Basic Probability Theory I

Dr. Tom Ilvento
FREC 408

A Probability puzzler!!

Our Strategy with Probability

- Generally, we want to get to an inference from a sample to a population.
- In this case the population is unknown, and we want to infer something from a known sample
- We want the sample to tell us something about the population

In probability

- With probability we will do the reverse:
  - We start with the notion that the population is known, i.e., the probabilities
  - And infer something about the chances of obtaining various samples from the population

Some Terms

- If we flip a coin and record the result – a tails
- The result we record is an observation
- The process of making an observation is called an experiment (Def4.1 P193)
  - More specifically, an experiment leads to a single outcome
  - Which can’t be predicted with certainty

Some terms

- The basic outcome of an experiment is called a sample point
- And the collection of outcomes is called the sample space (Def4.6 P199).
- Outcomes of experiments are called events (Def4.2 P194).
- Part of our strategy will be to identify all the sample points, or all the possible outcomes of the experiment
- That is, to identify the sample space
What Does Probability Mean?

- The probability of an event is a proportion between 0 and 1 that measures the likelihood that the outcome will occur when the experiment is performed.
- If we flip a coin, we expect that the probability of getting a heads is .5
  - Assuming a fair coin!

Probabilities

- It is a numerical measure of the likelihood that Outcome A will occur
  - \( P(A) \)
  - \( \text{Prob}(A) \)

<table>
<thead>
<tr>
<th></th>
<th>Certain</th>
<th>Impossible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

A closer look at probability

- Suppose you flip a coin
  - What’s the chance of getting a head?
  - What’s the chance of getting a tails?

Flipping a coin

- What does it mean when we state a coin flip as a probability?
  - Does it mean that every other time it will be a heads?
  - If we flip it 10 times, will we always get exactly five heads?
  - If we flip it 1,000 times, will we get exactly 500 heads?

Flipping a coin

- TRY IT IN CLASS
  - Flip a coin 10 times
  - Record the number of HEADs
  - Calculate a probability of a HEAD based on your data
  - Calculate the results: \# Head/10

The story of the Law of Averages

- John Kerrich was a South African mathematician who spent WWII in an internment camp – very unlucky!
- He had a lot of time on his hands, so he carried out experiments in probability theory
- One viewpoint about the Law of Averages was that if you tossed a coin a lot of times, the number of heads and the number of tails would eventually equal out
Law of Averages

- Not so, said Kerrich!
- The probability of every toss is still 50/50. No matter what happened before, each toss has an equal chance
- In 100 tosses of a coin, 1,000 tosses, or 10,000 tosses, the probability of the next toss is still the same.

Thus the difference between the number of heads and the number of tails could be quite large in 10,000 tosses
- We should expect 5,000 heads and 5,000 tails
- It could easily be 4,900 heads and 5,100 tails

But, in relative terms, the difference will get smaller and smaller as the number of tosses gets larger and larger
- As the number of tosses increases the percentage of heads will approach 50%
- Our expectation is 50% heads and tails, with some chance error.
- As we increase the number of tosses, the observed percentage of heads (or tails) approaches 50%
  - but it may never get there
  - And the absolute difference between heads and tails may be large

The proportion of heads approaches .5 as the number of tosses increases
- The basic definition of probability is:
  \[
  \text{Probability} = \frac{\text{# favorable events}}{\text{Total # outcomes}}
  \]
- It is a proportion!
Two Kinds of Probability

- **A priori** – given by a definition and before the fact. It is generally mathematically defined.
- **Example:** Rolling a die with equal probabilities for each outcome
  - There are 6 outcomes – the probabilities for rolling a one is \(\frac{1}{6}\)
  - A priori probabilities are based on the long run

Two kinds of probabilities

- **A posteriori** – derived empirically through repeated experiments
  - We observe the probabilities after the fact from an experiment or series of experiments
  - A survey approach would be a posteriori

An example of an experiment

- Flip two coins in succession, then observe the face of the two coins

Write out the sample space

- Sample space:
  - Observe H H
  - Observe H T
  - Observe T H
  - Observe T T

  \[ S: \{HH, HT, TH, TT\} \]

