Introduction to Probability

Kyle Richardson

June 2021

Probability and Dealing with Uncertainty

- **Playing with Dice**: Suppose we have a fair die with faces ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ that we decide to roll. **questions:**

Probability and Dealing with Uncertainty

- **Playing with Dice**: Suppose we have a fair die with faces ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ that we decide to roll. **questions:**

1. Will the die land on some face? **yes**

Probability and Dealing with Uncertainty

- **Playing with Dice**: Suppose we have a fair die with faces ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ that we decide to roll. **questions:**

1. Will the die land on some face? **yes**
2. Will the die land on ⚀?

Probability and Dealing with Uncertainty

- **Playing with Dice**: Suppose we have a fair die with faces ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ that we decide to roll. **questions:**

1. Will the die land on some face? **yes**
2. Will the die land on ⚀?

$$\textbf{maybe: } \frac{1}{6} \approx 0.16$$

Probability and Dealing with Uncertainty

- **Playing with Dice**: Suppose we have a fair die with faces ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ that we decide to roll. **questions:**

1. Will the die land on some face? **yes**
2. Will the die land on ⚀?

$$\textbf{maybe: } \frac{1}{6} \approx 0.16$$

3. Will the die land on a face with less than or equal to 4 dots?

Probability and Dealing with Uncertainty

- **Playing with Dice**: Suppose we have a fair die with faces ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ that we decide to roll. **questions:**

1. Will the die land on some face? **yes**
2. Will the die land on ⚀?

$$\textbf{maybe: } \frac{1}{6} \approx 0.16$$

3. Will the die land on a face with less than or equal to 4 dots?

$$\textbf{maybe: } \frac{4}{6} \approx 0.66$$

Probability and Dealing with Uncertainty

- **Playing with Dice**: Suppose we have a fair die with faces ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ that we decide to roll. **questions:**

1. Will the die land on some face? **yes**
2. Will the die land on ⚀?

$$\textbf{maybe: } \frac{1}{6} \approx 0.16$$

3. Will the die land on a face with less than or equal to 4 dots?

$$\textbf{maybe: } \frac{4}{6} \approx 0.66$$

Features of Questions 2-3: Do not have a yes/no answer, require some **measurement** or **(random) experimentation**.

Probability and Dealing with Uncertainty

- **Playing with Dice**: Suppose we have a fair die with faces ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ that we decide to roll. **questions:**

3. Will the die land on a face with less than or equal to 4 dots?

$$\textbf{maybe: } \frac{4}{6} \approx 0.66$$

Why is the **numerator** 4 and the **denominator** 6?

[1] Some examples and figures (and general presentation) taken from Chan (2021), https://engineering.purdue.edu/ChanGroup/eBook.html.

Probability and Dealing with Uncertainty

- **Playing with Dice**: Suppose we have a fair die with faces ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ that we decide to roll. **questions:**

3. Will the die land on a face with less than or equal to 4 dots?

$$\textbf{maybe: } \frac{4}{6} \approx 0.66$$

Why is the **numerator** 4 and the **denominator** 6?

$$\textbf{“event“ } E = \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}\}; |E| = 4$$

$$\textbf{“sample space“ } \Omega = \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\}; |\Omega| = 6$$

[1] Some examples and figures (and general presentation) taken from Chan (2021), `https://engineering.purdue.edu/ChanGroup/eBook.html`.

Probability and Dealing with Uncertainty

- **Playing with Dice**: Suppose we have a fair die with faces ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ that we decide to roll. **questions:**

3. Will the die land on a face with less than or equal to 4 dots?

maybe: $\frac{4}{6} \approx 0.66$

Why is the **numerator** 4 and the **denominator** 6?

"event" $E = \{$⚀, ⚁, ⚂, ⚃$\}; |E| = 4$

"sample space" $\Omega = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}; |\Omega| = 6$

One Definition: Probability is a measure of the size of a set.[1]

[1] Some examples and figures (and general presentation) taken from Chan (2021), https://engineering.purdue.edu/ChanGroup/eBook.html.

