The Data Driven Manager

4

Determining Probabilities

Basic Probability

Learning Objectives

- Describe the concept of probability
- Use the rules of probability to perform basic probability calculations
- Discriminate between marginal, joint and conditional probabilities
- Discriminate between mutually exclusive and non-mutually exclusive

Learning Objectives

- Discriminate between independent and dependent events
- Calculate marginal, joint, and conditional probability calculations under independent and dependent conditions

What is Probability?

- The chance that an event will or will not occur
- Typically expressed in fractions or decimals

Definitions

- **Experiment -** the process of obtaining an observation, outcome or simple event
- **Outcome -** result of an experiment (after the fact)

Definitions

- **Event -** one or more of the possible outcomes of a situation or experiment (a possibility, or possible occurrence) to which a probability is assigned (a set, or total number of outcomes)

- **Sample space** is the set of all possible outcomes from an experiment

Definitions

- **Events** are termed **mutually exclusive** when one and only one can take place at the same time
- **Collectively Exhaustive** refers to **lists** containing all of the possible events which may result from an experiment.

Classical Probability

The probability that an event will occur

$$P = \frac{N}{S}$$

where

P = Probability of an event

N = Number of outcomes where the event occurs

S = Total Number of possible outcomes; and where each of the possible outcomes are equally likely

Rules / Conditions of Probability

Typical conditions of concern:

- The case where one event **or** another will occur
- The situation with two or more events where **both** may occur

Marginal or Unconditional Probability

- $P(A)$ = the probability P of event A occurring
- Where a single probability is involved, only one event can take place

Example

A production lot of 100 parts contains one defective part. What is the P of selecting one part randomly from the lot, and drawing the defective?

$$P(D) = \frac{1}{100} = 0.01 = 1\%$$

Compound Events

- You have a compound event when there are 2 or more events

- If you have 2 or more events, you may want to know the probability of one event **or** the other (we'll call them A or B)

Probability of A or B

- If this is the question of interest, P(A or B), then we have to know if A and B are:
  - **Mutually exclusive** (the events can't happen at the same time) or
  - **Non-mutually exclusive** (the events A and B can happen at the same time).

Addition Rule

- The equation for this type of probability equation is:
  - P(A **or** B) = P(A) + P(B) - P(A and B)

- However, if they are mutually exclusive, the equation reduces to:
  - P(A **or** B) = P(A) + P(B)

Non-Mutually Exclusive Example

- We are going to flip two coins at the same time
- What is the probability that (A) the first coin is a head **or** the (B) second coin is a head?
- Note that these events are **non-mutually exclusive** because both can exist at the same time.

Non-Mutually Exclusive Example

- We have several possible outcomes:
  - **A happens and B does not happen**
  - **B happens and A does not happen**
  - **Both A and B happen**
  - Neither A nor B happens

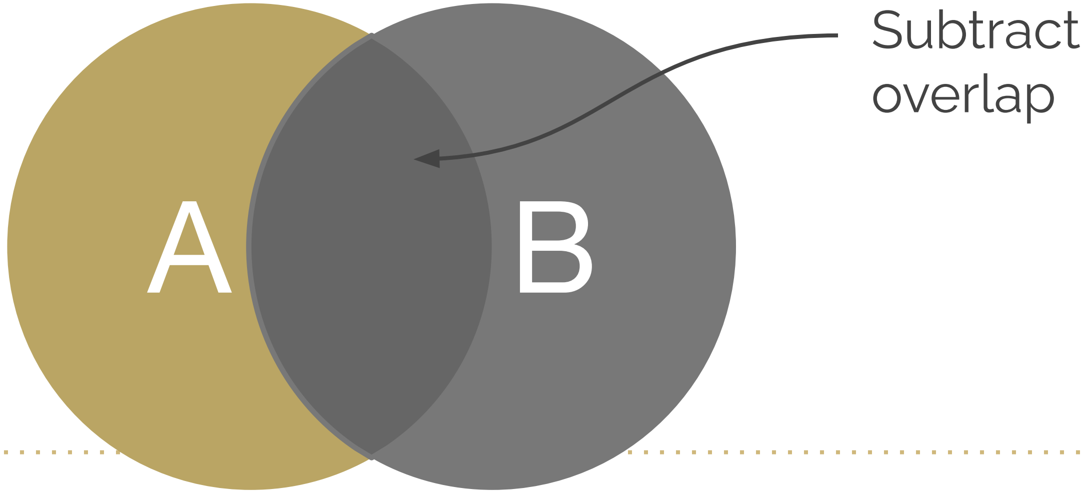
Non-Mutually Exclusive Example

- What are the associated probabilities?
  - The probability that we get a head on the first coin is ½.
  - The probability that we get a head on the second flip is ½.
  - The probability that we get both A and B is ½ x ½ = ¼.

