PSY 503: Foundations of Psychological Methods
Lecture 6: Basics of Probability Theory

Robin Gomila

Princeton

September 16, 2020

What is Probability theory?

Probability theory is the study of **random processes** (a.k.a., **random generative processes**, **random phenomena**).

Random processes: Intuition

- Let's flip a fair coin
- Can you tell me what the outcome will be?
- If we were to flip a fair coin many many times, would you be able to tell the proportion of times that we would obtain heads?
- If answer to first question is "NO"

AND

- Answer to second question is "YES"

THEN

- You are dealing with a random process

Definition

Random processes are **mechanisms** that produce (probabilistic) outcomes... from **a world of possible** outcomes... with some degree of **uncertainty** but with **regularity**.

Components of random generative processes

- **Sample space:** Ω
  - The set of all possible outcomes
  - "world of probabilistic outcomes"
  - e.g., heads and tails
- **Outcome:** ω
  - Possible realization of the random process
  - e.g., heads
- **Event:** A, B, C, etc.
  - A given outcome or set of outcomes
  - e.g., "tails did not happen"
- **Probability:** Proportion of times an event or set of events will occur if you keep repeating the random process

Examples of random processes

- Random assignment of N individuals to an experimental condition
- Random draw of a sample of n individuals from a population of N individuals
- Rolling a die

Illustration: Random assignment

- We randomly assigned an individual to a Treatment (T) vs. Control (C)
- Sample space? We could express Ω in the following ways:
  - $\Omega = \{\text{Treatment, Control}\}$
  - $\Omega = \{\text{T, C}\}$
- What if we assigned two individuals to Treatment (T) vs. Control (C)
  - $\Omega = \{\text{TT, TC, CT, CC}\}$

Events

An *event* is a subset of the sample space Ω and corresponds to the realization of one or more than one outcomes ω

Example

- Let $\Omega = \{\mathrm{TT},\ \mathrm{TC},\ \mathrm{CT},\ \mathrm{CC}\}$
- We could let A be **event** that both individuals are assigned to the same experimental condition
- We could write:
  - $A = \{TT,\ CC\}$
- Another example?

Notations

Syntax	Description
Ω	sample space
ω	a possible probabilistic outcome
$A \cup B$	A or B
$A \cap B$	A and B
A^C	not A
$A_1 \cup A_2 \cup \ldots \cup A_n$	at least one of $A_1, \ldots, A_n$
$A_1 \cap A_2 \cap \ldots \cap A_n$	all of $A_1, \ldots, A_n$
$A \cap B = \emptyset$	A and B are mutually exclusive

Practice with Events

- We randomly assign 8 participants to T vs. C
  - Possible outcome: $\omega = \mathrm{TTTTCCTC}$
  - Sample space: Set of all possible strings of length 8 of T's and C's

Practice with events

- Let's **randomly** generate a possible outcome ω_j in R

```r
sample(c("T", "C"),
       size = 8,
       replace = TRUE)
```

```
[1] "C" "T" "C" "T" "C" "T" "C" "T"
```

- In the background, does R draw from this sample space?
- NO: Keep in mind that R draws an outcome ω_j from $\Omega = \{T, C\}$ 8 times in a row with replacement

More practice with events

- Let C_1 be the event that the first participant is assigned to the control condition
    - e.g., $\{\text{CTCTCTTT}\}$, $\{\text{CTTTTTTC}\}$, or $\{\text{CCCCCTTC}\}$
- We could express this more generally:
    - $C_1 = \{(C,\ \omega_2,\ ...,\ \omega_8) : \omega_j \in \{\text{T}, \text{C}\} \text{ for } 2 \leq j \leq 8\}$

More practice with events

- Let C_i be the event that the ith participant is assigned to the control condition, for $i = 1, 2, 3, ..., 8$, and use C_i to define other events.
- Let A be the event, that at least one participant was assigned to control condition. We can write:

$$A = C_1 \cup C_2 \cup ... \cup C_8 \tag{1}$$

i.e., Participant 1 or participant 2 or . . . or participant 8 was assigned to control

- We could have written:

$$A = \bigcup_{i=1}^{8} C_i \tag{2}$$

More practice with events

$$B = \bigcap_{i=1}^{8} C_i \tag{3}$$

In plain English?

