UNIVERSITÄT LEIPZIG

10-207-0003: Introduction to Stochastics

Multiple Linear Regression

28.05.2025, Leipzig

Dr. Ing. Andreas Niekler

Computational Humanities UNIVERSITÄT LEIPZIG

UNIVERSITÄT LEIPZIG

SYLLABUS

1. Empirical research and scale levels
2. Univariate description and exploration of data
3. Graphical representation of characteristics / Explorative data analysis
4. Measures of data distribution
5. Multivariate Problems, Correlation
6. Regression
7. Multiple Linear Regression
8. ***Random Process, Probabilities***
9. Central Limit Theorem
10. Confidences
11. Statistical testing
12. Logistic regression
13. Bayes theorem

Additional: Entropy, Mutual Information, Maximum Likelihood Estimator, Mathy Stuff

TRANSITION FROM DESCRIPTIVE TO INDUCTIVE STATISTICS

- **Descriptive Statistics:** Description of a population using appropriate statistics.
  - Aggregated description of data through frequency distributions.
  - Graphical representation using histograms, box plots, scatter plots, mosaic plots, etc.
  - Quantification using measures of central tendency (mean, median, mode) and measures of dispersion (min, max, standard deviation, variance, quartiles), as well as a five-number summary.
  - Model-based description of relationships through regression models.
- **General Goal of Descriptive Statistics** Summarizing datasets through descriptive statistics, such as numerical values and graphical representations, for interpretation, i.e., reducing data to information.

CENTRAL QUESTION OF INDUCTIVE STATISTICS

- **Fundamental Problem:** We can usually only draw samples from a population.
- **Central Question of Inductive Statistics**: To what extent are samples informative about the population? Are the observed regularities, patterns, and relationships in a sample truly existent and generalizable, or are they purely "random"?
- **Apophenia:** is the human tendency to perceive meaningful patterns or connections in random or unrelated data, objects, or events where no such patterns actually exist. It's essentially "seeing connections that aren't there."

INDUCTIVE STATISTICS

👉 **Inductive statistics** allows us to draw conclusions about an entire population based on a sample. This involves generalizing from the specific (sample) to the general (population) using an inductive inference.

- Inductive inferences **generalize from the specific to the general** by assuming similarity between the sample and the population:
  - Inductive Inference to a Universal Statement:
    - Observation: I have seen a thousand white swans in my life.
    - Conclusion: → All swans are white.
  - Inductive Inference to a Statistical Statement:
    - Observation: Approximately 2% of my social circle votes for the AfD.
    - Conclusion: Only 2% of all people vote for this party.

THE CLASSICAL PROBLEM OF INDUCTION

👉 **The Problem of Induction:** A valid inductive inference presupposes the general assumption of similarity between things. While this **can be repeatedly observed as an empirical phenomenon**, its **generalization as a meta-principle** would itself be an inductive inference.

- **Problematic:** Ultimately, the validity of inductive inferences generally cannot be definitively proven.
- **Pragmatic:** To remain able to act, we nevertheless, for example, infer from the past to the future or generalize from a sample to the population.

