Lecture 1: Probability Concepts

Lecturer: Prof. F. Jalayer, 05 February 2021
Notes prepared by: Prof. F. Jalayer and Dr. H. Ebrahimian

1. Logical statements

$A \equiv$ a statement which can be TRUE (T) or FALSE (F), i.e., $A \equiv \text{T}$ or $A \equiv \text{F}$. For instance, *the seismicity rate of the zone is less than* 0.20.

Logical Statements are denoted with capital letters. The manipulation of logical statements can be performed via the *Boolean algebra*.

2. Logical Operations

The basic operations of Boolean algebra are as follows:

1. Logical Product (Conjunction)

   It is defined as $A \cdot B$ (read as A **AND** B). It is true only if both statements A and B are true.

2. Logical Sum (Disjunction)

   It is defined as $A + B$ (read as A **OR** B). It is true if at least one of the statements A or B are true.

3. Negation

   It is defined as $\overline{A}$ (read as **NOT** A). If $A \equiv \text{T}$ then $\overline{A} \equiv \text{F}$.

3. Properties of logical statements

(a) $A \cdot A = A$ (Idempotence)

(b) $A + A = A$ (Idempotence)

(c) $A \cdot B = B \cdot A$ (Commutativity)

(d) $A + B = B + A$ (Commutativity)

(e) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ (Associativity)

(f) $A + (B + C) = (A + B) + C$ (Associativity)

(g) $A \cdot (B + C) = (A \cdot B) + (A \cdot C)$ (Distributivity)

(h) $A + (B \cdot C) = (A + B) \cdot (A + C)$ (Distributivity)

(i) $\overline{A + B} = \overline{A} \cdot \overline{B}$ (Duality)

(j) $\overline{A \cdot B} = \overline{A} + \overline{B}$ (Duality)

4. From logical statement to probability

A mapping can be defined between a logical statement *A*, which might be true or false, and real numbers $\mathbb{R}^+$ in terms of a mathematical function defined as probability and denoted as $P(A) \geq 0$. In the next section, a logical derivation is performed in order to assign the upper and lower bounds associated with the probability $P(A)$ corresponding to the logical statement being T or F.

5. Rule of Product in Probability

The probability that the product of two statements is a TRUE statement, can be calculated as the probability that one of these statements is TRUE multiplied by (arithmetically) the probability of the other statement being TRUE knowing the fact that the former statement is TRUE:

$$P(A \cdot B) = P(A)P(B \mid A) = P(B)P(A \mid B) \tag{1}$$

where $P(B|A)$ is defined as the *conditional* probability of *B* given that (=knowing that) statement *A* is true. Similar definition can also be assigned to $P(A|B)$. It is noteworthy that in dealing with conditional probabilities, the logical operations are performed on the left hand side of the conditional symbol "|", and the statements to the right of the symbol are known to be TRUE. In the product rule, the probability of a logical product converts to the normal product of two probability terms. Hereafter, we show the probability of each statement A by the general form of $P(A|\mathbf{I})$ instead of $P(A)$. The logical statement $\mathbf{I}$ denotes the so-called "*background Information*" emphasizing that the probability is always conditional on the amount of information available; i.e., the probability represents the degree of belief in a certain statement based on the amount of information available. Thus, Eq. (1) can be re-written as follows:

$$P\left(A\cdot B\middle|\mathbf{I}\right)=P\left(A\middle|\mathbf{I}\right)P\left(B\mid A\cdot\mathbf{I}\right)=P\left(B\middle|\mathbf{I}\right)P\left(A\mid B\cdot\mathbf{I}\right) \tag{2}$$

For the sake of simplicity and brevity, the logical product is dropped from the both sides of the conditional operator. Consider the special case that A is TRUE in Eq. (2):

$$P\left(AB\middle|\mathbf{I}\right)=\underbrace{P\left(B\middle|\mathbf{I}\right)}_{A\equiv\text{TRUE}}=P\left(B\middle|\mathbf{I}\right)P\left(A\mid B\mathbf{I}\right) \tag{3}$$

