CAUSALITY CORRECTIONS IMPLEMENTED IN 2nd PRINTING

Updated 9/26/00

page v *insert* "TO RUTH" centered in middle of page

page xv *insert* in second paragraph "David Galles" after "Dechter"

page 2 *replace* "2000" with "2004" in 2nd paragraph, line 8 of 1.1.2

page 3 *insert* ", $\Rightarrow$" after "($\neg$)" in footnote 1
replace "and *not*" with "*not*, and *implies*," in footnote 1

page 19 *append* (continue italics) to end of Theorem 1.2.7, "(*We exclude* X_i *when speaking of its "nondescendants*".)"

page 30 *replace* in line 7 from top "mutually" with "jointly"
insert "parental" before "Markov" in first line of 2nd paragraph after Theorem 1.4.1
append to end of footnote 16 "but I am not aware of any nonparametric version."

page 52 *insert* "stable" after "IC*, that takes a" in 2nd paragraph after Theorem 2.6.2,
replace "sampled" with "stable" in Input line of IC* Algorithm.
append "(with respect to some latent structure)." to same line

page 68 *replace* in line 1 after Eq. (3.2), "mutually" with "jointly"

page 72 *replace* "(1990, 1999)" with "(1990, 2001)" on line 6

page 89 *replace* in paragraph starting "Indeed, if condition...". Should be "conditions require" not "condition require"

page 126 *insert* "and Robins (1997)" after "Pearl and Robins (1995)", line 2.

page 130 *replace* the "P" with "E" in the formula (second line of section 4.5.4.) Should read "$E(Y|\hat{x}, \widehat{pa}_{Y\setminus X})$"
replace "we can compute the difference" with "we should replace the controlled difference" in last line of page

page 131 *replace* from top of page through the end of section 4.5.4 with the following:

$$P(\text{admission}|\widehat{\text{male}}, \widehat{\text{dept}}) - P(\text{admission}|\widehat{\text{female}}, \widehat{\text{dept}})$$

with some average of this difference over all departments. This average should measure the increase in admission rate in a hypothetical experiment in which we instruct all female candidates to retain their department preferences but change their gender identification (on the application form) from female to male.

In general, the average direct effect is defined as the expected change in Y induced by changing X from x to x' while keeping the other parents of Y constant at whatever value they obtain under $do(x)$. This hypothetical change is what law makers instruct us to consider in race or sex discrimination cases: "The central question in any employment-discrimination case is whether the employer would have taken the same action had the employee been of a different race (age, sex, religion, national origin etc.) and everything else had been the same." (In Carson versus Bethlehem Steel Corp., 70 FEP Cases 921, 7th Cir. (1996)).

The formal expression for this hypothetical change involves probabilities of (nested) counterfactuals (see Section 7.1 for semantics and computation) that cannot be written in terms of the $do(x)$ operator.[9] Therefore, the average direct effect cannot in general be identified, even from data obtained under randomized control of all variables. However, if certain assumptions of "no confounding" are deemed valid,[10] then the average direct effect can be reduced to

$$\Delta_{x,x'}(Y) = \sum_{pa_{Y\setminus X}} [E(Y|\widehat{x'}, \widehat{pa}_{Y\setminus X}) - E(Y|\hat{x}, \widehat{pa}_{Y\setminus X})] P(pa_{Y\setminus X}|\hat{x}), \quad \textbf{(4.11)}$$

and the techniques developed in Section 4.4 for identifying control-specific plans, $P(y|\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)$, become applicable.

[9] Using the counterfactual notation of Section 7.1, the general expression for the average direct effect is

$$\Delta_{x,x'}(Y) = E(Y_{x'Z_x}) - E(Y_x)$$

where $Z = pa_{Y\setminus X}$. The subscript $x'Z_x$ represents the operation of setting X to x' and, simultaneously, setting Z to whatever value it would have obtained under the setting $X = x$. This general expression reduces to (4.11) if $Z_x \perp\!\!\!\perp Y_{x'z}$ holds for all z. Likewise, the average *indirect* effect is defined as $E(Y_{xZ_{x'}}) - E(Y_x)$.