A Simple Model of Probability

Components: 1) sample space Ω (set of all possible outcomes); 2) *event* E (some subset of Ω); *probability* (number that describes relative size of E and Ω)

$$\underbrace{\mathcal{P}}_{\text{measure}}\underbrace{[E]}_{\text{set}} = \underbrace{\frac{E}{\Omega}}_{\text{number} \in [0,1]}$$

A Simple Model of Probability

Components: 1) sample space Ω (set of all possible outcomes); 2) *event* E (some subset of Ω); *probability* (number that describes relative size of E and Ω)

$$\underbrace{\mathcal{P}}_{\text{measure}}\underbrace{[E]}_{\text{set}} = \underbrace{\frac{E}{\Omega}}_{\text{number} \in [0,1]}$$

1. What is the *probability* of obtaining a face that is less than 5 **and** an even number?

$$\frac{E = \{\epsdice{1},\epsdice{2},\epsdice{3},\epsdice{4}\} \cap \{\epsdice{2},\epsdice{4},\epsdice{6}\},}{\Omega = \{\epsdice{1},\epsdice{2},\epsdice{3},\epsdice{4},\epsdice{5},\epsdice{6}\}} = \frac{\{\epsdice{2},\epsdice{4}\}}{\{\epsdice{1},\epsdice{2},\epsdice{3},\epsdice{4},\epsdice{5},\epsdice{6}\}} \approx 0.33$$

A Simple Model of Probability

Components: 1) sample space Ω (set of all possible outcomes); 2) *event* E (some subset of Ω); *probability* (number that describes relative size of E and Ω)

$$\underbrace{\mathcal{P}}_{\text{measure}}\underbrace{[E]}_{\text{set}} = \underbrace{\frac{E}{\Omega}}_{\text{number} \in [0,1]}$$

1. What is the *probability* of obtaining a face that is less than 5 **and** an even number?

$$\frac{E = \{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}\} \cap \{\epsdice{2}, \epsdice{4}, \epsdice{6}\},}{\Omega = \{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}} = \frac{\{\epsdice{2}, \epsdice{4}\}}{\{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}} \approx 0.33$$

Important tool set theory; describing increasingly complex events.

A Simple Model of Probability

Components: 1) sample space Ω (set of all possible outcomes); 2) *event* E (some subset of Ω); *probability* (number that describes relative size of E and Ω)

$$\underbrace{\mathcal{P}}_{\text{measure}} \underbrace{[E]}_{\text{set}} = \underbrace{\frac{E}{\Omega}}_{\text{number} \in [0,1]}$$

1. What is the *probability* of obtaining a face that is less than 5 **and** an even number?

$$\frac{E = \{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}\} \cap \{\epsdice{2}, \epsdice{4}, \epsdice{6}\},}{\Omega = \{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}} = \frac{\{\epsdice{2}, \epsdice{4}\}}{\{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}} \approx 0.33$$

Important tool set theory; describing increasingly complex events.

2. What is the probability of heads twice $(\mathbf{H}, \mathbf{H})$ after two fair coin flips?

$$\frac{E = \{(\mathbf{H}, \mathbf{H})\}}{\Omega = \{(\mathbf{H}, \mathbf{H}), (\mathbf{H}, \mathbf{T}), (\mathbf{T}, \mathbf{T}), (\mathbf{T}, \mathbf{H})\}} = 0.25$$

A Simple Model of Probability

Components: 1) sample space Ω (set of all possible outcomes); 2) *event* E (some subset of Ω); *probability* (number that describes relative size of E and Ω)

$$\underbrace{\mathcal{P}}_{\text{measure}}\underbrace{[E]}_{\text{set}} = \underbrace{\frac{E}{\Omega}}_{\text{number} \in [0,1]}$$

1. What is the *probability* of obtaining a face that is less than 5 **and** an even number?

$$\frac{E = \{\text{⚀},\text{⚁},\text{⚂},\text{⚃}\} \cap \{\text{⚁},\text{⚃},\text{⚅}\},}{\Omega = \{\text{⚀},\text{⚁},\text{⚂},\text{⚃},\text{⚄},\text{⚅}\}} = \frac{\{\text{⚁},\text{⚃}\}}{\{\text{⚀},\text{⚁},\text{⚂},\text{⚃},\text{⚄},\text{⚅}\}} \approx 0.33$$

Important tool set theory; describing increasingly complex events.

