CHAPTER 3

# A Classification of Assignment Mechanisms

## 3.1 INTRODUCTION

As discussed in Chapter 1, the fundamental problem of causal inference is the presence of missing data – for each unit we can observe at most one of the potential outcomes. A key component in a causal analysis is, therefore, what we call the *assignment mechanism*: the process that determines which units receive which treatments, hence which potential outcomes are realized and thus can be observed, and, conversely, which potential outcomes are missing. In this chapter we introduce a taxonomy of assignment mechanisms that will serve as the organizing principle for this text. Formally, the assignment mechanism describes, as a function of all covariates and of all potential outcomes, the probability of any vector of assignments. We consider three basic restrictions on assignment mechanisms:

1. *Individualistic assignment*: This limits the dependence of a particular unit's assignment probability on the values of covariates and potential outcomes for other units.
2. *Probabilistic assignment*: This requires the assignment mechanism to imply a nonzero probability for each treatment value, for every unit.
3. *Unconfounded assignment*: This disallows dependence of the assignment mechanism on the potential outcomes.

Following Cochran (1965), we also make a distinction between experiments, where the assignment mechanism is both known and controlled by the researcher, and observational studies, where the assignment mechanism is not known to, or not under the control of, the researcher.

We consider three classes of assignment mechanisms, covered in Parts II, III, IV, V, and VI of this book. The first class, studied in Part II, corresponds to what we call *classical randomized experiments*. Here the assignment mechanism satisfies all three restrictions on the assignment process, and, moreover, the researcher knows and controls the functional form of the assignment mechanism. Such designs are well understood, and in such settings causal effects are often relatively straightforward to estimate, and, moreover, it is often possible to do finite sample inference.

We refer to the second class of assignment mechanisms, studied in Parts III and IV of this text, as *regular assignment mechanisms*. This class comprises assignment

mechanisms that, like classical randomized experiments, are individualistic, probabilistic, and unconfounded, but, in contrast to classical randomized experiments, the assignment mechanism need not be under the control of, or known by, the researcher. When the assignment mechanism is not under the control of the researcher, the restrictions on the assignment mechanism that make it regular are now usually assumptions, and they are typically not satisfied by design, as they are in classical randomized experiments. In general, we will not be sure whether these assumptions hold in any specific application, and in later chapters we will discuss methods for assessing their plausibility, as well as investigating the sensitivity to violations of them.

In practice, the regular observational study is a setting of great importance. It has been studied extensively from a theoretical perspective and is widely used in empirical work. Many, but not all, of the methods applicable to randomized experiments can be used, but often modifications to the specific methods are critical to enhance the credibility of the results. The simple methods that suffice in the context of randomized experiments tend to be more controversial when applied with regular assignment mechanisms. The concerns these simple methods raise are particularly serious if the covariate distributions under the various treatment regimes are substantially different, or *unbalanced* in our terminology. In that case, it can be very important, for the purpose of making credible causal inferences, to have an initial, what we call *design* stage of the study. In this design stage, the data on covariate values and treatment assignment (but, importantly, not the final outcome data) are analyzed in order to assemble samples with improved balance in covariate distributions, somewhat in parallel with the design stage of randomized experiments. Often in this setting, the number of pre-treatment variables is substantial, typically because, conditional on a large number of pre-treatment variables, unconfoundedness is more plausible. Although this creates no conceptual problems, it makes the practical problem of drawing credible causal inferences more challenging.

In Part V of the book we discuss methods for assessing the plausibility of the unconfoundedness assumption, and sensitivity analyses for assessing the implications of violations of it. In Part VI we analyze a number of assignment mechanisms where the assignment itself is regular, but the treatment received is not equal to the treatment assigned for all units. Thus, although the treatment assigned *is* unconfounded, the treatment received *is not* unconfounded, because the probability of receiving the active versus control treatment depends on potential outcomes. Such settings have arisen in the econometric literature to account for settings where individuals choose the treatment regime, at least partly based on expected benefits associated with the two treatment regimes. Although, as a general matter, such optimizing behavior is not inconsistent with regular assignment mechanisms, in some cases it suggests assignment mechanisms associated with so-called *instrumental variable* methods.

