4. Conditional Probability

$BT\ 1.3,\ 1.4$

$P(\text{dice roll}|\text{hand signal})$

CSE 312
Winter 2017
W.L. Ruzzo
Conditional Probability
Roll one fair die.

What is the probability that the outcome is 5?

\[\frac{1}{6}\] (5 is one of 6 equally likely outcomes)

What is the probability that the outcome is 5 given that the outcome is an even number?

0 (5 isn’t even)

What is the probability that the outcome is 5 given that the outcome is an odd number?

\[\frac{1}{3}\] (3 odd outcomes are equally likely; 5 is one of them)

Formal definitions and derivations below
**Conditional probability** of $E$ given $F$: probability that $E$ occurs given that $F$ has occurred.

"Conditioning on $F"

Written as $P(E|F)$

Means “$P(E$ has happened, given $F$ observed)"

Sample space $S$ reduced to those elements consistent with $F$ (i.e. $S \cap F$)

Event space $E$ reduced to those elements consistent with $F$ (i.e. $E \cap F$)

With equally likely outcomes:

$$P(E \mid F) = \frac{\# \text{ of outcomes in } E \text{ consistent with } F}{\# \text{ of outcomes in } S \text{ consistent with } F} = \frac{|EF|}{|SF|} = \frac{|EF|}{|F|}$$

$$P(E \mid F) = \frac{|EF|}{|F|} = \frac{|EF|/|S|}{|F|/|S|} = \frac{P(EF)}{P(F)}$$
Roll one fair die. What is the probability that the outcome is 5 given that it’s odd?

E = {5} event that roll is 5
F = {1, 3, 5} event that roll is odd

Way 1 (from counting):
\[ P(E | F) = \frac{|EF|}{|F|} = \frac{|E|}{|F|} = \frac{1}{3} \]

Way 2 (from probabilities):
\[ P(E | F) = \frac{P(EF)}{P(F)} = \frac{P(E)}{P(F)} = \frac{1/6}{1/2} = \frac{1}{3} \]

Way 3 (from restricted sample space):

All outcomes are equally likely. Knowing F occurred doesn’t distort relative likelihoods of outcomes within F, so they remain equally likely. There are only 3 of them, one being in E, so
\[ P(E | F) = \frac{1}{3} \]
Roll a fair die. What is the probability that the outcome is 5?

E = \{5\} \text{ (event that roll is 5) } \quad S = \{1, 2, 3, 4, 5, 6\} \quad \text{sample space}

P(E) = \frac{|E|}{|S|} = \frac{1}{6}

What is the prob. that the outcome is 5 given that it’s even?

G = \{2, 4, 6\}

Way 1 (counting):

P(E | G) = \frac{|EG|}{|G|} = \frac{|\emptyset|}{|G|} = \frac{0}{3} = 0

Way 2 (probabilities):

P(E | G) = \frac{P(EG)}{P(G)} = \frac{P(\emptyset)}{P(G)} = \frac{0}{\frac{1}{2}} = 0

Way 3 (restricted sample space):

Outcomes are equally likely. Knowing G occurred doesn’t distort relative likelihoods of outcomes within G; they remain equally likely. There are 3 of them, none being in E, so P(E | G) = \frac{0}{3}
Suppose you flip two coins & all outcomes are equally likely.

What is the probability that both flips land on heads if…

• The first flip lands on heads?
  
  Let \( B = \{ HH \} \) and \( F = \{ HH, HT \} \)
  
  \[
P(B|F) = P(BF)/P(F) = P(\{ HH \})/P(\{ HH, HT \})
  \]
  
  \[
  = (1/4)/(2/4) = 1/2
  \]

• At least one of the two flips lands on heads?

  Let \( A = \{ HH, HT, TH \} \)

  \[
P(B|A) = \frac{|BA|}{|A|} = 1/3
  \]

• At least one of the two flips lands on tails?

