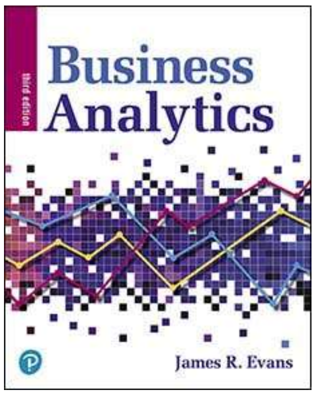
Business Analytics: Methods, Models, and Decisions

Third Edition



Chapter 5

Probability Distributions and Data Modeling

Part 1

 Slide - 1

Basic Concepts of Probability

- **Probability** is the likelihood that an outcome occurs. Probabilities are expressed as values between 0 and 1.
- An **experiment** is the process that results in an outcome.
- The **outcome** of an experiment is a result that we observe.
- The **sample space** is the collection of all possible outcomes of an experiment.

 Slide - 2

Combinations and Permutations

- Enumerating and counting the outcomes for an experiment can sometimes be difficult, particularly when the experiment consists of multiple steps.
- A bit of logic and visualization often helps.

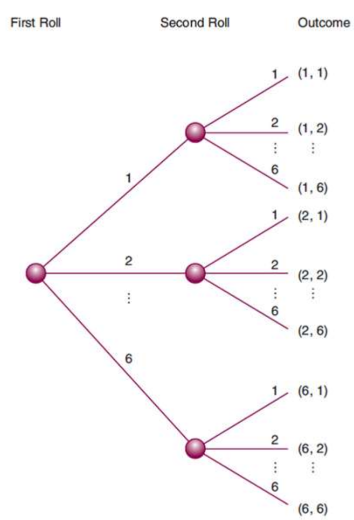
 Slide - 3

Example 5.1: Rolling Two Dice

- Tree diagram
- For an experiment with k steps, the number of outcomes is

$$n_1 \times n_2 \times \ldots \times n_k \qquad (5.1)$$

For rolling two dice, we have 6 outcomes for the first roll, and 6 outcomes for the second roll. The total number of outcomes is 6 x 6 = 36.



 Slide - 4

Example 5.2: Selecting *n* Objects from *N*

- In a group of five students, an instructor might wish to select three of them to make presentations.
- The first student selected can be student 1, 2, 3, 4, or 5. However, if student 1 is selected first, then the second student can be 2, 3, 4, or 5. Then, if the second student selected is student 4, the third student can be 2, 3, or 5.

Combinations

- When the order does not matter, we only want to count unique outcomes, which we call **combinations**.
- The number of combinations for selecting *n* objects from a set of *N* is

$$C(n, N) = \binom{N}{n} = \frac{N!}{n!(N-n)!} \qquad \textbf{(5.2)}$$

Example 5.3: Applying the Combinations Formula

- In Example 5.2, the number of ways of selecting three students from a group of five is

$$C(3, 5) = \binom{5}{3} = \frac{5!}{3!(5-3)!} = \frac{(5)(4)(3)(2)(1)}{3(2)(1) \times (2)(1)} = 10$$

Permutations

- If we want to select *n* objects from *N and* the order is important, then we call the outcomes **permutations**.
- The number of permutations of *n* objects selected from *N* is

$$P(n, N) = n!\binom{N}{n} = \frac{N!}{(N-n)!} \qquad \textbf{(5.3)}$$

Example 5.4: Applying the Permutations Formula

- Suppose we want to count the number of ways of selecting three students from a group of five where the order is important (for instance, knowing which student presents first, second, and third).

$$P(3, 5) = 3!\binom{5}{3} = \frac{5!}{(5-3)!} = \frac{(5)(4)(3)(2)(1)}{(2)(1)} = 60$$

Probability Definitions

Probabilities may be defined from one of three perspectives:

- *Classical definition*: probabilities can be deduced from theoretical arguments
- *Relative frequency definition*: probabilities are based on empirical data
- *Subjective definition*: probabilities are based on judgment and experience

Example 5.5 Classical Definition of Probability

Roll 2 dice

- 36 possible rolls $(1,1),(1,2),\ldots(6,5),(6,6)$
- Probability = number of ways of rolling a number divided by 36;

  e.g., probability of a 3 is 2/36 = 1/18

Suppose two consumers try a new product.

