Lesson 89 – Discrete Random Variables

HL2 Math - Santowski

Introductory Example

- Suppose Egypt's national football (soccer that is) team is playing against USA's national team and the game is here in Cairo.
- To generate data in our "experiment," we will go into the stadium and select three fans randomly and identify whether the fan supports Egypt (E) or the US team (U)
- The experiment gives us the following sample space:

$$\{(EEE),(EEU),(EUE),(UEE),(EUU),(UUE),(UEU),(UUU)\}$$

Introductory Example - continued

- Now, we will let X represent the number of fans selected who are supporting Egypt
- Therefore, the possible values of X are 0,1,2, or 3
- Now let's work out the probability that X takes on the value of 0 ➔
- And then let's work out the probability that X takes on the value of 1 ➔
- And then let's work out the probability that X takes on the value of 2 ➔
- And then let's work out the probability that X takes on the value of 3 ➔
- So we are effectively working out the probability that X takes on the value of x (where x = 0,1,2,3) ➔ we shall denote this idea as P(X = x)
- To determine the probabilities, let's say the game is in Cairo so P(E) = 0.9 and let's further say that our probabilities are independent ➔ so therefore ……..

Introductory Example - continued

- To determine the probabilities, let's say the game is in Cairo and P(E) = 0.9
- So to find P(X = 3) ➔ (i.e. all three fans we asked supported the Egyptian team:
- P(X = 3) = P(E) x P(E) x P(E) = 0.9 x 0.9 x 0.9= 0.729
- To find P(X = 2) ➔ that two of the three fans support Egypt: i.e ➔ EEU,EUE, UEE
- P(X = 2) = 0.9x0.9x0.1 + 0.9x0.1x0.9 + 0.1x0.9x0.9 = 0.243
- To find P(X = 1) ➔ that one of the three fans support Egypt: i.e ➔ EUU,UUE, UEU
- P(X = 1) = 0.9x0.1x0.1 + 0.1x0.1x0.9 + 0.1x0.9x0.1 = 0.027
- To find P(X = 0) ➔ that none of the three fans support Egypt: i.e ➔ UUU
- P(X = 0) = 0.1x0.1x0.1 = 0.001

Introductory Example - continued

- Several concluding comments:
- We will say that the probabilities in our experiment "behave well" in that (i) the probabilities are greater than zero: P(X = x) > 0 and (ii) the probability of our sample space is 1: P(X=3) + P(X=2) + P(X=1) + P(X=0) = 1
- Because the values that X takes on are random, the variable X has a special name ➔ its called a **random variable**!

Random Variable - Definition

- A **random variable** is a numerical measure of the outcome from a probability experiment, so its value is determined by chance. Random variables are denoted using letters such as *X*.

Random Variables – Two Types

- A **discrete random variable** is a random variable that has values that has either a finite number of possible values or a countable number of possible values.
- A **continuous random variable** is a random variable that has an infinite number of possible values that is not countable.

Discrete & Continuous Data

- A discrete space is one in which all possible outcomes can be clearly identified and counted:
  - Die Roll: S = {1,2,3,4,5,6}
  - Number of Boys among 4 children: S = {0, 1, 2, 3, 4}
  - Number of baskets made on two free throws: S = {0, 1, 2}
- Data that can has a clearly finite number of values is known as discrete data.
- A non-discrete space, which is called called a continuous interval/space, is one in which the outcomes are too numerous to identify every possible outcome:
- Height of Human Beings: S = [0.00 inches, 100.00 inches] → It is impractical to specify every possible height from 0.00 to 100.00.
- Muscle gain over 1 year of weight training: [0 pounds, 100+ pounds]

Discrete Random Variable Examples

Experiment	Random Variable	Possible Values
Make 100 Sales Calls	# Sales	0, 1, 2, ..., 100
Inspect 70 Radios	# Defective	0, 1, 2, ..., 70
Answer 33 Questions	# Correct	0, 1, 2, ..., 33
Count Cars at Toll Between 11:00 & 1:00	# Cars Arriving	0, 1, 2, ..., ∞

Continuous Random Variable Examples

Experiment	Random Variable	Possible Values
Weigh 100 People	Weight	45.1, 78, ...
Measure Part Life	Hours	900, 875.9, ...
Amount spent on food	$ amount	54.12, 42, ...
Measure Time Between Arrivals	Inter-Arrival Time	0, 1.3, 2.78, ...

