Class 7

Daniel B. Rowe, Ph.D.

Department of Mathematics, Statistics, and Computer Science
Agenda:

Recap Chapter 4.3

Lecture Chapter 4.4 - 4.6

Review 3.1- 4.2 for Exam 2.
Recap Chapter 4.3
4: Probability

4.3 Conditional Probability

Example: Draw card from deck. Let $A = \text{red card}$, $B = \text{heart}$.

$P(A) = ?$ vs $P(A|B) = ?$

$P(A) = 1/2$ vs. $P(A|B) = 1$

Conditional Probability Formula:

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

$P(A \text{ and } B) = 1/4$ \hspace{1cm} $P(B) = 1/4$

$$P(A | B) = 1$$

Figure from Johnson & Kuby, 2008.

Rowe, D.B.
Lecture Chapter 4.4 - 4.6
Chapter 4: Probability continued

Daniel B. Rowe, Ph.D.

Department of Mathematics, Statistics, and Computer Science
4: Probability
4.4 Rules of Probability - Probability of “Not $A$”

**Complimentary Events:** The compliment of $A$, $\bar{A}$ is the set of all sample points in the sample space that does not belong to event $A$.

**Compliment Rule:**
In words: probability of $A$ compliment $= \text{one} - \text{probability of } A$

In algebra: $P(\bar{A}) = 1 - P(A)$

From $P(A) + P(\bar{A}) = 1$  \hspace{1cm} (4.3)
4: Probability
4.4 Rules of Probability - Probability of “Not $A$”

Compliment: $S = \{A, \overline{A}\}$

Venn Diagram:

\[
P(\overline{A}) = 1 - P(A)
\]
4: Probability
4.4 Rules of Probability - Probability of “A or B”

**General Addition Rule**

Let $A$ and $B$ be two events defined in the sample space, $S$.

In words: probability of $A$ or $B = \text{probability of } A + \text{probability of } B - \text{probability of } A \text{ and } B$

In algebra: $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

(4.4)
4: Probability
4.4 Rules of Probability - Probability of “A or B”

Union: \( A \) or \( B \)

Venn Diagram:

\[
P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)
\]
4: Probability
4.4 Rules of Probability - Probability of “A and B”

General Multiplication Rule
Let A and B be two events defined in the sample space, S.

In words: probability of A and B = probability of A
× probability of B, knowing A

In algebra: \( P(A \text{ and } B) = P(A) \cdot P(B \mid A) \) (4.5)
4: Probability
4.4 Rules of Probability - Probability of “A and B”

Event Intersection: $A$ and $B$

Venn Diagram:

$$P(A \text{ and } B) = P(A) \cdot P(B \mid A)$$
Conditional Probability: Probability of event $A$ given that event $B$ has occurred is

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

the “$|$” is spoken as “given”
4: Probability
4.4 Rules of Probability

Union Example \((A \text{ or } B)\): Rolling a single die. 
\(A=\text{event } #1,2,3. \ B=\text{event odd number.}\)
4: Probability
4.4 Rules of Probability

**Union Example** \((A \text{ or } B)\): Rolling a single die.

\(A=\text{event } #1,2,3.\) \(B=\text{event odd number.}\)

\(A=\{1,2,3\}\)

\(S\)
4: Probability
4.4 Rules of Probability

**Union Example (A or B):** Rolling a single die. 
A = event #1,2,3. B = event odd number.

B = \{1,3,5\}
4: Probability
4.4 Rules of Probability

Union Example \((A \text{ or } B)\): Rolling a single die.

\(A=\text{event #1,2,3. }B=\text{event odd number.}\)

\((A \text{ or } B)\)

\[
P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)
\]

\(S\)
4: Probability
4.4 Rules of Probability

Intersection Example ($A$ and $B$): Rolling a single die. $A$=event #1,2,3. $B$=event odd number.

$S$

\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
\end{array}
4: Probability
4.4 Rules of Probability

**Intersection Example (A and B):** Rolling a single die.

A = event #1,2,3. B = event odd number.

\[ A = \{1,2,3\} \]

S

---

Rowe, D.B.
Intersection Example \((A \text{ and } B)\): Rolling a single die. 
\(A=\) event #1,2,3. \(B=\) event odd number. 
\(B=\{1,3,5\}\)
4: Probability
4.4 Rules of Probability

Intersection Example \((A \text{ and } B)\): Rolling a single die.

\(A=\text{event #1,2,3. } B=\text{event odd number.}\)

\((A \text{ and } B) =\{1,3\}\)

\[P(A \text{ and } B) = P(A) \cdot P(B \mid A)\]
4: Probability
4.5 Mutually Exclusive Events

Mutually exclusive events:
Events that share no common elements

In algebra: \[ P(A \text{ and } B) = 0 \]

In words:
1. If one event has occurred, the other cannot.
2. None of the elements in one is in other.
3. In Venn diagrams, no intersection.
4. Intersection of events has a probability of zero.
4: Probability
4.5 Mutually Exclusive Events

**Mutually Exclusive:** \( P(A \text{ and } B) = 0 \)

**Venn Diagram:**

\[ S \]

\[ A \]

\[ B \]
4: Probability
4.5 Mutually Exclusive Events

**Mutually Exclusive Example:** Rolling a single die. 
$A =$ event #1,2. $B =$ event #5,6.

$A \text{ and } B = \emptyset$

$P(A \text{ and } B) = 0$

null set
4: Probability
4.6 Independent Events

**Independent events**: Two events are independent if the occurrence or nonoccurrence of one gives us no information about the likeliness of occurrence for the other.

