Basic Probability

Learning Objectives

In this chapter, you learn:

- Basic probability concepts
- Calculating probabilities of events
- Joint events
- Conditional probability
- Decision trees

Basic Probability Concepts

- Probability – the chance that an uncertain event will occur (always between 0 and 1)

- Impossible Event – an event that has no chance of occurring (probability = 0)

- Certain Event – an event that is sure to occur (probability = 1)

Example

Find the probability of selecting a male taking statistics from the population described in the following table:

	Taking Stats	Not Taking Stats	Total
Male	84	145	229
Female	76	134	210
Total	160	279	439

Probability of male taking stats $= \dfrac{\text{number of males taking stats}}{\text{total number of people}} = \dfrac{84}{439} = 0.191$

Events

Each possible outcome of a variable is an event.

- Simple event
  - An event described by a single characteristic
  - e.g., A day in January from all days in 2013
- Joint event
  - An event described by two or more characteristics
  - e.g. A day in January that is also a Wednesday from all days in 2013
- Complement of an event A (denoted A')
  - All events that are not part of event A
  - e.g., All days from 2013 that are not in January

Mutually Exclusive Events

- Mutually exclusive events
  - Events that cannot occur simultaneously

Example: Randomly choosing a day from 2013

A = day in January; B = day in February

  - Events A and B are mutually exclusive

Collectively Exhaustive Events

- Collectively exhaustive events
  - One of the events must occur
  - The set of events covers the entire sample space

Example: Randomly choose a day from 2013

A = Weekday; B = Weekend; C = January; D = Spring;

- Events A, B, C and D are collectively exhaustive (but not mutually exclusive – a weekday can be in January or in Spring)
- Events A and B are collectively exhaustive and also mutually exclusive

Computing Joint and Marginal Probabilities

- The probability of a joint event, A and B:

$$P(A \text{ and } B) = \frac{\text{number of outcomes satisfying } A \text{ and } B}{\text{total number of elementary outcomes}}$$

- Computing a marginal (or simple) probability:

$$P(A) = P(A \text{ and } B_1) + P(A \text{ and } B_2) + \cdots + P(A \text{ and } B_k)$$

  - Where B_1, B_2, …, B_k are k mutually exclusive and collectively exhaustive events

Joint Probability Example

P(Jan. and Wed.)

$$= \frac{\text{number of days that are in Jan. and are Wed.}}{\text{total number of days in 2013}} = \frac{5}{365}$$

	Jan.	Not Jan.	Total
Wed.	5	47	52
Not Wed.	27	286	313
Total	32	333	365

Marginal Probability Example

P(Wed.)

$$= P(\text{Jan. and Wed.}) + \text{P(Not Jan. and Wed.)} = \frac{4}{365} + \frac{48}{365} = \frac{52}{365}$$

	Jan.	Not Jan.	Total
Wed.	4	48	52
Not Wed.	27	286	313
Total	31	334	365

Marginal & Joint Probabilities In A Contingency Table

Event	Event: B_1	Event: B_2	Total
A_1	$P(A_1 \text{ and } B_1)$	$P(A_1 \text{ and } B_2)$	$P(A_1)$
A_2	$P(A_2 \text{ and } B_1)$	$P(A_2 \text{ and } B_2)$	$P(A_2)$
Total	$P(B_1)$	$P(B_2)$	1