Naive probability of an event

Let A be an event with a finite sample space Ω. The *naive probability* of A is

$$P(A) = \frac{|A|}{|\Omega|} \tag{4}$$

in which |A| is the number of possible outcomes ω that satisfy A, and $|\Omega|$ is the total number of possible outcomes ω within Ω.

Wait, why is this naive?

- Requires Ω to be finite
- Requires each possible outcome ω to have the same weight
  - This can be misleading!
  - e.g., polls, attrition

Probability model: Definition

- The **probability model** of a random phenomenon is the mathematical representation of this phenomenon. It includes:
  - All of the possible outcomes included in the sample space
  - The probability of each possible probabilistic outcome ω included in the sample space
- This is all that there is to know about a random phenomenon
- Very powerful: Contain enough information to predict with certainty the percentage of times that an outcome ω will happen if we repeat the random generative process many (many, many, many) times

Probability model: Intuition

Overall, the probability of an outcome ω is the percentage of times that this outcome will happen if we repeat the random generative process:

- over and over again
- independently
- under the exact same conditions

Probability model: Example

- The probability model of rolling a fair die includes:
  - Its sample space: $\Omega = \{1, 2, 3, 4, 5, 6\}$
  - The probability of each possible outcome ω_j is: $\mathrm{P}(w_j) = \frac{1}{6}$

If we wonder what are the possible outcomes of rolling a fair die, we simply need to look at the probability model to realize that there are 6 possible outcomes.

If we wonder how likely it is that we will get a 6 after rolling a fair die, again, we can look at its probability model and learn that the probability of getting a 6 is $\frac{1}{6}$

Probability Rules

Probability rule # 1

- Probabilities take values between 0 and 1 (inclusive)
- For some event A:
$$0 \leq P(A) \leq 1$$
- Probability cannot be negative
- Probability cannot be greater than 1

Probability rule # 2

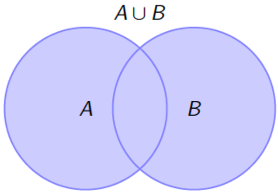
- Since Ω is the entire sample space,

$$P(\Omega) = 1$$

- e.g., What is the probability of getting an even or an odd number after rolling a fair die?

Probability rule # 3

- The probability that A **or** B occurs is the probability of the **union** of A and B
- Denoted by $P(A \cup B)$

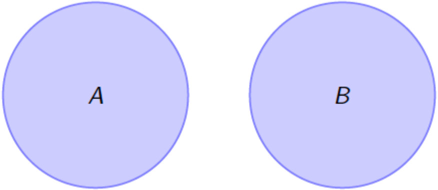


- Addition rule:

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

Mutually exclusive events

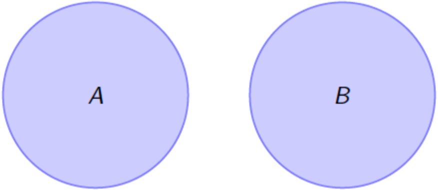
- Two events A_i and A_j are **mutually exclusive** (or **disjoint**) if they cannot happen at the same time



- For $i \neq j$, we have:

$$A_i \cap A_j = \emptyset$$

Probability rule # 3 for mutually exclusive events



- Under the addition rule:

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

- But what is $P(A \cap B)$? $P(A \cap B) = 0$
- Therefore,

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

Probability rule # 3 generalized to any number of mutually exclusive events

- Given any number of *mutually exclusive* events $A_1,\ A_2,\ ...,\ A_n$, the probability that one of these events will occur is the sum of their individual probabilities:
  - $P(A_1\ \cup\ A_2\ \cup\ ...\ \cup\ A_n) = P(A_1)\ + P(A_2)\ + ... +\ P(A_n)$
- Let F be the event of rolling a fair die and getting an even number
  - $F = \{2, 4, 6\} = 2\ \cup 4\ \cup 6$
  - $P(F) = P(2) + P(4) + P(6)$

Probability rule # 4

- The complement of event A is referred to as A^c
- By definition
$$P(A) + P(A^c) = 1$$
- This implies
$$P(A^c) = 1 - P(A)$$