The rest of this chapter is organized as follows. In the next section we introduce additional notation. In Section 3.3 we define the assignment mechanism, unit-level assignment probabilities, and the propensity score. In Section 3.4 we formally introduce the three general restrictions we consider imposing on assignment mechanisms. We then use those restrictions to define classical randomized experiments in Section 3.6. In Section 3.7 we define regular assignment mechanisms as a special class of observational studies. The next section, Section 3.8, discusses some non-regular assignment mechanisms. Section 3.9 concludes.

## 3.2 NOTATION

Continuing the potential outcomes discussion in Chapter 1, let us consider a population of $N$ units, indexed by $i = 1, \ldots, N$. The $i^{\text{th}}$ unit in this population is characterized by a $K$-component row vector of covariates (also referred to as pre-treatment variables or attributes), $X_i$, with $\mathbf{X}$ the $N \times K$ matrix of covariates in the population with $i^{\text{th}}$ row equal to $X_i$. In social science applications, the elements of $X_i$ may include an individual's age, education, socio-economic status, labor market history, pre-test scores, sex, and marital status. In biomedical applications, the covariates may also include measures of an individual's medical history, and family background information. Most important is that covariates are known *a priori* to be unaffected by the assignment of treatment.

For each unit there is also a pair of potential outcomes, $Y_i(0)$ and $Y_i(1)$, denoting its outcome values under the two values of the treatment: $Y_i(0)$ denotes the outcome under the control treatment, and $Y_i(1)$ denotes the outcome under the active treatment. Notice that when using this notation, we tacitly accept the Stable Unit Treatment Value Assumption (SUTVA) that treatment assignments for other units do not affect the outcomes for unit $i$, and that each treatment defines a unique outcome for each unit. The latter requirement implies that there is only a single version of the active and control treatments for each unit. Let $\mathbf{Y}(0)$ and $\mathbf{Y}(1)$ denote the $N$-component vectors (or the $N$-vectors for short) of the potential outcomes. More generally, the potential outcomes could themselves be multi-component row vectors, in which case $\mathbf{Y}(0)$ and $\mathbf{Y}(1)$ would be matrices with the $i^{\text{th}}$ rows equal to $Y_i(0)$ and $Y_i(1)$, respectively. Here, we largely focus on the situation where the potential outcomes are scalars, although in most cases extensions to vector-valued outcomes are conceptually straightforward.

Next, the $N$-component columns vector of treatment assignments is denoted by $\mathbf{W}$, with $i^{\text{th}}$ element $W_i \in \{0, 1\}$, with $W_i = 0$ if unit $i$ received the control treatment, and $W_i = 1$ if this unit received the active treatment. Let $N_{\text{c}} = \sum_{i=1}^{N} (1 - W_i)$ and $N_{\text{t}} = \sum_{i=1}^{N} W_i$ be the number of units assigned to the control and active treatment respectively, with $N_c + N_t = N$.

In Chapter 1 we defined the realized and possibly observed outcomes

$$Y_i^{\text{obs}} = Y_i(W_i) = \begin{cases} Y_i(0) & \text{if } W_i = 0, \\ Y_i(1) & \text{if } W_i = 1, \end{cases} \tag{3.1}$$

and the missing outcomes:

$$Y_i^{\text{mis}} = Y_i(1 - W_i) = \begin{cases} Y_i(1) & \text{if } W_i = 0, \\ Y_i(0) & \text{if } W_i = 1. \end{cases} \tag{3.2}$$

$\mathbf{Y}^{\text{obs}}$ and $\mathbf{Y}^{\text{mis}}$ are the corresponding $N$-vectors (or matrices in the case with multiple outcomes). We can invert these relations and characterize the potential outcomes in terms of the observed and missing outcomes:

$$Y_i(0) = \begin{cases} Y_i^{\text{mis}} & \text{if } W_i = 1, \\ Y_i^{\text{obs}} & \text{if } W_i = 0, \end{cases} \quad \text{and} \quad Y_i(1) = \begin{cases} Y_i^{\text{mis}} & \text{if } W_i = 0, \\ Y_i^{\text{obs}} & \text{if } W_i = 1. \end{cases} \tag{3.3}$$

This characterization illustrates that the causal inference problem is fundamentally a missing data problem: if we impute the missing outcomes, we "know" all the potential outcomes and thus the value of any causal estimand in the population of $N$ units.