Another example

- We conduct an experiment and roll a single die and record the face

Sample space of rolling a die

- Sample space:
  - Observe a 1
  - Observe a 2
  - Observe a 3
  - Observe a 4
  - Observe a 5
  - Observe a 6

  \[ S: \{1, 2, 3, 4, 5, 6\} \]
The sample space could be defined in several ways

- Experiment: roll a die and observe odd or even number
- Sample Space
  - Observe an even number (2, 4, 6)
  - Observe an odd number (1, 3, 5)
- S: [Even, Odd]

More examples

<table>
<thead>
<tr>
<th>EXPERIMENT</th>
<th>SAMPLE SPACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toss a Coin, Note Face</td>
<td>Head, Tail</td>
</tr>
<tr>
<td>Toss 2 Coins, Note Faces</td>
<td>HH, HT, TH, TT</td>
</tr>
<tr>
<td>Select 1 Card, Note Kind</td>
<td>2♦, 2♣, ..., A♠ (52)</td>
</tr>
<tr>
<td>Inspect a Part, Note Quality</td>
<td>Defective, OK</td>
</tr>
<tr>
<td>Interview and Observe Gender</td>
<td>Male, Female</td>
</tr>
</tbody>
</table>

Outcomes of Experiments

- We want outcomes (i.e., sample points) to be
  - Mutually Exclusive – two outcomes can’t occur at the same time (Def 4.4 P198)
  - Collectively Exhaustive – all possible outcomes are identified and no more possible outcomes are left out of sample space.

Some more terms

- Random Trials
  - The selection of any outcome is not predetermined; thus each outcome has a chance to be selected
  - A fixed uneven dice would not be random, and would be lead to outcomes that are biased

More terms

- Independent Trials
  - The outcome is not conditioned upon previous events. The outcome of previous event has no compact on the outcome of current event.
  - Flips of a coin are independent – the probability of a head on a second flip is not influenced by the outcome on the first flip

Probability Rules for Sample Points

- All sample point probabilities must lie between 0 and 1
  - A probability of 1 means CERTAINTY
  - The probabilities of all sample points within a sample space must sum to 1
### Roll of a single die

<table>
<thead>
<tr>
<th>Sample Points</th>
<th>Probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/6 = .1667</td>
</tr>
<tr>
<td>2</td>
<td>1/6 = .1667</td>
</tr>
<tr>
<td>3</td>
<td>1/6 = .1667</td>
</tr>
<tr>
<td>4</td>
<td>1/6 = .1667</td>
</tr>
<tr>
<td>5</td>
<td>1/6 = .1667</td>
</tr>
<tr>
<td>6</td>
<td>1/6 = .1667</td>
</tr>
<tr>
<td>TOTAL</td>
<td>6/6 = 1.000</td>
</tr>
</tbody>
</table>

### More Terms - Events
- An **event** is a specific collection of sample points
  - For example, on a roll of the dice
    - Sample points are \([1, 2, 3, 4, 5, 6]\)
    - The Event A: [Even numbers] contains the sample points 2, 4, 6
    - The Event B: [Odd numbers] contains the sample points 1, 3, 5
  - Events are subsets, which can overlap or be mutually exclusive

### Probability of an Event
- The **probability of an Event A**, denoted as \( P(A) \), is calculated by
  - summing the probabilities of the sample points in the sample space for Event A

### Steps for Calculating Probabilities of Events
1. **Define the experiment** – describe process of making an observation
2. **List the Sample Points**
3. **Assign probabilities to the Sample Points**
4. Determine the collection of **Sample Points** contained in the **Event** of interest
5. **Sum the Sample Points probabilities** to get the **Event** probability

### Rolling a single die
- **Event:** rolling a 2
  - Outcome 2
  - \( P(\text{rolling a 2}) = 1/6 \)
- **Event:** rolling 3 or higher
  - Outcomes 3, 4, 5, 6
  - \( P(\text{rolling 3+}) = 1/6 + 1/6 + 1/6 + 1/6 = 2/3 \)
  - \( P(\text{rolling 3+}) = .667 \)
Rolling a single die

- **Event:** rolling an even number
  - Outcomes 2, 4, 6
  - \( P \text{ (rolling even)} = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = .5 \)

Representing the Sample Space or Events

- We can represent the Sample Space and Events by
  - Listing of the Set
  - Venn Diagram
  - Contingency Table
  - Decision Tree Diagram

Venn Diagram

Experiment: Toss 2 Coins. Note Faces.

```
<table>
<thead>
<tr>
<th>Head</th>
<th>TH</th>
<th>HT</th>
<th>TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Event – at least one tail

\( S = \{HH, HT, TH, TT\} \) Sample Space

Contingency Table

Experiment: Toss 2 Coins. Note Faces.