Therefore,

$$P\left(A\mid B\mathbf{I}\right)=1 \tag{4}$$

On the other side, if A is FALSE,

$$P\left(AB\middle|\mathbf{I}\right)=\underbrace{P\left(A\middle|\mathbf{I}\right)}_{A\equiv\text{FALSE}}=P\left(A\middle|\mathbf{I}\right)P\left(B\mid A\mathbf{I}\right) \tag{5}$$

different possibilities can be imagined: one is the case of having $P(A|\mathbf{I}) = 0$ and the other is $P(A|\mathbf{I}) = +\infty$. As a result, two distinct intervals can be assigned to $P(A|\mathbf{I})$: [0, 1] or [1,+∞);

nevertheless, the interval of [0, 1] is universally considered as the domain of the probabilities, as shown in Figure 1.

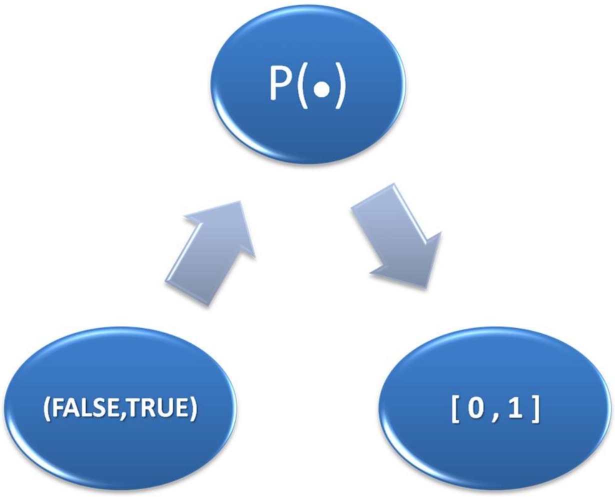

Figure 1: Assigning numerical values to logical statements

5.1. *The concept of being "Independent"*

If the statements A and B are independent, the knowledge on A (the fact that A is TRUE) has no effect on whether B is TRUE or not and vice versa; i.e., $P(B|A\mathbf{I})=P(B|\mathbf{I})$ and $P(A|B\mathbf{I})=P(A|\mathbf{I})$. Thus, Eq. (2) can be written as follows:

$$P\left(AB|\mathbf{I}\right)=\begin{cases}P\left(A|B\mathbf{I}\right)P\left(B\,|\,\mathbf{I}\right)=P\left(A|\mathbf{I}\right)P\left(B\,|\,\mathbf{I}\right)\\P\left(B|A\mathbf{I}\right)P\left(A\,|\,\mathbf{I}\right)=P\left(B|\mathbf{I}\right)P\left(A\,|\,\mathbf{I}\right)\end{cases} \tag{6}$$

As a result, the probability that the logical product of two independent statements is TRUE is equal to the product of their individual probabilities.

6. Rule of Sum in Probability

The rule of sum in probability can be defined based on the fact that the logical sum of a statement A and its negation is always equal to unity:

$$P\left(A+\overline{A}|\mathbf{I}\right)=1=P\left(A|\mathbf{I}\right)+P\left(\overline{A}\mid\mathbf{I}\right) \tag{7}$$

Thus:

$$P\left(A|\mathbf{I}\right)=1-P\left(\overline{A}\mid\mathbf{I}\right) \tag{8}$$

Therefore, using Eq. (8), the property (i) in Section 3, and the rule of product in Section 5, the probability of the logical sum of two statements can be written as:

$$P\left(A+B|\mathbf{I}\right)=1-P\left(\overline{A+B}\mid\mathbf{I}\right)=1-P\left(\overline{A}\cdot\overline{B}\mid\mathbf{I}\right)=1-P\left(\overline{A}\mid\mathbf{I}\right)P\left(\overline{B}\mid\overline{A}\mathbf{I}\right) \tag{9}$$

