Abbreviations and Notation

Frequently used abbreviations and notation are listed here.

A	Treatment or exposure
A-IPCW	Augmented inverse probability of censoring-weighted/weighting
A-IPW	Augmented inverse probability weighted/weighting
C	Censoring
CRT	Community randomized trial
i.i.d.	Independent and identically distributed
IPCW	Inverse probability of censoring-weighted/weighting
IPW	Inverse probability of weighted/weighting
LTMLE	Longitudinal targeted maximum likelihood estimation/estimator
MLE	Maximum likelihood substitution estimator of the g-formula *Not to be confused with nonsubstitution estimators using maximum likelihood estimation. MLE is also known as g-computation*
MSE	Mean squared error
O	Observed ordered data structure
P	Possible data-generating distribution
p	Possible density of data-generating distribution P_0
P_0	True data-generating distribution; $O \sim P_0$
p_0	True density of data-generating distribution P_0
P_n	Empirical probability distribution; places probability $1/n$ on each observed $O_i, i \ldots, n$
RCT	Randomized controlled trial
SCM	Structural causal model
SE	Standard error
SL	Super learner
TMLE	Targeted maximum likelihood estimation/estimator
W	Vector of covariates
Y	Outcome
Y_1, Y_0	Counterfactual outcomes with binary A

Uppercase letters represent random variables and lowercase letters are a specific value for that variable. If O is discrete, $p_0(o) = P_0(O = o)$ is the probability that O equals the value o, and if O is continuous, $p_0(o)$ denotes the Lebesgue density of P_0 at o. For simplicity and the sake of presentation, we will often treat O as discrete so that we can refer to $P_0(O = o)$ as a probability. For a simple example, suppose our data structure is $O = (W, A, Y) \sim P_0$ and O is discrete. For each possible value (w, a, y), $p_0(w, a, y)$ denotes the probability that (W, A, Y) equals (w, a, y).

$\mathcal{M}$	Statistical model; the set of possible probability distributions for P_0
$P_0 \in \mathcal{M}$	P_0 is known to be an element of the statistical model $\mathcal{M}$

In this text we often use the term *semiparametric* to include both nonparametric and semiparametric. When semiparametric excludes nonparametric, and we make additional assumptions, this will be explicit. A *statistical model* can be augmented with additional nonstatistical (e.g., causal) assumptions providing enriched interpretation, often represented as $\{P_\theta : \theta \in \Theta\}$ for some parameterization $\theta \rightarrow P_\theta$. We refer to this as a *model* (e.g., the probability distribution of the observed data $O = (W, A, Y = Y_A)$ could be represented as a missing data structure on counterfactual outcomes Y_0, Y_1 with missingness variable A, so that the probability distribution of O is indexed by the probability distribution of (W, Y_0, Y_1) and the conditional distribution of treatment A, given (W, Y_0, Y_1)).

$X = (X_j : j)$	Set of endogenous variables, $j = 1, \ldots, J$
$U = (U_{X_j} : j)$	Set of exogenous variables
$P_{U,X}$	Probability distribution for (U, X)
$p_{U,X}$	Density for (U, X)
$Pa(X_j)$	Parents of X_j among X
f_{X_j}	A function of $Pa(X_j)$ and an endogenous U_{X_j} for X_j
$f = (f_{X_j} : j)$	Collection of f_{X_j} functions that define the SCM
$\mathcal{M}^F$	Collection of possible $P_{U,X}$ as described by the SCM; includes non-testable assumptions based on real knowledge; $\mathcal{M}$ augmented with additional nonstatistical assumptions known to hold
$\mathcal{M}^{F*}$	Model under possible additional causal assumptions required for identifiability of target parameter
$P \rightarrow \Psi(P)$	Target parameter as mapping from a P to its value
$\Psi(P_0)$	True target parameter
$\hat{\Psi}(P_n)$	Estimator as a mapping from empirical distribution P_n to its value
$\psi_0 = \Psi(P_0)$	True target parameter value
ψ_n	Estimate of ψ_0

Consider $O = (L_0, A_0, \ldots, L_K, A_K, L_{K+1}) \sim P_0$.

L_k	Possibly time-varying covariate at $t = k$; alternate notation $L(k)$
A_k	Time-varying intervention node at $t = k$ that can include both treatment and censoring
$Pa(L_k)$	$=(\bar{A}_{k-1}, \bar{L}_{k-1})$
$Pa(A_k)$	$=(\bar{A}_{k-1}, \bar{L}_k)$

P_{0,L_k}	True conditional probability distribution of L_k, given $Pa(L_k)$, under P_0
P_{L_k}	Conditional probability distribution of L_k, given $Pa(L_k)$, under P
P_{n,L_k}	Estimate of conditional probability distribution P_{0,L_k} of L_k
P_{0,A_k}	True conditional probability distribution of A_k, given $Pa(A_k)$, under P_0
P_{A_k}	Conditional probability distribution of A_k, given $Pa(A_k)$, under P
P_{n,A_k}	Conditional probability distribution of A_k, given $Pa(A_k)$, under estimator P_n of P_0
ϵ	Fluctuation parameter
ϵ_n	Estimate of ϵ
$\{P_\epsilon : \epsilon\} \subset \mathcal{M}$	Submodel through P
H^*	Clever covariate
H^*_n	Estimate of H^*
$D(\psi)(O)$	Estimating function of the data structure O and parameters; shorthand $D(\psi)$
$D^*_0(O)$	Efficient influence curve; canonical gradient; alternate notation $D^*(P_0)(O)$, $D^*(P_0)$ or $D^*(O)$
$IC_0(O)$	Influence curve of an estimator at P_0, representing a function of O
$IC_n(O)$	Estimate of influence curve

We focus on the general data structure $O = (L_0, A_0, \ldots, L_K, A_K, L_{K+1}) \sim P_0$ in many chapters, introduced on the previous page. In this setting, the following specific notation definitions apply:

L_0	Baseline covariates
$\bar{L}$	$= (L_0, \ldots, L_{K+1})$
$\bar{A}$	$= (A_0, \ldots, A_K)$
L_d	Counterfactual outcome for regime d
d_0	Optimal regime depending on P_0
Q_{0,L_k}	True conditional probability distribution of L_k
Q_{L_k}	Possible conditional probability distribution of L_k
Q_{n,L_k}	Estimate of Q_{0,L_k}
Q	$= (Q_{L_0}, \ldots, Q_{L_{K+1}})$
$L(O, Q)$	Example of a loss function where it is a function of O and Q; alternate notation $L(Q)(O)$ or $L(Q)$
$\{Q_\epsilon : \epsilon\}$	Submodel through Q
$\bar{Q}_{L_k}$	Conditional mean of the probability distribution of L_k
G_{0,A_k}	True conditional probability distribution of A_k
G_{A_k}	Possible conditional probability distribution of A_k
G_{n,A_k}	Estimate of G_{0,A_k}
$\bar{Q}^0_n$	Initial estimate of $\bar{Q}_0$
$\bar{Q}^1_n$	First updated estimate of $\bar{Q}_0$
$\bar{Q}^k_n$	kth updated estimate of $\bar{Q}_0$
$\bar{Q}^*_n$	Targeted estimate of $\bar{Q}_0$ in TMLE procedure

$\Psi(Q_0)$	Alternate notation for true target parameter when it only depends on P_0 through Q_0
$\Psi(Q_n^*)$	Targeted estimator of parameter
Pf	Expectation of $f(O)$ under P, e.g., $P_0L(Q) = \int L(Q)(o)dP_0(o)$