Bayes Theorem

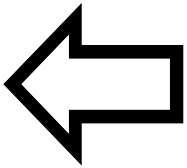
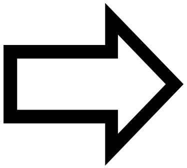
Applied To Data Analysis

Andy Grogan-Kaylor

University of Michigan

2022-03-09

How To Navigate This Presentation

- Use the  and  keys to move through the presentation.
- Press *o* for *panel overview*.

Derivation

	D=1	D=0
H=1	$P(D, H)$	$P(\mathrm{not} D, H)$
H=0	$P(D, \mathrm{not} H)$	$P(\mathrm{not} D, \mathrm{not} H)$

From the definition of conditional probability:

$$P(D|H) = P(D, H)/P(H)$$

$$P(H|D) = P(D, H)/P(D)$$

Then:

$$P(D|H)P(H) = P(D,H)$$

$$P(H|D)P(D) = P(D,H)$$

Then:

$$P(D|H)P(H) = P(H|D)P(D)$$

Bayes Theorem:

$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$

In Words:

$\text{posterior} \propto \text{likelihood} \times \text{prior}$

Example

Consider an example using 1,000 hypothetical studies. We imagine that only 10% of interventions are likely to have results. We adopt standard assumptions of adopting an α, or chance of detecting an effect when one is not there of 5%. We similarly assume 80% power β, or a 20% chance of failing to detect an effect when it is not present.

Data (D)	D=1 (effect)	D=0 (no effect)
Hypothesis (H)	100 effects	900 non-effects
H=1 (conclude effect)	80 true positives	45 false positives
H=0 (conclude no effect)	20 false negatives	855 true negatives

> With thanks to the Wikipedia article on this topic for inspiration for this example.

Visualization

Calculations

$$P(\text{effect}|\text{conclude effect}) = \frac{P(\text{conclude effect}|\text{effect})P(\text{effect})}{P(\text{conclude effect})}$$

$$= \frac{P(\text{conclude effect}|\text{effect})P(\text{effect})}{P(\text{conclude effect}|\text{effect})P(\text{effect}) + P(\text{conclude effect}|\text{no effect})P(\text{no effect})}$$

$$P(\text{effect}|\text{conclude effect}) = \frac{.8 \times .1}{.08 + .045} = .64$$

> See also Thinking Through Bayesian Ideas

Discussion

- Calculations suggest that a true effect is likely in 64% of the cases where one concludes the presence of an effect.
- Consequently, calculations suggest that 36% of the time, when one concludes there is an effect, there is actually no effect.
- Put another way, despite setting $\alpha = .05$, 36% of cases result in a false positive.
- Notice how the Bayesian approach lets us estimate the probability of different hypotheses given the data. In some cases, this may afford us the opportunity to accept our null hypothesis H_0.