- Therefore:
  P(A or B) = 0.5 + 0.5 - 0.25 = ¾ or 0.75

Addition Rule for Non-Mutually Exclusive Events

P(A or B) = P(A) + P(B) – P(A+B)



Addition Rule for Non-Mutually Exclusive Events Example

Given a mixed lot with the following characteristics:

Vendor	# Defective	# Not Defective
Vendor A	15	85
Vendor B	10	55

Addition Rule for Non-Mutually Exclusive Events Example

What is the probability, on a single random draw, of selecting a part from Vendor A **or** a defective part?

Addition Rule for Non-Mutually Exclusive Events Example

If we were to simply use P(A) + P(B), then

$$\frac{100}{165} + \frac{25}{165} = \frac{125}{165} = 0.7575$$

Addition Rule for Non-Mutually Exclusive Events Example

Note, however, that there are 15 more parts credited to the total than should be!

Vendor	# Defective	# Not Defective
Vendor A	**15**	85
Vendor B	10	55

Addition Rule for Non-Mutually Exclusive Events Example

P(A and B) = $\dfrac{15}{165}$

So, P(A or B) = P(A) + P(B) – P(A+B)

$$\frac{100}{165}+\frac{25}{165}-\frac{15}{165}=\frac{110}{165}=0.6666$$

Mutually Exclusive Example

You have a bag of candy. There are

- 3 green candies,
- 2 red candies,
- 7 yellow candies, and
- 3 blue candies

Mutually Exclusive Example

What is the probability of drawing one yellow **or** one green candy?

P(Yellow or Green) = 7/15 + 3/15 = 10/15 = ⅔ = 0.67

Probability of A and B

- If you have 2 or more events, you may want to know the probability of one event and the other (we'll call them A and B). This is called a **joint probability**
- A joint probability is the probability of two events occurring simultaneously or sequentially

Probability of A and B

- If this is the question of interest, P(A **and** B), then we have to know if A and B are:
  - **Independent** (the occurrence of A does not influence the occurrence of B, or vice versa) or
  - **Dependent** (the occurrence of A influences the occurrence of B, or vice versa)

Multiplication Rule

- The equation for this type of probability equation is:
  - $P(A \text{ and } B) = P(A) \times P(B|A)$

- However, if they are independent, the equation reduces to:
  - $P(A \text{ and } B) = P(A) \times P(B)$

Joint Probability - Independent Example

- We are going to flip a coin two times. What is the probability that the first flip is a head **and** the second flip is a head? Note that these are independent because the first flip is independent of the second flip.

Joint Probability - Independent Example

- We have several possible outcomes:
  - A happens and B does not happen
  - B happens and A does not happen
  - **Both A and B happen**
  - Neither A nor B happens

Joint Probability - Independent Example

- What are the associated probabilities?
  - The probability that we get a head on the first flip is ½.
  - The probability that we get a head on the second flip is ½.

Joint Probability - Independent Example

- Therefore:
- P(A and B) = 0.5 x 0.5 = 0.25

Joint Probability - Dependent Example

- We are going to sequentially draw 2 cards from a standard deck of 52 cards. What is the probability that the first card is a King **and** the second card is a Queen? Note that these are **dependent** because drawing a card reduces the total sample size, and affects the next sequential card draw.

Joint Probability - Dependent Example

- We have several possible outcomes:

First Draw	Second Draw
King	Not a Queen
Not a King	Queen
Not a King	Not a Queen
King	**Queen**

Joint Probability - Dependent Example

- What are the associated probabilities?
  - The probability that we get a King on the first draw is 4/52
  - The probability that we get a Queen on the second draw is 4/**51**

Joint Probability - Dependent Example

- Therefore:
  - $P(A \text{ and } B) = P(A) \times P(B|A)$
  - $P(A \text{ and } B) = 0.0769 \times 0.0784 = 0.006$

Conditional Probability

- We've now explored marginal probabilities and joint probabilities, under statistical independence and dependence. But what about **conditional** probabilities?
- A **conditional probability** is the probability of one event occurring in the presence of a second event

Conditional Probability

- P(B|A) = Probability of event B occurring, given that A has occurred.
- The equation associated with conditional probability is:

$$P(B|A) = \frac{P(BA)}{P(A)}$$

Conditional Probability - Independent Conditions

- P(B|A) = P(B)...because A and B are independent!

$$P(B|A) = \frac{P(BA)}{P(A)}$$

Conditional Probability - Dependent Conditions

- Let's consider the following example. In a standard deck of 52 cards, what is the probability of drawing a card that is a diamond (B), given that the card is red (A) ?

