2

Basic Concepts

Throughout the book the symbol N will denote a non-empty finite set of *variables*. The intended interpretation is that the variables correspond to primitive factors described by random variables. In Chapter 3 variables will be represented by nodes of a graph. The set N will also serve as the basic set for non-graphical tools of discrete mathematics introduced in this monograph (semi-graphoids, imsets etc.).

Convention 1. The following conventions will be used throughout the book. Given sets $A, B \subseteq N$ the juxtaposition AB will denote their union $A \cup B$. The following symbols will be reserved for sets of numbers: $\mathbb{R}$ will denote *real numbers*, $\mathbb{Q}$ *rational numbers*, $\mathbb{Z}$ *integers*, $\mathbb{Z}^+$ *non-negative integers* (including 0), $\mathbb{N}$ *natural numbers* (that is, positive integers excluding 0). The symbol $|A|$ will be used to denote the number of elements of a finite set A, that is, its *cardinality*. The symbol $|x|$ will also denote the *absolute value* of a real number x, that is, $|x| = \max\{x, -x\}$. $\diamondsuit$

2.1 Conditional independence

A basic notion of the monograph is a *probability measure over* N. This phrase will be used to describe the situation in which a measurable space $(\mathsf{X}_i, \mathcal{X}_i)$ is given for every $i \in N$ and a probability measure P is defined on the Cartesian product of these measurable spaces $(\prod_{i\in N} \mathsf{X}_i, \prod_{i\in N} \mathcal{X}_i)$. In this case I will use the symbol $(\mathsf{X}_A, \mathcal{X}_A)$ as a shorthand for $(\prod_{i\in A} \mathsf{X}_i, \prod_{i\in A} \mathcal{X}_i)$ for every $\emptyset \neq A \subseteq N$. The *marginal* of P for $\emptyset \neq A \subset N$, denoted by P^A, is defined by the formula

$$P^A(\mathsf{A}) = P(\mathsf{A} \times \mathsf{X}_{N\setminus A}) \quad \text{for} \quad \mathsf{A} \in \mathcal{X}_A \,.$$

Moreover, let us accept two natural conventions. First, the marginal of P for $A = N$ is P itself, that is, $P^N \equiv P$. Second, a fully formal convention is that the marginal of P for $A = \emptyset$ is a probability measure on a (fixed appended)

measurable space $(\mathsf{X}_\emptyset, \mathcal{X}_\emptyset)$ with a trivial σ-algebra $\mathcal{X}_\emptyset = \{\emptyset, \mathsf{X}_\emptyset\}$. Observe that a measurable space of this kind only admits one probability measure $P^\emptyset$.

To give the definition of conditional independence within this framework one needs a certain general understanding of the concept of conditional probability. Given a probability measure P over N and disjoint sets $A, C \subseteq N$, *conditional probability on* X_A *given* C (more specifically given $\mathcal{X}_C$) will be understood as a function of two arguments $P_{A|C} : \mathcal{X}_A \times \mathsf{X}_C \to [0,1]$ which ascribes an $\mathcal{X}_C$-measurable function $P_{A|C}(\mathsf{A}|\star)$ to every $\mathsf{A} \in \mathcal{X}_A$ such that

$$P^{AC}(\mathsf{A} \times \mathsf{C}) = \int_{\mathsf{C}} P_{A|C}(\mathsf{A}|x) \; \mathrm{d}P^C(x) \qquad \text{for every } \; \mathsf{C} \in \mathcal{X}_C \, .$$

Note that no restriction concerning the mappings $\mathsf{A} \mapsto P_{A|C}(\mathsf{A}|x)$, $x \in \mathsf{X}_C$ (often called the regularity requirement – see Section A.6.4, Remark A.1) is needed within this general approach. Let me emphasize that $P_{A|C}$ only depends on the marginal P^{AC} and that it is defined, for a fixed $\mathsf{A} \in \mathcal{X}_A$, uniquely within the equivalence P^C-almost everywhere (P^C-a.e.). Observe that, owing to the convention above, if $C = \emptyset$ then the conditional probability $P_{A|C}$ coincides, in fact, with the marginal for A, that means, one has $P_{A|\emptyset} \equiv P^A$ (because a constant function can be identified with its value).

Remark 2.1. The conventions above are in accordance with the following unifying perspective. Realize that for every $\emptyset \neq A \subset N$ the measurable space $(X_A, \mathcal{X}_A)$ is isomorphic to the space $(\mathsf{X}_N, \bar{\mathcal{X}}_A)$ where $\bar{\mathcal{X}}_A \subseteq \mathcal{X}_N$ is the coordinate σ-algebra representing the set A, namely

$$\bar{\mathcal{X}}_A = \{\mathsf{A} \times \mathsf{X}_{N \setminus A} \, ; \; \mathsf{A} \in \mathcal{X}_A\} = \{\mathsf{B} \in \mathcal{X}_N \, ; \; \mathsf{B} = \mathsf{A} \times \mathsf{X}_{N \setminus A} \quad \text{for} \;\; \mathsf{A} \subseteq \mathsf{X}_A\} \, .$$

Thus, $A \subseteq B \subseteq N$ is reflected by $\bar{\mathcal{X}}_A \subseteq \bar{\mathcal{X}}_B$ and it is natural to require that the empty set $\emptyset$ is represented by the trivial σ-algebra $\bar{\mathcal{X}}_\emptyset$ over X_N and N is represented by $\bar{\mathcal{X}}_N = \mathcal{X}_N$. Using this point of view, the marginal P^A corresponds to the restriction of P to $\bar{\mathcal{X}}_A$, and $P_{A|C}$ corresponds to the concept of conditional probability with respect to the σ-algebra $\bar{\mathcal{X}}_C$. Thus, the existence and the uniqueness of $P_{A|C}$ mentioned above follows from basic measure-theoretical facts. For details see the Appendix, Section A.6.4. $\triangle$

Given a probability measure P over N and pairwise disjoint subsets $A, B, C \subseteq N$ one says that A is *conditionally independent of* B *given* C *with respect to* P and writes $A \perp\!\!\!\perp B \,|\, C \; [P]$ if for every $\mathsf{A} \in \mathcal{X}_A$ and $\mathsf{B} \in \mathcal{X}_B$

$$P_{AB|C}(\mathsf{A} \times \mathsf{B}|x) = P_{A|C}(\mathsf{A}|x) \cdot P_{B|C}(\mathsf{B}|x) \qquad \text{for } \; P^C\text{-a.e. } \; x \in \mathsf{X}_C \, . \quad (2.1)$$

Observe that in case $C = \emptyset$ it collapses to a simple equality $P^{AB}(\mathsf{A} \times \mathsf{B}) = P^A(\mathsf{A}) \cdot P^B(\mathsf{B})$, that is, to a classic independence concept. Note that the validity of (2.1) does not depend on the choice of versions of conditional probability given C since these are determined uniquely within equivalence P^C-a.e.