1.3. THE PROBABILITY FUNCTION

DEFINITION

For an event $E \subseteq \Omega$, the **probability that** E **occurs** is written $\mathrm{P}(E)$.

[*Formally, we define probability triple* $(\Omega, \mathcal{A}, \mathrm{P})$ *where* Ω *is the sample space,* $\mathcal{A}$ *is a* σ*-algebra of subsets of* Ω*, and* P *is a probability "measure" operating on* $\mathcal{A}$.]

Interpretation : $\mathrm{P}(.)$ is a *set-function* that assigns "weight" to collections of possible outcomes of an experiment. There are many ways to think about precisely how this assignment is achieved;

CLASSICAL : "Consider equally likely sample outcomes ..."

FREQUENTIST : "Consider long-run *relative frequencies* ..."

SUBJECTIVE : "Consider personal degree of belief ..."

or merely think of $\mathrm{P}(.)$ as a set-function.

1.4. PROPERTIES OF P(.): THE SIMPLE AXIOMS OF PROBABILITY

Consider sample space Ω. Then probability function $\mathrm{P}(.)$ satisfies the following properties:

AXIOM 1 Let $E \subseteq \Omega$. Then $0 \leq \mathrm{P}(E) \leq 1$.

AXIOM 2 $\mathrm{P}(\Omega) = 1$.

AXIOM 3 If $E, F \subseteq \Omega$, with $E \cap F = \emptyset$, then $\mathrm{P}(E \cup F) = \mathrm{P}(E) + \mathrm{P}(F)$.

NOTE : Axiom 3 can be re-stated if we can consider an algebra $\mathcal{A}$ of subsets of Ω. If events $A_1, A_2, ...$ are disjoint elements of $\mathcal{A}$, then replace Axiom 3 by requiring that, for $n \geq 1$,

AXIOM 3* $$\mathrm{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathrm{P}(A_i).$$

Furthermore, if $\mathcal{A}$ is a σ-algebra, then Axiom 3* can be replaced by requiring that

AXIOM 3† $$\mathrm{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathrm{P}(A_i).$$

Note that, if $\mathcal{A}$ is a σ-algebra, then

$$\text{AXIOM } 3^{\dagger} \implies \text{AXIOM } 3^{*} \implies \text{AXIOM } 3$$

that is,

$$\text{COUNTABLE ADDITIVITY} \implies \text{FINITE ADDITIVITY} \implies \text{ADDITIVITY.}$$

<u>COROLLARIES</u> For events $E, F \subseteq \Omega$

$\mathrm{P}(E^{'}) = 1 - \mathrm{P}(E)$, and hence $\mathrm{P}(\emptyset) = 0$.

If $E \subseteq F$, then $\mathrm{P}(E) \leq \mathrm{P}(F)$.

In general, $\mathrm{P}(E \cup F) = \mathrm{P}(E) + \mathrm{P}(F) - \mathrm{P}(E \cap F)$.

$\mathrm{P}(E \cap F^{'}) = \mathrm{P}(E) - \mathrm{P}(E \cap F)$

Boole's Inequality

$$\mathrm{P}(E \cup F) \leq \mathrm{P}(E) + \mathrm{P}(F).$$

Bonferroni's Inequality

$$\mathrm{P}(E \cap F) \geq \mathrm{P}(E) + \mathrm{P}(F) - 1.$$

NOTE: The **general addition rule** for probabilities and Boole's Inequality extend to more than two events. Let $E_1, ..., E_n$ be events in Ω. Then

$$\mathrm{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i} \mathrm{P}(E_i) - \sum_{i<j} \mathrm{P}(E_i \cap E_j) + \sum_{i<j<k} \mathrm{P}(E_i \cap E_j \cap E_k) - ... + (-1)^n \, \mathrm{P}\left(\bigcap_{i=1}^{n} E_i\right)$$

and

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

To prove these results, construct the events $F_1 = E_1$ and $F_i = E_i \cap \left(\bigcup_{k=1}^{i-1} E_k\right)^{'}$ for $i = 2, 3, ..., n$.

Then $F_1, F_2, ...F_n$ are disjoint, and $\bigcup_{i=1}^{n} E_i = \bigcup_{i=1}^{n} F_i$, so

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

Now, by Corollary five above

$$\begin{aligned} \mathrm{P}(F_i) &= \mathrm{P}(E_i) - \mathrm{P}\left(E_i \cap \bigcup_{k=1}^{i-1} E_k\right) \qquad i = 2, 3, ..., n. \\ &= \mathrm{P}(E_i) - \mathrm{P}\left(\bigcup_{k=1}^{i-1} (E_i \cap E_k)\right) \end{aligned}$$

and the result follows by recursive expansion of the second term for $i = 2, 3, ...n$.

