Math 10

Part 2
Probability
© Maurice Geraghty 2017

Probability
- Classical probability
  - Based on mathematical formulas
- Empirical probability
  - Based on the relative frequencies of historical data.
- Subjective probability
  - "one-shot" educated guess.

Examples of Probability
- What is the probability of rolling a four on a 6-sided die?
- What percentage of De Anza students live in Cupertino?
- What is the chance that the Golden State Warriors will be NBA champions in 2017?
Part 2 Probability

Classical Probability

- **Event**
  - A result of an experiment
- **Outcome**
  - A result of the experiment that cannot be broken down into smaller events
- **Sample Space**
  - The set of all possible outcomes
- **Probability Event Occurs**
  - \( \frac{\# \text{ of elements in Event}}{\# \text{ Elements in Sample Space}} \)
- **Example** – flip two coins, find the probability of exactly 1 head.
  - \{HH, HT, TH, TT\}
  - \( P(1 \text{ head}) = \frac{2}{4} = 0.5 \)

Empirical Probability

- **Historical Data**
- **Relative Frequencies**
- **Example**: What is the chance someone rates their community as good or better?
  - \( 0.51 + 0.32 = 0.83 \)

Rule of Complement

- **Complement of an event**
  - The event does not occur
  - \( A' \) is the complement of \( A \)
  - \( P(A) + P(A') = 1 \)
  - \( P(A) = 1 - P(A') \)
Additive Rule

- The **UNION** of two events A and B is that either A or B occur (or both). (All colored parts)

- The **INTERSECTION** of two events A and B is that both A and B will occur. (Purple Part only)

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

Example

- In a group of students, 40% are taking Math, 20% are taking History.
- 10% of students are taking both Math and History.
- Find the Probability of a Student taking either Math or History or both.
  \[ P(M \text{ or } H) = 40\% + 20\% - 10\% = 50\% \]

Mutually Exclusive

- Mutually Exclusive
  - Both cannot occur
  - If A and B are mutually exclusive, then
    \[ P(A \text{ or } B) = P(A) + P(B) \]

- Example roll a die
  - A: Roll 2 or less       B: Roll 5 or more
  - \[ P(A) = \frac{2}{6} \quad P(B) = \frac{2}{6} \]
  - \[ P(A \text{ or } B) = P(A) + P(B) = \frac{4}{6} \]
Conditional Probability

- The probability of an event occurring GIVEN another event has already occurred.
- \( P(A|B) = \frac{P(A \text{ and } B)}{P(B)} \)

Example: Of all cell phone users in the US, 15% have a smartphone with AT&T. 25% of all cell phone users use AT&T. Given a selected cell phone user has AT&T, find the probability the user also has a smartphone.

\( A = \text{AT&T subscriber} \quad B = \text{Smart Phone User} \)

\[
\begin{align*}
P(A \text{ and } B) &= .15 \\
P(A) &= .25 \\
P(B|A) &= \frac{.15}{.25} = .60
\end{align*}
\]

Contingency Tables

- Two data items can be displayed in a contingency table.
- Example: auto accident during year and DUI of driver.

<table>
<thead>
<tr>
<th></th>
<th>Accident</th>
<th>No Accident</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>DUI</td>
<td>70</td>
<td>130</td>
<td>200</td>
</tr>
<tr>
<td>Non-DUI</td>
<td>30</td>
<td>770</td>
<td>800</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>900</td>
<td>1000</td>
</tr>
</tbody>
</table>

Given the Driver is DUI, find the Probability of an Accident.

\( A = \text{Accident} \quad D = \text{DUI} \)

\[
\begin{align*}
P(A \text{ and } D) &= .07 \\
P(D) &= .2 \\
P(A|D) &= \frac{.07}{.2} = .35
\end{align*}
\]
Marginal, Joint and Conditional Probability

- **Marginal Probability** means the probability of a single event occurring.
- **Joint Probability** means the probability of the union or intersection of multiple events occurring.
- **Conditional Probability** means the probability of an event occurring given that another event has already occurred.

Creating Contingency Tables

- You can create a hypothetical contingency table from reported cross tabulated data.
- First choose a convenient sample size (called a radix) like 10000.
- Then apply the reported marginal probabilities to the radix of one of the variables.
- Then apply the reported conditional probabilities to the total values of one of the other variable.
- Complete the table with arithmetic.

Example

Create a two-way table from the cross tabulation of gender from the 2016 election results (from CNN)
**Example**

First select a radix (sample size) of 10000

<table>
<thead>
<tr>
<th>GENDER</th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trump</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinton</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1000</td>
<td></td>
</tr>
</tbody>
</table>

**Example**

Then apply the marginal probabilities to the radix (53% female, 47% male)

<table>
<thead>
<tr>
<th>GENDER</th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trump</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinton</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5300</td>
<td>4700</td>
<td>10000</td>
</tr>
</tbody>
</table>

**Example**

Then apply the cross tabulated percentages for each gender. Make sure the numbers add up.

