MODULE 1

PROBABILITY

LECTURE 3

Topics

1.2 AXIOMATIC APPROACH TO PROBABILITY AND PROPERTIES OF PROBABILITY MEASURE

Theorem 2.3

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $E_1, E_2, \dots, E_n \in \mathcal{F}$ $(n \in \mathbb{N},\ n \geq 2\,)$. Then, under the notations of Theorem 2.2,

1.2.1.1 Boole's Inequality

$$S_{1,n} + S_{2,n} \leq P\left(\bigcup_{i=1}^{n} E_i\right) \leq S_{1,n};$$

1.2.1.2 Bonferroni's Inequality

$$P\left(\bigcap_{i=1}^{n} E_i\right) \geq S_{1,n} - (n-1).$$

Proof.

(i) We will use the principle of mathematical induction. We have

$$\begin{aligned} P(E_1 \cup E_2) &= \underbrace{P(E_1) + P(E_2)}_{S_{1,2}} \underbrace{-P(E_1 \cap E_2)}_{S_{2,2}} \\ &= S_{1,2} + S_{2,2} \\ &\leq S_{1,2}, \end{aligned}$$

where $S_{1,2} = P(E_1) + P(E_2)$ and $S_{2,2} = -P(E_1 \cap E_2) \leq 0$.

Thus the result is true for $n = 2$. Now suppose that the result is true for $n \in \{2, 3, \ldots, m\}$ for some positive integer $m\,(\geq 2)$, i.e., suppose that for arbitrary events $F_1, \ldots, F_m \in \mathcal{F}$

$$P\left(\bigcup_{i=1}^{k} F_i\right) \leq \sum_{i=1}^{k} P(F_i)\,, k = 2, 3, \ldots, m \tag{2.5}$$

and

$$P\left(\bigcup_{i=1}^{k} F_i\right) \geq \sum_{i=1}^{k} P(F_i) - \sum_{1 \leq i < j \leq k} P\big(F_i \cap F_j\big), \qquad k = 2, 3, \ldots, m. \tag{2.6}$$

Then

$$\begin{aligned} P\left(\bigcup_{i=1}^{m+1} E_i\right) &= P\left(\left(\bigcup_{i=1}^{m} E_i\right) \cup E_{m+1}\right) \\ &\leq P\left(\bigcup_{i=1}^{m} E_i\right) + P(E_{m+1}) \qquad \text{(using (2.5) for } k = 2\text{)} \\ &\leq \sum_{i=1}^{m} P(E_i) + P(E_{m+1}) \qquad \text{(using (2.5) for } k = m\text{)} \\ &= \sum_{i=1}^{m+1} P(E_i) = S_{1,m+1}. \qquad (2.7) \end{aligned}$$

Also,

$$\begin{aligned} P\left(\bigcup_{i=1}^{m+1} E_i\right) &= P\left(\left(\bigcup_{i=1}^{m} E_i\right) \cup E_{m+1}\right) \\ &= P\left(\bigcup_{i=1}^{m} E_i\right) + P(E_{m+1}) - P\left(\left(\bigcup_{i=1}^{m} E_i\right) \cap E_{m+1}\right) \quad \text{(using Theorem 2.2)} \\ &= P\left(\bigcup_{i=1}^{m} E_i\right) + P(E_{m+1}) - P\left(\bigcup_{i=1}^{m} (E_i \cap E_{m+1})\right). \qquad (2.8) \end{aligned}$$

Using (2.5), for $k = m$, we get

$$P\left(\bigcup_{i=1}^{m}(E_i \cap E_{m+1})\right) \le \sum_{i=1}^{m} P\,(E_i \cap E_{m+1}), \tag{2.9}$$

and using (2.6), for $k = m$, we get

$$P\left(\bigcup_{i=1}^{m} E_i\right) \ge S_{1,m} + S_{2,m}. \tag{2.10}$$

Now using (2.9) and (2.10) in (2.8), we get

$$\begin{aligned} P\left(\bigcup_{i=1}^{m+1} E_i\right) &\ge S_{1,m} + S_{2,m} + P(E_{m+1}) - \sum_{i=1}^{m} P(E_i \cap E_{m+1}) \\ &= \sum_{i=1}^{m+1} P(E_i) - \sum_{1 \le i < j \le m+1} P\left(E_i \cap E_j\right) \\ &= S_{1,m+1} + S_{2,m+1}. \end{aligned} \tag{2.11}$$

Combining (2.7) and (2.11), we get

$$S_{1,m+1} + S_{2,m+1} \le P\left(\bigcup_{1=1}^{m+1} E_i\right) \le S_{1,m+1,}$$

and the assertion follows by principle of mathematical induction.

