STAT 315-06: Conditioning and Bayes Formula (Section 2.8–2.10)

Luc Rey-Bellet

University of Massachusetts Amherst

luc@math.umass.edu

February 12, 2025

Two basic laws of probability

Addition rule

For any two events A and B

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

Proof: done earlier

Multiplication rule

For any two events A and B

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

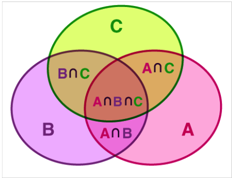
Proof:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \Longrightarrow P(A \cap B) = P(A)P(B|A)$$

Examples

- Suppose $P(A) = .5$ and $P(B) = .3$. What can you say about $P(A \cap B)$?
  What if $P(A) = .8$ and $P(B) = .5$?
- Suppose $P(A) = .5$, $P(B) = .3$ and $P(A \cap B) = .1$. Compute
  - $P(A|B)$
  - $P(B|A)$
  - $P(A|A \cup B)$
  - $P(A|A \cap B)$
  - $P(A \cap B|A \cup B)$

Addition rule for n events



- $n = 3$

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

- For general n one proves by induction

$$P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{i \neq j} P(A_i \cap A_j) + \sum_{i \neq j \neq k} P(A_i \cap A_j \cap A_k) + \cdots + (-1)^{n-1} P(\cap_{i=1}^{n} A_i)$$

Multiplication rule for n events

This can be generalized for more than 2 events.
Think of it as events occurring sequentially: A_1 occurs, and then A_2 given A_1, A_3 given A_1 and A_2, and so on...

Multiplication rule for n events

For any events $A_1, A_2, \cdots \cap A_n$

$$\begin{aligned} & P(A_1 \cap A_2 \cap \cdots A_n) \\ & = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \cdots A_{n-1}) \end{aligned}$$

Proof: For example $n = 4$

$$\begin{aligned} & P(\underline{A_1 \cap A_2 \cap A_3} \cap A_4) \\ & = P(\underline{A_1 \cap A_2} \cap A_3)P(A_4|A_1 \cap A_2 \cap A_3) \\ & = P(\underline{A_1} \cap A_2)P(A_3|A_1 \cap A_2)P(A_4|A_1 \cap A_2 \cap A_3) \\ & = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)P(A_4|A_1 \cap A_2 \cap A_3) \end{aligned}$$

Example: Sampling with or without replacement

Select 3 balls from an urn with 2 green balls, 3 red balls and 4 blue balls

- Probability that no balls drawn is green? Use conditioning

$$P(\text{no green}) = P(\text{no green on 1st})P(\text{no green on 2nd}|\text{no green on 1st}) \\ P(\text{no green on 3rd}|\text{no green on 1st and 2nd})$$

  - With replacement we have independence and $P(\text{no green}) = \frac{7}{9}\frac{7}{9}\frac{7}{9}$
  - Without replacement $P(\text{no green}) = \frac{7}{9}\frac{6}{8}\frac{5}{7}$

- Probability that you do not draw all colors? Use inclusion exclusion

$$P(\text{no green} \cup \text{no red} \cup \text{no blue}) = P(\text{no green}) + P(\text{no red}) + P(\text{no blue}$$
$$- P(\text{no green and red}) - P(\text{no green and blue}) - P(\text{no red and blue})$$
$$+ P(\text{no green red and blue})$$
$$= \frac{7}{9}\frac{6}{8}\frac{5}{7} + \frac{6}{9}\frac{5}{8}\frac{4}{7} + \frac{5}{9}\frac{4}{8}\frac{3}{7} - \frac{4}{9}\frac{3}{8}\frac{2}{7} - \frac{3}{9}\frac{2}{8}\frac{1}{7} - 0 + 0$$

Examples

- I know that of a list of ten passwords, one of the password opens your safe. I try password one at a time until the safe is open. Find the probability that you open the safe at the n^{th} trial.
- At the game of bridge 52 cards are distributed to four players (each player gets 13 cards) and the game is played two players against two players.
  Find the probability that each player receives an ace.
-

The conditioning method

Conditioning method

For any two events A and B

$$P(A) = P(A|B)P(B) + P(A|\overline{B})P(\overline{B})$$

Proof: Since $B \cup \overline{B} = S$

$$\begin{aligned} P(A) &= P(A \cap (B \cup \overline{B}) \\ &= P(A \cap B) + P(A \cap \overline{B}) \quad \text{(by the addition rule)} \\ &= P(A|B)P(B) + P(A|\overline{B})P(\overline{B}) \quad \text{(by the multiplication rule)} \end{aligned}$$

The conditioning method (cont'd)

The set $B_1, B_2, \cdots, B_n$ form a partition of S if

- $B_i \cap B_j = \emptyset$ if $i \neq j$, i.e. the sets B_i are mutually exclusive.
- $B_1 \cup B_2 \cup \cdots B_n = S$ i.e. the sets B_i cover S.

