Probability I

Patrick Breheny

February 17
People talk loosely about *probability* all the time: “What are the chances the Hawkeyes will win this weekend?” , “What’s the chance of rain tomorrow?”

For scientific purposes, we need to be more specific in terms of defining and using probabilities
A random process is a phenomenon whose outcome cannot be predicted with certainty.

An event is a collection of outcomes.

Examples:

<table>
<thead>
<tr>
<th>Random process</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flipping a coin</td>
<td>Obtaining heads</td>
</tr>
<tr>
<td>Child receives a vaccine</td>
<td>Child contracts polio</td>
</tr>
<tr>
<td>Patient undergoes surgery</td>
<td>Patient survives</td>
</tr>
</tbody>
</table>
Long-run frequency

- The probability of heads when flipping a coin is 50%
- The probability of rolling a 1 on a 6-sided die is 1/6
- Everyone agrees with these statements, but what do they really mean?
- The probability of an event occurring is defined as the fraction of time that it would happen if the random process occurs over and over again under the same conditions
- Therefore, probabilities are always between 0 and 1
Proportions are denoted with a $P(\cdot)$, as in $P(\text{Heads})$ or $P(\text{Child develops polio})$ or “Let $H$ be the event that the outcome of a coin flip is heads. Then $P(H) = 0.5$”

Example:
- The probability of being dealt a full house in poker is 0.0014
- If you were dealt 100,000 poker hands, how many full houses should you expect?
  
  $100,000(0.0014) = 140$

Note: It is important to distinguish between a probability of 0.0014 and a probability of 0.0014% (which would be a probability of 0.000014)
This works both ways:
- For the polio data, 28 per 100,000 children who got the vaccine developed polio.
- The probability that a child in our sample who got the vaccine developed polio is 28/100,000 = .00028.

Of course, what we really want to know is not the probability of a child in our sample developing polio, but the probability of a child in the population developing polio – we’re getting there.
We are often interested in events that are derived from other events:

- Rolling a 2 or 3
- Patient who receives a therapy is relieved of symptoms and suffers from no side effects

The event that $A$ does not occur is called the complement of $A$ and is denoted $A^C$.

The event that both $A$ and $B$ occur is called the intersection and is denoted $A \cap B$.

The event that either $A$ or $B$ occurs is called the union and is denoted $A \cup B$. 


These relations between events can be represented visually using Venn diagrams:
Introduction

- Let event $A$ denote rolling a 2 and event $B$ denote rolling a 3.
- What is the probability of rolling a 2 or a 3 ($A \cup B$)?
- It turns out to be

$$\frac{1}{6} + \frac{1}{6} = \frac{2}{6}$$

- On the surface, then, it would seem that

$$P(A \cup B) = P(A) + P(B)$$

- However, this is not true in general.
Let $A$ denote rolling a number 3 or less and $B$ denote rolling an odd number.

$P(A) + P(B) = 0.5 + 0.5 = 1$

Clearly, however, we could roll a 4 or a 6, which is neither $A$ nor $B$.

What’s wrong?
Double counting

- With a Venn diagram, we can get a visual idea of what is going wrong:

![Venn diagram](image)

- When we add $P(A)$ and $P(B)$, we count $A \cap B$ twice
- Subtracting $P(A \cap B)$ from our answer corrects this problem
In order to determine the probability of $A \cup B$, we need to know:

- $P(A)$
- $P(B)$
- $P(A \cap B)$

If we’re given those three things, then we can use the addition rule:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

This rule is always true for any two events.
Mutually exclusive events

- So why did $P(A \cup B) = P(A) + P(B)$ work when $A$ was rolling a 2 and $B$ was rolling a 3?
- Because $P(A \cap B) = 0$, so it didn’t matter whether we subtracted it or not.
- A special term is given to the situation when $A$ and $B$ cannot possibly occur at the same time: such events are called *mutually exclusive*. 

Patrick Breheny  
*Introduction to Biostatistics (BIOS 4120)*  13/20
Mutually exclusive events, example

- According to the National Center for Health Statistics, the probability that a randomly selected woman who gave birth in 1992 was aged 20-24 was 0.263.
- The probability that a randomly selected woman who gave birth in 1992 was aged 25-29 was 0.290.
- Are these events mutually exclusive?
- Yes, a woman cannot be two ages at the same time.
- Therefore, the probability that a randomly selected woman who gave birth in 1992 was aged 20-29 was $0.263 + 0.290 = 0.553$. 

Patrick Breheny
Introduction to Biostatistics (BIOS 4120)
Example: Failing to use the addition rule

- In the 17th century, French gamblers used to bet on the event that in 4 rolls of the die, at least one “ace” would come up (an ace is rolling a one).
- In another game, they rolled a pair of dice 24 times and bet on the event that at least one double-ace would turn up.
- The Chevalier de Méré, a French nobleman, thought that the two events were equally likely.
Example: Failing to use the addition rule

- His reasoning was as follows: letting $A_i$ denote the event of rolling an ace on roll $i$ and $AA_i$ denote the event of rolling a double-ace on roll $i$

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) = P(A_1) + P(A_2) + P(A_3) + P(A_4)$$

$$= \frac{4}{6} = \frac{2}{3}$$

$$P(AA_1 \cup AA_2 \cdots) = P(AA_1) + P(AA_2) + \cdots$$

$$= \frac{24}{36} = \frac{2}{3}$$
Example: Failing to use the addition rule

- Is the Chevalier using the addition rule properly?
- Are $A_1$ and $A_2$ mutually exclusive?
- No; it is possible to get an ace on roll #1 and roll #2, so you have to subtract $P(A_1 \cap A_2)$, $P(A_1 \cap A_3)$, \ldots
- We’ll calculate the real probabilities a little later
An article in the *American Journal of Public Health* reported that in a certain population, the probability that a child’s gestational age is less than 37 weeks is 0.142.

The probability that his or her birth weight is less than 2500 grams is 0.051.

The probability of both is 0.031.

Can we figure out the probability that either event will occur?

Yes: $0.142 + 0.051 - 0.031 = 0.162$
Because an event must either occur or not occur,
\[ P(A) + P(A^C) = 1 \]

Thus, if we know the probability of an event, we can always determine the probability of its complement:

\[ P(A^C) = 1 - P(A) \]

This simple but useful rule is called the complement rule.

Example: If the probability of getting a full house is 0.0014, then the probability of not getting a full house must be

\[ 1 - 0.0014 = 0.9986 \]
The probability of an event is the fraction of time that it happens (under identical repeated conditions)

- Know the meaning of complements ($A^C$), intersections ($A \cap B$), and unions ($A \cup B$)
- Addition rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- If (and only if!) $A$ and $B$ are mutually exclusive, we can ignore $P(A \cap B)$ in the addition rule
- Complement rule: $P(A^C) = 1 - P(A)$