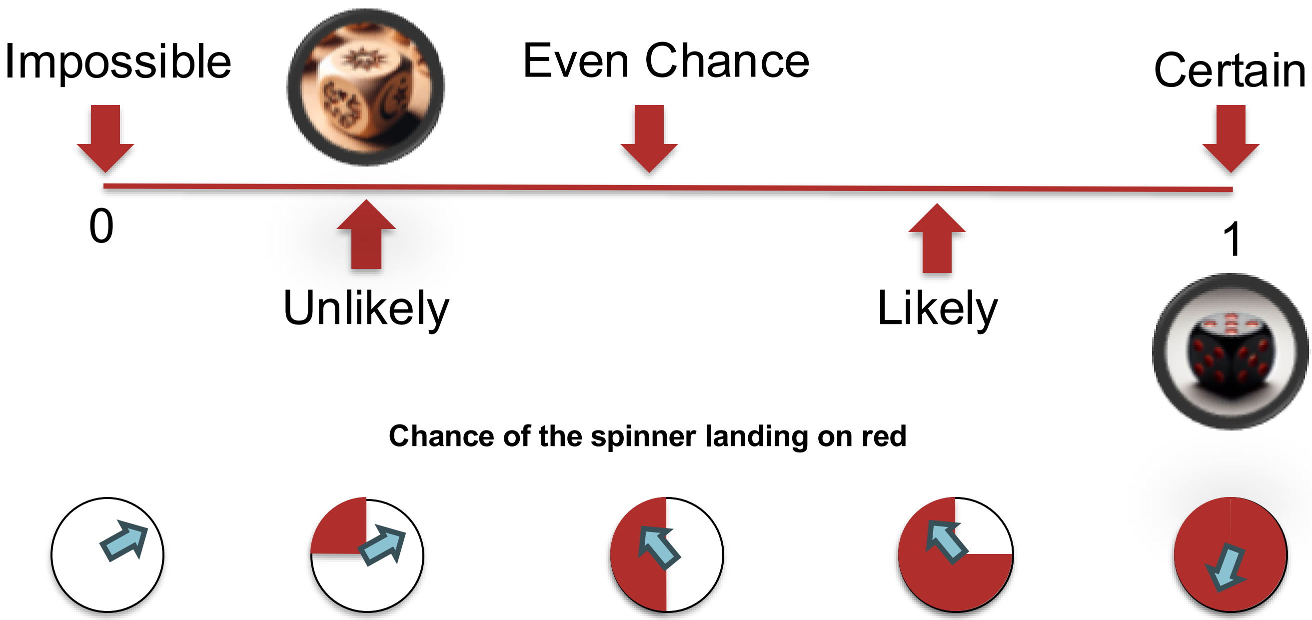
INDUCTIVE STATISTICAL INFERENCE

- In this course, we therefore do not ask whether an inductive conclusion is valid.
- Instead, we ask what a possible and plausible statistical inductive inference might look like, specifically:
  - The data from a panel study (n = 1021) show a positive correlation between the characteristics of education and social mobility.
    - **Conclusion:** Good education generally promotes social mobility.
  - According to a current election poll, about 3% of 1271 respondents would vote for a particular party in the next federal election.
    - **Conclusion:** This party will not enter the next Bundestag as a parliamentary group.

INDUCTIVE INFERENCE AND UNCERTAINTY

- Inductive conclusions generally provide statements that are neither demonstrably true nor false; they are mostly **uncertain**.
  - It's usually correct that a conclusion must be either true or false. Probabilities help us describe these uncertainties.
  - You can therefore understand probabilities as **degrees of truth**.
  - **How probable is something**?

PROBABILITY



DESCRIBING UNCERTAIN STATEMENTS WITH PROBABILITIES

- Forecast by Meteorologists based on current data:
  - It will rain tomorrow with an 80% probability.
- Forecast by Sociologists:
  - The probability that a child of an academic will pursue higher education is 60%, whereas for a child of a non-academic, it's only 20%.
- *Observation: The probabilities refer to statements or events where it's uncertain whether they are true or false, but it's meaningful to ask if they are true or false.*

EVENTS AS STATEMENTS

- An event (e) is a descriptive statement about the world for which it makes sense to ask whether it is true or false (i.e., it has a two-valued truth-value). Its truth value is also, in principle, verifiable or observable (empirical).
- These criteria are merely an attempt to refine and approximate the concept of an event as a two-valued, empirical, and descriptive statement.
  - Not descriptive: e = "One should help other people." (Normative)
  - Not empirical: e = "Harry Potter defeats Lord Voldemort." (Fictional)
  - Not two-valued truth-valued: e = "This sentence is false." (Liar paradox)

EVENTS AS STATEMENTS

- An event (e) is a descriptive statement about the world for which it makes sense to ask whether it is true or false (i.e., it has a two-valued truth-value). Its truth value is also, in principle, verifiable or observable (empirical).
- These criteria are merely an attempt to refine and approximate the concept of an event as a two-valued, empirical, and descriptive statement.
  - Not descriptive: e = "One should help other people." (Normative)
  - Not empirical: e = "Harry Potter defeats Lord Voldemort." (Fictional)
  - Not two-valued truth-valued: e = "This sentence is false." (Liar paradox)
- An event can thus, depending on the **possible state of the world**, either have occurred (true) or not occurred (false).
- Given a set of events, we are interested in $\{e_1, e_2, \dots, e_k\}$, we understand a **possible state of the world** s to be a sufficient functional description such that:
  - $s(e_i) = 1$, if event e_i has occurred in this state s.
  - $s(e_i) = 0$, if event e_i has not occurred in this state s.

EXAMPLE

- Let r be the event *"It rains tomorrow."*
- Let k be the event *"It's cold tomorrow at noon."*
- One possible state of the world, s_1: "Tomorrow it rains all day and stays overcast, with temperatures reaching a maximum of 5 degrees Celsius."
  $\rightarrow s_1(r) = 1$ and $s_1(k) = 1$
- Another possible state of the world, s2: "Tomorrow it stays dry, while the thermometer rises above 20 degrees Celsius at noon."
  $\rightarrow s_2(r) = 0$ and $s_2(k) = 0$

PROPOSITIONAL LOGIC OF EVENTS

- We can construct arbitrarily more complex events using three fundamental operations:
  - **Negation ($\neg e$):** Represents the **negation** of event e.
  - **Conjunction ($a \wedge b$):** Represents the event that both a **AND** b occur jointly.
  - **Disjunction ($a \vee b$):** Represents the event that a **OR** b occurs, or at least one of them occurs (inclusive OR).
- **Example:** Let r be the event "It rains tomorrow," and k be the event "It's cold tomorrow at noon."
- **Negation:** $\neg r$ means "It does **NOT** rain tomorrow."
- **Conjunction:** $r \wedge k$ means "Tomorrow it will rain **AND** be cold at noon."
- **Disjunction:** $r \vee k$ means "Tomorrow it will rain **OR** be cold at noon (or both)."

PROPOSITIONAL LOGIC OF EVENTS

- The truth values for complex events are determined by fixing the following so-called "truth tables":

$s(e)$	$s(\neg e)$
0	1
1	0

$s(a \wedge b)$		$s(b)$	
		0	1
$s(a)$	0	0	0
	1	0	1

$s(a \vee b)$		$s(b)$	
		0	1
$s(a)$	0	0	1
	1	1	1

- Note: These tables represent a formal way in which humans and machines can mechanistically evaluate the truth and falsity of statements.