## 3.3 ASSIGNMENT PROBABILITIES

To introduce the taxonomy of assignment mechanisms used in this text requires some formal mathematical terms. First, we define the assignment mechanism to be the function that assigns probabilities to all $2^N$ possible values for the $N$-vector of assignments $\mathbf{W}$ (each unit can be assigned to treatment or control), given the $N$-vectors of potential outcomes $\mathbf{Y}(0)$ and $\mathbf{Y}(1)$, and given the $N \times K$ matrix of covariates $\mathbf{X}$:

**Definition 3.1 (Assignment Mechanism)**
*Given a population of $N$ units, the assignment mechanism is a row-exchangeable function* $\Pr(\mathbf{W}|\mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1))$*, taking on values in* $[0, 1]$*, satisfying*

$$\sum_{\mathbf{W}\in\{0,1\}^N} \Pr(\mathbf{W}|\mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1)) = 1,$$

*for all* $\mathbf{X}$*,* $\mathbf{Y}(0)$*, and* $\mathbf{Y}(1)$*.*

The set $\mathbb{W} = \{0, 1\}^N$ is the set of all $N$-vectors with all elements equal to 0 or 1. By the assumption that the function $\Pr(\cdot)$ is row exchangeable, we mean that the order in which we list the $N$ units within the vectors or matrices is irrelevant. Note that this probability $\Pr(\mathbf{W}|\mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1))$ is *not* the probability of a particular unit receiving the treatment. Instead, it is the probability that a particular value for the full assignment – first two units treated, third a control, fourth treated, etc. – will occur. The definition requires that the probabilities across the full set of $2^N$ possible assignment vectors $\mathbf{W}$ sum to one. Note also that some assignment vectors $\mathbf{W}$ may have zero probability. For example, if we were to design a study to evaluate a new drug, it is likely that we would want to rule out the possibility that all subjects received the control drug. We could do so by assigning zero probability to the vector of assignments $\mathbf{W}$ with $W_i = 0$ for all $i$, or perhaps even assign zero probability to all vectors of assignments other than those with $\sum_{i=1}^N W_i = N/2$, for even values of the population size $N$.

In addition to the probability of joint assignment for the entire population, we are often interested in the probability of an individual unit being assigned to the active treatment:

**Definition 3.2 (Unit Assignment Probability)**
*The unit-level assignment probability for unit $i$ is*

$$p_i(\mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1)) = \sum_{\mathbf{W}:W_i=1} \Pr(\mathbf{W}|\mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1)).$$

Here we sum the probabilities across all possible assignment vectors $\mathbf{W}$ for which $W_i = 1$. Out of the set of $2^N$ different assignment vectors, half (that is $2^{N-1}$) have the property that $W_i = 1$. The probability that unit $i$ is assigned to the control treatment is $1 - p_i(\mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1))$. Note that according to this definition, the probability that unit $i$

receives the treatment can be a function of its own covariates $X_i$ and potential outcomes $Y_i(0)$ and $Y_i(1)$, and it generally is also a function of the covariate values, and potential outcomes, and treatment assignments of the other units in the population.

We are also often interested in the average of the unit-level assignment probabilities for subpopulations with a common value of the covariates, for example, $X_i = x$. We label this function the *propensity score* at $x$. In the finite population case the definition of the propensity score follows.

**Definition 3.3 (Finite Population Propensity Score)**
*The propensity score at x is the average unit assignment probability for units with* $X_i = x$,

$$e(x) = \frac{1}{N(x)} \sum_{i:X_i=x} p_i(\mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1))$$

*where* $N(x) = \#\{i = 1, \ldots, N | X_i = x\}$ *is the number of units with* $X_i = x$. *For values x with* $N(x) = 0$, *the propensity score is defined to be zero.*

To illustrate these definitions more concretely, consider four examples, the first three with with two units, and the last one with three units.