Think of $B_1, \cdots, B_n$ as n possible scenarios which are mutually exclusive and covers all the possibilities

Conditioning method: Law of total probability

If $B_1, B_2, \cdots, B_n$ form a partition of S then

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots P(A|B_n)P(B_n)$$

Proof: Same as with B and $\overline{B}$.

Examples

- Draw two cards: what is the probability the second card is an ace?
-

Example: Betting on Red at Roulette

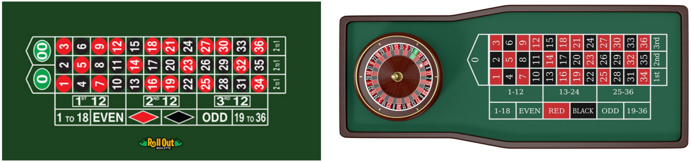
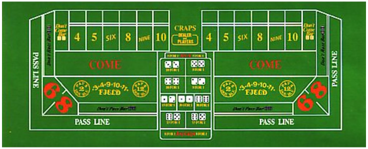


Figure: Left: Las Vegas. Right Monte-Carlo

- Las Vegas Roulette (left) has 38 numbers, 0, 00, 1 to 36. If you bet \$1 on RED you win \$1 if the number lands on a red number and lose \$1 otherwise.
- Monte-Carlo Roulette (right) has 37 numbers, 0, 1 to 36. If you bet \$1 on RED and land on 0 you go to prison. In the next round if you land on red you get \$1 back (win \$0) and if it lands on black or 0 you lose.
- Compute the probabilities to win for both cases.

Example: Craps



- Roll two regular dice:
  - If the sum is 7 or 11 you win.
  - If the sum 2, 3, or 12 you lose.
  - If the sum is 4, 5, 6 8,9, 10 this number is called "the point".
- If you get "the point" roll again until you get either "the point or a 7.
  - If you roll a 7 first you lose.
  - If you roll "the point" first you win.

Craps winning probability

Condition on the sum s in the first roll:

$$P(\mathrm{win}) = \sum_{n=2}^{12} P(\mathrm{win}|s)P(s)$$

On obtains

$$\begin{aligned}
P(\mathrm{win}) &= \underbrace{0 \times \frac{1}{36}}_{s=2} + \underbrace{0 \times \frac{2}{36}}_{s=3} + \underbrace{\frac{3}{3+6} \times \frac{3}{36}}_{s=4} + \underbrace{\frac{4}{4+6} \times \frac{4}{36}}_{s=5} \\
&\quad + \underbrace{\frac{5}{5+6} \times \frac{5}{36}}_{s=6} + \underbrace{1 \times \frac{6}{36}}_{s=7} + \underbrace{\frac{5}{5+6} \times \frac{5}{36}}_{s=8} \\
&\quad + \underbrace{\frac{4}{4+6} \times \frac{4}{36}}_{s=9} + \underbrace{\frac{3}{3+6} \times \frac{3}{36}}_{s=10} + \underbrace{1 \times \frac{2}{36}}_{s=11} + \underbrace{0 \times \frac{1}{36}}_{s=12} \\
&= \frac{244}{495} = 0.4929292\cdots
\end{aligned}$$

Bayes rule

Multiplication rule: $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$

Conditioning $P(A) = P(A|B)P(B) + P(A|\overline{B})P(\overline{B})$

Bayes' Theorem

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$
$$= \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\overline{B})P(\overline{B})}$$

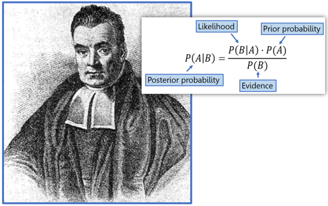
or more generally if the $B_1, \cdots, B_n$ form a partition

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + P(A|B_n)P(B_n)}$$

Bayes rule cont'd

Terminology:

- $P(B)$ = prior probability of the event B (before evidence in the form of the event A is gathered)
- $P(B|A)$ = posterior probability of the event B (after evidence in the form of the event A is gathered)
- $P(A|B)$ is the likelihood, $P(A)$ is the evidence and the ratio $\frac{P(A|B)}{P(A)}$ is the support A provided to B.