PROPOSITIONAL LOGIC OF EVENTS

- State of the World s: "Tomorrow it rains all day, but it will be warm throughout."
- → $s(r) = 1$ and $s(k) = 0$

$s(\mathrm{r})$	$s(\neg\mathrm{r})$
0	1
1	0

$s(\mathrm{r} \wedge \mathrm{k})$		$s(\mathrm{k})$	
		0	1
$s(\mathrm{r})$	0	0	0
	1	**0**	1

$s(r \vee \mathrm{k})$		$s(\mathrm{k})$	
		0	1
$s(\mathrm{r})$	0	0	1
	1	**1**	1

FROM EVENTS TO PROCESSES

- Let $S = \{s_1, s_2, \ldots\}$ be the set of all possible states, which describes our conception of what we consider to be possible
- The set $E = \{s \in S \mid s(e) = 1\}$, which encompasses all possible states of the world in which event e occurs, represents the so-called **extensional description** of e.
- So, while e as a statement **intensionally** defines what constitutes the event (i.e., its conceptual content), E as a set **extensionally** describes the event by listing the states in which it occurs (i.e., its conceptual scope).
- This set-theoretic approach, which is dual to the propositional logic approach, is utilized for describing random events. Here, events are not considered fundamental; instead, the **fundamental elements are the possible states of the world that result from a so-called random process**.

RANDOM PROCESS

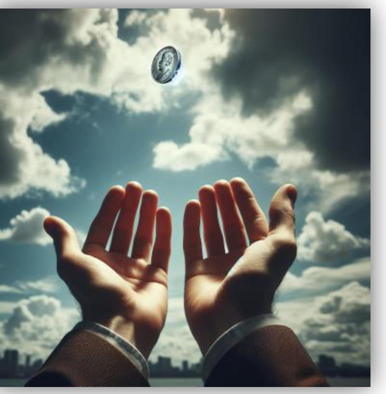
- The uncertainty of a statement, process, or observation is generally associated with the phenomenon of **randomness.** However, "random" can mean many things.
- 👉 A **random process** leads to one of several mutually exclusive outcomes. It's uncertain which outcome will be observed.
- Will it rain tomorrow or not?
- Uncertainty doesn't mean we don't know what *could* be observed:
  - You flip a coin: It will be either heads or tails.
  - You ask a "random" person about their highest educational qualification: It will be either no degree, a high school certificate, a college certificate, etc.
- We just don't know *which* of the possibilities will be observed.
  - Heads or tails?
  - Which educational qualification?
- In the context of a random process, randomness doesn't necessarily have to be understood as arbitrariness or an erratic phenomenon. Often, and more appropriately, it simply means **incomplete knowledge (epistemic uncertainty)**!

RANDOM VARIABLE

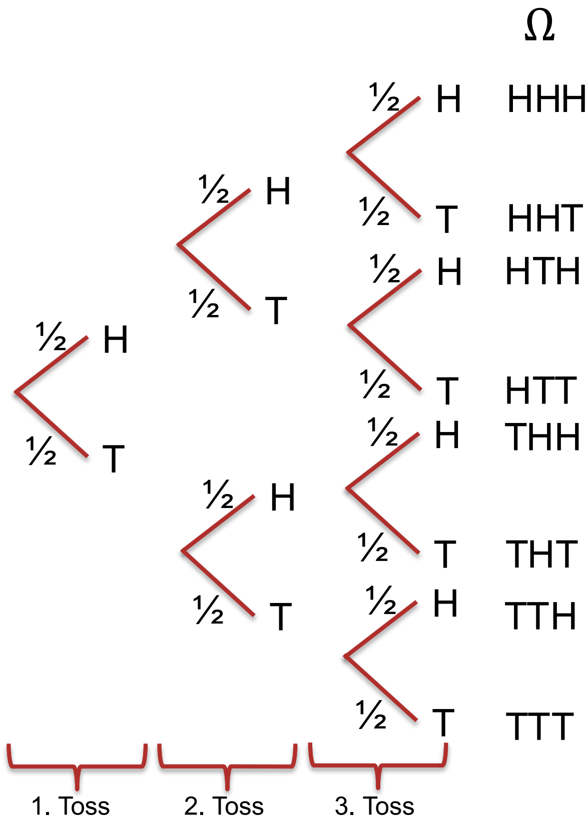
- A random process leads to one of several mutually exclusive **outcomes**. It's uncertain which outcome will be observed.
- **A random variable**, is a mathematical concept that represents a quantity or object whose value is determined by a random process. It **formalizes the uncertain or unpredictable nature** of such values.
- To formally describe random processes, we need suitable fundamental concepts and corresponding notation.
- **Formal Definition:** A random variable X is a measurable function $X: \Omega \rightarrow E$, where:
  - **Ω** is the (sample space), **a set of all possible outcomes** of the random process being modelled.
  - **E** is a measurable (state space), **representing the set of possible values** that the random variable can take on.
  - Can be discrete (e.g. True, False, Red, blue) or continuous (e.g. Real value between 0 and 10)
- **Example Dice (In this simple case Ω and E are equal):**
  - $\Omega = \{1,2,3,4,5,6\}$
  - $X = \{3,4,6,5,2,3,4,2,1,3,4,5,4,2,3,4,1,5,4,1,2,3,6,5,4,6,6\}$
    for Example: The score shown on the top face in each try