The expression for $P(A+B|\mathbf{I})$ in Eq. (9) can further be manipulated in order to derive it as the sum of probability of the two statements minus the probability of their logical product:

$$\begin{aligned}P\left(A+B|\mathbf{I}\right)&=1-P\left(\overline{A}\mid\mathbf{I}\right)\left[1-P\left(B\mid\overline{A}\mathbf{I}\right)\right]\\&=P\left(A\mid\mathbf{I}\right)+P\left(\overline{A}\mid\mathbf{I}\right)P\left(B\mid\overline{A}\mathbf{I}\right)\\&=P\left(A\mid\mathbf{I}\right)+P\left(B\overline{A}\mid\mathbf{I}\right)\\&=P\left(A\mid\mathbf{I}\right)+P\left(B\mid\mathbf{I}\right)P\left(\overline{A}\mid B\mathbf{I}\right)\\&=P\left(A\mid\mathbf{I}\right)+P\left(B\mid\mathbf{I}\right)\left[1-P\left(A\mid B\mathbf{I}\right)\right]\\&=P\left(A\mid\mathbf{I}\right)+P\left(B\mid\mathbf{I}\right)-P\left(B\mid\mathbf{I}\right)P\left(A\mid B\mathbf{I}\right)\end{aligned} \tag{10}$$

$$\Rightarrow P\left(A+B|\mathbf{I}\right)=P\left(A\mid\mathbf{I}\right)+P\left(B\mid\mathbf{I}\right)-P\left(A\cdot B\mid\mathbf{I}\right)$$

6.1. *The concept of being "Mutually Exclusive"*

Based on the rule of sum, if the logical product of two statements A and B is FALSE, they are Mutually Exclusive (ME), and $P(AB|\mathbf{I})=0$; thus, Eq. (10) can be re-written as:

$$P\left(A+B|\mathbf{I}\right)=P\left(A\,|\,\mathbf{I}\right)+P\left(B\,|\,\mathbf{I}\right) \tag{11}$$

For instance, a statement and its negation are always ME, i.e. $P\left(A\overline{A}\,|\,\mathbf{I}\right)=0$. It is also noteworthy that being ME and being independent are two different concepts. In order to address this issue, assume that A and B are independent while they are ME; thus, $P(AB|\mathbf{I})=P(A|\mathbf{I})P(B|\mathbf{I})=0$. Hence, the product of their probabilities must be equal to zero, which means that at least one of the two probabilities must be equal to zero (which clearly is a special case). Loosely speaking, the fact that two statements are exclusive creates some sort of "dependence" between them in the sense that if one of them is TRUE the other one has to be FALSE.

Generally, if $\{A_i, i=1:n\}$ is a set of ME statements, the following relations hold:

$$\begin{aligned} &P\left(A_iA_j\,|\mathbf{I}\right)=0\,,\; i\neq j \Leftrightarrow A_i\cdot A_j \equiv \text{FALSE given } \mathbf{I} \\ &P\left(A_1+A_2+\cdots+A_n\,|\mathbf{I}\right)=\sum_{i=1}^{n}P\left(A_i\,|\mathbf{I}\right) \end{aligned} \tag{12}$$

6.2. *The concept of being "Collectively Exhaustive"*

In case where the two statements A and B built up the space of all possibilities (i.e., A+B $\equiv$ TRUE, they are Collectively Exhaustive (CE); i.e.,

$$P\left(A+B|\mathbf{I}\right)=1.0 \tag{13}$$

According to Eq. (7), a statement and its negation are CE, i.e. $P\left(A+\overline{A}\,|\,\mathbf{I}\right)=1.0$. Generally, if $\{A_i, i=1:n\}$ is a set of CE statements:

$$P\left(A_1+A_2+\cdots+A_n\,|\mathbf{I}\right)=1 \Leftrightarrow A_1+A_2+\cdots+A_n \equiv \text{TRUE given } \mathbf{I} \tag{14}$$

6.3. *The concept of being "Mutually Exclusive and Collectively Exhaustive"*

Assume that $\{A_i,\ i=1:n\}$ is a set of ME and CE (MECE) statements, with reference to Eq. (12) and Eq. (14), we have:

$$P\left(A_iA_j\,|\mathbf{I}\right)=0\,,\ i\neq j$$
$$P\left(A_1+A_2+\cdots+A_n\,|\mathbf{I}\right)=\sum_{i=1}^{n}P\left(A_i\,|\mathbf{I}\right)=1.0 \tag{15}$$