- Four outcomes:
  1. like, like
  2. like, dislike
  3. dislike, like
  4. dislike, dislike
- Probability at least one dislikes product $= \frac{3}{4}$

Die Sum	Frequency
2	1
3	2
4	3
5	4
6	5
7	6
8	5
9	4
10	3
11	2
12	1
Sum	36

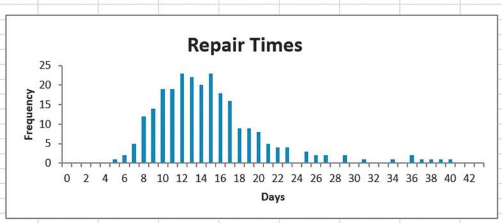
Pearson Slide - 11

Example 5.6: Relative Frequency Definition of Probability

- Use relative frequencies as probabilities.
- Probability a computer is repaired in 10 days = 0.076

Computer Repair Times

Days	Frequency	Relative Frequency	Cumulative Percentage
0	0	0.000	0.0%
1	0	0.000	0.0%
2	0	0.000	0.0%
3	0	0.000	0.0%
4	0	0.000	0.0%
5	1	0.004	0.4%
6	2	0.008	1.2%
7	5	0.020	3.2%
8	12	0.048	8.0%
9	14	0.056	13.6%
10	19	0.076	21.2%
11	19	0.076	28.8%
12	23	0.092	38.0%
13	22	0.088	46.8%



Pearson Slide - 12

Probability Rules and Formulas

- Label the *n* outcomes in a sample space as $O_1, O_2, \ldots, O_n$, where O_i represents the i^{th} outcome in the sample space. Let $P(O_i)$ be the probability associated with the outcome O_i.
- The probability associated with any outcome must be between 0 and 1.

$$0 \leq P(O_i) \leq 1 \text{ for each outcome } O_i \qquad \textbf{(5.4)}$$

- The sum of the probabilities over all possible outcomes must be equal to 1.

$$P(O_1) + P(O_2) + \cdots + P(O_n) = 1 \qquad \textbf{(5.5)}$$

 Slide - 13

Probabilities Associated with Events

- An **event** is a collection of one or more outcomes from a sample space.
- **Rule 1**. The probability of any event is the sum of the probabilities of the outcomes that comprise that event.

 Slide - 14

Example 5.7: Computing the Probability of an Event

Consider the events:

- Rolling 7 or 11 on two dice

  Probability $= \frac{6}{36} + \frac{2}{36} = \frac{8}{36}.$

- Repair a computer in 7 days or less

  Probability =

  $= O_1 + O_2 + O_3 + O_4 + O_5 + O_6 + O_7$

  $= 0.032$

Die Sum	Frequency
2	1
3	2
4	3
5	4
6	5
7	6
8	5
9	4
10	3
11	2
12	1
Sum	36

Computer Repair Times

Days	Frequency	Relative Frequency
0	0	0.000
1	0	0.000
2	0	0.000
3	0	0.000
4	0	0.000
5	1	0.004
6	2	0.008
7	5	0.020

Pearson Slide - 15

Complement of an Event

- If *A* is any event, the complement of *A*, denoted A^c, consists of all outcomes in the sample space not in *A*.
- **Rule 2**. The probability of the complement of any event *A* is $P(A^c) = 1 - P(A).$

Pearson Slide - 16

Example 5.8: Computing the Probability of the Complement of an Event

Dice Example:

- $A = \{7, 11\}$

  $P(A) = \dfrac{8}{36}$

- $A^c = \{2,3,4,5,6,8,9,10,12\}$
- Using Rule 2:

  $P(A^c) = 1 - 8/36 = 28/36$

Die Sum	Frequency
2	1
3	2
4	3
5	4
6	5
7	6
8	5
9	4
10	3
11	2
12	1
Sum	36

 Slide - 17

Union of Events

- The union of two events contains all outcomes that belong to either of the two events.
  - If *A* and *B* are two events, the probability that some outcome in either *A* or *B* (that is, the union of *A* and *B*) occurs is denoted as $P(A \text{ or } B)$.
- Two events are mutually exclusive if they have no outcomes in common.
- **Rule 3**. If events *A* and *B* are mutually exclusive, then $P(A \text{ or } B) = P(A) + P(B)$.

 Slide - 18

Example 5.9: Computing the Probability of Mutually Exclusive Events

Dice Example:

- $A = \{7, 11\} : P(A) = \dfrac{8}{36}$
- $B = \{2, 3, 12\} : P(B) = 4/36$
- $P(A \text{ or } B) = \text{Union of events A and B}$
  $$= P(A) + P(B)$$
  $$= \frac{8}{36} + \frac{4}{36} = \frac{12}{36}$$

Die Sum	Frequency
2	1
3	2
4	3
5	4
6	5
7	6
8	5
9	4
10	3
11	2
12	1
Sum	36

 Slide - 19

Non-Mutually Exclusive Events

- The notation (A and B) represents the intersection of events A and B – that is, all outcomes belonging to both A and B.
- **Rule 4**. If two events A and B are not mutually exclusive, then $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.

