4. Conditional Probability

\[ P( \quad | \quad ) \]
**Conditional probability** of E given F: probability that E occurs *given* that F has already occurred.

“Conditioning on F”

Written as $P(E|F)$

Means “$P(E, \text{given F already observed})$”

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,

$$P(E \mid F) = \frac{\# \text{ of outcomes in } E \text{ consistent with } F}{\# \text{ of outcomes in } S \text{ consistent with } F} = \frac{|EF|}{|SF|} = \frac{|EF|}{|F|}$$
General defn: \[ P(E \mid F) = \frac{P(EF)}{P(F)} \] where \( P(F) > 0 \)

Holds even when outcomes are \textit{not} equally likely.

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

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

**Implies:** \( P(EF) = P(E \mid F) \ P(F) \) \hspace{1cm} (chain rule)

General definition of \textbf{Chain Rule}:

\[
P(E_1 E_2 \cdots E_n) = \frac{P(E_1 \mid E_2) \ P(E_2 \mid E_1) \ P(E_3 \mid E_1, E_2) \ \cdots \ P(E_n \mid E_1, E_2, \ldots, E_{n-1})}{P(F)}
\]
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 \( F = \{HH, HT\} \)
  \[
P(B|F) = \frac{P(BF)}{P(F)} = \frac{P(\{HH\})}{P(\{HH, HT\})} = \frac{1/4}{2/4} = 1/2
  \]

• At least one of the two flips lands on heads?
  Let \( A = \{HH, HT, TH\} \)
  \[
P(B|A) = \frac{P(BA)}{P(A)} = \frac{P(\{HH\})}{P(\{HH, HT, TH\})} = \frac{1/4}{3/4} = 1/3
  \]
sending bit strings
Bit string with \( m \) 0's and \( n \) 1's sent on the network
All distinct arrangements of bits equally likely
\( E = \) first bit received is a 1
\( F = \) \( k \) of first \( r \) bits received are 1's
What's \( P(E|F) \)?

**Solution 1:**

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

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

\[
P(E | F) = \frac{P(EF)}{P(F)} = \frac{P(F | E)P(E)}{P(F)} = \frac{k}{r}
\]
Bit string with m 0’s and n 1’s sent on the network
All distinct arrangements of bits equally likely
E = first bit received is a 1
F = k of first r bits received are 1’s

Solution 2:
Observe:
\[ P(E|F) = P(\text{picking one of } k \text{ 1’s out of } r \text{ bits}) \]
So:
\[ P(E|F) = \frac{k}{r} \]
piling cards
Deck of 52 cards randomly divided into 4 piles
13 cards per pile
Compute $P(\text{each pile contains an ace})$
Solution:
$E_1 = \{ \text{in any one pile} \}$
$E_2 = \{ \text{and in different piles} \}$
$E_3 = \{ \text{different piles} \}$
$E_4 = \{ \text{all four aces in different piles} \}$

Compute $P(E_1 E_2 E_3 E_4)$
Piling cards

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

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

\[ P(E_1) = 1 \]
\[ P(E_2|E_1) = \frac{39}{51} \text{ (39 cards not in AH pile)} \]
\[ P(E_3|E_1E_2) = \frac{26}{50} \text{ (26 cards not in AS or AH piles)} \]
\[ P(E_4|E_1E_2E_3) = \frac{13}{49} \text{ (13 cards not in AS, AH, AD piles)} \]
E₁ = { in any one pile }

E₂ = { and in different piles }

E₃ = { different piles }

E₄ = { all four aces in different piles }

\[
P(E₁E₂E₃E₄) = P(E₁) \cdot P(E₂|E₁) \cdot P(E₃|E₁E₂) \cdot P(E₄|E₁E₂E₃)
\]
\[
= (39 \cdot 26 \cdot 13) / (51 \cdot 50 \cdot 49)
\]
\[
≈ 0.105
\]
law of total probability

E and F are events in the sample space S

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

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

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

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

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

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

This is a weighted average, conditioned on the event happening or not.
Bayes Theorem

Most common form:

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

Expanded form (using law of total probability):

\[
P(F \mid E) = \frac{P(E \mid F) P(F)}{P(E \mid F) P(F) + P(E \mid F^c) P(F^c)}
\]
When Microsoft Senior Vice President Steve Ballmer [now CEO] 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
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:

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

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

$$\approx 0.330$$
why it’s still good to get tested

<table>
<thead>
<tr>
<th></th>
<th>HIV+</th>
<th>HIV-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +</td>
<td>0.98 = P(E</td>
<td>F)</td>
</tr>
<tr>
<td>Test -</td>
<td>0.02 = P(E^c</td>
<td>F)</td>
</tr>
</tbody>
</table>

Let $E^c = \text{you test negative for HIV}$

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

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

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

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

$$\approx 0.0001$$
simple spam detection

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:

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

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

$$\approx 0.871$$
Say that 60% of email is spam
10% of spam has the word “Viagra”
1% of non-spam has the word “Viagra”
Let $V = \text{message contains the word “Viagra”}$
Let $J = \text{message is spam}$

What is $P(J|V)$ ?

Solution:

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

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

$$\approx 0.896$$
Child is born with \((A,a)\) gene pair (event \(B_{A,a}\))

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

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

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

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

**Solution:**

\[
P(M_1 \mid B_{Aa})
= \frac{P(B_{Aa} \mid M_1)P(M_1)}{P(B_{Aa} \mid M_1)P(M_1) + P(B_{Aa} \mid M_2)P(M_2)}
= \frac{1 \cdot p}{1 \cdot p + 0.5(1 - p)} = \frac{2p}{1 + p} > \frac{2p}{1 + 1} = p
\]
The *odds* of event E is \( \frac{P(E)}{P(E^c)} \)

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

\[
P(A) = \frac{3}{4}, \quad P(A^c) = \frac{1}{4}, \quad \text{so odds of A is 3 (or “3 to 1 in favor”)}
\]

Example: odds of having HIV:

\[
P(F) = .5\% \quad \text{so} \quad \frac{P(F)}{P(F^c)} = \frac{.005}{.995} \quad \text{(or 1 to 199 against)}
\]
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: \( P(F|E)/P(F^c|E) \) ?

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

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

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

(posterior odds = “Bayes factor” \cdot prior odds)
Let $E = \text{you test positive for HIV}$
Let $F = \text{you actually have HIV}$

What are the posterior odds?

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

(posterior odds = “Bayes factor” \cdot prior odds)

\[
= \frac{0.98}{0.01} \cdot \frac{0.005}{0.995}
\]

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

What is the ratio between $P(F|E)$ and $P(F^c|E)$?

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

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

$$= \frac{0.02}{0.99} \cdot \frac{0.005}{0.995}$$

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)
Say that 60% of email is spam
10% of spam has the word “Viagra”
1% of non-spam has the word “Viagra”

Let $V = \text{message contains the word “Viagra”}$
Let $J = \text{message is spam}$

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

Solution:

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

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

$$\begin{align*}
15 &= \frac{0.10}{0.01} \cdot \frac{0.6}{0.4}
\end{align*}$$