In algebra: \[ P(A) = P(A \mid B) = P(A \mid \text{not } B) \]

In words:
1. Prob of \( A \) unaffected by knowledge that \( B \) has occurred, not occurred, or no knowledge.
2. …
3. …
Two events $A$ and $B$ are independent if the probability of one is not “influenced” by the occurrence or nonoccurrence of the other.

Two Events $A$ and $B$ are independent if:

1. $P(A) = P(A \mid B)$
2. $P(B) = P(B \mid A)$
3. $P(A \text{ and } B) = P(A) \cdot P(B)$

Examples:
Dependent events: Events that are not independent. That is, occurrence of one event does have an effect on the probability of occurrence of the other event.

In algebra: $P(A) \neq P(A \mid B)$
4: Probability
4.6 Independent Events - Special multiplication rule

Special multiplication rule:
Let $A$ and $B$ be two independent events defined in a sample space $S$.

In words: The probability of $A$ and $B = \text{probability of } A \times \text{probability of } B$

In algebra: $P(A \text{ and } B) = P(A) \cdot P(B)$ \hfill (4.7)

More generally

$P(A \text{ and } B \text{ and } C \text{ and } D \text{ and } E) = P(A) \cdot P(B) \cdot P(C) \cdot P(D) \cdot P(E)$
4: Probability
4.7 Are Mutually Exclusive and Independence Related?

Read this section on your own.
4: Probability

Questions?

Homework: Chapter 4 # 57, 61, 63, 67, 87, 91, 95, 103, 105, 111
Review Chapters 3.1 – 4.2
(Exam 2 Chapters)

Just the highlights!
3: Descriptive Analysis and Bivariate Data

3.2 Bivariate Data: two qualitative

**Cross-tabulation tables or contingency tables**

![Table 3.1](image)

**TABLE 3.3**

Cross-Tabulation of Gender and Major (frequencies)

<table>
<thead>
<tr>
<th>Gender</th>
<th>LA</th>
<th>BA</th>
<th>T</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>F</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Col. Total</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>30</td>
</tr>
</tbody>
</table>

LA = liberal arts  
BA = business admin  
T = technology

Figure from Johnson & Kuby, 2008.

Rowe, D.B.
3: Descriptive Analysis and Bivariate Data

3.2 Bivariate Data: one qualitative and one quantitative

Example:

<table>
<thead>
<tr>
<th>Table 3.7</th>
<th>Stopping Distances (in feet) for Three Tread Designs [TA03-07]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design A</td>
<td>Design B</td>
</tr>
<tr>
<td>(n = 6)</td>
<td>(n = 6)</td>
</tr>
<tr>
<td>37 36 38</td>
<td>33 35 38</td>
</tr>
<tr>
<td>34 40 32</td>
<td>34 42 34</td>
</tr>
</tbody>
</table>

Figures from Johnson & Kuby, 2008.

Vertical box-and-whiskers

© 2007 Thomson Higher Education

Rowe, D.B.
3: Descriptive Analysis and Bivariate Data

3.2 Bivariate Data: two quantitative, Scatter Diagram

**TABLE 3.10**

<table>
<thead>
<tr>
<th>Student</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Push-ups, $x$</td>
<td>27</td>
<td>22</td>
<td>15</td>
<td>35</td>
<td>30</td>
<td>52</td>
<td>35</td>
<td>55</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Sit-ups, $y$</td>
<td>30</td>
<td>26</td>
<td>25</td>
<td>42</td>
<td>38</td>
<td>40</td>
<td>32</td>
<td>54</td>
<td>50</td>
<td>43</td>
</tr>
</tbody>
</table>

*(x,y)* ordered pairs.

**Input variable** called independent variable, **Output variable** called dependent variable,

**Scatter Diagram:** A plot of all the ordered pairs of bivariate data. The input variable, $x$, the horizontal axis and the output variable, $y$, is plotted on the vertical axis.

Figures from Johnson & Kuby, 2008.

Rowe, D.B.
3: Descriptive Analysis and Bivariate Data
3.3 Linear Correlation

Linear Correlation, \( r \), is a measure of the strength of a linear relationship between two variables \( x \) and \( y \).