EXAMPLE 1 Suppose we have two units. Then there are four ($2^2$) possible values for $\mathbf{W}$,

$$\mathbf{W} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$$

We conduct a randomized experiment where all treatment assignments have equal probability. Then the assignment mechanism is equal to

$$\Pr(\mathbf{W}|\mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1)) = 1/4, \quad \text{for } \mathbf{W} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}. \tag{3.4}$$

In this case the unit assignment probability $p_i(\mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1))$ is equal to $1/2$ for both units $i = 1, 2$. In a randomized experiment with no covariates, the propensity score is equal to the unit assignment probabilities, here all equal to $1/2$. □

EXAMPLE 2 We conduct a randomized experiment with two units where only those assignments with exactly one treated and one control unit are allowed. Then the assignment mechanism is

$$\Pr(\mathbf{W}|\mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1)) = \begin{cases} 1/2 & \text{if } \mathbf{W} \in \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}, \\ 0 & \text{if } \mathbf{W} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}. \end{cases} \tag{3.5}$$

This does not change the unit-level assignment probabilities, which remains equal to $1/2$ for both units, and so does the propensity score. □

EXAMPLE 3 A third, more complicated, assignment mechanism with two units is the following. The unit with more to gain from the active treatment (using a coin toss in the

case of a tie) is assigned to the treatment group, and the other to the control group. This leads to

$$\Pr(\mathbf{W}|\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1)) = \begin{cases} 1 & \text{if } Y_2(1)-Y_2(0) > Y_1(1)-Y_1(0) \text{ and } \mathbf{W} = \begin{pmatrix}0\\1\end{pmatrix}, \\ 1 & \text{if } Y_2(1)-Y_2(0) < Y_1(1)-Y_1(0) \text{ and } \mathbf{W} = \begin{pmatrix}1\\0\end{pmatrix}, \\ 1/2 & \text{if } Y_2(1)-Y_2(0) = Y_1(1)-Y_1(0) \text{ and } \mathbf{W} \in \left\{\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\right\}, \\ 0 & \text{if } \mathbf{W} \in \left\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}\right\}, \\ 0 & \text{if } Y_2(1)-Y_2(0) < Y_1(1)-Y_1(0) \text{ and } \mathbf{W} = \begin{pmatrix}0\\1\end{pmatrix}, \\ 0 & \text{if } Y_2(1)-Y_2(0) > Y_1(1)-Y_1(0) \text{ and } \mathbf{W} = \begin{pmatrix}1\\0\end{pmatrix}. \end{cases} \tag{3.6}$$

In this example the unit-level treatment probabilities $p_i(\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1))$ are equal to zero, one, or a half, depending whether the gain for unit $i$ is smaller or larger than for the other unit, or equal. Given that there are no covariates, the propensity score remains a constant, equal to $1/2$ in this case. This is a type of assignment mechanism that we often rule out when attempting to infer causal effects. □

EXAMPLE 4 A sequential randomized experiment allows for dependence of the assignment mechanism on the potential outcomes, thus violating some of the assumptions we consider later. In this example, there are three units, and thus eight possible values for $\mathbf{W}$:

$$\mathbf{W} \in \left\{ \begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix} \right\}.$$

Suppose there is a covariate $X_i$ measuring the order in which the units entered the experiment, $X_i \in \{1,2,3\}$. Without loss of generality, let us assume that $X_i = i$. For the first unit, with $X_i = 1$, a fair coin toss determines the treatment. The second unit, with $X_i = 2$, is assigned to the alternative treatment. Let the observed outcomes for the first and second unit be $Y_1^{\text{obs}}$ and $Y_2^{\text{obs}}$. The third unit, with $X_i = 3$, is assigned to the active or control treatment that appears better, based on a comparison of observed outcomes by treatment status for the first two units. If both treatments appear equally beneficial, the third unit is assigned to the active treatment. For example, if $W_1 = 0$, $W_2 = 1$, and $Y_1^{\text{obs}} > Y_2^{\text{obs}}$, then the third unit gets assigned to the control group; if $W_1 = 0$, $W_2 = 1$, and $Y_1^{\text{obs}} \leq Y_2^{\text{obs}}$, the third units gets assigned to the treatment group; and similarly given the alternative

assignments for the first two units. Formally:

$$\Pr(\mathbf{W}|\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1),\mathbf{X}) = \left\{ \begin{array}{ll} 1/2 & \text{if } Y_1(0) > Y_2(1), \text{ and } \mathbf{W} = \begin{pmatrix}0\\1\\0\end{pmatrix}, \\ 1/2 & \text{if } Y_1(1) \geq Y_2(0), \text{ and } \mathbf{W} = \begin{pmatrix}1\\0\\1\end{pmatrix}, \\ 1/2 & \text{if } Y_1(0) \leq Y_2(1), \text{ and } \mathbf{W} = \begin{pmatrix}0\\1\\1\end{pmatrix}, \\ 1/2 & \text{if } Y_1(1) < Y_2(0), \text{ and } \mathbf{W} = \begin{pmatrix}1\\0\\0\end{pmatrix}. \end{array} \right\} \tag{3.7}$$