Reverend Thomas Bayes (1701-1761)

Bayes' theorem "*is to the theory of probability what the Pythagorean theorem is to geometry*". (Sir Harold Jeffreys)

Examples: Spam filter

Spam filters in your email use (repeated) Bayesian computations (see https://en.wikipedia.org/wiki/Naive_Bayes_spam_filtering)
Consider the event

$$W = \{\text{subject line contains the word } BLABLA\}$$

and we are interested in computing $P(SPAM|C)$.
It is known that 60% of all emails are SPAM, and that 2% of SPAM email contains the word *BLABLS* versus only 1% of non-SPAM emails.

$$\begin{aligned} P(SPAM|C) &= \frac{P(C|SPAM)P(SPAM)}{P(C|SPAM)P(SPAM) + P(C|notSPAM)P(notSPAM)} \\ &= \frac{\frac{2}{100}\frac{60}{100}}{\frac{2}{100}\frac{60}{100} + \frac{1}{100}\frac{40}{100}} = \frac{3}{4} \end{aligned}$$

The probbaility that the message is SPAM if it contains the words BLABLA is .75 compared to .6 for all messages.

Covid-19 tests   See the paper: https://doi.org/10.1093/ajcp/aqaa141

Various tests currently used to detect COVID-19, in 3 categories

- Molecular test detects active virus.
- Antigen tests detects associated proteins and are much faster.
- Serologic tests detect (present and past) infections markers.

How accurate a test is determined by two numbers

- Sensitivity (PPA) = Prob to test positive if infected.
- Specificity (PNA) = Prob to test negative if healthy.

FDA standards: (in ideal laboratory conditions).

- PPA $\geq$ %90
- PNA $\geq$ %95

Baseline PPA and PNA (FIND)[13]

	Molecular	Antigen	Antibody
PPA (sensitivity), %	86.14	61.70	68.44
PNA (specificity), %	95.84	98.26	95.6
Index sample type, No.	10	4	6
Company names, No.	33	3	54
Test names, No.	35	4	74
Test formats, No.	3	2	6
Targets, No.	4	4	5

PNA, percent negative agreement; PPA, percent positive agreement.

Real data from early 2020 (although there exists much more accurate tests now.)

Covid-19 tests, cont'd

More important criteria for public health

- Positive predictive value (PPV) = Prob. that a positive is infected.
- Negative predictive value (NPV)= Probab. that a negative is healthy

Define Events

- Pos = positive test. $Neg = \overline{Pos}$ = negative test
- I= infected $H = \bar{I}$= not infected (healthy)

Quantities of Interest

- Sensitivity = $P(Pos|I)$
- Specificity = $P(Neg|H)$
- PPV = $P(I|Pos)$
- NPV= $P(H|Neg)$

Covid-19 test cont'd

PPV:

$$P(I|Pos) = \frac{P(Pos|I)P(I)}{P(Pos|I)P(I) + \underbrace{P(Pos|H)}_{=1-P(Neg|H)} P(H)}$$

Assume $P(I) = 0.05$ (five percent infected)

- Molecular $PPV = \frac{.8614 \times .05}{.8614 \times .05 + .0416 \times .95} = .5214$
- Antigen $PPV = \frac{.617 \times .05}{.617 \times .05 + .0174 \times .95} = 0.6511$

Assume $P(I) = 0.2$ (twenty percent infected)

- Molecular PPV$= \frac{.8614 \times .2}{.8614 \times .2 + .0416 \times .8} = .8381$
- Antigen PPV$= \frac{.617 \times .2}{.617 \times .2 + .0174 \times .8} = 0.4699$

Covid-19 test cont'd

Lesson learned:

- A positive test is always followed by a second test!
- If $P(I)$ is "small" then even high specificity might not be good enough.
- Same issues occur with NPV. Lots of false negatives.