EXAMPLE COIN

- Ω is the sample space, E or X is the state space
- You flip a coin once:
  - Outcome for X: either Heads (H) or Tails (T).
  - Sample Space: $\Omega = \{H, T\}$
- You flip the coin twice:
  - Outcome X: either $(H, H), (H, T), (T, H)$, or (T, T).
  - Sample Space: $\Omega = \{(H, H), (H, T), (T, H), (T, T)\}$
- You flip the coin n times:
  - Outcome X: either $(H, \dots, H), \dots, (T, \dots, T)$.
  - Sample Space: $\Omega = \{(H, \dots, H), \dots, (T, \dots, T)\}$



EXAMPLE COIN

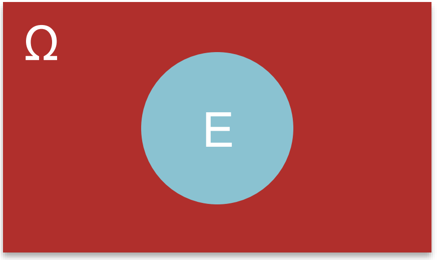


MORE EXAMPLES

- Number of correct answers in a multiple-choice exam with 80 questions:
  - **Outcome X**: Either 0, 1, 2, 3, ..., 80.
  - **Sample Space** $\Omega : \{0,1,\dots,80\}$
- You ask a person for their income (in whole Euros):
  - **Outcome X**: Either 0, 1, 2, ...
  - **Sample Space**: $\Omega = \{0,1,2,\dots\} = N_0$ (set of non-negative integers)
- You measure a person's height (in cm, with arbitrary precision or quasi-continuously):
  - **Outcome X**: Any positive real number.
  - **Sample Space**: $\Omega = \{s \in R \mid s > 0\}$

RANDOM EVENT

- A random process is completely described by a **sample space** Ω. The outcomes $X \in \Omega$ of a random process thus form the elementary states of the world and allow events to be expressed extensionally.
- A **random event** E is a collection of outcomes; that is, $E \subseteq \Omega$.
- This formal description of events as sets of outcomes allows for a compact representation of events and an operational event calculus.



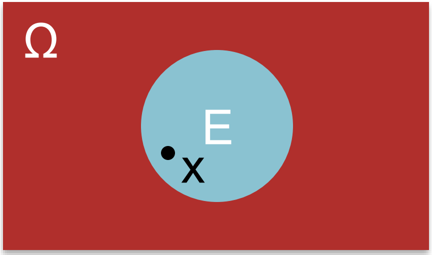
2 COIN FLIPS

- Sample Space: $\Omega = \{(H,H),(H,T),(T,H),(T,T)\}$
- Event Description:

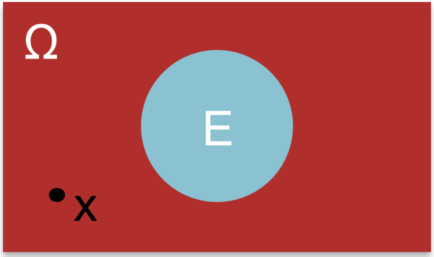
Description of e	**Set** E
First Heads, than Tail	$\{(H,T)\}$
One Tail and one Head	$\{(H,T),(T,H)\}$
First Heads	$\{(H,H),(H,T)\}$
Same Side	$\{(H,H),(T,T)\}$
At least one Head	$\{(H,H),(H,T),(T,H)\}$
No Head	$\{(T,T)\}$

OCCURRENCE OF AN EVENT

- If a random process results in a particular outcome $\mathrm{X} \in S$:
- An event E has **occurred** if $X \in E$.



- An event E has **not occurred** if $\mathrm{X} \notin E$.



2 COIN FLIPS

- We observed first Heads than Tails: $x = (H, T)$

Description of e	**Set E**	$\boldsymbol{x} \in E$
First Heads, than Tail	$\{(H, T)\}$	True
One Tail and one Head	$\{(H, T), (T, H)\}$	True
First Heads	$\{(H, H), (H, T)\}$	True
Same Side	$\{(H, H), (T, T)\}$	False
At least one Head	$\{(H, H), (H, T), (T, H)\}$	True
No Head	$\{(T, T)\}$	False

ELEMENTARY EVENT

- An **elementary event** is any (random) event that comprises only a single outcome.
- So, elementary events are of the form $E = \{s\}$, where $s \in \Omega$.
- To distinguish between s and $\{s\}$:
  - s denotes a **possible outcome** of a random process.
  - $\{s\}$ is the **event** that this particular outcome occurs.
- **Event vs. Outcome**
  - Outcomes describe which events occur and which do not.

SURE EVENT

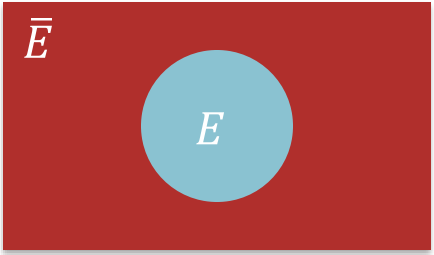
- The sample space Ω, as a set of outcomes, can itself be understood as an event and is then referred to as the **sure event** (or certain event).
  - Some $s \in \Omega$ is sure to occur.
  - Ω is the only event we can be certain or say for sure will occur.
- **Example: Coin Flip**
  - $\Omega = \{H, T\}$ is the event "Heads or Tails," which is *a priori* certain.