In this case the unit assignment probability is equal to $1/2$ for the first two units,

$$p_2(\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1)) = p_2(\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1)) = 1/2,$$

and, for unit 3, equal to

$$p_3(\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1)) = \begin{cases} 0 & \text{if } Y_1(0) > Y_2(1) \text{ and } Y_1(1) < Y_2(0), \\ 1 & \text{if } Y_1(1) \geq Y_2(0) \text{ and } Y_1(0) \leq Y_2(1), \\ 1/2 & \text{otherwise.} \end{cases}$$

Because the covariates identify the unit, the propensity score is equal to the unit assignment probabilities. Thus, for $x = 1$ and $x = 2$ the propensity score is equal to $1/2$. If $x = 3$, the propensity score is equal to $p_3(\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1))$. □

## 3.4 RESTRICTIONS ON THE ASSIGNMENT MECHANISM

Before classifying the various types of assignment mechanisms that are the basis of the organization of this text, we present three general properties that assignment mechanisms may satisfy. These properties restrict the dependence of the unit-level assignment probabilities on values of covariates and potential outcomes for other units, or restrict the range of values of the unit-level assignment probabilities, or restrict the dependence of the assignment mechanism on potential outcomes.

The first property we consider is *individualistic assignment*, which limits the dependence of the treatment assignment for unit $i$ on the outcomes and assignments for other units:

**Definition 3.4 (Individualistic Assignment)**
*An assignment mechanism* $\Pr(\mathbf{W}|\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1))$ *is individualistic if, for some function* $q(\cdot)\in[0,1]$,

$$p_i(\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1)) = q(X_i,Y_i(0),Y_i(1)), \quad \text{for all } i=1,\ldots,N,$$

*and*

$$\Pr(\mathbf{W}|\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1)) = c\cdot\prod_{i=1}^{N} q(X_i,Y_i(0),Y_i(1))^{W_i}\,(1-q(X_i,Y_i(0),Y_i(1)))^{1-W_i},$$

*for* $(\mathbf{W},\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1))\in\mathbb{A}$, *for some set* $\mathbb{A}$, *and zero elsewhere (c is the constant that ensures that the probabilities sum to unity).*

Individualistic assignment is violated in sequential experiments such as Example 4. Given individualistic assignment, the propensity score simplifies to:

$$e(x) = \frac{1}{N_x}\sum_{i:X_i=x} q(X_i,Y_i(0),Y_i(1)).$$

Next, we define *probabilistic assignment*, which requires every unit to have positive probability of being assigned to treatment level 0 and to treatment level 1:

**Definition 3.5 (Probabilistic Assignment)**
*An assignment mechanism* $\Pr(\mathbf{W}|\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1))$ *is probabilistic if the probability of assignment to treatment for unit i is strictly between zero and one:*

$$0 < p_i(\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1)) < 1, \quad \text{for each possible } \mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1),$$

*for all* $i=1,\ldots,N$.

Note that this merely requires that every unit has the possibility of being assigned to the active treatment and the possibility of being assigned to the control treatment.

The third property is a restriction on the dependence of the assignment mechanism on potential outcomes:

**Definition 3.6 (Unconfounded Assignment)**
*An assignment mechanism is unconfounded if it does not depend on the potential outcomes:*

$$\Pr(\mathbf{W}|\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1)) = \Pr(\mathbf{W}|\mathbf{X},\mathbf{Y}'(0),\mathbf{Y}'(1)),$$

*for all* $\mathbf{W}$, $\mathbf{X}$, $\mathbf{Y}(0)$, $\mathbf{Y}(1)$, $\mathbf{Y}'(0)$, *and* $\mathbf{Y}'(1)$.

If an assignment mechanism is unconfounded, we can drop the two potential outcomes as arguments and write the assignment mechanism as $\Pr(\mathbf{W}|\mathbf{X})$. The assignment mechanisms in Examples 1 and 2 are, but those in in Examples 3 and 4 are not, unconfounded.