  Let \( G = \{ TH, HT, TT \} \)

  \[
P(B|G) = P(BG)/P(G) = P(\emptyset)/P(G) = 0/P(G) = 0
  \]
slicing up the spam
24 emails are sent, 6 each to 4 users. 10 of the 24 emails are spam. All possible outcomes equally likely.

\[ E = \text{user \#1 receives 3 spam emails} \]

What is \( P(E) \) ?

\[
P(E) = \frac{|E|}{|S|} = \frac{\binom{10}{3} \binom{14}{3} \binom{18}{6} \binom{12}{6} \binom{6}{6}}{\binom{24}{6} \binom{18}{6} \binom{12}{6} \binom{6}{6}} \approx 0.3245
\]
24 emails are sent, 6 each to 4 users. 10 of the 24 emails are spam. All possible outcomes equally likely

\( E = \) user #1 receives 3 spam emails \( F = \) user #2 receives 6 spam emails

What is \( P(E|F) \)? [and do you expect it to be larger than \( P(E) \), or smaller?]

\[
P(E \mid F) = \frac{|EF|}{|F|} = \frac{\binom{10}{3} \binom{4}{3} \binom{14}{6} \binom{12}{6} \binom{6}{6}}{\binom{10}{6} \binom{18}{6} \binom{12}{6} \binom{6}{6}} \approx 0.0784
\]
24 emails are sent, 6 each to 4 users.
10 of the 24 emails are spam.
All possible outcomes equally likely

E = user #1 receives 3 spam emails
F = user #2 receives 6 spam emails
G = user #3 receives 5 spam emails

What is \( P(G|F) \)?

\[
P(G | F) = \frac{|GF|}{|F|} = \frac{\binom{10}{6} \binom{4}{5} \binom{14}{1} \binom{12}{6} \binom{6}{6}}{\binom{10}{6} \binom{18}{6} \binom{12}{6} \binom{6}{6}} = 0
\]
Conditional probability - general definition

General defn: $P(E \mid F) = \frac{P(EF)}{P(F)}$ where $P(F) > 0$

Holds even when outcomes are not equally likely.

Example: $S = \{\# \text{ of heads in 2 coin flips}\} = \{0, 1, 2\}$
NOT equally likely outcomes: $P(0) = P(2) = 1/4$, $P(1) = 1/2$

Q. What is prob of 2 heads (E) given at least 1 head (F)?
A. $P(EF)/P(F) = P(E)/P(F) = (1/4)/(1/4 + 1/2) = 1/3$

Same as earlier formulation of this example (of course!)
conditional probability: the chain rule

General defn: \[ P(E \mid F) = \frac{P(EF)}{P(F)} \] where \( P(F) > 0 \)

Holds even when outcomes are not equally likely.

What if \( P(F) = 0 \)?

\( P(E\mid F) \) undefined: (you can’t observe the impossible)

**Implies (when \( P(F) > 0 \)):** \( P(EF) = P(E\mid F) \cdot P(F) \) ("the chain rule")

General definition of Chain Rule:

\[
P(E_1 E_2 \cdots E_n) = P(E_1) P(E_2 \mid E_1) P(E_3 \mid E_1, E_2) \cdots P(E_n \mid E_1, E_2, \ldots, E_{n-1})
\]
chain rule example - piling cards
Deck of 52 cards randomly divided into 4 piles
13 cards per pile
Compute $P($each pile contains an ace$)$

Solution:

$E_1 = \{ \text{ in any one pile } \}$

$E_2 = \{ \text{ & in different piles } \}$

$E_3 = \{ \text{ in different piles } \}$

$E_4 = \{ \text{ all four aces in different piles } \}$

Compute $P(E_1 \cap E_2 \cap E_3 \cap E_4)$
E_1 = \{ \text{in any one pile} \}
E_2 = \{ \text{in different piles} \}
E_3 = \{ \text{in different piles} \}
E_4 = \{ \text{all four aces in different piles} \}

P(E_1E_2E_3E_4)
= P(E_1) P(E_2|E_1) P(E_3|E_1E_2) P(E_4|E_1E_2E_3)
piling cards

\[ E_1 = \{ \text{in any one pile} \} \]
\[ E_2 = \{ \text{in different piles} \} \]
\[ E_3 = \{ \text{in different piles} \} \]
\[ E_4 = \{ \text{all four aces in different piles} \} \]

\[ P(E_1E_2E_3E_4) = P(E_1) \cdot P(E_2|E_1) \cdot P(E_3|E_1E_2) \cdot P(E_4|E_1E_2E_3) \]

\[ P(E_1) = \frac{52}{52} = 1 \quad (A\heartsuit \text{ can go anywhere}) \]

\[ P(E_2|E_1) = \frac{39}{51} \quad (39 \text{ of } 51 \text{ slots not in } A\heartsuit \text{ pile}) \]

\[ P(E_3|E_1E_2) = \frac{26}{50} \quad (26 \text{ not in } A\heartsuit, A\spadesuit \text{ piles}) \]

\[ P(E_4|E_1E_2E_3) = \frac{13}{49} \quad (13 \text{ not in } A\heartsuit, A\spadesuit, A\diamondsuit \text{ piles}) \]

