STA 291
Lecture 8

• Probability
  – Probability Rules
  – Joint and Marginal Probability
Union and Intersection

• Let $A$ and $B$ denote two events.
• The union of two events:
  
  \[ A \cup B \]

• The intersection of two events:
  
  \[ A \cap B \]
Complement

• Let $A$ denote an event.
• The complement of an event $A$: $A^C$

Law of Complements:

$$P(A) = 1 - P(A^C)$$
Additive Law of Probability

Let A and B be two events in a sample space S. The probability of the union of A and B is

\[ P(A \cup B) = P(A) + P(B) - P(A \cap B). \]

Let A and B be two events in a sample space S. The probability of the union of two disjoint (mutually exclusive) events A and B is

\[ P(A \cup B) = P(A) + P(B). \]
**Using Additive Law of Probability**

**Example:** At State U, all first-year students must take chemistry and math. Suppose 15% fail chemistry, 12% fail math, and 5% fail both. Suppose a first-year student is selected at random. What is the probability that student selected failed at least one of the courses? What is the probability that student pass both?
Disjoint Events

• Let $A$ and $B$ denote two events.
• Disjoint (mutually exclusive) events:

$$A \cap B = \emptyset$$

• No overlap
Probability tables

- Simple table: One row of outcomes, one row of corresponding probabilities.

- $R \times C$ probability tables: when the outcomes are classified by two features
• Gender and support President Obama?

• Smoker? And Lung disease?

• Age group and support Obama?
Example: Smoking and Lung Disease

<table>
<thead>
<tr>
<th></th>
<th>Lung Disease</th>
<th>No Lung Disease</th>
<th>Marginal (smoke status)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoker</td>
<td>0.12</td>
<td>0.19</td>
<td>0.31</td>
</tr>
<tr>
<td>Nonsmoker</td>
<td>0.03</td>
<td>0.66</td>
<td>0.69</td>
</tr>
<tr>
<td>Marginal (disease status)</td>
<td>0.15</td>
<td>0.85</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Frequency table and probability table

<table>
<thead>
<tr>
<th></th>
<th>Lung Disease</th>
<th>No Lung Disease</th>
<th>(total) Marginal (smoke status)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoker</td>
<td>120</td>
<td>190</td>
<td></td>
</tr>
<tr>
<td>Nonsmoker</td>
<td>30</td>
<td>660</td>
<td></td>
</tr>
<tr>
<td>(total) Marginal (disease status)</td>
<td></td>
<td></td>
<td>1000</td>
</tr>
</tbody>
</table>
• Equivalent to a table with 4 entries:

<table>
<thead>
<tr>
<th>Category</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoker &amp; Lung Disease</td>
<td>0.12</td>
</tr>
<tr>
<td>Smoker &amp; Not Lung Disease</td>
<td>0.19</td>
</tr>
<tr>
<td>Nonsmoker &amp; Lung Disease</td>
<td>0.03</td>
</tr>
<tr>
<td>Nonsmoker &amp; Not Lung Disease</td>
<td>0.66</td>
</tr>
</tbody>
</table>

But the R x C table reads much better
• From the R x C table we can get a table for smoker status alone, or disease status alone.

• Those are called marginal probabilities
It’s a one way street

• Given the joint probability table, we can figure out the marginal probability

• Given the marginal, we may not determine the joint: there can be several different joint tables that lead to identical marginal.
Example: Smoking and Lung Disease

<table>
<thead>
<tr>
<th></th>
<th>Lung Disease</th>
<th>Not Lung Disease</th>
<th>Marginal (smoke status)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoker</td>
<td>0.02</td>
<td>0.29</td>
<td>0.31</td>
</tr>
<tr>
<td>Nonsmoker</td>
<td>0.13</td>
<td>0.56</td>
<td>0.69</td>
</tr>
<tr>
<td>Marginal (disease status)</td>
<td>0.15</td>
<td>0.85</td>
<td></td>
</tr>
</tbody>
</table>

Same marginal, different joint.
Using the table

• $P(\text{smoker and lung disease}) = 0.02$

• $P(\text{smoker or lung disease}) = 0.44$

(either by looking at the table
Or using the additive rule for probability)
Independence of events

• May not always hold.

• If and when it hold: With independence, one way street becomes two way street.

• Smoking and lung disease are obviously not independent in reality.
Independence

- If events $A$ and $B$ are independent, then the events $A$ and $B$ have no influence on each other.
- So, the probability of $A$ is unaffected by whether $B$ has occurred.
Multiplication rule of probability

If $A$ and $B$ are two independent events, then

$$P(A \cap B) = P(A)P(B).$$

• i.e. joint prob. = product of two marginal prob.
Conditional Probability

\[ P(A|B) = \frac{P(A \cap B)}{P(B)} \], provided \( P(B) \neq 0 \)

- Note: \( P(A/B) \) is read as “the probability that \( A \) occurs given that \( B \) has occurred.”
Independent Events

Multiplication Rule for Independent Events: Let A and B be two independent events, then

\[ P(A \cap B) = P(A)P(B). \]

Mathematically, if A is independent of B, then: \( P(A/B) = P(A) \)

Examples:
• Flip a coin twice. What is the probability of observing two heads?
• Flip a coin twice. What is the probability of getting a head and then a tail? A tail and then a head?
• In general, if events $A$ and $B$ are not independent, then the multiplication rule becomes

$$P(A \cap B) = P(A)P(B \mid A)$$
Terminology

- $P(AnB) = P(A \text{ and } B)$
  
  Joint probability of $A$ and $B$
  
  (of the intersection of $A$ and $B$)

- $P(A/B)$
  
  Conditional probability of $A$ given $B$

- $P(A)$
  
  (Marginal) probability of $A$
• If we have the probability table, then everything can be figured out from the table. NO need to use the rules.

• Only when no table is available, then we may be able to find out some probabilities from some given/known probabilities (a partial table) using rules.
• In homework/exam, you may be given a probability table, and are asked to verify certain rules.

Or

• Given a partial table, you are asked to use various rules to find the missing probabilities in the table.
Examples
Attendance Survey Question

• On a 4”x6” index card
  – Please write down your name and section number
  – Today’s Question:

  – Is A independent of B in reality?
A={Stock market go up today};
B={snow > 3 inch in New York today}