Two Types of Random Variables

- **Discrete random variables**
  - Number of sales
  - Number of calls
  - Shares of stock
  - People in line
  - Mistakes per page

- **Continuous random variables**
  - Length
  - Depth
  - Volume
  - Time
  - Weight

Notations used with DRV

- We use capital letter, like X, to denote the random variable and use small letter to list the possible values of the random variable.
- Example. A single die is cast, X represent the number of pips showing on the die and the possible values of X are x=1,2,3,4,5,6.
- The statement P(X = x) signifies the probability that the random variable X takes on possible value of x

Probability Distributions for DRV

- The **probability distribution** of a discrete random variable is a graph, table or formula that specifies the probability associated with each possible outcome the random variable can assume.
  - $p(x) \geq 0$ for all values of x
  - $\Sigma p(x) = 1$

13

Discrete Probability Distribution Example

Experiment: Toss 2 coins. Count number of tails.

Probability Distribution

Values, x	**Probabilities, $p(x)$**
0	1/4 = .25
1	2/4 = .50
2	1/4 = .25

Visualizing Discrete Probability Distributions

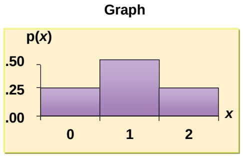
Listing

{ (0, .25), (1, .50), (2, .25) }

Table

# Tails	f(x) Count	p(x)
0	1	.25
1	2	.50
2	1	.25

Graph



Formula (to be explained later)

$$p(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

Example #1 – Using a Formula

- A discrete random variable, X, has a probability distribution function given by $f(x) = P(X = x) = kx^2$ where x = 1,2,3,4
- (a) Prepare a probability distribution table
- (b) Hence or otherwise, determine the value of k.
- (c) Prepare a bar chart/histogram to represent f(x)
- (d) Calculate $P(X = 3)$
- (e) Calculate $P(X = 3 \text{ or } X = 4)$
- (f) Calculate $P(1 \leq X \leq 3)$

Example #2 – Given Probabilities

- The probabilities that a customer selects 1, 2, 3, 4, or 5 items at a convenience store are 0.32, 0.12, 0.23, 0.18, and 0.15 respectively.
  - (a) Construct a probability distribution (table) for the data and draw a probability distribution histogram.
  - (b) Find $P(X > 3.5)$
  - (c) Find $P(1.0 \leq X < 3.0)$
  - (d) Find $P(X < 5)$
  - (e) Calculate $P(X = 3)$
  - (f) Calculate $P(X = 3 \text{ or } X = 4)$
  - (g) How probable is it that a customer selects at least 2 items?

Example #3 – Contextual

- From past experience, a company has found that in carton of transistors, 92% contain no defective transistors, 3% contain one defective transistor, 3% contain two defective transistors, and 2% contain three defective transistors.
- (a) Construct a probability distribution.
- (b) Draw a probability distribution histogram.
- (c) Draw a cumulative probability distribution histogram
- (d) Calculate the mean, variance, and standard deviation for the defective transistors

Example #4 - Contextual

- There are 8 red socks and 6 blue socks in a drawer. A sock is taken out, its color noted and returned to the drawer. This procedure is repeated 4 times.
- Let R be the random variable "number of red socks taken."
- (a) Determine the probability distribution of R and represent it in a chart and a graph
- (b) The procedure is repeated, but this time each sock is NOT returned after it has been taken out.
- (c) The procedure is repeated in the DARK. How many socks must be taken out so that you have a matching pair?

Summary Measures

1. Expected Value (Mean of probability distribution)
   - Weighted average of all possible values
   - $\mu = E(x) = \Sigma x\, p(x)$
2. Variance
   - Weighted average of squared deviation about mean
   - $\sigma^2 = E[(x - \mu)^2] = \Sigma\, (x - \mu)^2\, p(x)$
3. Standard Deviation
   - $\sigma = \sqrt{\sigma^2}$

Formulas

- If $\mu = \frac{\sum_{i=1}^{k} f_i x_i}{n}$ show that $\frac{\sum_{i=1}^{k} f_i (x_i - \mu)^2}{n} = \frac{\sum_{i=1}^{k} f_i x_i^2}{n} - \mu^2$