Probability Theory for EOR

Conditioning is a useful tool

Conditioning as a problem-solving tool

Example: a simple game

You play a game where you start with 1 dollar. You flip a three sided fair die.

1. Throw 1: you lose one dollar.
2. Throw 2: you keep your dollar.
3. Throw 3: you win one dollar.

What is the probability that this is a never-ending game (NEG)?

Solution via conditioning: conditional on the information you know

- The complement of (NEG) is the probability that you lose all your money.
- Define D as NEG^C.
- At the start of the game, you have one dollar. Options
  - Lose the dollar, then D occurs
  - Keep the dollar, then the exact same game repeats
  - Win a dollar, take a moment!

$$\begin{aligned} P(D) =& P(D|\{lose\ the\ dollar\})P(\{lose\ the\ dollar\}) \\ &+ P(D|\{keep\ the\ dollar\})P(\{keep\ the\ dollar\}) \\ &+ P(D|\{win\ the\ dollar\})P(\{win\ the\ dollar\}). \end{aligned}$$

$$P(D) = \frac{1}{3} + \frac{1}{3}P(D) + \frac{1}{3}P(D)^2.$$

Solving for $P(D)$ yields $P(D) = 1$, so $P(NEG) = 0$.
Sometiems, it is easier to work with conditional probability.

Pitfalls and Paradoxes: I

Problem: the effect of studying econometrics on wages

You want to know the effect of studying econometrics on wages.

- MSc in Econometrics (E, $\{$MSc in Econometrics$\}$);
- Two types of wages:
  high (H, $\{$earn high salary$\}$) and low (H^C, $\{$low salary$\}$).

Research shows: $P(H) = \frac{6}{100}$, $P(H|E) = \frac{60}{100}$, $P(E) = \frac{5}{100}$.

- So, should you study econometrics to get a high wage?

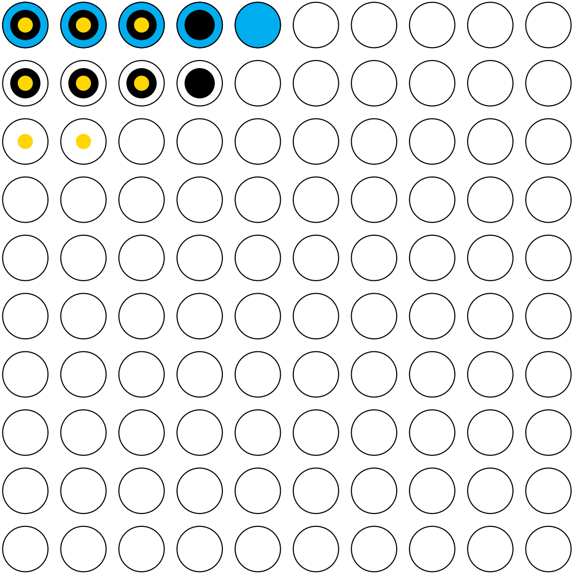
Problem: econometrics leads to a high wage?

You want to know the effect of studying econometrics on wages.

- MSc in Econometrics (E, {MSc in Econometrics});
- Two types of wages:
  high (H, {earn high salary}) and low (H^C, {low salary}).

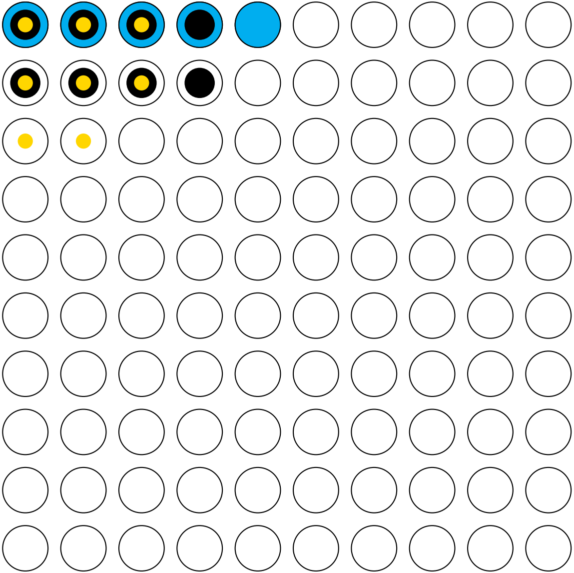
Research shows: $P(H) = \frac{6}{100}$, $P(H|E) = \frac{60}{100}$, $P(E) = \frac{5}{100}$.