```
<table>
<thead>
<tr>
<th>1st Coin</th>
<th>2nd Coin</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>HH</td>
<td>HH, HT</td>
</tr>
<tr>
<td>Tail</td>
<td>TH</td>
<td>TH, TT</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>S</td>
</tr>
</tbody>
</table>
```

\( S = \{HH, HT, TH, TT\} \) Sample Space

Tree Diagram

Experiment: Toss 2 Coins. Note Faces.

```
First Toss       Second Toss
H                H
H                T
H                H
T                TH
```

\( S = \{HH, HT, TH, TT\} \)

Problem example

- I have a jar containing five marbles, 2 of which are blue and 3 of which are red.
- I randomly draw two marbles
- What is the probability of drawing 2 blue marbles?
Marble problem solution
- Define the experiment
- List or draw out the sample points
- Assign probabilities to the sample points
- Determine the collection of sample points contained in an event of interest
- Sum the sample points probabilities to get the event probability

Marble problem
- Sample Points: B for blue, R for Red
  - B₁ B₂
  - B₁ R₁
  - B₁ R₂
  - B₁ R₃
  - B₂ R₁
  - B₂ R₂
  - B₂ R₃
  - R₁ R₂
  - R₁ R₃
  - R₂ R₃

Marble Problem
- In this problem the order of the marbles is not important to us
- So there is only one combination of drawing two blue marbles
- Unless otherwise known, each sample point has an equal probability
- Thus each has a 1/10 chance of being drawn

Express sample points as combinations of blue and red marbles
- Sample Points: Two Blue, Blue & Red, Two Red
- Sample Points: Probability
  - Two Blue: 1/10
  - Blue & Red: 1/10
  - Two Red: 1/10

Marble problem
- What is the probability that two blue marbles are drawn?
  - 1/10
- What is the probability that a blue and a red marble are drawn?
  - 6/10 OR 3/5
- What is the probability that two red marbles are drawn?
  - 3/10
Express sample points as combinations of blue and red marbles

- Sample Points | Probability
- Two Blue       | 1/10
- Blue & Red     | 6/10
- Two Red        | 3/10

**Combinatorial formula**

- To find the number of samples of $N$ things taken $n$ at a time

\[ \binom{N}{n} = \frac{N!}{n!(N-n)!} \]

- Where $n!$ is

\[ n! = n(n-1)(n-2)\ldots(1) \]

\[ 5! = 5(5-1)(5-2)(5-3)(5-4) = 120 \]

- Note: $0! = 1$ and $1! = 1$

**Survey Problem**

- A researcher wanted to find the primary reason for a company to engage in diversity training.
- She surveyed businesses to determine the primary reason for diversity training, offering five mutually exclusive, and exhaustive options.
- Listed the percentages for each.

**Diversity survey problem**

<table>
<thead>
<tr>
<th>Reason</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comply with personnel policies (CPP)</td>
<td>7%</td>
</tr>
<tr>
<td>Increase productivity (IP)</td>
<td>47%</td>
</tr>
<tr>
<td>Stay competitive (SC)</td>
<td>38%</td>
</tr>
<tr>
<td>Social responsibility (SR)</td>
<td>4%</td>
</tr>
<tr>
<td>Other (O)</td>
<td>4%</td>
</tr>
</tbody>
</table>

The percentages can be read as probabilities.

**Diversity problem**

- What is the probability that the primary reason for diversity training is business related; i.e., related to competition or productivity?

- What is the probability that social responsibility is not a primary reason for diversity training?
Diversity Problem

- What is the probability that the primary reason for diversity training is business related; i.e., related to competition or productivity?
  
P(B) = P(IP+SC) = .47 + .38 = .85

- What is probability that social responsibility is not a primary reason for diversity training?
  
P(Not SR) = P(CPP+IP+SC+O) = .07 + .47 + .38 + .04 = .96

For the last answer
  - We could have thought of this as the compliment of SR
  - Denoted as SR^c
  
SR^c = 1 – P(SR)
SR^c = 1 - .04 = .96