Topics Covered in SMMDA

Lecture #1:

- Data Science
- Set Theory
  - Set-Builder Notation
  - Proper subset
  - Equal set
  - Intervals
  - Set Operations
    - Union
    - Intersection
  - Disjoint
  - Difference / Compliment
  - Universal set
  - De Morgan's law
  - Cardinality of a set
  - Infinite, finite & powerset
  - Cartesian Products
    - Relations

Lecture #2:

- Statistics
- Data Sets
  - populations → parameters
  - Samples → Statistics
- Descriptive vs. Inferential Stat
- Types of Data
  - Qualitative
  - Quantitative
- Level of measurements
  - Nominal (Name/ Label/ unordered) / Qualitative
  - Ordinal (order / ranked) Qualitative or Quantitative
  - Interval (order / scalable / ranged) Quantitative
  - Ratio (data comparison (twice, thrice) Quantitative
- Key Terms of Data types
  - Continuous → interval / float / numeric
  - Discrete → random / integer / count / numeric

Key Terms of Data types
- → Categorical → Gender/Color/Chocolate types/...
- → Binary → (0/1, T/F)
- → Ordinal → ordered

LECTURE 3:

- ↳ Probability
- ↳ Sample Space / ~~Sample points~~
- ↳ Events / trial / Experiment
- ↳ Classical Probability = $P(E) = \frac{\text{No. of total outcomes contains in event E}}{\text{Total No. of outcomes in the sample space(S)}}$

  $0 \leq P(A) \leq 1$

  $P(\phi) = 0$

  $P(S) = 1$

  Probability → $\boxed{0 \to 1}$ $P(A) \neq -ve \leq 1$
- ↳ Complement of event A → A' or $\bar{A}$ or A^c
- ↳ Empirical Probability → without sample space

  $P(E) \Rightarrow \frac{\text{Frequency of E}}{\text{Sum of the frequencies}}$
- ↳ Law of large Numbers
  - → More no. of experiments → more accuracy in results of expected outcome
  - but in die toss it does not work due to 50% Probability of H & T.

LECTURE #4:

- ↳ Subjective Probability → educated guess, estimate, opinion, inexact info. based.
- ↳ Sample Space → make from → Tree Diagram → Detail examples
  - → Tables → Detail examples

LECTURE #5:

- ↳ Intersection → $A \cap B$
- ↳ Union → $A \cup B$ → P(A or B) → P(in A, or in B, or in both)
- ↳ Mutually exclusive or Disjoint → $A \cap B = \phi = \{\}$
- ↳ Addition Rules I (1, 2, 3, 4, 5) examples, II (1, 2, 3, 4, 5, 6, 7, 8, 9) examples
  - ↳ Mutually exclusive / disjoint events

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

    or

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

Scanned with CamScanner

→ Addition Rule II
↳ Not mutually exclusive
P(A or B) = P(A) + P(B) − P(A ∩ B)

↳ Error made in Addition Rule points
① Disjoint event → double count → error
② Probability exceed range (0–1)

Lecture #6: Conditional Probability

↳ Independent and Dependent Events

Independent:
P(B/A) = P(B)
or
P(A/B) = P(A)
↓
A occurrence don't effect B

Dependent:
→ P(A and B) = P(A) × P(B/A) (first / given)
→ A occurrence depends / effect B
P(B/A) = P(A and B) / P(A) → A and B have in common $\boxed{A \cap B}$

↳ Replacement vs. without replacement

don't picked vs. picked up/selected
remain same probability vs. probability decreased

⇒ Also see Lect# 6 last Slide to clarify properly.
Slide# 60

Lecture# 7:

↳ Random Variable
↳ Discrete vs Continuous Sample Space
↳ Discrete vs Continuous Random variables
↳ Discrete
↳ Probability Distributions
- Discrete Probability Distribution
- Continuous Probability Distribution

↳ Discrete Probability Distribution
- Binomial D___
- Hypergeometric D___
- Poisson D___
- Geometric D___
- Negative binomial D___
- Discrete Uniform D

