Notes_2_9_Law_of_Total_Probability_and_Bayes_Rule

Tuesday, August 30, 2016 9:47 AM

Notes_2_9_
Law_of_T...

CHALLENGE: Come up with a random experiment and three events which are **pairwise independent**, but not mutually independent.

COMMON SETTING: Many experiments consist of a sequence of n trials that are viewed as independent (e.g., flipping a coin 10 times). If A_i denotes the event associated with the ith trial, and the trials are **independent**, then

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i).$$

Example 2.26. An unbiased die is rolled six times. Let $A_i = \{i \text{ appears on roll } i\}$, for $i = 1, 2, ..., 6$. Then, $P(A_i) = 1/6$, and assuming independence,

$$P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6) = \prod_{i=1}^{6} P(A_i) = \left(\frac{1}{6}\right)^6.$$

Suppose that if A_i occurs, we will call it "a match." What is the probability of at least one match in the six rolls?

SOLUTION: Let B denote the event that there is at least one match. Then, $\overline{B}$ denotes the event that there are no matches. Now,

$$P(\overline{B}) = P(\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3 \cap \overline{A}_4 \cap \overline{A}_5 \cap \overline{A}_6) = \prod_{i=1}^{6} P(\overline{A}_i) = \left(\frac{5}{6}\right)^6 = 0.335.$$

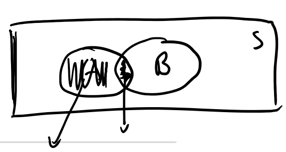
Thus, $P(B) = 1 - P(\overline{B}) = 1 - 0.335 = 0.665$, by the complement rule.

EXERCISE: Generalize this result to an n sided die. What does this probability converge to as $n \to \infty$? $\square$

2.9 Law of Total Probability and Bayes Rule

SETTING: Suppose A and B are events in a nonempty sample space S. We can express the event A as follows

$$\underbrace{A = (A \cap B) \cup (A \cap \overline{B})}_{\text{union of disjoint events}}.$$



$A \cap \overline{B}$ $A \cap B$

$P(A) = P(A \cap B) + P(A \cap \overline{B})$

$= P(A|B)P(B) + P(A|\overline{B})P(\overline{B})$

$$P(E_1|E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)}$$

By the third Kolmolgorov axiom,

$$
\begin{aligned}
P(A) &= P(A \cap B) + P(A \cap \overline{B}) \\
&= P(A|B)P(B) + P(A|\overline{B})P(\overline{B}),
\end{aligned}
$$

where the last step follows from the multiplication law of probability. This is called the **Law of Total Probability** (LOTP). The LOTP is helpful. Sometimes $P(A|B)$, $P(A|\overline{B})$, and $P(B)$ may be easily computed with available information whereas computing $P(A)$ directly may be difficult.

NOTE: The LOTP follows from the fact that B and $\overline{B}$ **partition** S; that is,

(a) B and $\overline{B}$ are disjoint, and

(b) $B \cup \overline{B} = S$.

Example 2.27. An insurance company classifies people as "accident-prone" and "non-accident-prone." For a fixed year, the probability that an accident-prone person has an accident is 0.4, and the probability that a non-accident-prone person has an accident is 0.2. The population is estimated to be 30 percent accident-prone. (a) What is the probability that a new policy-holder will have an accident?

SOLUTION:
Define $A = \{$policy holder has an accident$\}$ and $B = \{$policy holder is accident-prone$\}$. Then, $P(B) = 0.3$, $P(A|B) = 0.4$, $P(\overline{B}) = 0.7$, and $P(A|\overline{B}) = 0.2$. By the LOTP,

$$
\begin{aligned}
P(A) &= P(A|B)P(B) + P(A|\overline{B})P(\overline{B}) \\
&= (0.4)(0.3) + (0.2)(0.7) = 0.26. \square
\end{aligned}
$$

(b) Now suppose that the policy-holder does have an accident. What is the probability that he was "accident-prone?"

SOLUTION: We want $P(B|A)$. Note that

$$
P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)} = \frac{(0.4)(0.3)}{0.26} = 0.46. \square
$$

$P(A \cap B) = P(A|B)\,P(B)$
$\quad \parallel$
$P(B \cap A) = P(B|A)\,P(A)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(\bar{A}) = 1 - P(A)$

$P(A \cap B) = P(A|B)P(B)$
$= P(B|A)P(A)$

$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$

NOTE: From this last part, we see that, in general,

$$
P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\overline{B})P(\overline{B})}.
$$

This is a form of **Bayes Rule**.

Example 2.28. A lab test is 95 percent effective at detecting a certain disease when it is present (sensitivity). When the disease is not present, the test is 99 percent effective at declaring the subject negative (specificity). If 8 percent of the population has the disease (prevalence), what is the probability that a subject has the disease given that (a) his test declaring the subject negative (specificity). If 8 percent of the population has the disease (prevalence), what is the probability that a subject has the disease given that (a) his test is positive? (b) his test is negative?