The combination of unconfoundedness and individualistic assignment plays a very important role. In that case,

$$\Pr(\mathbf{W}|\mathbf{X},\mathbf{Y}(0),\mathbf{Y}(1)) = c\cdot\prod_{i=1}^{N} q(X_i)^{W_i}\cdot(1-e(X_i))^{1-W_i}. \tag{3.8}$$

so that

$$e(x) = q(x),$$

so that the assignment mechanism is the product of the propensity scores. Note that, under unconfoundedness, the propensity score is no longer just the average assignment probability for units with covariate value $X_i = x$; it can also be interpreted as the unit-level assignment probability for such units.

Given individualistic assignment, the combination of probabilistic and unconfounded assignment is referred to as *strongly ignorable treatment assignment* (Rosenbaum and Rubin, 1983a). More generally, *ignorable treatment assignment* refers to the weaker restriction where the assignment mechanism can be written in terms of $\mathbf{W}$, $\mathbf{X}$, and $\mathbf{Y}^{\text{obs}}$ only, without dependence on $\mathbf{Y}^{\text{mis}}$ (Rubin, 1978).

## 3.5 ASSIGNMENT MECHANISMS AND SUPER-POPULATIONS

In part of this text we view our sample of size $N$ as a random sample from an infinite super-population. In that case we employ slightly different formulations of the restrictions on the assignment mechanism. Sampling from the super-population generates a joint sampling distribution on the quadruple of unit-level variables $(Y_i(0), Y_i(1), W_i, X_i)$, $i = 1,\ldots,N$. More explicitly, we assume the $(Y_i(0), Y_i(1), W_i, X_i)$ are independently and identically distributed draws from a distribution indexed by a global parameter. We write this in factored form as

$$f_{W|Y(0),Y(1),X}(W_i|Y_i(0),Y_i(1),X_i,\phi)\cdot f_{Y(0),Y(1)|X}(Y_i(0),Y_i(1)|X_i,\theta)\cdot f_X(X_i|\psi), \tag{3.9}$$

where the parameters are in their respective parameter spaces, and the full parameter vector is $(\phi,\theta,\psi)$, where each of these components is generally a function of the global parameter.

In this setting we define the propensity score as

**Definition 3.7 (Super-Population Propensity Score)**
*The propensity score at x is the population average unit assignment probability for units with $X_i = x$,*

$$e(x) = \mathbb{E}_{\text{SP}}\left[f_{W|Y(0),Y(1),X}(1|Y_i(0),Y_i(1),X_i,\phi)f_{Y(0),Y(1)|X}(Y_i(0),Y_i(1)|X_i,\theta)\Big|\, X_i = x\right],$$

*for all x in the support of $X_i$; $e(x)$ is here a function of $\phi$, a dependence that we usually suppress notationally.*

The "SP" subscript on the expectations operator indicates that the expectation is taken over the distribution generated by random sampling. In this case the expectation is taken over the potential outcomes $(Y_i(0), Y_i(1))$. By iterated expectations the propensity score in the super-population setting is also equal to $\Pr(W_i = 1|X_i = x, \phi, \theta)$ where the probability is taken both over the assignment mechanism and over the random sampling.

Note that with our definition of super-populations the assignment mechanism is automatically individualistic (of course, given $(\phi, \theta)$).

**Definition 3.8 (Super-Population Probabilistic Assignment)**
*An assignment mechanism is super-population probabilistic if the probability of assignment to treatment for unit i is strictly between zero and one:*

$$0 < f_{W|Y(0),Y(1),X}(1|Y_i(0), Y_i(1), X_i, \phi) < 1, \quad \text{for each possible } X_i, Y_i(0), Y_i(1).$$

**Definition 3.9 (Super-Population Unconfounded Assignment)**
*An assignment mechanism is super-population unconfounded if it does not depend on the potential outcomes:*

$$f_{W|Y(0),Y(1),X}(w|y_0, y_1, x, \phi) = f_{W|Y(0),Y(1),X}(w|y_0', y_1', x, \phi),$$

*for all* $y_0$, $y_1$, $x$, $y_0'$, $y_1'$, $\phi$, *and for* $w = 0, 1$.

## 3.6 RANDOMIZED EXPERIMENTS

Part II of this text deals with the inferentially most straightforward class of assignment mechanisms, randomized assignment. Randomized experimental designs have traditionally been viewed as the most credible basis for causal inference, as reflected in the typical reliance of the U.S. Food and Drug Administration on such experiments in its approval process for pharmaceutical treatments.