Binomial Distribution

$b(x; n, p) = {}^nC_x p^x q^{n-x}$ $x = 0, 1, 2, 3, \ldots n$

$${}^nC_x = \frac{n!}{x!(n-x)!}$$

n = total no. of trials
x = no. of successes $(0, 1, 2, 3, \ldots n)$
p = Probability of success
q = Probability of failure
$p + q = 1$

Lecture #8:

- Binomial Distribution 4 conditions
  - Each trial has only 2 outcomes
  - Fixed (trials)
  - Outcomes of each trial (Independent)
  - Success (probability same for each trial)
- Sampling with replacement
  - 2 parameters n and p.
  - Mean $= np$
  - Variance $= npq$
  - Standard Deviation $= \sqrt{npq}$
  - $p = q \rightarrow$ symmetrical distribution
  - $p < 0.5 \rightarrow$ -vely skewed
  - $p > 0.5 \rightarrow$ +vely skewed
- Normal Distribution / Gaussian distribution
  - no. of trials $\rightarrow$ large $= n$
  - $p, q \rightarrow$ small
  - So, $np \geq 10$ and $np(1-p) \geq 10$
- Expected Value
  - $\rightarrow E(x) = \mu = \sum x P(x)$ ★ (∵ Although Probability never -ve) (but expected value of random variable can be -ve)

Continue ...

Properties of Binomial Distri...

1) Expected value = mean = np
2) Variance = $\sigma^2 = npq$
3) Standard Deviation = S.D = $\sqrt{npq}$

Difference Bernolli & Binomial

Bernolli $\downarrow$ 1 coin or 1 die

Binomial $\downarrow$ n 10 books 110 pens

LECTURE

Topic: Binomial Distribution

$P[X=x] = {}^nC_x p^x q^{n-x}$

n : finite value, total ex. 1000 books, 10 pens, 15 bulbs

p : success } $p+q=1$

q : Failure

X : Quotient → last line of question

Example 1: If 10% of pens manufactured by company are defective. Find probablity that a box containing 12 pens contains

i) Exactly 2 defective pens ⇒ 2

ii) At least 2 defective pens ≤ 2

Solution:

Step 1

1) $n = \text{total} = 12$ 10%

2) $p = \text{defective} = \frac{10}{100} = \frac{1}{10}$

3) $q = 1 - p = 1 - \frac{1}{10} = \frac{9}{10}$

Step 2 Condition to find

i) $P(X=2) = {}^nC_x p^x q^{n-x}$

$= {}^{12}C_2 \left(\frac{1}{10}\right)^2 \left(\frac{9}{10}\right)^{12-2}$

$= 66 \times \left(\frac{1}{100}\right)\left(\frac{9}{10}\right)^{10}$

$= 66 \times \frac{1}{100} \times 0.34$

$= 0.2244$

ii) $P[X \geq 2] = 1 - P[X < 2]$

$= 1 - \left[P(X=0) + P(X=1)\right]$

$P(X=0) = {}^{12}C_0 \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^{12} = 1 \cdot 1 \cdot 0.28$

$P(X=0) \Rightarrow 0.282$

$P(X=1) = {}^{12}C_1 \left(\frac{1}{10}\right)^1 \left(\frac{9}{10}\right)^{11} =$

$P(X=1) = 12 \times 0.1 \times 0.31$

$P(X=1) \Rightarrow 0.376$

$= 1 - [(0.282) + (0.376)]$

$= 1 - (0.658)$

$= 0.3414$

Example 2: Probability that at any moment one telephone line out of 10 will be busy is 0.2

→ Find probability that 5 lines are busy.

Solution: First we have to find parameters

n — total — 10

p — busy = 0.2

q — $1 - p = 0.8$

→ $P[X=5] = {}^nC_x p^x q^{n-x}$

$= {}^{10}C_5 (0.2)^5 (0.8)^{10-5}$

$= 252 \times 0.00032 \times 0.327$

$= 0.0264$