SOLUTION: Let $D = \{\text{disease is present}\}$ and $\maltese = \{\text{test is positive}\}$. We are given that $P(D) = 0.08$ (prevalence), $P(\maltese|D) = 0.95$ (sensitivity), and $P(\overline{\maltese}|\overline{D}) = 0.99$ (specificity). In part (a), we want to compute $P(D|\maltese)$. By Bayes Rule,

$$\begin{aligned} P(D|\maltese) &= \frac{P(\maltese|D)P(D)}{P(\maltese|D)P(D) + P(\maltese|\overline{D})P(\overline{D})} \\ &= \frac{(0.95)(0.08)}{(0.95)(0.08) + (0.01)(0.92)} \approx 0.892. \end{aligned}$$

In part (b), we want $P(D|\overline{\maltese})$. By Bayes Rule,

$$\begin{aligned} P(D|\overline{\maltese}) &= \frac{P(\overline{\maltese}|D)P(D)}{P(\overline{\maltese}|D)P(D) + P(\overline{\maltese}|\overline{D})P(\overline{D})} \\ &= \frac{(0.05)(0.08)}{(0.05)(0.08) + (0.99)(0.92)} \approx 0.004. \end{aligned}$$

$P(\overline{D}|\overline{\maltese}) = 1 - .004 = .996$

Table 2.1: *The general Bayesian scheme.*

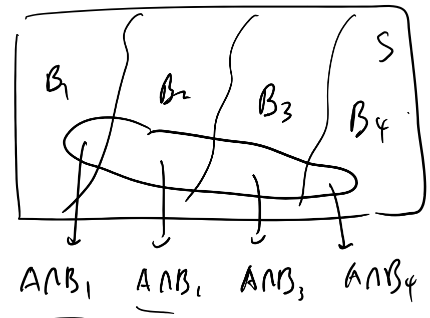
Measure before test		Result		Updated measure
$P(D)$		F		$P(D\|F)$
0.08	$\longrightarrow$	$\maltese$	$\longrightarrow$	0.892
0.08	$\longrightarrow$	$\overline{\maltese}$	$\longrightarrow$	0.004

NOTE: We have discussed the LOTP and Bayes Rule in the case of the partition $\{B, \overline{B}\}$. However, these rules hold for any partition of S.

TERMINOLOGY: A sequence of sets $B_1, B_2, ..., B_k$ is said to form a **partition** of the sample space S if

(a) $B_1 \cup B_2 \cup \cdots \cup B_k = S$ (exhaustive condition), and

(b) $B_i \cap B_j = \emptyset$, for all $i \neq j$ (disjoint condition).



LAW OF TOTAL PROABILITY (*restated*): Suppose that $B_1, B_2, ..., B_k$ form a partition of S, and suppose $P(B_i) > 0$ for all $i = 1, 2, ..., k$. Then,

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i).$$

$$P(A) = \sum_{i=1}^{4} P(A \cap B_i) = \sum_{i=1}^{4} P(A|B_i)P(B_i)$$

Proof. Write

$$A = A \cap S = A \cap (B_1 \cup B_2 \cup \cdots \cup B_k) = \bigcup_{i=1}^{k} (A \cap B_i).$$

Thus,

$$P(A) = P\left[\bigcup_{i=1}^{k} (A \cap B_i)\right] = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(A|B_i)P(B_i). \quad \square$$

BAYES RULE (*restated*): Suppose that $B_1, B_2, ..., B_k$ form a partition of S, and suppose that $P(A) > 0$ and $P(B_i) > 0$ for all $i = 1, 2, ..., k$. Then,

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}.$$

$$P(B_j|A) = \frac{P(B_j \cap A)}{P(A)} = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{P(A)}$$

Proof. Simply apply the definition of conditional probability and the multiplication law of probability to get

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)}.$$

Then, just apply LOTP to $P(A)$ in the denominator to get the result. $\square$

REMARK: Bayesians will call $P(B_j)$ the **prior probability** for the event B_j; they call $P(B_j|A)$ the **posterior probability** of B_j, given the information in A.

Example 2.29. Suppose that a manufacturer buys approximately 60 percent of a raw material (in boxes) from Supplier 1, 30 percent from Supplier 2, and 10 percent from

Supplier 3. For each supplier, defective rates are as follows: Supplier 1: 0.01, Supplier 2: 0.02, and Supplier 3: 0.03. The manufacturer observes a defective box of raw material.

(a) What is the probability that it came from Supplier 2?

(b) What is the probability that the defective did not come from Supplier 3?

SOLUTION: (a) Let $A = \{\text{observe defective box}\}$. Let B_1, B_2, and B_3, respectively, denote the events that the box comes from Supplier 1, 2, and 3. The prior probabilities (ignoring the status of the box) are

$$P(B_1) = 0.6$$
$$P(B_2) = 0.3$$
$$P(B_3) = 0.1.$$

Note that $\{B_1, B_2, B_3\}$ partitions the space of possible suppliers. Thus, by Bayes Rule,

$$\begin{aligned} P(B_2|A) &= \frac{P(A|B_2)P(B_2)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3)} \\ &= \frac{(0.02)(0.3)}{(0.01)(0.6) + (0.02)(0.3) + (0.03)(0.1)} = 0.40. \end{aligned}$$

This is the updated (posterior) probability that the box came from Supplier 2 (updated to include the information that the box was defective).

SOLUTION: (b) First, compute the posterior probability $P(B_3|A)$. By Bayes Rule,

$$\begin{aligned} P(B_3|A) &= \frac{P(A|B_3)P(B_3)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3)} \\ &= \frac{(0.03)(0.1)}{(0.01)(0.6) + (0.02)(0.3) + (0.03)(0.1)} = 0.20. \end{aligned}$$

Thus,

$$P(\overline{B}_3|A) = 1 - P(B_3|A) = 1 - 0.20 = 0.80,$$

by the complement rule. □

NOTE: Read Sections 2.11 (Numerical Events and Random Variables) and 2.12 (Random Sampling) in WMS.