Conditional Probability - Dependent Conditions

- The probability of getting a red **and** diamond card is going to be 13/52, or ¼ (0.25) since there are 13 diamond cards that are also red

- The probability of getting a red card is 26/52, or ½ (0.5) since half the deck of cards is red (hearts and diamonds)

Conditional Probability - Dependent Conditions

- Putting this into the equation
- 0.25 / 0.5 = 0.5

$$P(B|A) = \frac{P(BA)}{P(A)}$$

Conditional Probability - Dependent Conditions

- Assuming a randomly selected part is from Vendor A, what is the P that it is also defective?

Vendor	# Defective	# Not Defective	Total
Vendor A	15	85	100
Vendor B	10	55	65
Total	25	140	**165**

Conditional Probability - Dependent Conditions

$$P(Def|A) = \frac{P(Def\ and\ A)}{P(A)}$$

$$P(Def|A) = \frac{15/_{165}}{100/_{165}} = 0.15$$

Conditional Probability - Dependent Conditions

- Note: This is the same as observing that there are 15 defectives out of 100 Vendor A parts, then $^{15}/_{100} = 0.15$

Vendor	# Defective	# Not Defective	Total
Vendor A	**15**	85	100
Vendor B	10	55	65
Total	25	140	**165**

Conditional Probability - Dependent Conditions

- Note also that the P(Defective and Vendor A) constitutes a joint probability under statistical dependence. Creating a table of joint P values for the sample space:

Conditional Probability - Dependent Conditions

Event	P	Fraction
Vendor A and Defective	0.0909	$\frac{15}{165}$
Vendor A and Not Defective	0.5151	$\frac{85}{165}$
Vendor B and Defective	0.0606	$\frac{10}{165}$
Vendor B and Not Defective	0.3333	$\frac{55}{165}$

Conditional Probability - Dependent Conditions

- As a second example, assume that a non-defective part has been drawn. What is the P that it is from Vendor B?

$$P(B|ND) = \frac{P(B\ and\ ND)}{P(ND)} = \frac{0.3333}{0.8484} = 0.3929$$

Conditional Probability - Dependent Conditions

- Note that should a non-defective part have been selected, the P of it being a part from Vendor B is

$$\frac{55}{140} = 0.3929 \ \ or \ P(B|ND)$$

Practice Problem

A statistics professor uses her Microsoft Surface Pro to print to her wireless printer at home. Approximately 40% of the time she attempts to print a pdf, the print job fails from her computer.

Practice Problem

Additionally, she recently installed an Eero 6 wireless mesh router to which the printer is "connected". Approximately 15% of the time, the printer isn't connected to the router (or the router doesn't recognize it).

Practice Problem

What is the probability that, on any given pdf print job, a print job will be successful (that is, computer and printer are working correctly and a document is created)?

What is the probability that, on any given attempt, that a print job will be **unsuccessful**?

Probability Distributions

Learning Objectives

- Describe the concept of a probability distribution
- Discriminate between discrete and continuous probability distributions
- Identify the probability distributions most commonly used in decision making

Probability Distributions

- Probability distributions are theoretical frequency distributions which are collectively exhaustive

Probability Distributions

- For example, suppose we have historical evidence to show that a particular vendor will provide a defective part to us 20 times out of 100. Therefore, the P of receiving a Defective part (D) is:

$$P(D) = \frac{20}{100} = 0.20$$

Probability Distributions

- Let us determine the probabilities associated with any two parts randomly drawn from a large production lot.

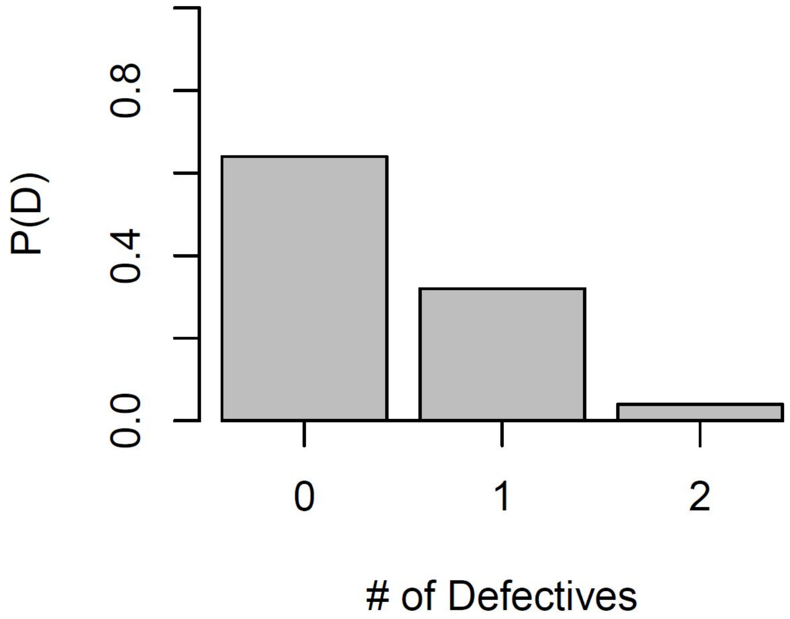
1st Part	2nd Part	# Def. @ 2 parts	P
D (0.20)	ND (0.80)	1	0.16
D (0.20)	D (0.20)	2	0.04
ND (0.80)	D (0.20)	1	0.16
ND (0.80)	ND (0.80)	0	0.64
		Total	1.00