**Definition 3.10 (Randomized Experiment)**
*A randomized experiment is an assignment mechanism that*

*(i) is probabilistic, and*
*(ii) has a known functional form that is controlled by the researcher.*

In Part II of this text we will be concerned with a special case – what we call classical randomized experiments:

**Definition 3.11 (Classical Randomized Experiment)**
*A classical randomized experiment is a randomized experiment with an assignment mechanism that is*

*(i) individualistic, and*
*(ii) unconfounded.*

The definition of a classical randomized experiment rules out sequential experiments as in Example 4. In sequential experiments, the assignment for units assigned in a later

stage of the experiment generally depends on observed outcomes for units assigned earlier in the experiment.

A leading case of a classical randomized experiment is a *completely randomized experiment*, where, *a priori*, the number of treated units, $N_t$, is fixed (and thus the number of control units $N_c = N - N_t$ is fixed as well). In such a design, $N_t$ units are randomly selected, from a population of $N$ units, to receive the active treatment, with the remaining $N_c$ assigned to the control group. In this case, each unit has unit assignment probability $q = N_t/N$, and the assignment mechanism equals

$$\Pr(\mathbf{W}|\mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1)) = \begin{cases} 1 \Big/ \binom{N}{N_t} & \text{if } \sum_{i=1}^{N} W_i = N_t, \\ 0 & \text{otherwise,} \end{cases}$$

where the number of distinct values of the assignment vector with $N_t$ units out of $N$ assigned to the active treatment is

$$\binom{N}{N_t} = \frac{N!}{N_t! \cdot (N - N_t)!}, \qquad \text{with } J! = J(J-1)\ldots 1.$$

Other prominent examples of classical randomized experiments include stratified randomized experiments and paired randomized experiments, discussed in Chapters 9 and 10.

## 3.7 OBSERVATIONAL STUDIES: REGULAR ASSIGNMENT MECHANISMS

In Parts III and IV of this text, we discuss cases where the exact assignment probabilities may be unknown to the researcher, but the researcher still has substantial information concerning the assignment mechanism. For instance, a leading case is where the researcher knows the set of variables that enters into the assignment mechanism but does not know the functional form of the dependence. Such information will generally come from subject-matter knowledge. For example, medical decisions in some situations are made solely using patients' medical records, but precisely how may be unknown. In general we refer to designs with unknown assignment mechanisms as *observational studies*:

**Definition 3.12 (Observational Study)**
*An assignment mechanism corresponds to an observational study if the functional form of the assignment mechanism is unknown.*

The special case of an assignment mechanism that is the focus of Part III of the book is a *regular assignment mechanism*:

**Definition 3.13 (Regular Assignment Mechanism)**
*An assignment mechanism is regular if*

*(i) the assignment mechanism is individualistic,*
*(ii) the assignment mechanism is probabilistic, and*
*(iii) the assignment mechanism is unconfounded.*

If, in addition, the functional form of a regular assignment mechanism is known, the assignment mechanism corresponds to a classical randomized experiment. If the functional form is not known, the assignment mechanism corresponds to an observational study with a regular assignment mechanism.

In Part III of this book we focus on the design stage of studies where the assumption of a regular assignment mechanism is viewed as plausible. In this design stage we focus on the data on treatment assignment and pre-treatment variables only, without seeing the outcome data. The concern at this stage is balance in the covariate distributions between treated and control groups. In completely and stratified randomized experiments, balance is guaranteed by design, but in observational studies this needs to be done by special analyses. We assess balance, and in cases where initially there is insufficient balance, we develop methods for improving balance.

In Part IV we discuss methods of analysis for causal inference with regular assignment mechanisms in some detail. Even if in many cases it may appear too strong to assume that an assignment mechanism is regular, we will argue that, in practice, it is a very important starting point for many studies. There are two main reasons for this. The first is that in many well-designed observational studies, researchers have attempted to record all the relevant covariates, that is, all the variables that may be associated with both outcomes and assignment to treatment. If they have been successful in this endeavor, or at least approximately so, a regular assignment mechanism may be a reasonable approximation to the true assignment mechanism. The second reason is that specific alternatives to regular assignment mechanisms are typically even less credible. Under a regular assignment mechanism, it will be sufficient to adjust appropriately for differences between treated and control units' covariate values to draw valid causal inferences. Any alternative method involves causal interpretations of comparisons of units with different treatments who also are observed to *differ* systematically in their values for covariates. It is relatively uncommon to find a convincing argument in support of such alternatives, although there are some notable exceptions, such as instrumental variables analyses discussed in Part VI of the book. More details of these arguments are presented in Chapter 12.

