets, the pot is divided equally among the k winners. Suppose you bought a winning ticket in a week in which $2n$ tickets are sold and the pot is n dollars.

(a) What is the probability q that a random ticket will be a winner? $q = \frac{1}{{}_{46}C_6}$

(b) What is the PMF of K_n, the number of other (besides your own) winning tickets?

$W_n = \frac{n}{K_n + 1}$

(c) What is the expected value of W_n, the prize you collect for your winning ticket?

$S_{K_n} = \{0, 1, 2, \ldots, 2n-1\}$

$P[K_n = 0] = (1-q)^{2n-1}$

$P[K_n = x] = {}_{2n-1}C_x \, q^x (1-q)^{2n-1-x}$.

$P[K_n = 1] = {}_{2n-1}C_1 \, q (1-q)^{2n-2}$

Problem 2.8.10

Let random variable X have PMF $P_X(x)$. We wish to guess the value of X before performing the actual experiment. If we call our guess $\hat{x}$, the expected square of the error in our guess is

$$e(\hat{x}) = E\left[(X - \hat{x})^2\right]$$

$(X-\hat{x})^2$

Show that $e(\hat{x})$ is minimized by $\hat{x} = E[X]$.

First order necessary condition requires that

$$\left.\frac{de(\hat{x})}{d\hat{x}}\right|_{\hat{x}=\hat{x}^*} = 0$$

$$e(\hat{x}) = E\left[(X-\hat{x})^2\right] = E\left[X^2 + \hat{x}^2 - 2\hat{x}X\right] = E[X^2] + \hat{x}^2 - 2\hat{x}E[X]$$

$$\frac{de(\hat{x})}{d\hat{x}} = 2\hat{x} - 2E[X].$$

Set to 0

$\Rightarrow \hat{x} = E[X]$

If N is discrete Uniform$(1,100)$:

$$P_{N|B}(n) = P[N=n \mid B]$$

$$= \begin{cases} \dfrac{P[N=n]}{P[N \geq 20]} & n \in \{20, 21, \ldots, 100\} \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \dfrac{1/100}{81/100} = 1/81 & n \in \{20, \ldots, 100\} \\ 0 & \text{otherwise} \end{cases}$$

Suppose X is the no. of additional tests needed.

$$P[X=1] = 1/81$$

$$P[X=2] = P[N=21 \mid N \geq 20] = \frac{1}{81}$$

$$\vdots$$

$$\Rightarrow X \sim \text{Discrete Unif}(1, 81)$$

When $N \sim \text{Geom}(p)$, $X \sim \text{Geom}(p)$.

This property is called the memoryless property.

Geometric is the only memoryless distribution among discrete RVs.

Problem 2.9.6

Select integrated circuits, test them in sequence until you find the first failure, and then stop. Let N be the number of tests. All tests are independent with probability of failure $p = 0.1$. Consider the condition $B = \{N \geq 20\}$.

(a) Find the PMF $P_N(n)$.

(b) Find $P_{N|B}(n)$, the conditional PMF of N given that there have been 20 consecutive tests without a failure.

(c) What is $E[N|B]$, the expected number of tests given that there have been 20 consecutive tests without a failure?

$$\frac{(1-p)^{n-1}p}{\sum_{n=20}^{\infty}(1-p)^{n-1}p} = (1-p)^{n-20}p$$

$$P_{N|B}(n) = P[N=n|B]$$

$$= \frac{P[N=n, N\geq 20]}{P[N\geq 20]} = \begin{cases} \frac{P[N=n]}{P[N\geq 20]} & n\geq 20 \uparrow n\in B \\ 0 & \text{otherwise} \end{cases}$$

Problem 2.9.7

Every day you consider going jogging. Before each mile, including the first, you will quit with probability q, independent of the number of miles you have already run. However, you are sufficiently decisive that you never run a fraction of a mile. Also, we say you have run a marathon whenever you run at least 26 miles.

(a) Let M equal the number of miles that you run on an arbitrary day. What is $P[M > 0]$? Find the PMF $P_M(m)$.

(b) Let r be the probability that you run a marathon on an arbitrary day. Find r.

(c) Let J be the number of days in one year (not a leap year) in which you run a marathon. Find the PMF $P_J(j)$. This answer may be expressed in terms of r found in part (b).