2. What is the probability of heads twice $(\mathbf{H},\mathbf{H})$ after two fair coin flips?

$$\frac{E = \{(\mathbf{H},\mathbf{H})\}}{\Omega = \{(\mathbf{H},\mathbf{H}),(\mathbf{H},\mathbf{T}),(\mathbf{T},\mathbf{T}),(\mathbf{T},\mathbf{H})\}} = 0.25$$

More complex (random) **"experiment"**, interested in the outcomes of two separate events (aka random variables) (X_1, X_2).

A Simple Model of Probability (continued)

3. Assuming a shuffled deck of 52 cards that is distributed to 4 players, 13 cards each, what is the probability that each player gets an ace?

A Simple Model of Probability (continued)

3. Assuming a shuffled deck of 52 cards that is distributed to 4 players, 13 cards each, what is the probability that each player gets an ace?

$$\frac{|E| = 13^4}{|\Omega| = \binom{52}{4}} \approx 0.438$$

A Simple Model of Probability (continued)

3. Assuming a shuffled deck of 52 cards that is distributed to 4 players, 13 cards each, what is the probability that each player gets an ace?

$$\frac{|E| = 13^4}{|\Omega| = \binom{52}{4}} \approx 0.438$$

Important tool: combinatorics, how to count.

A Simple Model of Probability (continued)

3. Assuming a shuffled deck of 52 cards that is distributed to 4 players, 13 cards each, what is the probability that each player gets an ace?

$$\frac{|E| = 13^4}{|\Omega| = \binom{52}{4}} \approx 0.438$$

Important tool: combinatorics, how to count.

4. What is the probability of having to wait for a bus? (**continuous interval**)

$$\Omega = \{t \mid 0 \leq t \leq 30\text{min.}\}$$

A Simple Model of Probability (continued)

3. Assuming a shuffled deck of 52 cards that is distributed to 4 players, 13 cards each, what is the probability that each player gets an ace?

$$\frac{|E| = 13^4}{|\Omega| = \binom{52}{4}} \approx 0.438$$

Important tool: combinatorics, how to count.

4. What is the probability of having to wait for a bus? (**continuous interval**)

$$\Omega = \{t \mid 0 \leq t \leq 30\text{min.}\}$$

5. What is the probability of Scottish independence?

A Simple Model of Probability (continued)

3. Assuming a shuffled deck of 52 cards that is distributed to 4 players, 13 cards each, what is the probability that each player gets an ace?

$$\frac{|E| = 13^4}{|\Omega| = \binom{52}{4}} \approx 0.438$$

Important tool: combinatorics, how to count.

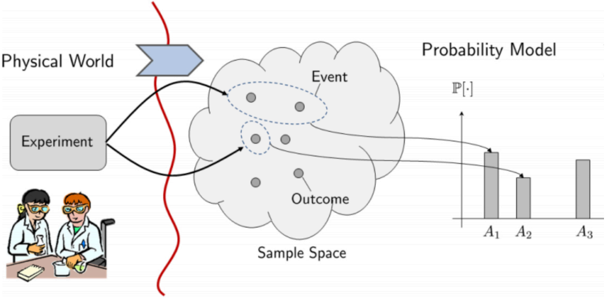
4. What is the probability of having to wait for a bus? (**continuous interval**)

$$\Omega = \{t \mid 0 \leq t \leq 30\text{min.}\}$$

5. What is the probability of Scottish independence?

Subjective probability vs. *relative frequentist* interpretation of probability (our mini theory of probability so far).

The General Picture: Elements of a Probabilistic Model



Components: Sample Space, Events, Probability measure $\mathcal{P}$

Mathematical Tools: set theory, combinatorics, (measure theory), ...

Starting Point: Probability is a measure of the size of a set.

What are the (standard) rules of probability?

- **Kolmogorov axioms**, relate to rules that govern our measurement function $\mathcal{P}$ (how does this work?):

What are the (standard) rules of probability?

- **Kolmogorov axioms**, relate to rules that govern our measurement function $\mathcal{P}$ (how does this work?):

  1. For every event E, $\mathcal{P}[E] \geq 0$. (**non-negativity**)

What are the (standard) rules of probability?

