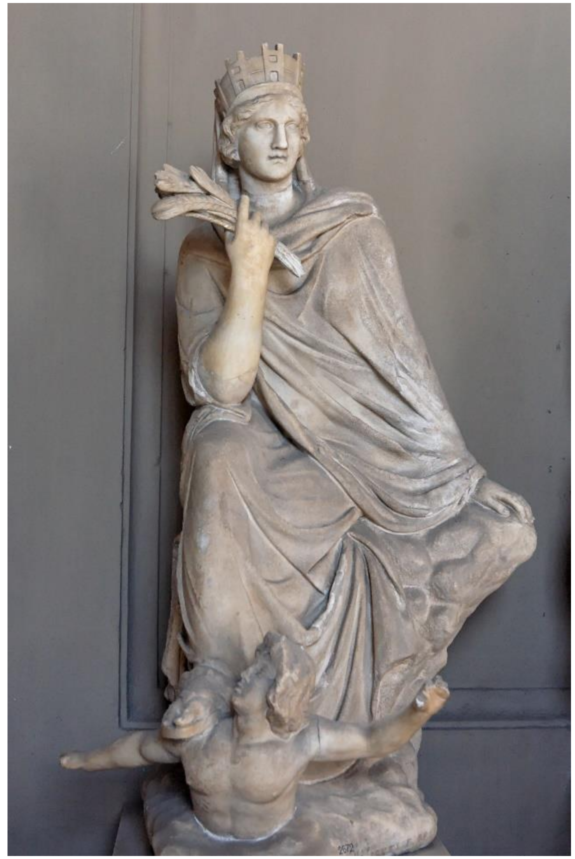
Review: Conditional Probability



The Tyche of Antioch, Roman copy of a bronze by Eutychides

SML310: Research Projects in Data Science, Fall 2019

Michael Guerzhoy

Random variables and events

- Random Variable (R.V.): a quantity we observe
  - Heads/Tails if we toss a coin
  - Amount of precipitation during a day
  - …
- Event: a set* of observations of values of R.V.s
  - The coin came up Heads more than 3 times *and* the amount of precipitation was between 2mm and 3mm
  - The student got exactly one A+ during the semester
- We (sometimes) can model the probability of events
  - P(Heads)
  - P(precip < 2mm and Tails)

Confusing notation

- We sometimes write P(E) to mean “the probability of event E happening” and sometimes P(X=x) to mean “the probability of the event ‘the value of the R.V. X is x’ happening”
  - Sometimes, we write P(X) to mean p(x) = P(X=x), the function that takes in X as a parameter and gives us the probability
- Notation we’ll try to stick to: X is a random variable (a quantity to be measured). x is a measured value/realization

Probability and events: practice

- Two R.V.s:
  - X1: the outcome of coin toss #1
  - X2: the outcome of coin toss #2
- In terms of X1 and X2, define the event “no coin came up Heads”
- What is the probability of the event?
  - Assume that the coin is fair and the tosses are independent

Conditional probability

- P(A|B): the probability that event A happened, if we know that event B happened
- We choose a Princeton student at random
- R.V.s:
  - H: height of the student
  - S: the student's favourite sport
- $P(H > 6' \mid S = \text{basketball})$
- $P(H > 6' \mid S = \text{tennis})$

events

Conditional probability

- Two R.V.s:
  - X1: the outcome of coin toss #1
  - X2: the outcome of coin toss #2
- $P(X1+X2 = 2) = ?$
- $P(X1+X2 = 2 \mid X1 = 1) = ?$
- https://en.wikipedia.org/wiki/Boy_or_Girl_paradox

Independence

- The R.V.'s X1 and X2 are independent if for all values x1 and x2

$$P(X1=x1|X2=x2) = P(X1=x1)$$

- The observed value of X2 does not influence our probability estimates for X1
  - X2 doesn't provide us with information about the value of X1

Bayes' Rule and Law of Total Probability

- $P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(A,B)}{P(B)}$
- $P(B=b) = \sum_{i=1\ldots K} P(B=b|A=a_i)P(A=a_i)$, if the possible values of A are $a_1, \ldots, a_K$