**positive relationship:**
as \( x \) increases so does \( y \)

**negative relationship:**
as \( x \) increases \( y \) decreases

\(-1 \leq r \leq 1\)

No correlation \( r \approx 0 \)
Positive \( r \approx 0.5 \)
High positive \( r \approx 0.8 \)
Negative \( r \approx -0.5 \)
High negative \( r \approx -0.8 \)

Figure from Johnson & Kuby, 2008.
### 3: Descriptive Analysis and Bivariate Data

#### 3.3 Linear Correlation

**Example:**

<table>
<thead>
<tr>
<th>Push-ups, $x$</th>
<th>27</th>
<th>22</th>
<th>15</th>
<th>35</th>
<th>30</th>
<th>52</th>
<th>35</th>
<th>55</th>
<th>40</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sit-ups, $y$</td>
<td>30</td>
<td>26</td>
<td>25</td>
<td>42</td>
<td>38</td>
<td>40</td>
<td>32</td>
<td>54</td>
<td>50</td>
<td>43</td>
</tr>
</tbody>
</table>

\[
SS(x) = \sum_{i=1}^{n} x_i^2 - \frac{1}{n} \left( \sum_{i=1}^{n} x_i \right)^2 = 13717 - \frac{(351)^2}{10} = 1396.9
\]

\[
SS(y) = \sum_{i=1}^{n} y_i^2 - \frac{1}{n} \left( \sum_{i=1}^{n} y_i \right)^2 = 15298 - \frac{(380)^2}{10} = 858.0
\]

\[
SS(xy) = \sum_{i=1}^{n} x_i y_i - \frac{1}{n} \left( \sum_{i=1}^{n} x_i \right) \left( \sum_{i=1}^{n} y_i \right) = 14257 - \frac{(351)(380)}{10} = 919.0
\]

\[
r = \frac{SS(xy)}{\sqrt{SS(x)SS(y)}} = \frac{919.0}{\sqrt{(1396.9)(858.0)}} = 0.84
\]

Questions?
3: Descriptive Analysis and Bivariate Data

3.4 Linear Regression

We try different lines until we find the “best” one, \( \hat{y} = b_0 + b_1 x \)

\( b_0 \) is estimated y-intercept
\( b_1 \) is estimated slope.

Move line until sum of the squared residuals is a minimum.

\[ \sum_{i=1}^{n} \varepsilon_i^2 \]
3: Descriptive Analysis and Bivariate Data

3.4 Linear Regression

(x,y) pairs: (1,1),(3,2),(2,3),(4,4)

Plotted points.

The line goes through $(\bar{x}, \bar{y})$.

The slope is $b_1=0.8$.

The y-intercept $b_0=0.5$.

Two points (2.5,2.5) and (0,5).

\[
SS(xy) = \frac{1}{n} \left( \sum_{i=1}^{n} x_i y_i \right) - \frac{1}{n} \left( \sum_{i=1}^{n} x_i \right) \left( \sum_{i=1}^{n} y_i \right)
\]

\[
SS(x) = \frac{1}{n} \left( \sum_{i=1}^{n} x_i^2 \right) - \frac{1}{n} \left( \sum_{i=1}^{n} x_i \right)^2
\]

\[
b_1 = \frac{SS(xy)}{SS(x)}
\]

\[
b_0 = \bar{y} - b_1 \bar{x}
\]
3: Descriptive Analysis and Bivariate Data

3.4 Linear Regression

Example: Our data!

Height vs. Weight

\( b_1 = \frac{SS(xy)}{SS(x)} = \frac{8088.8}{1585.3} = 5.1 \)

\( b_0 = (148.1) - (5.1)(66.9) = -193.1 \)

units of lbs/in

point-slope formula

\( \bar{x} \)

\( \bar{y} \)

\( (\bar{x}, \bar{y}) \)
4: Probability
4.2 Probability of Events

An **experiment** is a process by which a measurement is taken or observations is made. i.e. flip coin or roll die

An **outcome** is the result of an experiment. i.e. Heads, or 3

**Sample space** is a listing of possible outcomes. i.e. \( S=\{H,T\} \)

An **event** is an outcome or a combination of outcomes. i.e. even number when rolling a die
4: Probability

4.2 Probability of Events

Property 1: \[ 0 \leq P(A_i) \leq 1 \]

Property 2: \[ \sum_{i=1}^{n} P(A_i) = 1 \]

Approaches to probability.

(1) Empirical (AKA experimental)

empirical probability of \( A \) = \[ \frac{\text{number of times } A \text{ occurred}}{\text{number of trials}} \]

(2) Theoretical (AKA classical or equally likely)

theoretical probability of \( A \) = \[ \frac{\text{number of times } A \text{ occurs in sample space}}{\text{number of elements in the sample space}} \]
4: Probability - Empirical

4.2 Probability of Events – Law of large numbers

Had computer flip a single coin 1000 times.

Flip # on $x$ axis

$P'(H)$ on $y$ axis.

This shows convergence to true value of 1/2.
4: Probability - Theoretical
4.2 Probability of Events
So let's flip a coin twice.

Can flip three times.

\[
S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}
\]

\[
P(HHH) = \frac{\text{# times } HHH \text{ occurs in } S}{\text{# elements in } S}
\]