- **Kolmogorov axioms**, relate to rules that govern our measurement function $\mathcal{P}$ (how does this work?):

    1. For every event E, $\mathcal{P}[E] \geq 0$. (**non-negativity**)
    2. $\mathcal{P}[\Omega] = 1$

What are the (standard) rules of probability?

- **Kolmogorov axioms**, relate to rules that govern our measurement function $\mathcal{P}$ (how does this work?):

  1. For every event E, $\mathcal{P}[E] \geq 0$. (**non-negativity**)
  2. $\mathcal{P}[\Omega] = 1$
  3. For disjoint events $\{E_1, E_2, ..., E_n\}$ (i.e., where for each E_i, E_j, $E_i \cap E_j = \emptyset$), (**additivity**)

$$\mathcal{P}[E_1 \cup E_2 \cup ... \cup E_n] = \sum_{i=1}^{n} \mathcal{P}[E_i]$$

Consequences of the Axioms

- The probability of events are *less than or equal to* 1, i.e., $0 \leq \mathcal{P}[E] \leq 1$

  Hint: Axiom 1 (the first part of inequality) combined with Axiom 2.

- **monotonicity:** If $E \subseteq E'$ then $\mathcal{P}[E] \leq \mathcal{P}[E']$
- $\mathcal{P}[\emptyset] = 0$

  $\mathcal{P}[E] = \mathcal{P}[E \cup \emptyset]$(since any $E \cup \emptyset = E$) $= \mathcal{P}[E] + \mathcal{P}(\emptyset)$ (axioms 3)

- If $E_1 \cap E_2 = \emptyset$, $\mathcal{P}[E_1 \cup E_2] = \mathcal{P}[A_1] + \mathcal{P}[A_2]$
- For event E and $\mathcal{P}[E]$, $\mathcal{P}[E^c] = 1 - \mathcal{P}[E]$

  Hint: Apply above, $E = E_1$, $E^c = E_2$ since $E \cap E^c = \emptyset$

Do the rules make sense? Can we justify them?

- **Question:** What is the probability of a coin landing on heads (**H**)?

[2] Examples from Jeffrey (2004), see for an overview of this view.

Do the rules make sense? Can we justify them?

- **Question:** What is the probability of a coin landing on heads (**H**)?

  Another view: Probability reflects how much we are willing to wager in a bet; our **beliefs** about the world.[2]

---

[2] Examples from Jeffrey (2004), see for an overview of this view.

Do the rules make sense? Can we justify them?

- **Question:** What is the probability of a coin landing on heads (**H**)?

  Another view: Probability reflects how much we are willing to wager in a bet; our **beliefs** about the world.[2]

Worth \$1 if heads	Worth \$1 if heads
Price \$0.5	Price \$0.01
Probability 0.5	Probability 0.01
belief: fair coin	**belief:** biased coin

[2] Examples from Jeffrey (2004), see for an overview of this view.

Do the rules make sense? Can we justify them?

- **Question:** What is the probability of a coin landing on heads (**H**)?

  Another view: Probability reflects how much we are willing to wager in a bet; our **beliefs** about the world.[2]

Worth \$1 if heads	Worth \$1 if heads
Price \$0.5	Price \$0.01
Probability 0.5	Probability 0.01
belief: fair coin	**belief:** biased coin

> Viewed in this way, we can then start thinking about subjective probabilities (e.g., prob. of Scottish Independence)
> Rules of probability still play a very important role!

[2] Examples from Jeffrey (2004), see for an overview of this view.

Do the rules make sense? Can we justify them?

- **Question:** What is the probability of a coin landing on heads (**H**)?

Worth \$1 if heads	Worth \$1 if heads
Price \$0.5	Price \$0.01
Probability 0.5	Probability 0.01
belief: fair coin	**belief:** biased coin

Scenario: [3] We bet on the proposition **Heads** for a stake $S = \$1$, and an agent's *betting quotient* q (amount lost if loss); optionally: s for $\neg$**H**.

General Setting

Lands H	**Payoff**
True	$\$1 - \$q$
False	$-\$q$

[3] For more details, see Vineberg (2016).
https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=dutch-book

Do the rules make sense? Can we justify them?