 Slide - 20

Example 5.10: Computing the Probability of Non-Mutually Exclusive Events

Dice Example:

- $A = \{2,3,12\} : P(A) = \frac{4}{36}$
- $B = \{\text{even number}\} : P(B) = \frac{18}{36}$
- $(A \text{ and } B) = \{2,\ 12\} : P(A \text{ and } B) = \frac{2}{36}$
- $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

$$= \frac{4}{36} + \frac{18}{36} - \frac{2}{36}$$

$$= \frac{20}{36}$$

Die Sum	Frequency
2	1
3	2
4	3
5	4
6	5
7	6
8	5
9	4
10	3
11	2
12	1
Sum	36

Joint and Marginal Probability

- The probability of the intersection of two events is called a **joint probability.**
- The probability of an event, irrespective of the outcome of the other joint event, is called a **marginal probability.**

Conditional Probability

- **Conditional probability** is the probability of occurrence of one event *A*, given that another event *B* is known to be true or has already occurred.

Pearson Slide - 23

Conditional Probability Formula

- The conditional probability of an event *A* given that event *B* is known to have occurred is

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} \qquad \textbf{(5.6)}$$

- We read the notation $P(A \mid B)$ as "the probability of *A* given *B*."

Pearson Slide - 24

Example 5.14: Using the Conditional Probability Formula

Joint Probability Table	Brand 1	Brand 2	Brand 3	Grand Total
Female	0.09	0.06	0.22	0.37
Male	0.25	0.17	0.21	0.63
Grand Total	**0.34**	**0.23**	**0.43**	**1**

- $P(B_1 \mid M) = \dfrac{P(B_1 \text{ and } M)}{P(M)} = \dfrac{(0.25)}{(0.63)} = 0.397$
- $P(B_1 \mid F) = \dfrac{P(B_1 \text{ and } F)}{P(F)} = \dfrac{(0.09)}{(0.37)} = 0.243$
- Summary of conditional probabilities:

***P*(Brand|Gender)**	**Brand 1**	**Brand 2**	**Brand 3**
Male	0.397	0.270	0.333
Female	0.243	0.162	0.595

- Applications in marketing and advertising.

Independent Events

- Two events *A* and *B* are **independent** if $P(A \mid B) = P(A)$.
- *Energy Drink Survey* example: the probability of preferring a brand depends on gender.
- Thus, we may say that brand preference and gender are not independent.

Example 5.16: Determining If Two Events Are Independent

- Are Gender and Brand Preference Independent?
- $P(B_1) = 0.34$

Joint Probability Table	Brand 1	Brand 2	Brand 3	Grand Total
Female	0.09	0.06	0.22	0.37
Male	0.25	0.17	0.21	0.63
Grand Total	0.34	0.23	0.43	1

- $P(B_1 \mid M) = 0.397$

P(Brand|Gender)	Brand 1	Brand 2	Brand 3
Male	0.397	0.270	0.333
Female	0.243	0.162	0.595

- Because $0.397 \neq 0.34$, Gender and Brand Preference are not independent.

Random Variables

- A **random variable** is a numerical description of the outcome of an experiment.
- A **discrete random variable** is one for which the number of possible outcomes can be counted.
- A **continuous random variable** has outcomes over one or more continuous intervals of real numbers.

Example 5.18: Discrete and Continuous Random Variables

Examples of discrete random variables:
- outcomes of dice rolls
- whether a customer likes or dislikes a product
- number of hits on a Web site link today

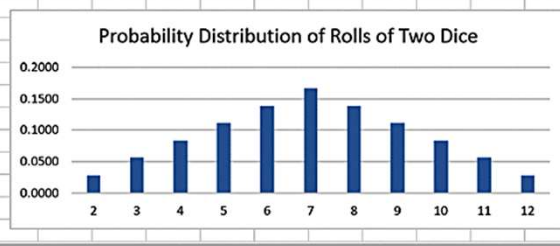
Examples of continuous random variables:
- weekly change in DJIA
- daily temperature
- time between machine failures

Probability Distributions

- A **probability distribution** is a characterization of the possible values that a random variable may assume along with the probability of assuming these values.
- We may develop a probability distribution using any one of the three perspectives of probability: classical, relative frequency, and subjective.