Probability Distributions

- We can now create a probability distribution conforming to our theoretical expectation for two parts so that:

# of Defectives	Draws	P(D)
0	(ND, ND)	0.64
1	(ND, D) + (D, ND)	0.32
2	(D,D)	0.04

Probability Distributions



Types of Probability Distributions

- **Discrete** – A discrete probability distribution is one where there are a limited number of possible values

- **Continuous** – A continuous probability distribution has relatively unlimited possibilities for variable values

Random Variables

- A random variable is one which can take on different values as a result of the outcomes of a random experiment.

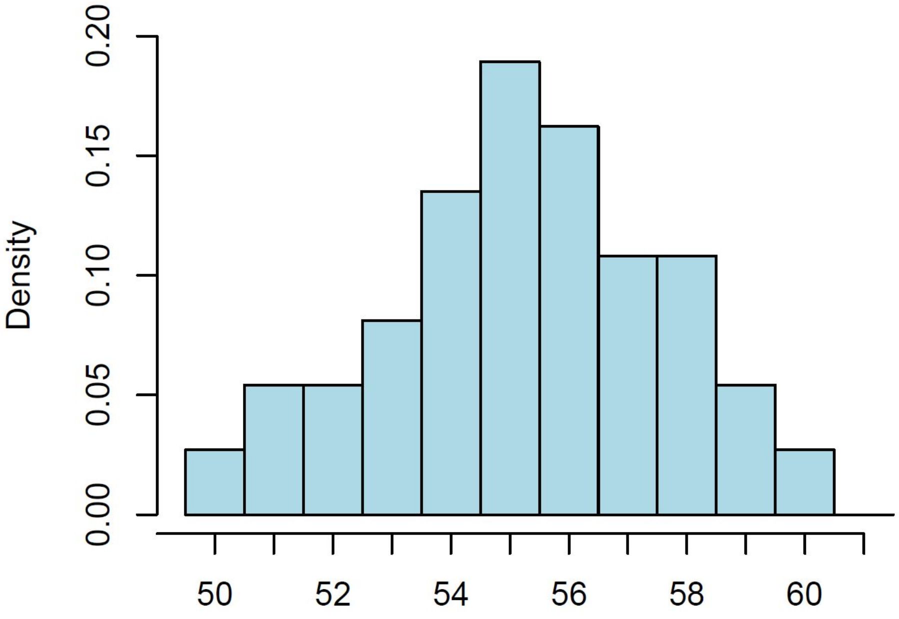
- Random variables, further, can be either discrete or continuous

Probability Distribution for Discrete Random Variable

- Assume that an automated process produces between 50 and 60 parts per day. During a two month production period, daily production levels (DP) were noted and the following data were generated:

Daily Production (DP)	# of Days	P(DP)
50	1	0.027
51	2	0.054
52	2	0.054
53	3	0.081
54	5	0.135
55	7	0.189
56	6	0.162
57	4	0.108
58	4	0.108
59	2	0.054
60	1	0.027
	$\Sigma f = 37$	1.000

Probability Distribution for Discrete Random Variable



Expected Value of a Discrete Random Variable

- One of the most important factors related to any probability distribution is the ability to define the expected value of a random variable

Expected Value of a Discrete Random Variable

- The expected value of a discrete random variable is the weighted average of the expected outcomes

Daily Production (DP)	P	Weighted P Value (DP x P)
50	0.027	1.351
51	0.054	2.757
52	0.054	2.811
53	0.081	4.297
54	0.135	7.297
55	0.189	10.405
56	0.162	9.081
57	0.108	6.162
58	0.108	6.270
59	0.054	3.189
60	0.027	1.621
	Sum	**55.243**

Expected Value of a Discrete Random Variable

- Therefore, E(DP) = 55.243 - equal to the mean

- Of course, we would like to avoid performing this type of activity each time we encounter a probability distribution.

Expected Value of a Discrete Random Variable

- This is particularly true if the data is known to occur in some previously understood form

- In such a case, we may employ a series of algebraic formulas and theoretical probability models to accomplish the analysis of our data

Commonly-Employed Probability Distributions

Discrete

- **Binomial**
- **Poisson**
- Hypergeometric
- Geometric

Continuous

- **Normal**
- **Exponential**
- Weibull Family
- Johnson Family
- Other Distributions