IMPOSSIBLE EVENT

- The empty set $\emptyset = \{\}$ is referred to as the **impossible event**.
- There is no $x \in \Omega$ such that $\mathrm{X} \in \emptyset$.
- $\emptyset$ is the only event we are certain will not occur.
- The impossible event is the complement of the sure event.
  - For example, in the coin flip scenario, $\emptyset$ can be imagined as the proposition "the coin lands on both heads and tails," which is *a priori* false or impossible.

COMPLEMENTARY EVENT

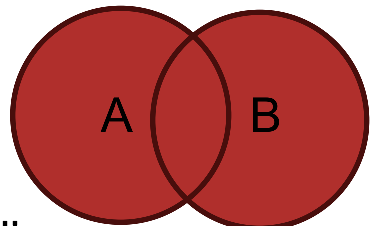
- For any event E, $\bar{E}$ represents its **complementary event**, which occurs precisely when E does not occur.
- $\bar{E} \coloneqq S \setminus E = \{s \in \Omega \mid s \notin E\}$



- **Example: Two Coin Flips**
  - $E = \{(H,H),(H,T)\} \rightarrow \bar{E} = \{(T,H),(T,T)\}$
    - E means "First flip is Heads."
    - $\bar{E}$ means "First flip is Tails."
  - $E = \{(H,H),(T,T)\} \rightarrow \bar{E} = \{(H,T),(T,H)\}$
    - E means "Both flips are the same."
    - $\bar{E}$ means "The flips are different."

COMPOSITION OF EVENTS - UNION

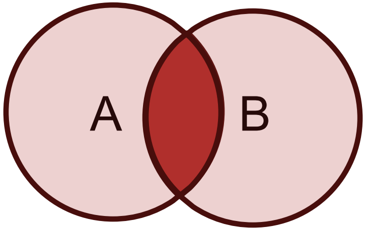
- **Union of Two Events (Disjunction of Two Events)**
- The event $A \cup B$ occurs if and only if A or B occurs.
- $A \cup B \coloneqq x \in \Omega \mid x \in A \; or \; x \in B\}$
- $A \cup B$ is the event "A or B":



- **Example: Two Coin Flips**
- $\{(H,H)\} \cup \{(H,T)\} = \{(H,H),(H,T)\}$
  - This means "First Heads, then Heads again OR First Heads, then Tails" which simplifies to "First flip is Heads."

COMPOSITION OF EVENTS - INTERSECTION

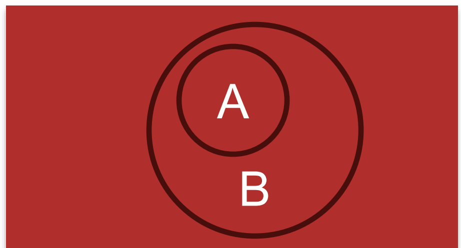
- Intersection of Two Events (Conjunction of Two Events)
- The event $A \cap B$ occurs if and only if both A and B occur.
  - $A \cap B \coloneqq \{x \in \Omega \mid x \in A \; and \; x \in B\}$
  - $A \cap B$ is the event "A and B":



- **Example: Two Coin Flips**
- $\{(H,H),(T,T)\} \cap \{(H,H),(H,T)\} = \{(H,H)\}$
  - This means "Both sides are the same AND first flip is Heads," which simplifies to "Two Heads."

RELATIONSHIP BETWEEN EVENTS - IMPLICATION

- An event A implies an event B if $A \subseteq B$.
- $A \subseteq B$: For every $x \in A$, it also holds that $x \in B$.
  $\Rightarrow$ So, if A occurs, B must also occur:



- Example of two coin tosses: $\{(H,H)\} \subseteq \{(H,H),(T,T)\}$
  - The occurrence of the event 'Two Heads' implies the occurrence of the event 'Same Sides'.
- **Remark:** Implication as a relation, i.e., $A \subseteq B$ is true, is equivalent to the condition $A \cap B = A$ or $A \cup B = B$.

PROBABILITY MEASURE

- Fundamental concepts like random process, sample space, and event enable a formal description of what is uncertain.
- The **probability of an event** describes how uncertain or certain the occurrence of an event is.

👉 **Probability Measure:** A **probability measure** $P(\cdot)$ assigns a real number $P(X = E)$ to each event E, which describes how certain the occurrence of an event is.

- The value $P(X = E)$ is called the **probability of event $\boldsymbol{E}$**.

INTERPRETING PROBABILITY

- I'm flipping a coin: The probability of heads is 50%.
- What exactly does that mean?
- There are different interpretations of the concept of probability.
  - **Subjective Interpretation of Probability:** This measures the degree of subjective belief that the event "heads" will occur.
  - **Frequentist Interpretation of Probability:** In 50% of all cases, the event "heads" will occur.

SUBJECTIVE PROBABILITIES

- **Subjective Perspective:** For an individual, a random process represents a situation of epistemic uncertainty.
  - **Coin Toss Example:** I'm uncertain whether it will be heads or tails.
  - **Weather Example:** Someone is uncertain whether it will rain tomorrow or not.

👉 **Subjective Probability:** Probability expresses a degree of personal conviction regarding how certain the occurrence of an event is.