(d) Define $K = M - 26$. Let A be the event that you have run a marathon. Find $P_{K|A}(k)$.

Assignments

- Clearly state all the RV(s), range space of the RV(s), PMF, event space that you may use to solve the problem
  - Note that event of interest must lie in the event space/ range space
  - It is fine to start with a layman description of the event. However, you must specify the event using proper mathematical notation
- Problems 2.2.6 & 2.2.7
- Problem 2.3.7
- Problem 2.3.11
- Problem 2.3.13
- Problems 2.5.1 – 2.5.8
- 2.5.9
- 2.4.1 and 2.6.1
- 2.6.5, 2.6.6, 2.7.5, 2.7.6
- 2.7.9
- 2.8.9
- 2.9.6
- 2.9.7

Probability and Random Processes

Continuous Random Variables

Sanjit Kaul

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Continuous Random Variables

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Intervals

- (x_1, x_2)
  - Includes neither x_1 nor x_2 but all real numbers in between
- $[x_1, x_2)$
- $(x_1, x_2]$
- $[x_1, x_2]$

Pick randomly $x \in (0, L)$

$P[X = x] = ?$

Note that X is a continuous RV since its range space $S_X = (0, L)$, which is continuum.

Create discrete RV Y by discretizing $(0, L)$ into n exclusive parts.

$$P[Y = y] = \begin{cases} \frac{1}{n} & y \in \{1, 2, \ldots, n\} \\ 0 & \text{otherwise} \end{cases}$$

$P[X = x] = ?$

When $\{X = x\}$ occurs, $\{Y - \ldots\}$ occurs

$$Y = \cancel{\text{ceil}\left(x + L/n\right)}$$

$$Y = \text{ceil}\left(\frac{x}{L/n}\right)$$

$$\{Y = \text{ceil}(x/(L/n))\} \supset \{X = x\}$$

$$P\left[\{Y = \text{ceil}(x/(L/n))\}\right] \geq P[X = x]$$

Taking the limit $n \to \infty$, we get $P[X = x] = 0$.

A Complete Description of a Continuous RV

A Very Narrow Carrom Board

- A very narrow board of length L can be approximated by a line
  - This is a terrible attempt on my part to emulate the example in RY

A Very Narrow Carrom Board

- A very narrow board of length L can be approximated by a line
  - This is a terrible attempt on my part to emulate the example in RY
- You shoot the striker (red)
- Your strike is powerful
- This ensures that the striker stops at any point along the thin board with equal probability
- The striker stops at a point in [0,L]

A Very Narrow Carrom Board

- Let X be the RV that describes the final location of the pointer
- We have $S_X = [0,L]$
- The range is a continuum. Thus X is a CRV
- What is $P[X = x]$ where $x \in S_x$?

A Very Narrow Carrom Board

- Let's think in discrete terms

- We split the board into n equal intervals
- Define the RV Y with $S_Y = \{1,2,\ldots,n\}$
- Y is a discrete uniform RV
- Note that in the discrete problem all values of X in interval k correspond to Y=k.

$$P[Y = k] = \boxed{}$$

A Very Narrow Carrom Board

- Let's think in discrete terms

- We split the board into n equal intervals
- Define the RV Y with $S_Y = \{1,2,\ldots,n\}$
- Y is a discrete uniform RV
- Note that in the discrete problem all values of X in interval k correspond to Y=k.

$$P[Y = k] = P[k - 1 \leq X < k]$$

A Very Narrow Carrom Board

If $X = x$ then $Y =$

A Very Narrow Carrom Board

If $X = x$ then $Y = \lceil x/(L/n) \rceil = \lceil nx/L \rceil$

- Also
- Therefore

$$P[X = x] \quad P[Y = \lceil nx/L \rceil] = 1/n$$

A Very Narrow Carrom Board

- Getting as close as possible to the continuous case

$$P[X = x] \leq \lim_{n \to \infty} 1/n = 0$$

- Therefore

$$P[X = x] = 0$$

- Every outcome {X=x} has a probability 0!
  - Note that this is not true for an interval on the board
- The PMF is not a useful way of describing a continuous RV
- What about the CDF?