<table>
<thead>
<tr>
<th>GENDER</th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trump</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinton</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5300</td>
<td>4700</td>
<td>10000</td>
</tr>
</tbody>
</table>
Part 2 Probability

Example

Finally, complete the table using arithmetic.

<table>
<thead>
<tr>
<th>GENDER</th>
<th>VOTED FOR</th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trump</td>
<td>2323</td>
<td>2444</td>
<td></td>
<td>4817</td>
</tr>
<tr>
<td>Clinton</td>
<td>2862</td>
<td>1927</td>
<td></td>
<td>4789</td>
</tr>
<tr>
<td>Other</td>
<td>265</td>
<td>329</td>
<td></td>
<td>594</td>
</tr>
<tr>
<td>Total</td>
<td>5300</td>
<td>4700</td>
<td></td>
<td>10000</td>
</tr>
</tbody>
</table>

Multiplicative Rule

- \( P(A \text{ and } B) = P(A) \times P(B|A) \)
- \( P(A \text{ and } B) = P(B) \times P(A|B) \)
- Example: A box contains 4 green balls and 3 red balls. Two balls are drawn. Find the probability of choosing two red balls.
  - \( A=\text{Red Ball on 1st draw} \quad B=\text{Red Ball on 2nd Draw} \)
  - \( P(A)=3/7 \quad P(B|A)=2/6 \)
  - \( P(A \text{ and } B) = (3/7)(2/6) = 1/7 \)

Multiplicative Rule – Tree Diagram
Independence

- If A is not dependent on B, then they are **INDEPENDENT** events, and the following statements are true:
  - \( P(A|B) = P(A) \)
  - \( P(B|A) = P(B) \)
  - \( P(A \text{ and } B) = P(A) \times P(B) \)

Example

<table>
<thead>
<tr>
<th></th>
<th>Accident</th>
<th>No Accident</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>DUI</td>
<td>70</td>
<td>130</td>
<td>200</td>
</tr>
<tr>
<td>Non-DUI</td>
<td>30</td>
<td>770</td>
<td>800</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>900</strong></td>
<td><strong>1000</strong></td>
</tr>
</tbody>
</table>

A: Accident  
D: DUI Driver

\( P(A) = .10 \quad P(A|D) = .35 \) (70/200)

Therefore A and D are **DEPENDENT** events as \( P(A) < P(A|D) \)

Example

<table>
<thead>
<tr>
<th></th>
<th>Accident</th>
<th>No Accident</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domestic Car</td>
<td>60</td>
<td>540</td>
<td>600</td>
</tr>
<tr>
<td>Import Car</td>
<td>40</td>
<td>360</td>
<td>400</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>900</strong></td>
<td><strong>1000</strong></td>
</tr>
</tbody>
</table>

A: Accident  
D: Domestic Car

\( P(A) = .10 \quad P(A|D) = .10 \) (60/600)

Therefore A and D are **INDEPENDENT** events as \( P(A) = P(A|D) \)

Also \( P(A \text{ and } D) = P(A) \times P(D) = (.1)(.6) = .06 \)
Random Sample

- A **random sample** is where each member of the population has an equally likely chance of being chosen, and each member of the sample is **INDEPENDENT** of all other sampled data.

Tree Diagram method

- **Alternative Method of showing probability**
- Example: Flip Three Coins
- Example: A Circuit has three switches. If at least two of the switches function, the Circuit will succeed. Each switch has a 10% failure rate if all are operating, and a 20% failure rate if one switch has already failed. Find the probability the circuit will succeed.

Circuit Problem

- A
- B
- C
- B'
- C'
- Pr(Good) = .81 + .072 + .064 = .946
Switching the Conditionality

- Often there are questions where you desire to change the conditionality from one variable to the other variable.
- First construct a tree diagram.
- Second, create a Contingency Table using a convenient radix (sample size).
- From the Contingency table it is easy to calculate all conditional probabilities.

Example

- 10% of prisoners in a Canadian prison are HIV positive.
- A test will correctly detect HIV 95% of the time, but will incorrectly “detect” HIV in non-infected prisoners 15% of the time (false positive).
- If a randomly selected prisoner tests positive, find the probability the prisoner is HIV+.

Example

A=Prisoner is HIV+
B=Test is Positive for HIV

\[
\begin{array}{c|c|c|c}
A & A' & B & B' \\
\hline
.95 & .05 & .095 & .95 \\
.05 & .95 & .005 & .95 \\
.15 & .85 & .135 & .85 \\
.85 & .15 & .765 & .85 \\
\end{array}
\]
### Example

<table>
<thead>
<tr>
<th></th>
<th>HIV+</th>
<th>HIV-</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test+</td>
<td>950</td>
<td>1350</td>
<td>2300</td>
</tr>
<tr>
<td>Test-</td>
<td>50</td>
<td>7650</td>
<td>7700</td>
</tr>
<tr>
<td>Total</td>
<td>1000</td>
<td>9000</td>
<td>10000</td>
</tr>
</tbody>
</table>

\[
P(A \mid B) = \frac{950}{2300} \approx 0.413
\]