Inductive arguments, inductive fallacies and related biases.
Part II: probability and causality

Margherita Benzi
Università del Piemonte Orientale
Topics

We will discuss three problems

• Probabilistic fallacies
  • The gambler’s fallacy
  • The conjunction fallacy

• Causal fallacies
  • Simpson’s paradox
But before starting...

Has anyone found about the Baader Meinh of effect?
What is the associated heuristic?
Other biases?
• I. The gambler’s fallacy
To introduce the subject, let’s look a video. It is a fragment of the film “Rosenkrantz and Guildestern are dead” (1990); it is directed by Tom Stoppard, who wrote the the tragicomedy (staged for the first time in 1966). The actors are Tim Roth and Gary Oldman.

https://www.youtube.com/watch?v=KchhSIIVwMdY
Questions

• Why is Guildenstern worrying? (or: is there anything strange and a bit disturbing?)

• Why does Guildenstern lend two coins to Rosenkrantz?

• GUIL: “A weaker man might be moved to re-examine his faith, if in nothing else at least in the law of probability”

• But he quotes two “laws of probabilities” that point to opposite directions. What laws are they?
“Law” I (The “law of averages”)

• GUIL (understanding): Games. (Flips a coin.) The law of averages, if I have got this right, means that if six monkeys were thrown up in the air for long enough they would land on their tails about as often as they would land on their –

• (Comment: hilarious confusion of 3 different ‘chance set-up’s: coins, dice and monkeys in the “Infinite monkeys theorem”)

[Image of a chimpanzee typing at a typewriter]
‘Law’ 2 (equal chance)

- GUIL: Time has stopped dead, and the single experience of one coin being spun once has been repeated 156 times... (He flips a coin, looks at it, tosses it to ROS.) On the whole, doubtful. Or: a spectacular vindication of the principle that each individual coin spun individually (he spins one) is as likely to come down heads as tails and therefore should cause no surprise that each individual time it does.

- Why there is a tension between the “law of averages” and this law?
- Which one is wrong?
• Answer: the s.c. “law of averages” is a fallacy: the gambler’s fallacy, aka “The gambler’s ruin”, or “Montecarlo’s fallacy”, or “The fallacy of the maturity of chances”:

• This fallacy is the erroneous belief that if a possible outcome of a random sequence does not occur for a long time, then it is more likely that it will occur in the next trial, or at least once in the next few trials.

• Why is it a fallacy?
• Kolmogorov’s Axioms

• Let $A$ be an event (or a sentence describing an event)-

• For any $A$, $0 \leq \Pr(A) \leq 1$
• If $A$ is certain, $\Pr(A) = 1$
• If $A$ and $B$ are incompatible, $\Pr(A \text{ or } B) = \Pr(A) + \Pr(B)$
• Conditional probability
  \[ \text{Pr} (A \mid B) \text{ “The probability of } A \text{ given } B \text{”} \]

• Probabilistic Independence
  \[ A \text{ is probabilistically independent of } B \text{ if } \text{Prob}(A \mid B) = \text{Prob} (A) \]

• The probability of \( A \) - assumed that \( B \) - is not different from the probability of \( A \) in general.

• In other words, \( B \) does not alter the probability of \( A \).
If we consider a genuine chance set up, i.e. something that produces genuine random events, like tossing an unbiased coin, one outcome does not influence the next one. Therefore, they are probabilistically independent.

(by contrast, knowing that a child is six years old influences the likelihood that she can read; therefore, the age of the child and her ability to read are not probabilistically independent)
• If a coin, or a roulette, or a deck of cards are unbiased, one outcome does not influences the next one. Let us a consider the tossing of a coin.
• $H = \text{head}, \ T = \text{tail}$.
• If the coin is fair
  \[ \Pr(T) = \Pr(H) = \frac{1}{2} \]
  \[ \Pr(T | H) = \Pr(T) = \frac{1}{2} \]
• We have also
  \[ \Pr(H | H) = \frac{1}{2}, \]
• And
  \[ \Pr(H | \text{HHHHHHHHHHHHHHH}) = \Pr(T | \text{HHHHHHHHHHHHHHH}) = \frac{1}{2}. \]
• In other terms, the probability of $\text{HHHHHHHHHHHHHHH}$ is the same as the probability of $\text{HHHHHHHHHHHHHT}$
• If we know for certain that the coin is fair, the fact of having obtained a long sequence of heads should not lower our degree of belief that the next toss will have the outcome ‘head’.