Example 5.19: Probability Distribution of Dice Rolls

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Dice Roll Probabilities												
2													
3	Outcome	2	3	4	5	6	7	8	9	10	11	12	Total
4	Number of Ways	1	2	3	4	5	6	5	4	3	2	1	36
5	Probability	0.0278	0.0556	0.0833	0.1111	0.1389	0.1667	0.1389	0.1111	0.0833	0.0556	0.0278	1



Empirical Probability Distributions

- We can calculate the relative frequencies from a sample of empirical data to develop a probability distribution. Because this is based on sample data, we usually call this an **empirical probability distribution**.
- An empirical probability distribution is an approximation of the probability distribution of the associated random variable, whereas the probability distribution of a random variable, such as one derived from counting arguments, is a theoretical model of the random variable.

Empirical Probability Distribution Example

Computer Repair Times

Days	Frequency	Relative Frequency	Cumulative Percentage
0	0	0.000	0.0%
1	0	0.000	0.0%
2	0	0.000	0.0%
3	0	0.000	0.0%
4	0	0.000	0.0%
5	1	0.004	0.4%
6	2	0.008	1.2%
7	5	0.020	3.2%
8	12	0.048	8.0%
9	14	0.056	13.6%
10	19	0.076	21.2%
11	19	0.076	28.8%
12	23	0.092	38.0%
13	22	0.088	46.8%

 Slide - 33

Discrete Probability Distributions

- For a discrete random variable X, the probability distribution of the discrete outcomes is called a **probability mass function** and is denoted by a mathematical function, $f(x)$.
  - The symbol x_i represents the i^{th} value of the random variable X and $f(x_i)$ is the probability.
- Properties:
  - the probability of each outcome must be between 0 and 1
  - the sum of all probabilities must add to 1

$$0 \le f(x_i) \le 1 \quad \text{for all } i \qquad \textbf{(5.10)}$$

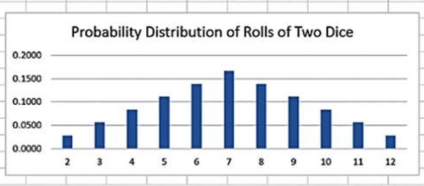
$$\sum_i f(x_i) = 1 \qquad \textbf{(5.11)}$$

 Slide - 34

Example 5.21: Probability Mass Function for Rolling Two Dice

- x_i = values of the random variable *X*, which represents sum of the rolls of two dice
  - $x_1 = 2, x_2 = 3, \ldots, x_{10} = 11, x_{11} = 12$
  - $f(x_1) = \frac{1}{36} = 0.0278; f(x_2) = \frac{2}{36} = 0.0556$, etc.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Dice Roll Probabilities**												
2													
3	**Outcome**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Total**
4	**Number of Ways**	1	2	3	4	5	6	5	4	3	2	1	36
5	**Probability**	0.0278	0.0556	0.0833	0.1111	0.1389	0.1667	0.1389	0.1111	0.0833	0.0556	0.0278	1



Cumulative Distribution Function

- A **cumulative distribution function**, $F(x)$, specifies the probability that the random variable *X* assumes a value less than or equal to a specified value, *x*; that is,

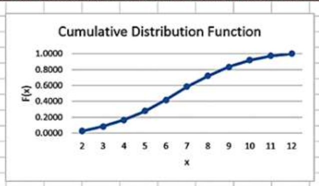
$$F(x) = P(X \le x)$$

Example 5.22: Using the Cumulative Distribution Function

- Probability of rolling a 6 or less $= F(6) = 0.1667$
- Probability of rolling between 4 and 8:

$$= P(4 \le X \le 8) = P(3 < X \le 8) = P(X \le 8) - P(X \le 3)$$
$$= 0.7222 - 0.0833 = 0.6389$$

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Dice Roll Probabilities												
2													
3	Outcome	2	3	4	5	6	7	8	9	10	11	12	Total
4	Number of Ways	1	2	3	4	5	6	5	4	3	2	1	36
5	Probability	0.0278	0.0556	0.0833	0.1111	0.1389	0.1667	0.1389	0.1111	0.0833	0.0556	0.0278	1
6	Cumulative Probability	0.0278	0.0833	0.1667	0.2778	0.4167	0.5833	0.7222	0.8333	0.9167	0.9722	1.0000	