- **Weather Forecast Example:**
  - My probability for rain is 50%.
  - A meteorologist's (presumably better informed) probability is 80%.

OPERATIONALIZING SUBJECTIVE PROBABILITY

- The concept of subjective probability can also be operationally defined (de Finetti, Ramsey):
  - A **bet** $\boldsymbol{B}[\boldsymbol{E}]$ guarantees a win of 1 Euro if event **E** occurs, and 0 Euro if it doesn't.
  - An **agent** states their fair price for this bet as 0.71 Euro. This means they're willing to both **accept** and **offer** the bet at this price.
  - This price, $\boldsymbol{pr}[\boldsymbol{E}] = \mathbf{0.71}$, also referred to as the **prevision of E**, can be interpreted as the agent's **subjective probability** of 71%.
  - An agent is considered **rational** if their previsions for different events are **consistent** or **coherent**. This means that a "Dutch book" isn't possible, where a system of various bets would lead the agent to a (nearly) certain loss.
  - **A Consequence:**

$$pr[E] + pr[\bar{E}] = pr[S] = 1$$

FREQUENTIST PROBABILITIES

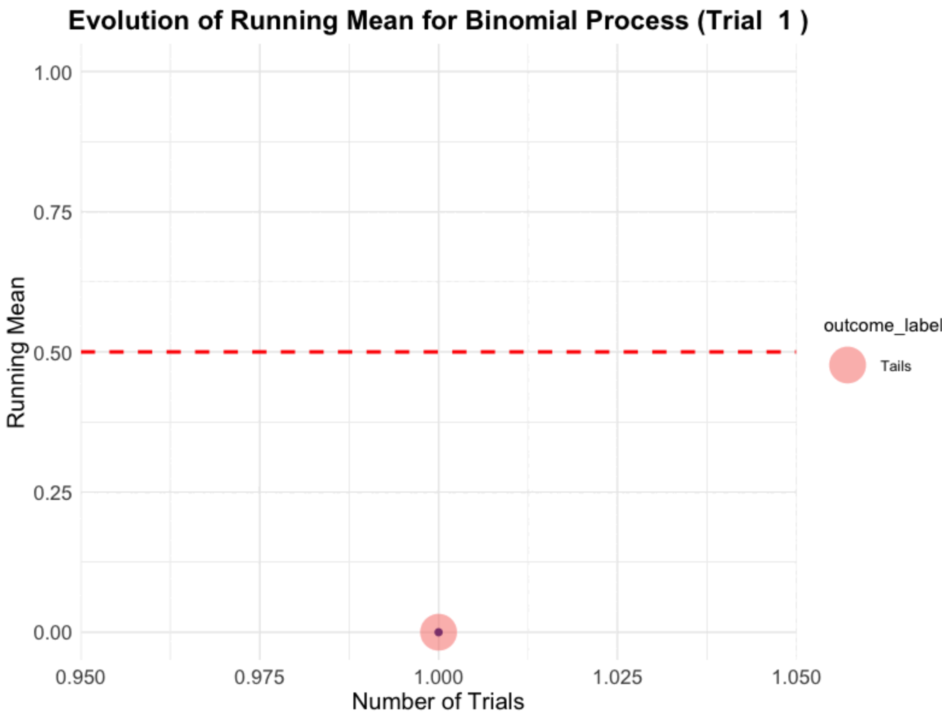
- **Another Perspective:** The hypothetical repetition of a random process doesn't necessarily lead to the same outcome.
  - **Coin Toss Example:** No causal reasons are discernible why a toss would result in either heads or tails. If the coin is tossed multiple times, both heads and tails appear.
- **Random Experiment:** A **random experiment** is a random process that can be repeated (at least in principle) any number of times under the same conditions.
  - The concept of a random experiment is problematic in that all random processes are **singular**—they can only happen once.
  - Thus, a true random experiment can't exist. However, it's a **useful concept** that captures the phenomenon of seemingly repeatable random processes.
  - Perhaps it would be better to speak of a **random phenomenon** in contrast to sequences of inherently singular random processes.
  - In practice, however, sequences of very similar random processes, such as coin tosses or draws from a population, are **approximately understood as the realizations of a random experiment**.

FREQUENTIST PROBABILITIES

👉**Frequentist Concept of Probability:** The **probability of an event** is the expected **relative frequency** with which a random experiment leads to the occurrence of that event, if the random experiment is **repeated infinitely often**.

- **Coin Toss Example:** P(Heads) = 50% means that in half of all cases, the event "heads" would occur if you could toss the same coin an infinite number of times under otherwise identical conditions.
- **Problematic:** It's not necessarily guaranteed that an infinite sequence of events will converge to a **stable frequency**.
- **Coin Example:** The sequence H,T,T,H,H,H,H,T,T,T,T,T,T,T,T... does not converge to a stable frequency.

