RANDOM VARIABLES

A random variable is a variable which has its value determined by a probability experiment.

It takes any value from a sample space.

e.g if you flip a coin once, how many tails can you get.

T is our random variable of tails possible from our probability experiment.

S = { H, T} so T is either 0 or 1

Random Variables

- A variable is said to be random if the sum of the events probabilities is one.
- **Random variable**
- A random variable is a variable that assigns a value to each outcome of an experiment. Letters such as X , Y and Z are used to denote a random variable.
- Examples
  1. Experiment: Select a mutual fund; X = the number of companies in the fund portfolio.
  The values of X are 2, 3, 4, ...
- 2. Experiment: Select a soccer player; Y = the number of goals the player has scored during the season.
  The values of Y are 0, 1, 2, 3, ...
- 3. Experiment: Survey a group of 10 soccer players; Z = the average number of goals scored by the players during the season.
  The values of Z are 0, 0.1, 0.2, 0.3, ..., 1.0, 1.1, ...

Discrete random variable

- Assumes a finite or countable number of distinct values. [Denoted by X]
- It can be defined on both countable and uncountable sample spaces.
- $P(X = x) = p(x)$ is the prob that X takes a value x
- $p(x)$ is the prob function for X.
- Probability distribution for a discrete random variable X is represented by a Formula, a Table or a Graph. So, $f(x)p(x) = P(X = x)$ for each x within a range of X is called Probability distribution of X.
- Tables show probabilities of various outcomes in an experiment.

Properties of the Probability distribution for a discrete random variable

1. $f(x) \geq 0$ for each value within a domain.
2. $\sum_x f(x) = 1$ for all values within its domain.

Example

Experiment of tossing 3 coins at the same time.

S = {HHH HHT HTH THH HTT THT TTH TTT}

Suppose the number of tails is the random variable X

$$X = \{x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7 \quad x_8\}$$

$$\{0 \quad 1 \quad 1 \quad 1 \quad 2 \quad 2 \quad 2 \quad 3\}$$

$$P(X=x) = \{ \qquad \}$$

Probability Distributions of Random Variables

- Given a random variable X and that X = 2. We mean the event consisting of all outcomes that have an assigned X-value of 2. For example: Throw a pair of fair dice and take X to be the sum of the numbers facing up. Then
- The event that X = 2 is {(1, 1)}. The event that you throw a 2
  The event that X = 3 is {(2, 1), (1, 2)}. The event that you throw a 3
  The event that X = 4 is {(3, 1), (2, 2), (1, 3)}. The event that you throw a 4
  Each of these events has a certain probability.

P(X = 4) = 3/36 = 1/12

Thus the event in question consists of three of the 36 possible equally likely outcomes.

A coin is tossed 4 times:

Suppose the number of heads is the random variable X. Complete the table below:

Sample space	Probability	Value for random variable X
HHHH	1/16	

Continuous random variable

- A variable that takes any value between two specified values. The probability distribution is continuous as well.
- f(x) is a continuous function of x

[Make a comparison between discrete and continuous random variables]

Cumulative probability

- Probability that the value of a random variable falls within a specified range.
- Flipping a coin 2 times. What is the prob that this would result in one or fewer heads. This is an example of a cumulative probability $P(X \leq x)$ ie the results would be prob of 0 heads + prob of one head:
- $P(X \leq 1) = P(X = 0) + P(X = 1) = ¼ + ½ = ¾$

Table for the coin experiment

Number of heads: x	Probability: P(X = x)	Cumulative probability: P(X ≤ x)
0	¼	¼
1	½	¾
2	¼	1

Uniform distribution

- All values of a random variable occur with equal probability. Toss a die. What is the prob that the die will land on a 6?

S = {1,2,3,4,5,6}. Each possible outcome is a random variable X and each outcome is equally likely, thus uniform distribution.

$P(X = 6) = 1/6$

- In the same experiment what is the prob that the die will land a number less that 5?
- Thus cumulative probability:
- $P(X< 5) = P(X=1) + P(X=2) + P(X=3) + P(X=4) = 2/3$

- **A probability distribution for a *discrete random variable* X *consists of:***

✓ **Possible values.**

✓ **Corresponding probabilities.**

Some Distributions

Binomial Probability Distribution /A Bernoulli trial

Binomial means there are two discrete mutually exclusive outcomes of a trial.

e.g success or failure

heads or tails

on or off

sick or healthy

A series of trails n will follow a Binomial distribution if:

a) The probability of success p is constant

b) trails are independent of one another

Binomial

p = probability of success

q = probability of failure = (1-p)

For a Binomial random variable

B(X, n, p)

P(x) = $\frac{n!}{x!(n-x)!} * p^x q^{(n-x)}$

Mean = np

Variance = npq

Standard deviation = $\sqrt{npq}$

Where n is the total number of observations.

Example: lets say you have 10 multiple choice questions with answers A B C D.

Calculate the probability of getting ‘ No answer’ correct?

exercise

A die is tossed 5 times. What is the prob of getting exactly 2 fours?

Poisson probability distribution

- Properties of Poison probability distribution:
- Outcomes are successes or failures.
- Average number of successes (μ) in a specified region is known.
- The prob that a success will occur is proportional to the size of the region.
- The prob that a success will occur in a small region is 0.

Poison formula

- $P(x, \mu) = (e^{-\mu}) * (\mu^x)/x!$
- e: constant = 2.71828
- μ: Mean number of successes that occur in a specified region.
- X: Actual number of successes that occur in a specified region.
- P(x, μ): Poison prob that exactly x successes occur in the Poison experiment for a mean number of successes μ.

$$\text{Variance} = \mu$$
$$\text{Standard deviation} = \sqrt{\mu}$$

example

- The average number of homes sold by HCC is 2 homes per day. What is the prob that exactly 3 homes will be sold tomorrow?
- $P(x, \mu) = (e^{-\mu}) * (\mu^x)/x!$
- $\mu = 2$
- $x = 3$
- $e = 2.71828$

exercise

- Suppose 100 pages of the book are randomly selected . Mean = 1.5. What is the probability that there are ‘no typos’?

Range, Mean, Mode, Median, variance and Standard Dev

Range: measure of dispersion.

Range = Largest obser – Smallest obser.

Mean: sum the sampled values divided by the number of items in the sample.

[Define mode and Median]

Variation: Expectation of the squared deviation of a random variable from its mean. Sum of squared deviations of a variable.

Variance: Measure of the spread of data/measure of deviation from the mean. Variance – Averaging, Population variance, Sample variance.

Standard Dev: Square root of variance**.**

Coefficient of variation: *is the standard deviation divided by the mean.*

Mean = $\frac{\sum x}{n}$

Variance σ^2 = $\Sigma(xi-\bar{x})^2$ /N-1

Standard deviation = $\sqrt{\frac{\Sigma(xi-\bar{x})^2}{N-1}}$

exersice

1. { 7, 11, 11, 15, 20, 20, 28}

Examples

1. Find the variance and standard deviation of the following scores on an exam: 92, 95, 85, 80, 75, 50 .
2. Find the standard deviation of the average temperatures recorded over a five-day period last winter: 18, 22, 19, 25, 12.