(ii) We have

$$\begin{aligned} P\left(\bigcap_{\mathrm{i}=1}^{n} E_i\right) &= 1 - P\left(\left(\bigcap_{\mathrm{i}=1}^{n} E_i\right)^c\right) \\ &= 1 - P\left(\bigcup_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{E}_{\mathrm{i}}^c\right) \\ &\ge 1 - \sum_{1=1}^{n} P\,({E_i}^c) \qquad \text{(using Boole's inequality)} \end{aligned}$$

$$= 1 - \sum_{i=1}^{n} \big(1 - P(E_i)\big)$$

$$= \sum_{i=1}^{n} P(E_i) - (n-1). \blacksquare$$

Remark 2.4

Under the notation of Theorem 2.2 we can in fact prove the following inequalities:

$$\sum_{j=1}^{2k} S_{j,n} \leq P\left(\bigcup_{j=1}^{n} E_j\right) \leq \sum_{j=1}^{2k-1} S_{j,n}\,, k = 1,2, \dots, \left[\frac{n}{2}\right],$$

where $\left[\frac{n}{2}\right]$ denotes the largest integer not exceeding $\frac{n}{2}$. $\blacksquare$

Corollary 2.1

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $E_1, E_2, \dots, E_n \in \mathcal{F}$ be events. Then

(i) $P(E_i) = 0,\ i = 1, \dots, n \Leftrightarrow P(\bigcup_{i=1}^{n} E_i) = 0;$
(ii) $P(E_i) = 1,\ i = 1, \dots, n \Leftrightarrow P(\bigcap_{i=1}^{n} E_i) = 1.$

Proof.

(i) First suppose that $P(E_i) = 0, i = 1, \dots, n.$ Using Boole's inequality, we get

$$0 \leq P\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} P(E_i) = 0.$$

It follows that $P(\bigcup_{i=1}^{n} E_i) = 0.$
Conversely, suppose that $P\big(\bigcup_{j=1}^{n} E_j\big) = 0$. Then $E_i \subseteq \bigcup_{j=1}^{n} E_j\,, i = 1, \dots, n$, and therefore,

$$0 \leq P(E_i) \leq P\left(\bigcup_{j=1}^{n} E_j\right) = 0, \qquad i = 1, \dots, n,$$

i.e., $P(E_i) = 0,\ i = 1, \dots, n.$

(ii) We have
$P(E_i) = 1, i = 1, \dots, n \Leftrightarrow P(E_i^c) = 0, i = 1, \dots, n$

$$\Leftrightarrow P\left(\bigcup_{i=1}^{n} E_i^c\right) = 0 \quad \text{(using (i))}$$

$$\Leftrightarrow P\left(\left(\bigcup_{i=1}^{n} E_i^c\right)^c\right) = 1,$$

$$\Leftrightarrow P\left(\bigcap_{i=1}^{n} E_i\right) = 1. \blacksquare$$

Definition 2.4

A countable collection $\{E_i : i \in \Lambda\}$ of events is said to be exhaustive if $P(\bigcup_{i\in\Lambda} E_i) = 1.$ ■

1.2.2 Equally Likely Probability Models

Example 2.2

Consider a probability space $(\Omega, \mathcal{F}, P)$. Suppose that, for some positive integer $k \geq 2$, $\Omega = \bigcup_{i=1}^{k} C_i$, where $C_1, C_2, \ldots, C_k$ are mutually exclusive, exhaustive and equally likely events, i.e., $C_i \cap C_j = \phi$, if $i \neq j$, $P\left(\bigcup_{i=1}^{k} C_i\right) = \sum_{i=1}^{k} P(C_i) = 1$ and $P(C_1) = \cdots = P(C_k) = \frac{1}{k}$. Further suppose that an event $E \in \mathcal{F}$ can be written as

$$E = C_{i_1} \cup C_{i_2} \cup \cdots \cup C_{i_r},$$

where $\{i_1, \ldots, i_r\} \subseteq \{1, \ldots, k\}$, $C_{i_j} \cap C_{i_k} = \phi$, $j \neq k$ and $r \in \{2, \ldots, k\}$. Then

$$P(E) = \sum_{j=1}^{r} P\left(C_{i_j}\right) = \frac{r}{k}.$$

Note that here k is the total number of ways in which the random experiment can terminate (number of partition sets $C_1, \ldots, C_k$), and r is the number of ways that are favorable to $E \in \mathcal{F}$.