FREQUENTIST PROBABILITIES



FURTHER CONCEPTS OF PROBABILITY

- Many other (controversial) interpretations, refinements, and distinctions of the concept of probability exist.
  - **Propensity (Popper):**
    - Probabilities express a certain **causal tendency towards realization**. (However, Popper doesn't truly explicate the concept further.)
    - Example: A loaded die with more weight on the 2 will be more likely to result in a 5 than a fair die.
  - **Qualitative Probabilities:**
    - Probability is a **consistent relation between events**.
    - Example: If event A is more probable than B, and B is more probable than C, then A should be more probable than C.
- Less controversial, however, are the purely **formal properties of probabilities**.
  - There's even a proposal to **not strive for any interpretation of probability** at all, but rather to leave probability as a **primitive fundamental concept** and only investigate its formal properties. This is known as **"mathematical probability."**

FORMAL AXIOMS OF PROBABILITY

- These formal properties are summarized as axioms, from which further properties or rules are deduced:
  - **Axiom I:** The probability of an event lies between 0 and 1:

$$0 \leq P(E) \leq 1$$

  - **Axiom II:** The probability of a certain event is 1:

$$P(S) = 1$$

  - **Axiom III:** The probability of the union of two disjoint events is equal to the sum of the probabilities of the individual events:

$$A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B)$$

CALCULATION RULES

– Further important properties and calculation rules for probabilities arise from the axioms.

👉 **The Complement Rule**

$$P(\bar{E}) = 1 - P(E)$$

– This holds true because: $E \cap \bar{E} = \emptyset$ (E and its complement are disjoint) and $E \cup \bar{E} = S$ (The union of E and its complement is the certain event)

– **Consequence:**

$$P(E) + P(\bar{E}) \stackrel{\text{def}}{=} P(E \cup \bar{E}) \stackrel{\text{def}}{=} 1$$

CALCULATION RULES

- **Impossible is Maximally Improbable**

$$P(\emptyset) = 0$$

  - The impossible event never occurs, and is thus maximally improbable.
  - However, for an event $E \neq \emptyset$ with $P(E) = 0$, it's not necessarily implied that E will never occur. It's only expected to *almost certainly* not occur.

- **General Addition Rule**

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

- If A and B are disjoint (i.e., $A \cap B = \emptyset$), this simplifies to Axiom III:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) + 0$$

- Axiom III is therefore also called the Addition Rule.

CALCULATION RULES

- **Multiplication Rule for independent events:** If events A and B are **independent**, then the probability of both A and B occurring is the product of their individual probabilities:

$$P(A \cap B) = P(A) \cdot P(B)$$

- **Example:** If you flip a fair coin twice, the probability of getting heads on both flips is:

$$P(Heads\ on\ 1st\ flip \cap Heads\ on\ 2nd\ flip)$$
$$= P(Heads\ on\ 1st\ flip) \cdot P(Heads\ on\ 2nd\ flip) = 0.5 \cdot 0.5 = 0.25$$

- This rule applies specifically when the occurrence of one event **does not influence** the probability of the other event occurring.

CALCULATION RULES

👉Decomposition

$$P(A) = P(A \cap B) + P(A \cap \bar{B})$$

– Decomposition means:

$$A = (A \cap B) \cup (A \cap \bar{B})$$

- **Monotonicity**

$$A \subseteq B \Rightarrow P(A) \leq P(B)$$

– Monotonicity here means: If event A implies event B, then B is at least as probable as A.

DETERMINING A PROBABILITY MEASURE

- For a **discrete sample space S**, it's sufficient to define the probabilities for the **elementary events**.
- **Notation:** p_s is short for $P(\{s\})$, meaning:

$$p_s := P(\{s\})$$

- It then suffices to check whether:

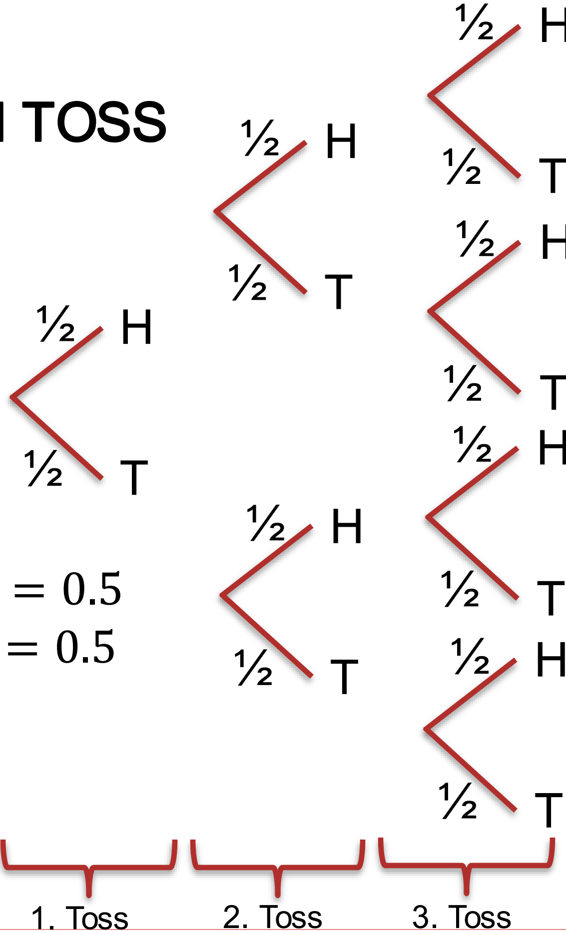
$$0 \leq p_s \leq 1, s \in S$$

$$\sum_{s \in S} p_s = 1$$

- The probability for any arbitrary event E is then:

$$P(E) = \sum_{s \in E} p_s$$

EXAMPLE COIN TOSS



HHH *½ x ½ x ½ =1/8*

HHT *½ x ½ x ½ =1/8*

HTH *½ x ½ x ½ =1/8*

HTT *½ x ½ x ½ =1/8*

THH *½ x ½ x ½ =1/8*

THT *½ x ½ x ½ =1/8*

TTH *½ x ½ x ½ =1/8*

TTT *½ x ½ x ½ =1/8*

$$P(X = H) = 0.5$$
$$P(X = T) = 0.5$$

EXAMPLE COIN TOSS

Result s	**(HHH)**	**(HHT)**	**(HTH)**	**(HTT)**	**(THH)**	**(THT)**	**(TTH)**	**(TTT)**
Probability p_s	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