A conceptual trick: what's randomized?

a) randomize cards, deal sequentially into 4 piles

b) sort cards, aces first, deal randomly into empty slots among 4 piles.
piling cards

\[ E_1 = \{ \text{in any one pile} \} \]

\[ E_2 = \{ \text{in different piles} \} \]

\[ E_3 = \{ \text{in different piles} \} \]

\[ E_4 = \{ \text{all four aces in different piles} \} \]

\[
P(E_1E_2E_3E_4) = P(E_1) \cdot P(E_2|E_1) \cdot P(E_3|E_1E_2) \cdot P(E_4|E_1E_2E_3) \]

\[ = \frac{52}{52} \cdot \frac{39}{51} \cdot \frac{26}{50} \cdot \frac{13}{49} \]

\[ \approx 0.105 \]
Conditional Probability is Probability
"P(- | F)" is a probability law, i.e., satisfies the 3 axioms

Proof:
the idea is simple—the sample space contracts to F; dividing all (unconditional) probabilities by P(F) correspondingly re-normalizes the probability measure; additivity, etc., inherited – see text for details; better yet, try it!

Ex: \( P(A \cup B) \leq P(A) + P(B) \)
\[
\therefore P(A \cup B | F) \leq P(A | F) + P(B | F)
\]

Ex: \( P(A) = 1 - P(A^C) \)
\[
\therefore P(A | F) = 1 - P(A^C | F)
\]

etc.
Another Example
sending bit strings
Bit string with \( m \) 1’s and \( n \) 0’s sent on the network

   All distinct arrangements of bits equally likely

\( E = \) first bit received is a 0

\( F = k \) of first \( r \) bits received are 0’s

What’s \( P(E|F) \)?

**Solution 1 (“restricted sample space”):**

Observe:

\[ P(E|F) = P(\text{picking one of } k \text{ 0's out of } r \text{ bits}) \]

So:

\[ P(E|F) = \frac{k}{r} \]
Bit string with \( m \) 1’s and \( n \) 0’s sent on the network

All distinct arrangements of bits equally likely

E = first bit received is a 0

F = \( k \) of first \( r \) bits received are 0’s

What’s \( P(E|F) \)?

Solution 2 (counting):

\[ EF = \{ \text{(n+m)-bit strings} \mid \text{1}^{\text{st}} \text{bit} = 0 \text{ & (k-1)0's in the next (r-1)} \} \]

\[ |EF| = \binom{r-1}{k-1} \binom{n+m-r}{n-k} \]

\[ |F| = \binom{r}{k} \binom{n+m-r}{n-k} \]

\[ P(E|F) = \frac{|EF|}{|F|} = \frac{\binom{r-1}{k-1} \binom{n+m-r}{n-k}}{\binom{r}{k} \binom{n+m-r}{n-k}} = \frac{k}{r} \]

One of the many binomial identities
sending bit strings

Bit string with m 1’s and n 0’s sent on the network

All distinct arrangements of bits equally likely
E = first bit received is a 0
F = k of first r bits received are 0’s

What’s P(E|F)?

Solution 3 (more fun with conditioning):

\[
P(E) = \frac{n}{m+n}
\]

\[
P(F | E) = \frac{(n-1) \binom{m}{r-k}}{\binom{m+n-1}{r-1}}
\]

\[
P(F) = \frac{\binom{n}{k} \binom{m}{r-k}}{\binom{m+n}{r}}
\]

\[
P(E | F) = \frac{P(EF)}{P(F)} = \frac{P(F | E)P(E)}{P(F)} = \cdots = \frac{k}{r}
\]

Above eqns, plus the same binomial identity twice.