Joint Probabilities

Marginal (Simple) Probabilities

Probability Summary So Far

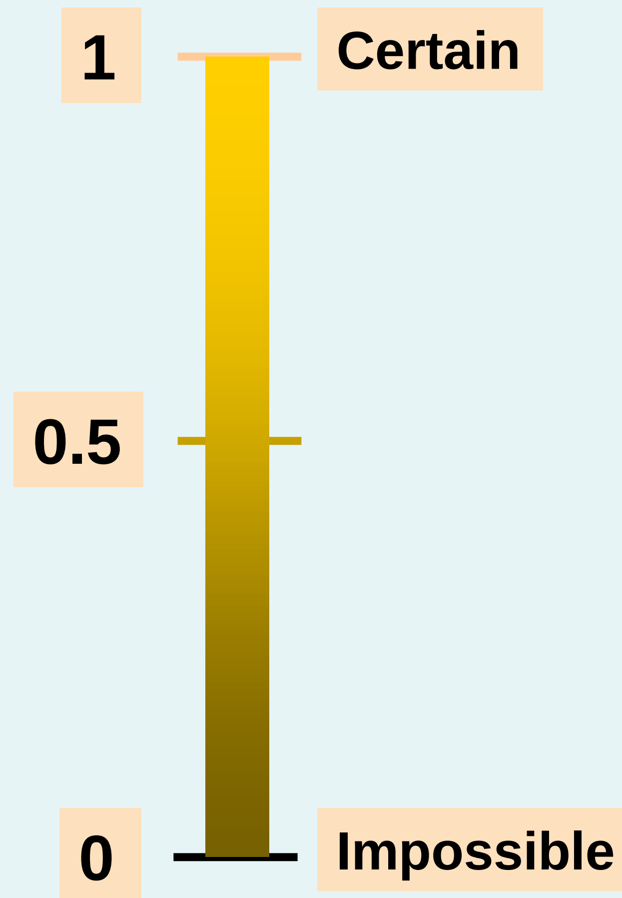
- Probability is the numerical measure of the likelihood that an event will occur
- The probability of any event must be between 0 and 1, inclusively

$0 \leq P(A) \leq 1$ For any event A

- The sum of the probabilities of all mutually exclusive and collectively exhaustive events is 1

$$P(A) + P(B) + P(C) = 1$$

If A, B, and C are mutually exclusive and collectively exhaustive



General Addition Rule

General Addition Rule:

P(A or B) = P(A) + P(B) - P(A and B)

If A and B are mutually exclusive, then

P(A and B) = 0, so the rule can be simplified:

P(A or B) = P(A) + P(B)

For mutually exclusive events A and B

General Addition Rule Example

P(Jan. or Wed.) = P(Jan.) + P(Wed.) - P(Jan. and Wed.)

= 32/365 + 52/365 - $\boxed{\text{5/365}}$ = 79/365

Don't count the five Wednesdays in January twice!

	Jan.	Not Jan.	Total
Wed.	5	47	52
Not Wed.	27	286	313
Total	32	333	365

Computing Conditional Probabilities

- A conditional probability is the probability of one event, given that another event has occurred:

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

The conditional probability of A given that B has occurred

$$P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)}$$

The conditional probability of B given that A has occurred

Where P(A and B) = joint probability of A and B

P(A) = marginal or simple probability of A

P(B) = marginal or simple probability of B

Conditional Probability Example

- Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a GPS. 20% of the cars have both.

- What is the probability that a car has a GPS, given that it has AC ?

i.e., we want to find P(GPS | AC)

Conditional Probability Example

(continued)

- Of the cars on a used car lot, **70%** have air conditioning (AC) and **40%** have a GPS and **20%** of the cars have both.

	GPS	No GPS	Total
AC	0.2	0.5	0.7
No AC	0.2	0.1	0.3
Total	0.4	0.6	1.0

$$P(\text{GPS} \mid \text{AC}) = \frac{P(\text{GPS and AC})}{P(\text{AC})} = \frac{0.2}{0.7} = 0.2857$$

Conditional Probability Example

(continued)

- Given AC, we only consider the top row (70% of the cars). Of these, 20% have a GPS. 20% of 70% is about 28.57%.