- $E = \{(HHH), (TTT)\}$
- $P(E) = p_{HHH} + p_{TTT} = \frac{1}{8} + \frac{1}{8} = \frac{2}{8} = 25\%$

- We can also determin the probability measure:
  - P(X=(HHH)) = 1, and set the rest to zero as long as they sum to 1!

SYMMETRY CONSIDERATIONS IN PROBABILITY

- Symmetry considerations can help us determine probabilities.

👉 **Principle of Indifference (Principle of Insufficient Reason):** If there are no known reasons why one of two events, A and B, is more certain than the other, then the two events should be considered **equally probable**:

$$P(A) = P(B)$$

- This principle is **normative**: one *should* assume equiprobability.
  - Equiprobability is meant to express an **epistemic state of greatest uncertainty** (maximally uninformed).

LAPLACE EXPERIMENT

- The Principle of Indifference is attributed, among others, to Laplace and leads to the "classical" concept of probability known as **Laplace Probability**.
- **Laplace Experiment:** A Laplace experiment is a random process with a **finite number of outcomes**, i.e., $|S| < \infty$, such that all elementary events are **equally probable**:

$$p_s = \frac{1}{|S|}$$

- **Laplace Probability:** In Laplace experiments, the probability of an event E is equal to the number of outcomes favorable to this event divided by the total number of possible outcomes.
- **Laplace Formula (Counting Rule):**

$$P(E) = \frac{|E|}{|S|}$$

THE NUMBER OF HEADS AFTER 3 TOSSES

Outcome	X=Number of Heads
HHH	3
HHT	2
HTH	2
HTT	1
THH	2
THT	1
TTH	1
TTT	0

- Ω = {0,1,2,3}
- But this time the outcomes are **NOT** all equally likely.
- P(X = 3) = 1/8
- P(X = 2) = 3/8
- P(X = 1) = 3/8
- P(X = 0) = 1/8

- $P(E) = P(X = 2) = \frac{|E|}{|S|} = \frac{|\{(HHT),(HTH),(THH)\}|}{|\{(HHH),(HHT),(HTH),(HTT),(THH),(THT),(TTH),(TTT)\}|} = \frac{3}{8}$

- What would happen if we reformulate X as “The number of Heads larger 0 after 3 tosses”?

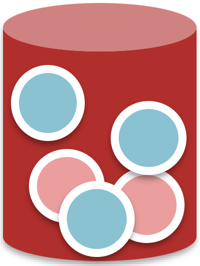
URN EXAMPLE

- A ball is drawn "randomly" from the urn.
- What color is it?
- **Modeling using the Sample Space**

$$\Omega = \{\boldsymbol{b}, \boldsymbol{r}\}:$$

- **To be determined:**

  $p_b = P(\{b\})$ and $p_r = 1 - p_b = P(\{r\})$



- The assumption of a Laplace experiment seems implausible, as it appears more certain that a blue ball will be drawn than a red ball: $p_b > p_r$

URN EXAMPLE

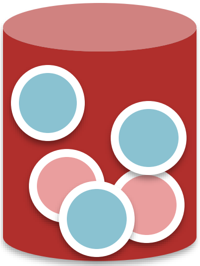
- **Improved Sample Space for Symmetry:**

$$\widetilde{\Omega} = \{b_1, b_2, b_3, r_1, r_2\}$$

- **Assumption of a Laplace experiment:**

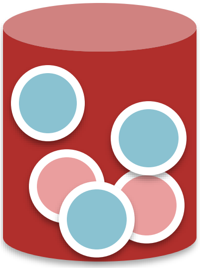
$$p_{\tilde{s}} = \frac{1}{5}, \tilde{s} \in \widetilde{\Omega}$$

$$p_b = p_{b_1} + p_{b_2} + p_{b_3} = \frac{5}{3}, p_r = 1 - p_b$$
$$= p_{r_1} + p_{r_2} = \frac{1}{5}$$



URN EXAMPLE

- Repeated drawing from an urn is a crucial abstract model for **sampling from a population**.
- A population survey, a topic selection, a book selection in a Linbrary where we "randomly" select and interview individuals, customers or text samples, can be modeled as drawing from an urn.
- **Example Election:** Does a person vote for the SPD (modeled as a red ball) or not (modeled as a blue ball)?



SEE YA‘LL NEXT WEEK!

Dr. Ing. Andreas Niekler

Computational Humanities

Paulinum, Augustusplatz 10, Raum P 616, 04109 Leipzig

T +49 341 97-32239

andreas.niekler@uni-leipzig.de

https://www.uni-leipzig.de/personenprofil/mitarbeiter/dr-andreas-niekler