• But even if we know this fact (“the coin – or the roulette wheel – “has no memory”): in the movie, also Rosenkrantz has no memory) we often feel strongly inclined to think that, the longer the past sequence of heads has been, the more probable is the outcome ‘tail of the next toss’.
• Perhaps the most famous example of the gambler’s fallacy occurred in a game of roulette at the Monte Carlo Casino on August 18, 1913, when the ball fell in black 26 times in a row. This was an extremely uncommon occurrence, with a probability of around 1 in 136.8 million. Gamblers lost millions of francs betting against black, reasoning incorrectly that the streak was causing an imbalance in the randomness of the wheel, and that it had to be followed by a long streak of red. [Wikipedia]
• But why we think that

$\Pr (HHHHHHHHHHHHHHHT) > \Pr (HHHHHHHHHHHHHHHHH)$?

A plausible answer is that the first sequence is more similar to our notion of ‘random sequence’; if this explanation is correct, we have, again, an instance of the availability heuristics.
II. The conjunction fallacy
• Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations when she was university student.

• Choose the most probable hypotheses:
  (a) Linda is a bank teller
  (b) Linda is a bank teller and is active in the feminist movement

• Tversky and Kahneman discovered that people were strongly oriented to judge (b) more probable than (a) (in particular, in a sample of 142 students who were asked to assess the comparative probability of b) vs a), 85% judged b) more probable than a).
  This answer is false, according to probability calculus.
• Conjunction rule

for any pair of events A and B the probability of their conjunction can never be higher than the probability of any of them alone.

\[ \Pr (A \land B) \leq \Pr (A), \Pr (B). \]
• According to Prob. Calc. the probability of the conjunction of A and B can never be greater than the probability of any of its conjuncts

• examples the probability that I met a Frenchman today cannot be smaller than the probability that I meet a Frenchman wearing sun glasses;

• the probability that going out from this building we meet a lady with a dog cannot be lower than the probability that we meet a lady with a dog that is a labrador).

• Why do people answer that (b) (FT) is more probable than (a)(T)?
• What heuristics could be the cause of the conjunction fallacy, and why?
• Do you think that the more frequent answer is irrational, or completely wrong?
• If not, why?
• Violations of rationality
• Psychology, philosophy, economy
• More widespread hypothesis: Availability Heuristics strikes again
• Controversy: «Linda industry»
Gigerenzer

The law is not violated, however, if participants in these studies understand the word “probability” in a sense different from the one assigned to it by modern probability theory. There is similarly no violation if B is interpreted to mean $T \land \neg F$, or is interpreted in any way other than as a conjunct of $T \land F$. 
“According to Hertwig and Gigerenzer (1999, p. 278), subjects are even urged to choose a non-mathematical interpretation because of Paul Grice’s relevance maxim. This maxim is a conversational rule usually taken to be followed by participants in a dialogue. It says, roughly, that one’s contributions have to be relevant to the topic and goal of the conversation. Applied to the case at hand, this means that every part of the experimenters’ instruction is relevant. On the mathematical interpretation of ‘probable’, however, the description of Linda is irrelevant to the question subjects are supposed to answer. Hence, people might choose a reading of ‘probable’ not captured by probability theory because, otherwise, the personality sketch had to be considered idle.” [Mark Siebel. "There’s something about Linda" Siebel, Mark. "There’s something about Linda: Probability, coherence and rationality." In First Salzburg workshop on paradigms of cognition, Salzburg. 2002.]
• Likelihood vs probability

• Pr \( (E|H) \) vs Pr\( (H|E) \)

• Confirmation vs Probability

(Crupi, tentori e Russo 2013)

(The portrait confirms ‘feminist’ more than bank teller)

The story continues
• A very short visual break
• Other statistical fallacies: false association
False statistical associations
http://www.tylervigen.com/spurious-correlations
http://www.tylervigen.com/spurious-correlations

Number people who drowned while in a swimming-pool correlates with Power generated by US nuclear power plants
Per capita consumption of mozzarella cheese correlates with Civil engineering doctorates awarded
Per capita cheese consumption correlates with the number of people who died by becoming tangled in their bedsheets.