Thus, for any $E \in \mathcal{F}$,

$$P(E) = \frac{\text{number of cases favorable to } E}{\text{total number of cases}} = \frac{r}{k},$$

which is the same as classical method of assigning probabilities. Here the assumption that $C_1, \ldots, C_k$ are equally likely is a part of probability modeling. ■

For a finite sample space Ω, when we say that an experiment has been performed at random we mean that various possible outcomes in Ω are equally likely. For example when we say that two numbers are chosen at random, without replacement, from the set $\{1,2,3\}$ then $\Omega = \{\{1,2\},\{1,3\},\{2,3\}\}$ and $P(\{1,2\}) = P(\{1,3\}) = P(\{2,3\}) = \frac{1}{3}$, where $\{i,j\}$ indicates that the experiment terminates with chosen numbers as i and $j, i, j \in \{1,2,3\}, i \neq j$.

Example 2.3

Suppose that five cards are drawn at random and without replacement from a deck of 52 cards. Here the sample space Ω comprises of all $\binom{52}{5}$ combinations of 5 cards. Thus number of favorable cases= $\binom{52}{5} = k$, say. Let $C_1, \dots, C_k$ be singleton subsets of Ω.Then $\Omega = \bigcup_{i=1}^{k} C_i$ and $P(C_1) = \cdots = P(C_k) = \frac{1}{k}$. Let E_1 be the event that each card is spade. Then

$$\text{Number of cases favorable to } E_1 = \binom{13}{5}.$$

Therefore,

$$P(E_1) = \frac{\binom{13}{5}}{\binom{52}{5}}\cdot$$

Now let E_2 be the event that at least one of the drawn cards is spade. Then E_2^c is the event that none of the drawn cards is spade, and number of cases favorable to $E_2^c = \binom{39}{5}\cdot$ Therefore,

$$P(E_2^c) = \frac{\binom{39}{5}}{\binom{52}{5}},$$

and $P(E_2) = 1 - P(E_2^c) = 1 - \frac{\binom{39}{5}}{\binom{52}{5}}\cdot$

Let E_3 be the event that among the drawn cards three are kings and two are queens. Then number of cases favorable to $E_3 = \binom{4}{3}\binom{4}{2}$ and, therefore,

$$P(E_3) = \frac{\binom{4}{3}\binom{4}{2}}{\binom{52}{5}}\cdot$$

Similarly, if E_4 is the event that among the drawn cards two are kings, two are queens and one is jack, then

$$P(E_4) = \frac{\binom{4}{2}\binom{4}{2}\binom{4}{1}}{\binom{52}{5}}. \blacksquare$$

Example 2.4

Suppose that we have $n\ (\geq 2)$ letters and corresponding n addressed envelopes. If these letters are inserted at random in n envelopes find the probability that no letter is inserted into the correct envelope.

Solution. Let us label the letters as $L_1, L_2, \dots, L_n$ and respective envelopes as $A_1, A_2, \dots, A_n$. Let E_i denote the event that letter L_i is (correctly) inserted into envelope A_i, $i = 1, 2, \dots, n$. We need to find $P(\bigcap_{i=1}^{n} E_i^c)$. We have

$$P\left(\bigcap_{i=1}^{n} E_i^c\right) = P\left(\left(\bigcup_{i=1}^{n} E_i\right)^c\right) = 1 - P\left(\bigcup_{i=1}^{n} E_i\right) = 1 - \sum_{k=1}^{n} S_{k,n},$$

where, for $k \in \{1, 2, \dots, n\}$,

$$S_{k,n} = (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} P\left(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}\right).$$

Note that n letters can be inserted into n envelopes in $n!$ ways. Also, for

$1 \leq i_1 < i_2 < \cdots < i_k \leq n, E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}$ is the event that letters $L_{i_1}, L_{i_2}, \dots, L_{i_k}$ are inserted into correct envelopes. Clearly number of cases favorable to this event is $(n-k)!$. Therefore, for $1 \leq i_1 < i_2 < \cdots < i_k \leq n$,

$$P\left(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}\right) = \frac{(n-k)!}{n!}$$

$$\Rightarrow S_{k,n} = (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \frac{(n-k)!}{n!}$$

$$= (-1)^{k-1} \binom{n}{k} \frac{(n-k)!}{n!}$$

$$= \frac{(-1)^{k-1}}{k!}$$

$$\Rightarrow P\left(\bigcap_{i=1}^{n} E_i^c\right) = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots + \frac{(-1)^n}{n!}. \blacksquare$$