Conditional Probability and Independence

Baye's Theorem [Ross S3.3]

Law of Total Probability:

Let $E, F \subset S$.

Then $\quad E = ES = E(F \cup F^c) = EF \cup EF^c$

and $$\begin{aligned} P[E] &= P[EF] + P[EF^c] \\ &= P[E|F]P[F] + P[E|F^c]P[F^c] \end{aligned}$$

Example 6.1: The probability of an insurance claim is

- 0.4 for 30% of persons (type 1),
- 0.5 for 70% of persons (type 2).

What is the probability that a random person has a claim?

Solution:

Let $$\begin{aligned} F &= \{\text{type 1 person}\} \\ F^c &= \{\text{type 2 person}\} \\ E &= \{\text{there is a claim}\} \end{aligned}$$

Then

$$\begin{aligned} P[E] &= P[E|F]P[F] + P[E|F^c]P[F^c] \\ &= 0.4 \times 0.3 + 0.5 \times 0.7 = 0.47 \end{aligned}$$

Let $F_1, ..., F_n$ partition S.

Then
$$E = ES = E\left(\bigcup_{i=1}^{n} F_i\right)$$
$$= \bigcup_{i=1}^{n} (EF_i)$$

So
$$P[E] = P[\cup_{i=1}^{n}(EF_i)]$$
$$= \sum_{i=1}^{n} P[EF_i]$$
$$= \sum_{i=1}^{n} P[E|F_i]P[F_i] \qquad \text{[Law of total probability]}$$

Example 6.2: You roll a 4-sided die. If result is ≤ 2, you roll once more and otherwise stop. What is probability that the sum ≥ 4?

Solution: Let $F_i = \{\text{first roll } = i\}$, $\quad E = \{\text{sum } \geq 4\}$.

$P[F_i] = 1/4$

$P[E|F_1] = P[\text{second roll is 3 or 4}] = 1/2$

$P[E|F_2] = P[\text{second roll is 2, 3 or 4}] = 3/4$

$P[E|F_3] = 0$

$P[E|F_4] = 1$

$$\Rightarrow P[E] = P[E|F_0]P[F_0] + P[E|F_1]P[F_1] + P[E|F_2]P[F_2] + P[E|F_3]P[F_3]$$
$$= 1/2 \times 1/4 + 3/4 \times 1/4 + 0 \times 1/4 + 1 \times 1/4$$
$$= 9/16.$$

Baye's Theorem and Inference:

Let $F_1, F_2, \dots, F_n$ partition S.

Say we know $P[E|F_j]$. We want to compute $P[F_j|E]$:

$$\begin{aligned} P[F_j|E] &= \frac{P[EF_j]}{P[E]} \\ &= \frac{P[E|F_j]P[F_j]}{P[E|F_1]P[F_1] + P[E|F_2]P[F_2] + \cdots + P[E|F_n]P[F_n]} \end{aligned} \tag{6.1}$$

This is Baye's theorem/rule.

Application to inference:

Before any partial information is revealed (i.e., observing E occurs), the probabilities are:

$P[F_1], P[F_2], \dots, P[F_n]$ $\Big\}$ "prior probabilities"

After observing E occur, they are revised as:

$P[F_1|E], P[F_2|E], \dots, P[F_n|E]$ $\Big\}$ "posterior probabilities"

according to (6.1).

Posterior probabilities are key to practical inference (e.g., classification, pattern recognition, detection, etc.)

Example 6.3: A 3-card deck has

1) one card with red on both sides
2) one card with black on both sides
3) one card with red on one side + black on the other.

One side of 1 card is picked at random. It is red. What is the probability that other side is black?

Solution: Let $S = \{RR, RB, BB\}$, $\quad R = \{\text{side shown is red}\}$.

$$\begin{aligned}
P[RB|R] &= \frac{P[RB \cap R]}{P[R]} \\
&= \frac{P[R|RB]P[RB]}{P[R|RR]P[RR] + P[R|RB]P[RB] + P[R|BB]P[BB]} \\
&= \frac{\frac{1}{2}\frac{1}{3}}{1 \times \frac{1}{3} + \frac{1}{2}\frac{1}{3} + 0 \times \frac{1}{3}} \\
&= 1/3
\end{aligned}$$

Why?

If you see red, there are 3 different ways this could happen (one side of RB + two sides of RR). But only 1 results in black on the other side.

Example 6.4: A blood test has 95% prob of detecting a desease when it is present. It has a 1% false positive rate when it is not present. 0.5% of people have the desease.

a) If a random person tests positive, what is prob. that desease is present?

b) If a random person tests negative, what is prob. that desease is present?

Solution: $E = \{\text{positive result}\}, \quad F = \{\text{desease present}\}$

a)
$$\begin{aligned} P[F|E] &= \frac{P[EF]}{P[E]} \\ &= \frac{P[E|F]P[F]}{P[E|F]P[F] + P[E|F^c]P[F^c]} \\ &= \frac{0.95 \times 0.005}{0.95 \times 0.005 + 0.01 \times 0.995} \\ &\approx 0.323 \end{aligned}$$

b)
$$\begin{aligned} P[F|E^c] &= \frac{P[E^cF]}{P[E^c]} \\ &= \frac{P[E^c|F]P[F]}{P[E^c|F]P[F] + P[E^c|F^c]P[F^c]} \\ &= \frac{0.05 \times 0.005}{0.05 \times 0.005 + 0.99 \times 0.995} \\ &\approx 0.000254 \end{aligned}$$