Plot twist: there is an additional event ("lurking in the background"): you like mathematics (M) or not (M^C).



Sample: 100 students: E, M, H.

Sample: 100 students



$$P(H) = \frac{8}{100}$$

$$P(H|E) = \frac{3}{5}$$

$$P(H|M) = \frac{6}{8}$$

$$P(H|M,E) = \frac{3}{4}$$

$$P(H|M^C) = \frac{2}{92}$$

$$P(H|M^C,E) = 0$$

In numbers

- High income: $P(H) = \frac{8}{100}$.
- High income | E: $P(H|E) = \frac{3}{5}$
- High income | M: $P(H|M) = \frac{3}{4}$.
- High income | E and M $P(H|M, E) = \frac{3}{4}$.
- High income | not M: $P(H|M^C) = \frac{2}{92}$
- High income | E, but not M, $P(H|M^C, E) = 0$

Lessons:

- H is not independent of E.
- **H is independent of E given M. Conditional independence!**
- H is not independent of E given M^C.

- *What fits is best.*
  - *If M, it doesn't matter whether you E.*
  - *If M^C, you are better of E^C.*

Pitfalls and Paradoxes: II

Prosecutor's fallacy

In 1998, Sally Clark was tried for murder after two of her sons died shortly after birth. During the trial, an expert witness for the prosecution testified that the probability of a newborn dying of sudden infant death syndrome (SIDS) was $1/8500$, so the probability of two deaths due to SIDS in one family was $(1/8500)^2$, or about one in 73 million. Therefore, he continued, the probability of Clark's innocence was one in 73 million.

What do you think of this line of reasoning?

$P(A|B) \neq P(B|A)$!

- The expert witness says

$$P(D_1|I) \approx \frac{1}{8500}, \quad P(D_1 \cap D_2|I) \approx \frac{1}{8500^2}$$

- Then he states that *therefore*

$$P(I|D_1 \cap D_2) \approx \frac{1}{8500^2}$$

- Bayes says no!

$$P(I|D_1 \cap D_2) = \frac{P(D_1 \cap D_2|I)P(I)}{P(D_1 \cap D_2|I)P(I) + P(D_1 \cap D_2|I^C)P(I^C)}$$

If *Presumption of innocence* (namely, $P(I^C) \approx 0$), then $P(I|D_1 \cap D_2) \approx 1$.

Pitfalls and Paradoxes: III

Simpson's paradox

You have a job that consists of two tasks. You know that at both tasks, you are outperforming your colleague Jim, who is sleeping half of the time. At one point, your boss calls you in and says:

"*Bert, you're not doing well. Even Jim, who's sleeping half of the time, scores better than you. We have to let you go.*"

What happened?

Job Performance	You (Bert)		Jim	
	Task 1	Task 2	Task 1	Task 2
Fails	40	1	10	18
Successes	80	22	15	100

Denote by

- S the event that a task is a success,
- T_i be the event that the task is task i, with $i = \{1, 2\}$,
- B the event that you are performing the task,
  B^C is the event that Jim performs the task.

We see that the probability that you succeed at Task 1 and Task 2 is higher than for Jim, i.e.

$$P(S|T_1, B) \approx \frac{80}{120} > \frac{15}{25} \approx P(S|T_1, B^C)$$
$$P(S|T_2, B) \approx \frac{22}{23} > \frac{100}{118} \approx P(S|T_2, B^C)$$

Why you lost your job

Job Performance	You (Bert)		Jim	
	Task 1	Task 2	Task 1	Task 2
Fails	40	1	10	18
Successes	80	22	15	100

Aggregating across the two tasks however

$$P(S|B) \approx \frac{102}{143} < \frac{115}{143} \approx P(S|B^C)$$

Using the LOTP

$$P(S|B) = P(S|T_1, B)P(T_1|B) + P(S|T_2, B)P(T_2|B)$$
$$P(S|B^C) = P(S|T_1, B^C)P(T_1|B^C) + P(S|T_2, B^C)P(T_2|B^C)$$

- Be careful with conditional probabilities.
    I. Econometrics $\longrightarrow$ high wage: omitted conditioning variable.
    (Microeconometrics: endogeneity issue.)
    Careful, conditional probability v.s. causality.
    II. Mistaking A conditional on B with B conditional on A
    III. Averaging over events that are not comparable.