Exercise 5 — Bayes’ Theorem & Confusion Matrix

Exam-style question: Cancer vs. Symptoms

Notation. Let A = “Cancer”, ¬A = “No cancer”, B = “Symptoms present”, ¬B = “No symptoms”. Bayes: **P(A|B) = P(B|A) P(A) / P(B)**. Total probability: **P(B) = P(B|A)P(A) + P(B|¬A)P(¬A)**.

Given (from the exam statement)

Study size	n = 1000
Cancer cases	#A = 2 (both have symptoms)
Symptoms without cancer	$\#(\neg A \cap B) = 8$
No symptoms and no cancer	$\#(\neg A \cap \neg B) = 990$

Counts table (confusion-matrix style)

	Cancer (A)	No cancer (¬A)	Row total
Symptoms (B)	2	8	10
No symptoms (¬B)	0	990	990
Column total	2	998	1000

(a) How many subjects did NOT have cancer?

Complement of the cancer cases:
#(¬A) = n – #A = 1000 – 2 = 998.

(b) Prevalence (prior) P(A)

Prevalence is the fraction with cancer:
P(A) = #A / n = 2 / 1000 = 0.002 (0.2%).

(c) Sensitivity P(B|A)

Both cancer subjects have symptoms, hence:
P(B|A) = $\#(A \cap B)$ / #A = 2 / 2 = 1.

(d) Show that P(B) = 1%

Compute the symptom rate among non-cancer subjects: **P(B|¬A) = 8/998**. Also **P(¬A)=1–P(A)=0.998**.

Total probability:
P(B) = P(B|A)P(A) + P(B|¬A)P(¬A)
= (1)(0.002) + (8/998)(0.998) = 0.002 + 0.008 = **0.01** = **1%**.

Sanity check: 10 symptomatic out of 1000 → 10/1000 = 0.01.

(e) Posterior P(A|B) = P(Cancer | Symptoms)

Bayes’ theorem:
P(A|B) = P(B|A)P(A) / P(B)

= (1)(0.002) / 0.01 = **0.2** (20%).

Sanity check from counts: among 10 with symptoms, 2 have cancer → 2/10 = 0.2.

Key takeaway

Posterior probabilities can differ greatly from the prior. Here, symptoms raise the cancer probability from 0.2% to 20% within this study cohort.