Lecture 03 : Random Variables

Presented By:
Dr. Vinay Kulkarni

Aegis
SCHOOL OF BUSINESS
SCHOOL OF DATA SCIENCE
SCHOOL OF TELECOMMUNICATION

Before we start …

- Review of last week’s exercise

SCHOOL OF BUSINESS
SCHOOL OF DATA SCIENCE
SCHOOL OF TELECOMMUNICATION

RANDOM VARIABLES & PROBABILITY DISTRIBUTIONS

Exercise

- For the 10000 numbers generated:
  - Truncate the numbers to 1 decimal place
  - Find out the frequency of each number
  - Calculate the probability of occurrence of each number in the set
  - Plot probability v/s number
  - Find out the cumulative sum of the probabilities
  - Plot the cumulative probability v/s number

Probability Model

- A Probability Model outlines:
  - The possible outcomes of an experiment
  - Associated probability of each outcome
- Types of Probability
  - Empirical Probability: Experimental
  - Classical Probability: Theoretical
  - Subjective Probability: Judgmental

Random Variables

- **Random experiment**
  - Process of measurement or observation in which the outcome **cannot** be completely determined in advance
- **Sample space**
  - All possible outcomes of a random experiment
- **Random Variable**
  - A real-valued quantity, or numerical measure, whose value depends on the outcomes of a random experiment
  - Can be **Discrete** or **Continuous**

Random Variable

- The probability that a Random Variable may assume a particular value is governed by a Probability Function:
  - For Discrete variables
    - **Probability Mass Function (PMF)**
  - For Continuous variables
    - **Probability Density Function (PDF)**

  - For Discrete / Continuous variables
    - **Cumulative Distribution Function**

Random Variable and Probabilities

- Let X be the Random Variable
- Let x be one of its possible values
- Let P(x) be the probability that X=x
- Then

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

$$\sum_{all\ values\ of\ x} P(x) = 1$$

Aegis
SCHOOL OF BUSINESS
SCHOOL OF DATA SCIENCE
SCHOOL OF TELECOMMUNICATION

Random Variable: Expected Value

- Expected value : E(X)
  - Weighted average, of all possible values, considering their probabilities

- For Discrete random variable

$$E(X) = \mu_X = \sum_{\text{all values of } x} [x.p(x)]$$

- For Continuous random variable

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

Aegis
SCHOOL OF BUSINESS
SCHOOL OF DATA SCIENCE
SCHOOL OF TELECOMMUNICATION

Random Variable: Variance and Standard Dev

- Variance of a Random Variable

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

- Standard Deviation of a Random Variable

$$\sigma_x = +\sqrt{\sigma_X^2}$$

- Where

$$E(X^2) = \sum_{all\, values\, of\, x} [x^2.p(x)]$$, in the discrete case, and

$$E(X^2) = \int_{-\infty}^{\infty} x^2.f(x)dx$$, in the continuous case

- Variance and Standard Deviation reflects the extent to which the Random variable is close to its mean

Aegis
SCHOOL OF BUSINESS
SCHOOL OF DATA SCIENCE
SCHOOL OF TELECOMMUNICATION

Exercise

- Which of the following tables represent valid discrete probability distributions?

(a)

X	p(x)
1	0.2
2	0.35
3	0.12
4	0.40
5	-0.07

(b)

x	p(x)
1	0.2
2	0.25
3	0.10
4	0.14
5	0.49

(c)

x	p(x)
1	0.2
2	0.25
3	0.10
4	0.15
5	0.30

- For valid distributions, calculate:
  – The expected value , variance and standard deviation

Aegis
SCHOOL OF BUSINESS
SCHOOL OF DATA SCIENCE
SCHOOL OF TELECOMMUNICATION

Exercise - Answers

- E(X) = 3.10
- Variance = 2.39
- Standard Deviation = 1.54596

Sample – to – Population

- Goal of all these definitions:
  - Given the nature of the phenomena
    - Probability distributions
  - And a sample with certain deductions
    - Measured observations
  - Predict additional properties and confidence levels
- We therefore need to
  - Understand generic phenomena
  - And the probabilistic nature of their events

Probability Distributions

- Binomial Distribution
- Poisson Distribution
- Students T-Distribution
- Chi-Square Distribution
- F-Distribution
- Normal Distribution
- Log normal Distribution
- Other Distributions
  - Bernoulli
  - Geometric
  - Hypergeometric
  - Multinomial
  - Exponential
  - Beta
  - Gamma