- **Question:** What is the probability of a coin landing on heads (**H**)?

Worth \$1 if heads	Worth \$1 if heads
Price \$0.5	Price \$0.01
Probability 0.5	Probability 0.01
belief: fair coin	**belief:** biased coin

Scenario: [3] We bet on the proposition **Heads** for a stake $S = \$1$, and an agent's *betting quotient* q (amount lost if loss); optionally: s for $\neg$**H**.

General Setting

Lands H	Payoff
True	$\$1 - \$q$
False	$-\$q$

Fair Coin Wager $q = \$0.5$

Lands H	Payoff $S = \$1$
True	\$0.5
False	$-\$0.5$

[3] For more details, see Vineberg (2016).
https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=dutch-book

Do the rules make sense? Can we justify them?

- **Question:** What is the probability of a coin landing on heads (**H**)?

Worth $1 if heads	Worth $1 if heads
Price $0.5	Price $0.01
Probability 0.5	Probability 0.01
belief: fair coin	**belief:** biased coin

Scenario: [3] We bet on the proposition **Heads** for a stake $S = \$1$, and an agent's *betting quotient* q (amount lost if loss); optionally: s for $\neg$**H**.

General Setting

Lands H	Payoff
True	$\$1 - \$q$
False	$-\$q$

Fair Coin Wager $q = \$0.5$

Lands H	Payoff $S = \$1$
True	$0.5
False	−$0.5

Expected value: $(0.5 * 0.5) + (-0.5 * 0.5) = 0$

[3] For more details, see Vineberg (2016).
https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=dutch-book

Do the rules make sense? Can we justify them?

- **Question:** What is the probability of a coin landing on heads (**H**)?

Worth \$1 if heads	Worth \$1 if heads
Price \$0.5	Price \$0.01
Probability 0.5	Probability 0.01
belief: fair coin	**belief:** biased coin

Scenario: [3] We bet on the proposition **Heads** for a stake $S = \$1$, and an agent's *betting quotient* q (amount lost if loss); optionally: s for ¬**H**.

General Setting

Lands H	Payoff
True	$\$1 - \$q$
False	$-\$q$

Fair Coin Wager $q = \$0.5$

Lands H	Payoff $S = \$1$
True	\$0.5
False	$-\$0.5$

Irrational Wager $q + s > 1$

Lands H	Payoff $S = \$1$
True	$\$1 - \$(q + s)$
False	$\$1 - \$(q + s)$

[3] For more details, see Vineberg (2016).
https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=dutch-book

Do the rules make sense? Can we justify them?

- **Question:** What is the probability of a coin landing on heads (**H**)?

Worth $1 if heads	Worth $1 if heads
Price $0.5	Price $0.01
Probability 0.5	Probability 0.01
belief: fair coin	**belief:** biased coin

Scenario: [3] We bet on the proposition **Heads** for a stake $S = \$1$, and an agent's *betting quotient* q (amount lost if loss); optionally: s for $\neg$**H**.

General Setting

Lands H	Payoff
True	$\$1 - \$q$
False	$-\$q$

Fair Coin Wager $q = \$0.5$

Lands H	Payoff $S = \$1$
True	$\$0.5$
False	$-\$0.5$

Irrational Wager $q + s > 1$

Lands H	Payoff $S = \$1$
True	$\$1 - \$(q + s)$
False	$\$1 - \$(q + s)$

Dutch book argument: making wagers in way that flouts probability axioms (e.g., axiom 2) yields irrational behavior (someone always loses).

[3] For more details, see Vineberg (2016).
https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=dutch-book

Conclusion

- Probability theory: A mathematical tool we can use to reason about uncertainty; key concepts: probability as a measure of the size of sets, **(random) events**, **sample space**, Kolmogorov axioms.
- Can serve as a good scientific inductive logic, **subjective probability**, *flouting rules can lead you astray*, (standard) probability theory matters!

Thank you.

References I

Chan, S. H. (2021). *Introduction to Probability for Data Science*. Michigan Publishing.

Jeffrey, R. (2004). *Subjective probability: The real thing*. Cambridge University Press.

Vineberg, S. (2016). Dutch Book Arguments. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2016 edition.