A generally useful trick:
Reversing conditioning (more to come)
Law of Total Probability
E and F are events in the sample space S

\[ E = EF \cup EF^c \]

\[ EF \cap EF^c = \emptyset \]

\[ \Rightarrow P(E) = P(EF) + P(EF^c) \]
law of total probability

\[ P(E) = P(\text{EF}) + P(\text{EF}^c) \]

\[ = P(E|F) \cdot P(F) + P(E|F^c) \cdot P(F^c) \]

\[ = P(E|F) \cdot P(F) + P(E|F^c) \cdot (1 - P(F)) \]

More generally, if \( F_1, F_2, ..., F_n \) partition \( S \) (mutually exclusive, \( \bigcup_i F_i = S, P(F_i) > 0 \)), then

\[ P(E) = \sum_i P(E|F_i) \cdot P(F_i) \]

(weighted average, conditioned on event \( F \) happening or not.)

(weighted average, conditioned on which event \( F_i \) happened)

(Analogous to reasoning by cases; both are very handy.)
Sally has 1 elective left to take: either Phys or Chem. She will get an A with probability 3/4 in Phys, with prob 3/5 in Chem. She flips a coin to decide which to take.

What is the probability that she gets an A?

Phys, Chem partition her options (mutually exclusive, exhaustive)

\[
P(A) = P(A \cap \text{Phys}) + P(A \cap \text{Chem})
\]

\[
= P(A|\text{Phys})P(\text{Phys}) + P(A|\text{Chem})P(\text{Chem})
\]

\[
= (3/4)(1/2) + (3/5)(1/2)
\]

\[
= 27/40
\]

Note that conditional probability was a means to an end in this example, not the goal itself. One reason conditional probability is important is that this is a common scenario: break a complex problem into simpler cases.
Example: Gambler’s Ruin
2 Gamblers: Alice & Bob.
A has $i$ dollars; B has $(N-i)$
Flip a coin. Heads – A wins $1; Tails – B wins $1
Repeat until A or B has all N dollars

What is $P(A$ wins$)$?

Let $E_i = \text{event that A wins starting with } i$

Approach: Condition on 1st flip

\[ p_i = \frac{1}{2} (p_{i+1} + p_{i-1}) \]

\[ 2p_i = p_{i+1} + p_{i-1} \]

\[ p_{i+1} - p_i = p_i - p_{i-1} \]

\[ p_2 - p_1 = p_1 - p_0 = p_1, \text{ since } p_0 = 0 \]

So: \[ p_2 = 2p_1 \]

\[ \ldots \]

\[ p_i = ip_1 \]

\[ p_N = Np_1 = 1 \]

\[ p_i = i/N \]
Bayes Theorem
Bayes Theorem

6 balls in an urn, some red, some white

Probability of drawing 3 red balls, given 3 in urn?

\[ w = 3 \]
\[ r = 3 \]

Probability of 3 red balls in urn, given that I drew three?

\[ w = ?? \]
\[ r = ?? \]
the theory that would not die
how bayes’ rule cracked the enigma code, hunted down russian submarines & emerged triumphant from two centuries of controversy
sharon bertsch mcgrayne

Yale University Press, 2011
When Microsoft Senior Vice President [later CEO] Steve Ballmer first heard his company was planning a huge investment in an Internet service offering... he went to Chairman Bill Gates with his concerns...

Gates began discussing the critical role of “Bayesian” systems...

source: http://www.ar-tiste.com/latimes_oct-96.html
Bayes Theorem

Most common form:

\[ P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E)} \]

Expanded form (using law of total probability):

\[ P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid F^c)P(F^c)} \]

Proof:

\[ P(F \mid E) = \frac{P(EF)}{P(E)} = \frac{P(E \mid F)P(F)}{P(E)} \]
Most common form:

\[ P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E)} \]

Expanded form (using law of total probability):

\[ P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid F^c)P(F^c)} \]

Why it’s important:

Reverse conditioning

\[ P(\text{model} \mid \text{data}) \sim P(\text{data} \mid \text{model}) \]

Combine new evidence (E) with prior belief (P(F))

Posterior vs prior
An urn contains 6 balls, either 3 red + 3 white or all 6 red. You draw 3; all are red. Did urn have only 3 red?

Can’t tell!

Suppose it was 3 + 3 with probability $p = \frac{3}{4}$. Did urn have only 3 red?