	GPS	No GPS	Total
AC	0.2	0.5	0.7
No AC	0.2	0.1	0.3
Total	0.4	0.6	1.0

$$P(\text{GPS} \mid \text{AC}) = \frac{P(\text{GPS and AC})}{P(\text{AC})} = \frac{0.2}{0.7} = 0.2857$$

Independence

- Two events are independent if and only if:

$$\boxed{P(A|B) = P(A)}$$

- Events A and B are independent when the probability of one event is not affected by the fact that the other event has occurred

Multiplication Rules

- Multiplication rule for two events A and B:

$$\boxed{P(A \text{ and } B) = P(A|B)P(B)}$$

Note: If A and B are independent, then $P(A|B) = P(A)$ and the multiplication rule simplifies to

$$P(A \text{ and } B) = P(A)P(B)$$

Bayes Formula

Bayes' Theorem

- Bayes' Theorem is used to revise previously calculated probabilities based on new information.

- Developed by Thomas Bayes in the 18^{th} Century.

- It is an extension of conditional probability.

Example

There exists a test for a certain viral infection. It is 95% reliable for infected patients and 99% reliable for the healthy ones. That is, if a patient has the virus (event V), the test shows that (event S) with probability $P\{S \mid V\} = 0.95$, and if the patient does not have the virus, the test shows that with probability $P\{\text{not } S \mid \text{not } V\} = 0.99$. Suppose that 4% of all the patients are infected with the virus, $P\{V\} = 0.04$. When the test shows positive result, find the (conditional) probability that the patient is infected.

Bayes' Theorem

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

How to find P(A)?

Marginal Probability (The Law of Total Probability)

- Marginal probability for event A:

$$P(A) = P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_k)P(B_k)$$

  - Where B_1, B_2, …, B_k are k mutually exclusive and collectively exhaustive events

Bayes' Theorem

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_k)P(B_k)}$$

- where:

  B_i = i^{th} event of k mutually exclusive and collectively exhaustive events

  A = new event that might impact $P(B_i)$

Bayes' Theorem Example

- A drilling company has estimated a 40% chance of striking oil for their new well.
- A detailed test has been scheduled for more information. Historically, 60% of successful wells have had detailed tests, and 20% of unsuccessful wells have had detailed tests.
- Given that this well has been scheduled for a detailed test, what is the probability that the well will be successful?

Bayes' Theorem Example

(continued)

- Let S = successful well

  U = unsuccessful well
- P(S) = 0.4 , P(U) = 0.6 (prior probabilities)
- Define the detailed test event as D
- Conditional probabilities:

  P(D|S) = 0.6 P(D|U) = 0.2
- Goal is to find P(S|D)

Bayes' Theorem Example

(continued)

Apply Bayes' Theorem:

$$P(S|D) = \frac{P(D|S)P(S)}{P(D|S)P(S) + P(D|U)P(U)}$$

$$= \frac{(0.6)(0.4)}{(0.6)(0.4) + (0.2)(0.6)}$$

$$= \frac{0.24}{0.24 + 0.12} = 0.667$$

So the revised probability of success, given that this well has been scheduled for a detailed test, is 0.667

Bayes' Theorem Example

(continued)

- Given the detailed test, the revised probability of a successful well has risen to 0.667 from the original estimate of 0.4

Event	Prior Prob.	Conditional Prob.	Joint Prob.	Revised Prob.
S (successful)	0.4	0.6	(0.4)(0.6) = 0.24	0.24/0.36 = 0.667
U (unsuccessful)	0.6	0.2	(0.6)(0.2) = 0.12	0.12/0.36 = 0.333

Sum = 0.36

Exercises

- All athletes at the Olympic games are tested for performance-enhancing steroid drug use. The imperfect test gives positive results (indicating drug use) for 90% of all steroid-users but also (and incorrectly) for 2% of those who do not use steroids. Suppose that 5% of all registered athletes use steroids. If an athlete is tested positive, what is the probability that he/she uses steroids?

Chapter Summary

In this chapter we discussed:

- Basic probability concepts
  - Sample spaces and events, contingency tables, simple probability, and joint probability
- Basic probability rules
  - General addition rule, addition rule for mutually exclusive events, rule for collectively exhaustive events
- Conditional probability
  - Statistical independence, marginal probability, decision trees, and the multiplication rule
- Bayes Formula and the Law of Total Probability