## 3.8 OBSERVATIONAL STUDIES: IRREGULAR ASSIGNMENT MECHANISMS

In Part VI of this book, we discuss another class of assignment mechanisms. We focus on settings where assignment to treatment may differ for some units from the receipt of treatment. We assume that assignment to treatment itself is unconfounded, but allow receipt of treatment to be confounded. This class of assignment mechanisms includes noncompliance in randomized experiments and sometimes utilizes *instrumental variables* analyses. Often in these designs, the receipt of treatment can be viewed as "latently regular" – that is, it would be regular given some additional covariates that are not fully observed. To conduct inference in such settings, it is often useful to invoke additional conditions, in particular *exclusion restrictions*, which rule out the presence of particular causal effects.

The remainder of this text provides more detailed discussion of methods of causal inference given each of these types of assignment mechanisms. In the next part of the book, Chapters 4–11, we start with classical randomized experiments.

## 3.9 CONCLUSION

This chapter presented the taxonomy of assignment mechanisms that serves as the organizing principle for this text. Using three restrictions on the assignment mechanism – individualistic assignment, probabilistic assignment, and unconfoundedness – we define regular assignment mechanisms and the special case of classical randomized experiments. In the next part of the book, we study classical randomized experiments, followed in Parts III and IV by the study of observational studies with regular assignment mechanisms. In Parts V and VI of the text we analyze some additional assignment mechanisms where receipt of treatment is confounded.

## NOTES

Of the restrictions on assignment mechanisms we discuss in the current chapter, the first one, individualistic assignment, is often made implicitly, but the term is new. The notion of probabilistic assignment is often stated formally, although it is rarely given a formal label. The term unconfoundedness was coined by Rubin (1990a). It is sometimes referred to as the *conditional independence assumption* (Lechner, 2001; Angrist and Pischke, 2009). In the econometrics literature it is also closely related to the notion of *exogeneity* (Manski, Sandefur, McLanahan, and Powers, 1992), although formal definitions of exogeneity do not coincide with unconfoundedness (see Imbens, 2004, for some discussion). The combination of probabilistic assignment and unconfoundedness is referred to as *Strong Ignorability* or *Strongly Ignorable Treatment Assignment* by Rosenbaum and Rubin (1984). There is a close link between some of the assumptions used in the context of causal inference and the terminology in missing data problems. In the missing data literature, strong ignorability is closely linked with *Missing at Random* missingness mechanisms (Rubin, 1976c; Little and Rubin, 2002; Frumento, Mealli, Pacini, and Rubin, 2012).

*Instrumental variables* methods originate in the econometrics literature and go back to the 1920s and 1940s (P. Wright, 1928; S. Wright 1921, 1923; Tinbergen, 1928; Haavelmo, 1943). For a historical perspective, see Stock and Trebbi (2003) and Imbens (2014). For modern approaches see Imbens and Angrist (1994), and Angrist, Imbens, and Rubin (1996). For textbook discussions, see Wooldridge (2010) and Angrist and Pischke (2008).

Some methods for assignment mechanisms not covered in this edition of the book include *Principal Stratification*, *Regression Discontinuity Designs, Difference In Differences* methods, and case-control designs. The notion of *Principal Stratification* generalizes the binary-treatment version of instrumental variables. It was introduced by Frangakis and Rubin (2002). *Regression discontinuity designs* originate in the

psychology literature (Thistlewaite and Campbell, 1960). See for a historical overview Cook (2008), and for recent surveys Imbens and Lemieux (2008) and Lee and Lemieux (2010). *Difference in Differences* (DID) methods are another set of methods intended for irregular designs. DID methods are widely used in the econometric literature. See Angrist and Pischke (2008) for a general discussion and references. Case-control designs, more accurately called case-noncase designs, are commonly used in epidemiology, especially when looking for exposures that lead to rare diseases (i.e., the cases).