$M =$ urn has 3 red + 3 white

$D =$ I drew 3 red

\[
P(D \mid M) = \frac{\binom{3}{3}}{\binom{6}{3}} = \frac{1}{20}
\]

\[
P(M \mid D) = \frac{P(D \mid M)P(M)}{P(D \mid M)P(M) + P(D \mid M^c)P(M^c)}
\]

\[
= \frac{\left(\frac{1}{20}\right)\left(\frac{3}{4}\right)}{\left(\frac{1}{20}\right)\left(\frac{3}{4}\right) + (1)(1 - \frac{3}{4})} = \frac{3}{23}
\]

prior = $\frac{3}{4}$; posterior = $\frac{3}{23}$
Say that 60% of email is spam
90% of spam has a forged header
20% of non-spam has a forged header
Let $F =$ message contains a forged header
Let $J =$ message is spam
What is $P(J|F)$ ?

Solution:

$$P(J | F) = \frac{P(F | J)P(J)}{P(F | J)P(J) + P(F | J^c)P(J^c)}$$

$$= \frac{(0.9)(0.6)}{(0.9)(0.6) + (0.2)(0.4)}$$

$$\approx 0.871$$

prior = 60% 
posterior = 87%
Say that 60% of email is spam
10% of spam has the word “Viagra”
1% of non-spam has the word “Viagra”
Let \( V \) = message contains the word “Viagra”
Let \( J \) = message is spam
What is \( P(J|V) \) ?

Solution:

\[
P(J \mid V) = \frac{P(V \mid J)P(J)}{P(V \mid J)P(J) + P(V \mid J^c)P(J^c)}
\]

\[
= \frac{(0.1)(0.6)}{(0.1)(0.6) + (0.01)(1 - 0.6)}
\]

\[
\approx 0.9375
\]
Child is born with \((A,a)\) gene pair (event \(B_{A,a}\))

Mother has \((A,A)\) gene pair

Two possible fathers: \(M_1 = (a,a), M_2 = (a,A)\)

\[ P(M_1) = p, \quad P(M_2) = 1-p \]

What is \(P(M_1 \mid B_{A,a})\)?

Solution:

\[
P(M_1 \mid B_{A,a}) = \frac{P(B_{Aa} \mid M_1)P(M_1)}{P(B_{Aa} \mid M_1)P(M_1) + P(B_{Aa} \mid M_2)P(M_2)}
\]

\[= \frac{1 \cdot p}{1 \cdot p + 0.5(1 - p)} = \frac{2p}{1 + p} \geq \frac{2p}{1 + 1} = p\]

E.g., \(1/2 \rightarrow 2/3\)

i.e., the given data about child raises probability that \(M_1\) is father

Exercises:
- What if \(M_2\) were \((A,A)\)?
- What if child were \((A,A)\)?
Suppose an HIV test is 98% effective in detecting HIV, i.e., its “false negative” rate = 2%. Suppose furthermore, the test’s “false positive” rate = 1%.

0.5% of population has HIV

Let E = you test positive for HIV
Let F = you actually have HIV

What is P(F|E) ?

Solution:

\[
P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid F^c)P(F^c)}
\]

\[
= \frac{(0.98)(0.005)}{(0.98)(0.005) + (0.01)(1 - 0.005)}
\]

\[
\approx 0.330
\]

P(E) \approx 1.5%

Note difference between conditional and joint probability: P(F|E) = 33% ; P(FE) = 0.49%
Let $E^c = \text{you test negative for HIV}$
Let $F = \text{you actually have HIV}$

What is $P(F|E^c)$?

$$P(F | E^c) = \frac{P(E^c | F)P(F)}{P(E^c | F)P(F) + P(E^c | F^c)P(F^c)}$$

$$= \frac{(0.02)(0.005)}{(0.02)(0.005) + (0.99)(1 - 0.005)}$$

$$\approx 0.0001$$
Odds
The *probability* of event $E$ is $P(E)$.