 Slide - 37

Expected Value of a Discrete Random Variable

- The expected value of a random variable corresponds to the notion of the mean, or average, for a sample.
- For a discrete random variable *X*, the expected value, denoted $E[X]$, is the weighted average of all possible outcomes, where the weights are the probabilities:

$$E[X] = \sum_{i=1}^{\infty} x_i f(x_i) \qquad \textbf{(5.12)}$$

 Slide - 38

Example 5.23: Computing the Expected Value

- Rolling two dice
  - $E[X] = 2(0.0278) + 3(0.0556) + 4(0.0833) + 5(0.1111) + 6(0.1389) + 7(0.1667) + 8(0.1389) + 9(0.1111) + 10(0.0833) + 11(0.0556) + 12(0.0278) = 7$

	A	B	C
1	Expected Value Calculations		
2			
3	Outcome, x	Probability, f(x)	x*f(x)
4	2	0.0278	0.0556
5	3	0.0556	0.1667
6	4	0.0833	0.3333
7	5	0.1111	0.5556
8	6	0.1389	0.8333
9	7	0.1667	1.1667
10	8	0.1389	1.1111
11	9	0.1111	1.0000
12	10	0.0833	0.8333
13	11	0.0556	0.6111
14	12	0.0278	0.3333
15		Expected value	7.0000

 Slide - 39

Example 5.24: Expected Value on Television

The Apprentice

- Teams were required to select an artist (mainstream or avant-garde) and sell their art for the most money possible. A simple expected value calculation would have easily predicted the winner.

Deal or No Deal

- Contestant had 5 briefcases left with \$100, \$400, \$1000, \$50,000 or \$300,000 in them.
- Expected value of briefcases is \$70,300.
- Banker offered contestant \$80,000 to quit, which was higher than the expected value. The probability of choosing the \$300,000 briefcase was only 0.2, so the decision should have been easy to make.

 Slide - 40

Expected Value and Decision Making

- The expected value is a "long-run average" and is appropriate for decisions that occur on a repeated basis.
- For one-time decisions, however, you need to consider the downside risk and the upside potential of the decision.

 Slide - 41

Example 5.25: Expected Value of a Charitable Raffle

- Cost of raffle ticket is $50.
- 1000 raffle tickets are sold.
- Winning prize is $25,000.
- $E[X] = -\$25$
- If you played this game repeatedly over the long run, you would lose an average of $25.00 each time you play.
- However, for any one game, you would either lose $50 or win $24,950.
  - Is the risk of losing $50 worth the potential of winning $24,950?

x	$f(x)$
−$50	0.999
$24,950	0.001

 Slide - 42

Example 5.26: Airline Revenue Management

- Full and discount airfares are available for a flight.
- Full-fare ticket costs $560.
- Discount ticket costs $400.
- *X* = ticket price paid
- *p* = 0.75 (the probability of selling a full-fare ticket)
- $E[X] = 0.75(\$560) + 0.25(0) = \$420$
- The airline should not discount full-fare tickets because the expected value of a full-fare ticket is greater than the cost of a discount ticket.
- Break-even point: $\$400 = p(\$560)$ or $p = 0.714$

 Slide - 43

Variance of a Discrete Random Variable

- The variance, $\mathrm{Var}[X]$, of a discrete random variable *X* is a weighted average of the squared deviations from the expected value:

$$\mathrm{Var}\,[X] = \sum_{i=1}^{\infty} (x_i - E[X])^2 f(x_i) \qquad \textbf{(5.13)}$$

 Slide - 44

Example 5.27: Computing the Variance of a Random Variable

- Rolling two dice

	A	B	C	D	E	F
1	Variance Calculations					
2						
3	Outcome, x	Probability, f(x)	x*f(x)	(x - E[X])	(x - E[X])^2	(x - E[X])^2*f(x)
4	2	0.0278	0.0556	-5.0000	25.0000	0.6944
5	3	0.0556	0.1667	-4.0000	16.0000	0.8889
6	4	0.0833	0.3333	-3.0000	9.0000	0.7500
7	5	0.1111	0.5556	-2.0000	4.0000	0.4444
8	6	0.1389	0.8333	-1.0000	1.0000	0.1389
9	7	0.1667	1.1667	0.0000	0.0000	0.0000
10	8	0.1389	1.1111	1.0000	1.0000	0.1389
11	9	0.1111	1.0000	2.0000	4.0000	0.4444
12	10	0.0833	0.8333	3.0000	9.0000	0.7500
13	11	0.0556	0.6111	4.0000	16.0000	0.8889
14	12	0.0278	0.3333	5.0000	25.0000	0.6944
15		Expected value	7.0000		Variance	5.8333

