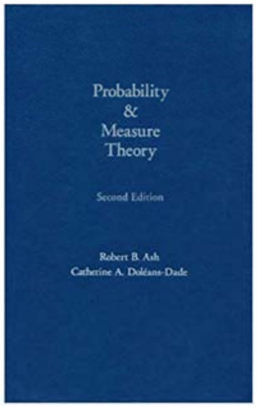
Real Analysis

Chapter 4. Basic Concepts of Probability

4.5. Conditional Probability—Proofs of Theorems



Theorem 4.5.1

Theorem 4.5.1.

(a) Let $P(A) > 0$. Events A and B are independent if and only if $P(B \mid A) = P(B)$.

(b) Let $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$. Then

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots$$

$$\cdots P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

Proof. (a) By definition, $P(A \cap B) = P(B \mid A)P(A)$. If A and B are independent then $P(A \cap B) = P(A)P(B)$ so that we have $P(B) = P(B \mid A)$, as claimed. If $P(B \mid A) = P(B)$, then we have $P(A \cap B) = P(A)P(B)$ and so A and B are independent, as claimed.

(b) Since $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$ then by monotonicity of measure, $P(A_1 \cap A_2 \cap \cdots \cap A_k) > P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$ for $k = 1, 2, \ldots, n-1$.

Theorem 4.5.1 (continued)

Proof (continued). Inductively, we have

$$\begin{aligned}
& P((A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \cap A_n) \\
= \; & P((A_1 \cap A_2 \cap \cdots \cap A_{n-2}) \cap A_{n-1})P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \\
= \; & P((A_1 \cap A_2 \cap \cdots \cap A_{n-2})P(A_{n-1} \mid A_1 \cap A_2 \cap \cdots \cap A_{n-2}) \\
& \times P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \\
& \vdots \\
= \; & P((A_1 \cap A_2 \cap \cdots \cap A_i)P(A_{i+1} \mid A_1 \cap A_2 \cap \cdots \cap A_i) \\
& \times P(A_{i+2} \mid A_1 \cap A_2 \cap \cdots \cap A_{i+1}) \cdots P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \\
& \vdots \\
= \; & P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots P(A_{n-1} \mid A_1 \cap A_2 \cap \cdots \cap A_{n-2}) \\
& \times P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1})
\end{aligned}$$

□

Theorem 4.5.2

Theorem 4.5.2. Theorem of Total Probability.
Let $B_1, B_2, \ldots$ form a finite or countably infinite family of mutually exclusive and exhaustive events; that is, $\uplus_i B_i = \Omega$.

(a) If A if any event, then $P(A) = \sum_i P(A \cap B_i)$. Thus $P(A)$ is calculated by making a list of mutually exclusive exhaustive ways in which A can happen, and adding the individual probabilities.

(b) $P(A) = \sum_i P(B_i)P(A \mid B_i)$ where the sum is taken over those i for which $P(B_i) > 0$. Thus $P(A)$ is a weighted average of the conditional probabilities $P(A \mid B_i)$.

Proof. (a) We have by countable additivity:

$$P(A) = P(A \cap \Omega) = P(A \cap (\uplus_i B_i)) = P(\uplus_i (A \cap B_i)) = \sum_i P(A \cap B_i).$$

Theorem 4.5.2 (continued)

Proof (continued). (b) By part (a), $P(A) = \sum_i P(A \cap B)i)$. If $P(B_i) = 0$ for some i then by monotonicity $P(A \cap B_i) \leq P(B_i) = 0$. If $P(B_i) > 0$ then by definition of conditional probability, $P(A \cap B_i) = P(B_i)P(A \mid B_i)$. So

$$P(A) = \sum_i P(A \cap B_i) = \sum_{i'} P(B_{i'} P(A \mid B_{i'})$$

where i' ranges over all values for which $P(B_{i'}) > 0$, as claimed. □