Correlation: 94.71% (r=0.947091)

Data sources: U.S. Department of Agriculture and Centers for Disease Control & Prevention

http://www.tylervigen.com/spurious-correlations
• Part II

Causal fallacies
Causal fallacies are inductive fallacies that result in wrong causal judgements.

many subspecies

‘post hoc ergo propter hoc’

‘cum hoc ergo propter hoc’

‘non causa pro causa’

a popular fallacy about a fallacy – the term is often used as a generalization over the preceding two, but the fallacy of ‘non causa pro causa’ was not born as an inductive fallacy, but as a logical one.
• Paradosso di Simpson
Nel campione generale la proporzione di guariti tra chi prende la medicina è più alta della percentuale di guariti tra chi non prende la medicina. La med. sembra efficace

<table>
<thead>
<tr>
<th>Popolaz.</th>
<th>Guariti</th>
<th>Non Guariti</th>
<th>Totale</th>
<th>% guariti</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicina</td>
<td>20</td>
<td>20</td>
<td>40</td>
<td>50%</td>
</tr>
<tr>
<td>No Medicina</td>
<td>16</td>
<td>24</td>
<td>40</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>44</td>
<td>80</td>
<td></td>
</tr>
</tbody>
</table>
Tra le donne la proporzione di guarite tra chi prende la medicina è più *bassa* della percentuale di guarite tra chi non prende la medicina. La med. *non* sembra efficace

<table>
<thead>
<tr>
<th>Femmine</th>
<th>Guarite</th>
<th>Non Guarite</th>
<th>Totale</th>
<th>% guarite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicina</td>
<td>2</td>
<td>8</td>
<td>10</td>
<td>20%</td>
</tr>
<tr>
<td>No Medicina</td>
<td>9</td>
<td>21</td>
<td>30</td>
<td>30%</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>29</td>
<td>40</td>
<td></td>
</tr>
</tbody>
</table>
Tra i maschi la proporzione di guariti tra chi prende la medicina è più *bassa* della percentuale di guariti tra chi non prende la medicina. La med. *non* sembra efficace

<table>
<thead>
<tr>
<th>Maschi</th>
<th>Guariti</th>
<th>Non Guariti</th>
<th>Totale</th>
<th>% guariti</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medicina</strong></td>
<td>18</td>
<td>12</td>
<td>30</td>
<td>60%</td>
</tr>
<tr>
<td><strong>No Medicina</strong></td>
<td>7</td>
<td>3</td>
<td>10</td>
<td>70%</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>15</td>
<td>40</td>
<td></td>
</tr>
</tbody>
</table>
Simpson’s paradox

A drug is inactive in the sub-populations

But *looks* beneficial wrt the whole population
Why?

Different proportions and dimensions

Man recover no matter how

(gender as confounder)
La *graduate school* di Berkeley fu accusata di attuare una politica di ammissione che discriminava le donne. Di conseguenza, si poneva la seguente questione: ‘Il fatto di essere una donna è realmente una causa di non ammissione a Berkeley?’. L’accusa appariva fondata sulla base dei dati probabilistici: la probabilità di essere accettati era molto più alta per i maschi di quanto non lo fosse per le femmine. Tuttavia, Bickel, Hammel e O’Connell [...] esaminarono i dati con maggiore attenzione e scoprirono che questo cessava di valere quando ripartivano [le ammissioni] per dipartimento. Nella maggior parte degli ottantacinque dipartimenti la probabilità che una donna fosse ammessa era esattamente la stessa di un uomo, ed in alcuni dipartimenti era addirittura più alta. [Cartwright 1983, p.37]
Other inductive causal fallacies

• “jumping” from associations or from temporal successions to causal relations.
• *cum hoc ergo propter hoc* (constant conjunction)
• every time Mario comes to the picnic, it rains;
• *post hoc ergo propter hoc*
• the rooster causes the sunrise because it starts singing just before it.
• *inverse causation*
• Scratching is the cause of the itching of my hand
• *Selection bias:*
• we tend to think that some activities, as being an Olympic swimmer, or a professional dancer, have a very strong effect on the physical constitution of the swimmers or the dancers, without considering that if one has not a certain type of physical constitution it is very difficult that he or she becomes an Olympic swimmer or a professional dancer.
• men tended to apply to the departments that are the hardest to get into

• men tended to apply to departments that were easier to get into

• So women were rejected more than men.