7. Total Probability Theorem (TPT)

Let $\{A_i,\ i=1:n\}$ be a set of MECE statements. The sum of statements $\{A_i,\ i=1:n\}$ is "TRUE" according to the CE property shown in Eq. (14); hence, the statement B can be written as the product of B and the TRUE statement ($A_1+A_2+\cdots+A_n$), and $P(B|\mathbf{I})$ can be expanded as follows:

$$P\left(B|\mathrm{I}\right)=P\left[B\cdot\underbrace{\left(A_1+A_2+\cdots+A_n\right)}_{\text{TRUE}}|\mathbf{I}\right]=P\left(BA_1+BA_2+\cdots+BA_n\,|\mathbf{I}\right) \tag{16}$$

The set $\{BA_i,\ i=1:n\}$ in Eq. (16) are ME since with reference to the first expression in Eq. (12):

$$P\left((BA_i)\cdot(BA_j)|\mathrm{I}\right)=P\left((B\cdot B)\cdot(A_i\cdot A_j)|\mathrm{I}\right)=P\left(B\cdot\mathrm{F}|\mathrm{I}\right)=0 \tag{17}$$

Thus, with reference to the second expression in Eq. (12) and the rule of product in Section 5:

$$P\left(B|\mathbf{I}\right)=\sum_{i=1}^{n}P\left(BA_i\,|\mathbf{I}\right)=\sum_{i=1}^{n}P\left(B|A_i\mathbf{I}\right)P\left(A_i\,|\mathbf{I}\right) \tag{18}$$

8. The Bayesian Theorem

Assume the following definitions: D denotes a statement regarding observed data, and $\mathrm{H}=\{\mathrm{H}_j,\ j=1:N\}$ is a set of MECE hypotheses statements. This means that we can define the "*likelihood*" as the probability $P(D|\mathrm{H}_i\mathbf{I})$ that D is TRUE given the hypothesis H_i, where $i\in\{1:N\}$, and background information **I**. The probability that hypothesis H_i is TRUE given data D and background information **I** can be calculated as:

$$P\left(\mathrm{H}_i \middle| D\mathbf{I}\right) = \frac{P\left(D \cdot \mathrm{H}_i \middle| \mathbf{I}\right)}{P\left(D \middle| \mathbf{I}\right)} = \frac{P\left(D \middle| \mathrm{H}_i \mathbf{I}\right) P\left(\mathrm{H}_i \middle| \mathbf{I}\right)}{P\left(D \middle| \mathbf{I}\right)} = \frac{P\left(D \middle| \mathrm{H}_i \mathbf{I}\right) P\left(\mathrm{H}_i \middle| \mathbf{I}\right)}{\sum_{j=1}^{N} P\left(D \middle| \mathrm{H}_j \mathbf{I}\right) P\left(\mathrm{H}_j \middle| \mathbf{I}\right)} \tag{19}$$

where $P(\mathrm{H}_i|D\mathbf{I})$ is called the "*posterior probability*" or "*a posteriori*" which indicates how plausible is H_i given that we know the observed data D and the background information I; and $\mathrm{P}(\mathrm{H}_i|\mathbf{I})$ is the "*prior probability*" or "*a priori*" that indicates the plausibility of H_i given I. Substituting $P(D|\mathbf{I})$ with c, the Bayesian expression in Eq. (19) can be simplified as follows:

$$P\left(\mathrm{H}_i \middle| D\mathbf{I}\right) = c^{-1} P\left(D \middle| \mathrm{H}_i \mathbf{I}\right) P\left(\mathrm{H}_i \middle| \mathbf{I}\right) \tag{20}$$

The reason for leaving the nominator as a constant is that it acts as a normalising constant and is equal to the sum of the product of likelihood and prior for all the plausible MECE hypotheses H_i.