NOTE: Suppose Axiom $3^\dagger$ holds for a given Ω, and let $A_1, A_2, ... \subseteq \Omega$ be a non-decreasing sequence of events. Then, if A is the limit event for the monotone sequence, we have

$$
\begin{aligned}
\mathrm{P}(A) &= \mathrm{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \mathrm{P}\left(A_1 \cup \bigcup_{i=2}^{\infty}(A_i \cap A_{i-1}^{'})\right) \\
&= \mathrm{P}(A_1) + \sum_{i=2}^{\infty} \mathrm{P}(A_i \cap A_{i-1}^{'}) \qquad \text{by Axiom } 3^\dagger \\
&= \mathrm{P}(A_1) + \lim_{n\to\infty}\left\{\sum_{i=2}^{n}[\mathrm{P}(A_i) - \mathrm{P}(A_{i-1})]\right\} \\
&= \lim_{n\to\infty} \mathrm{P}(A_n)
\end{aligned}
$$

as the sequence $A_1,\ A_2 \cap A_1^{'},\ A_3 \cap A_2^{'},\ A_4 \cap A_3^{'}$ and so on is a disjoint sequence of events whose union is identical to A.

Hence $\mathrm{P}(A) = \mathrm{P}\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} \mathrm{P}(A_n)$, and thus P is a *continuous* function.

NOTE: A similar result holds for non-increasing sequences.

NOTE: The converse of this result is also true, that is, if P is continuous, then Axiom $3^\dagger$ holds, and P is countably additive.

1.5. CONDITIONAL PROBABILITY

For events $E, F \subseteq \Omega$ the **conditional probability** that F occurs **given** that E occurs is written $\mathrm{P}(F|E)$, and is defined by

$$\mathrm{P}(F|E) = \frac{\mathrm{P}(E \cap F)}{\mathrm{P}(E)}$$

if $\mathrm{P}(E) > 0$.

NOTE: $\mathrm{P}(E \cap F) = \mathrm{P}(E)\mathrm{P}(F|E)$, and in general, for events $E_1, ..., E_k$,

$$\mathrm{P}\left(\bigcap_{i=1}^{k} E_i\right) = \mathrm{P}(E_1)\mathrm{P}(E_2|E_1)\mathrm{P}(E_2|E_1 \cap E_2)...\mathrm{P}(E_k|E_1 \cap E_2 \cap ... \cap E_{k-1}).$$

This result is known as the CHAIN or MULTIPLICATION RULE.

DEFINITION

Events E and F are **independent** if

$$\mathrm{P}(E|F) = \mathrm{P}(E) \text{ so that } \mathrm{P}(E \cap F) = \mathrm{P}(E)\mathrm{P}(F)$$

Extension : Events $E_1, ..., E_k$ are independent if, for **every** subset of events of size $l \leq k$, indexed by $\{i_1, ..., i_l\}$, say,

$$\mathrm{P}\left(\bigcap_{j=1}^{l} E_{i_j}\right) = \prod_{j=1}^{l} \mathrm{P}(E_{i_j}).$$

1.6. THE THEOREM OF TOTAL PROBABILITY

THEOREM

Let $E_1, ..., E_k$ be a partition of Ω, and let $F \subseteq \Omega$. Then

$$\mathrm{P}(F) = \sum_{i=1}^{k} \mathrm{P}(F|E_i)\mathrm{P}(E_i)$$

Proof

$E_1, ..., E_k$ form a partition of Ω, and $F \subseteq \Omega$, so

$$\begin{aligned} F &= (F \cap E_1) \cup ... \cup (F \cap E_k) \\ \Longrightarrow \quad \mathrm{P}(F) &= \sum_{i=1}^{k} \mathrm{P}(F \cap E_i) \qquad \text{(by AXIOM } 3^*\text{, as } E_i \cap E_j = \emptyset) \\ &= \sum_{i=1}^{k} \mathrm{P}(\ F \mid E_i\ )\mathrm{P}(E_i) \end{aligned}$$

Extension: If we assume that Axiom $3^\dagger$ holds, that is, that P is countably additive, then the theorem still holds, that is, if $E_1, E_2, ...$ are a partition of Ω, and $F \subseteq \Omega$, then

$$\mathrm{P}(F) = \sum_{i=1}^{\infty} \mathrm{P}(\ F \cap E_i\ ) = \sum_{i=1}^{\infty} \mathrm{P}(\ F \mid E_i\ )\mathrm{P}(E_i)$$

if $\mathrm{P}(E_i) > 0$ for all i.

1.7. BAYES THEOREM

THEOREM

Suppose $E, F \subseteq \Omega$, with $\mathrm{P}(E), \mathrm{P}(F) > 0$. Then

$$\mathrm{P}(E|F) = \frac{\mathrm{P}(F|E)\mathrm{P}(E)}{\mathrm{P}(F)}$$

Proof

$$\mathrm{P}(E|F)\mathrm{P}(F) = \mathrm{P}(E \cap F) = \mathrm{P}(F|E)\mathrm{P}(E), \text{ so } \mathrm{P}(E|F)\mathrm{P}(F) = \mathrm{P}(F|E)\mathrm{P}(E).$$

Extension: If $E_1, ..., E_k$ are disjoint, with $\mathrm{P}(E_i) > 0$ for $i = 1, ..., k$, and form a partition of $F \subseteq \Omega$, then

$$\mathrm{P}(\ E_i \mid F\ ) = \frac{\mathrm{P}(\ F \mid E_i\ )\mathrm{P}(E_i)}{\sum_{i=1}^{k} \mathrm{P}(\ F \mid E_i\ )\mathrm{P}(E_i)}$$

Extension: The extension to the countably additive (infinite) case also holds.