The *odds* of event $E$ is $P(E)/(P(E^c))$

**Example:** $A$ = any of 2 coin flips is H:  

$P(A) = 3/4$, $P(A^c) = 1/4$, so odds of $A$ is 3  
(or “3 to 1 in favor”)

**Example:** odds of having HIV:  

$P(F) = .5\%$ so $P(F)/P(F^c) = .005/.995$  
(or 1 to 199 *against*; this is close, but not equal to,  
$P(F)=1/200$)
Odds

Probabilities and Odds are interconvertible:

\[ Odds(E) = \frac{P(E)}{1 - P(E)} \]

\[ P(E) = \frac{Odds(E)}{1 + Odds(E)} \]
posterior odds from prior odds

F = some event of interest (say, “HIV+”)

E = *additional* evidence (say, “HIV test was positive”)

**Prior odds** of F: $P(F)/P(F^c)$

What are the **Posterior odds** of F: $P(F|E)/P(F^c|E)$?

$$P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E)}$$

$$P(F^c \mid E) = \frac{P(E \mid F^c)P(F^c)}{P(E)}$$

$$\frac{P(F \mid E)}{P(F^c \mid E)} = \frac{P(E \mid F)}{P(E \mid F^c)} \cdot \frac{P(F)}{P(F^c)}$$

$\left( \text{posterior odds} \right) = \left( \text{“Bayes factor”} \right) \cdot \left( \text{prior odds} \right)$

There’s nothing new here, versus prior results, but the simple form, and the simple interpretation are convenient.
Let $E = \text{you test positive for HIV}$

Let $F = \text{you actually have HIV}$

What are the posterior odds?

\[
\frac{P(F \mid E)}{P(F^c \mid E)} = \frac{P(E \mid F)}{P(E \mid F^c)} \cdot \frac{P(F)}{P(F^c)}
\]

(posterior odds = “Bayes factor” · prior odds)

\[
\begin{array}{c|c|c}
 & HIV+ & HIV- \\
\hline
\text{Test +} & 0.98 = P(E \mid F) & 0.01 = P(E \mid F^c) \\
\text{Test -} & 0.02 = P(E^c \mid F) & 0.99 = P(E^c \mid F^c) \\
\end{array}
\]

More likely to test positive if you are positive, so Bayes factor >1; positive test increases odds, 98-fold in this case, to 2.03:1 against (vs prior of 199:1 against)
Let $E^c = \text{you test negative for HIV}$
Let $F = \text{you actually have HIV}$

What are the posterior odds (ratio between $P(F|E^c)$ and $P(F^c|E^c)$)?

\[
\frac{P(F \mid E^c)}{P(F^c \mid E^c)} = \frac{P(E^c \mid F) \cdot P(F)}{P(E^c \mid F^c) \cdot P(F^c)}
\]

(posterior odds = “Bayes factor” · prior odds)

\[
= \frac{0.02 \cdot 0.005}{0.99 \cdot 0.995}
\]

Unlikely to test negative if you are positive, so Bayes factor $< 1$; negative test decreases odds 49.5-fold, to 9850:1 against (vs prior of 199:1 against)
simple spam detection

Say that 60% of email is spam
10% of spam has the word “Viagra”
1% of non-spam has the word “Viagra”
Let $V =$ message contains the word “Viagra”
Let $J =$ message is spam

What are posterior odds that a message containing “Viagra” is spam?

Solution:

$$\frac{P(J \mid V)}{P(J^c \mid V)} = \frac{P(V \mid J)}{P(V \mid J^c)} \frac{P(J)}{P(J^c)}$$

(posterior odds = “Bayes factor” · prior odds)

$$15 = \frac{0.10}{0.01} \cdot \frac{0.6}{0.4}$$
Summary
Conditional probability

\[ P(E|F): \text{Conditional probability that } E \text{ occurs given that } F \text{ has occurred.} \]

Reduce event/sample space to points consistent w/ \( F \) (\( E \cap F \); \( S \cap F \))

\[
P(E \mid F) = \frac{P(EF)}{P(F)} \quad (P(F) > 0)
\]

\[
P(E \mid F) = \frac{|EF|}{|F|}, \text{ if equiprobable outcomes.}
\]

\[ P(EF) = P(E|F) P(F) \quad (\text{“the chain rule”}) \]

“\( P(- \mid F) \)” is a probability law, i.e., satisfies the 3 axioms

\[ P(E) = P(E|F) P(F) + P(E|F^c) (1-P(F)) \quad (\text{“the law of total probability”}) \]

Bayes theorem

\[
P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E)}
\]

prior, posterior, odds, prior odds, posterior odds, Bayes factor