Probability fundamentals

Hong Ge

Oct, 2025

Key Concepts

- probability basics
  - Example: Medical diagnosis
  - joint, conditional and marginal probabilities
  - the two rules of probability: sum and product rules
  - Bayes rule
- Bayesian inference and prediction with finite regression models
  - likelihood and prior
  - posterior and predictive distribution
- the marginal likelihood
  - Bayesian model selection
  - Example: How Bayes avoids overfitting

Medical inference (diagnosis)

Breast cancer facts:

- 1% of scanned women have breast cancer
- 80% of women with breast cancer get positive mammography scans
- 9.6% of women without breast cancer also get positive mammography scans

Question: A woman gets a scan, and it is positive; what is the probability that she has breast cancer?

1. less than 1%
2. around 10%
3. around 90%
4. more than 99%

Medical inference

Breast cancer facts:

- 1% of scanned women have breast cancer
- 80% of women with breast cancer get positive mammography scans
- 9.6% of women without breast cancer also get positive mammography scans

Define: C = presence of breast cancer; $\bar{C}$ = no breast cancer.
M = scan is positive; $\bar{M}$ = scan is negative.
The probability of cancer for scanned women is $p(C) = 1\%$
If there is cancer, the probability of a positive mammography is $p(M|C) = 80\%$
If there is no cancer, we still have $p(M|\bar{C}) = 9.6\%$
The question is what is $p(C|M)$?

Medical inference

What is $p(C|M)$?
Consider 10000 subjects of screening

- $p(C) = 1\%$, therefore 100 of them have cancer, of which
  - $p(M|C) = 80\%$, therefore 80 get a positive mammography
  - 20 get a negative mammography
- $p(\bar{C}) = 99\%$, therefore 9900 of them do not have cancer, of which
  - $p(M|\bar{C}) = 9.6\%$, therefore 950 get a positive mammography
  - 8950 get a negative mammography

	M	$\bar{M}$
C	80	20
$\bar{C}$	950	8950

What is $p(C|M)$?

	M	$\bar{M}$
C	80	20
$\bar{C}$	950	8950

$p(C|M)$ is obtained as the proportion of all positive mammographies for which there actually is breast cancer

$$p(C|M) = \frac{p(C, M)}{p(C, M) + p(\bar{C}, M)} = \frac{p(M|C)p(C)}{p(M)} = \frac{80}{80 + 950} \simeq 7.8\%$$

This is an example of Bayes' rule:

$$p(A|B)p(B) = p(A, B) = p(B|A)p(A),$$

which is just a consequence of the definition of *conditional probability*

$$p(A|B) = \frac{p(A, B)}{p(B)}, \quad (\text{where } p(B) \neq 0).$$

Just two rules of probability theory

Astonishingly, the rich theory of probability can be derived using just two rules:
The *sum rule* states that

$$p(A) = \sum_B p(A, B), \quad \text{or} \quad p(A) = \int_B p(A, B) dB,$$

for discrete and continuous variables. Sometimes called *marginalization*.
The *product rule* states that

$$p(A, B) = p(A|B)p(B).$$

It follows directly from the definition of conditional probability, and leads directly to Bayes' rule

$$p(A|B)p(B) = p(A, B) = p(B|A)p(A) \Rightarrow p(A|B) = \frac{p(B|A)p(A)}{p(B)}.$$

Special case:
if A and B are *independent*, $p(A|B) = p(A)$, and thus $p(A, B) = p(A)p(B)$.