[10] See details in Technical Report R-273 posted on www.cs.ucla.edu/∼judea.

page 164 *replace* ... "$do(x, y, w)$" with "$do(x, z, w)$" in line 8 after Definition 5.4.3.

page 165 *replace* last two sentences of section 5.4.2 with:
The expressions corresponding to these policies are $P(y|do(x),\ do(z))$ and $P(y|do(x))$, and this pair of distributions fully represents the policy implications of indirect effects. Similar conclusions have been expressed by Robins and Greenland (1992). (But see Chapter 4, footnote 9, page 131.)

page 177 *delete* "tormented" in paragraph 3, line 2

page 184 *append* to end of Definition 6.2.1 (continue italics - except 'unbiased'): "*If* (6.10) *holds, we say that* $P(y|x)$ *is* unbiased."

page 236 *replace* "& $(X \rightarrow Y|$" with "& $(X \not\rightarrow Y|$" in first formula of Theorem 7.3.8

page 240 *replace* the last sentence in the last paragraph of section 7.4.1 with:
However, this effectiveness is partly acquired by limiting the counterfactual antecedent to conjunction of elementary propositions. Disjunctive hypotheticals, such as "if Bizet and Verdi were compatriots," usually lead to multiple solutions and hence to nonunique probability assignments.

page 246 *insert* in footnote 26 after "(see Section 5.4.3)." "Epidemiologists refer to (7.46) as "no-confounding" (see (6.10))."

page 255 *replace* in the 2nd line "pregnant" with "nonpregnant"

page 259 *insert* close parentheses after "(Sections 3.2 and 7.1", line 2 of Preface

page 284 *replace* "Michie in press" with "Michie 1999" in the last line of paragraph 4

page 329 *replace* "(1999)" with "(2000)" in last line of page

page 332 *replace* in paragraph starting "Even an erratic and ...". Change "role" to "roll"

page 354 line 2 from bottom, *replace* "mediated by tar deposits" with "unmediated by tar deposits"

page 361 *update* Dawid 1997 citation. *replace* "To appear ..." with "Also [with discussion] in *Journal of the American Statistical Association* 95:407–48, 2000."

page 363 *append* to Halpern (1998) citation, "Also, *Journal of Artificial Intelligence Research* 12:317–37, 2000."
update Halpern and Pearl (1999) citation. *replace* "(1999)" with "(2000)", *replace* "Actual causality." with "Causes and explanations.", and *append* "www.cs.ucla.edu/∼judea/"

page 364 *update* Hoover 1999 citation. *replace* "(1999)" with "(2001)"

page 365 *update* Kuroki citation. *append* “29: 105–17.” after “*Journal of the Japan Statistical Society*”.

page 366 *update* Michie citation. *replace* “(in press)” with “(1999)” and *insert* “pp. 60-86” before “vol. 15”

page 368 *update* Pearl 1999 citation. *remove* “To appear in” and replace “121” with “121:93–149.”

page 369 *insert* in Robins 1997 citation, “M. Berkane (Ed.),” before “*Latent Variable Modeling*...”

page 370 *add* to Shipley 1997 citation, “Also in *Structural Equation Modelling*, 7:206–18, 2000.”

page 381 *insert* “27–8”, after “(examples) price and demand” and before “215-17”
replace “245” with “245–7”, at end of “(exogeneity) controversies regarding...245”
combine “explanation” and “explanations” to read “explanation, 25, 58, 221–3, 285, 308–9”

page 382 *insert* “131” after “indirect effects,” and before “165”

ADDENDUM TO CORRECTIONS IMPLEMENTED IN 2nd PRINTING

Updated 12/14/00

page 28 *replace* "income (Z)" with "income (I)" in the caption of Figure 1.5

page 48 *replace* in line before Definition 2.4.1, "when one of the coins becomes slightly biased." with "when the coins become slightly biased."

page 51 *append* to line 7, Rule R_4 to read:
Orient a–b into $a \to b$ whenever there are two chains a–$c \to d$ and $c \to d \to b$ such that c and b are nonadjacent and a and d are adjacent.

page 231 *replace* Definition 7.3.4 and 2 lines following to read:
Definition 7.3.4 (Recursiveness)
Let X and Y be singleton variables in a model, and let $X \to Y$ stand for the inequality $Y_{xw}(u) \neq Y_w(u)$ for some values of x, w, and u. A model M is recursive if, for any sequence $X_1, X_2, \ldots, X_k$, we have

$$X_1 \to X_2,\ X_2 \to X_3, \ldots, X_{k-1} \to X_k \Rightarrow X_k \not\to X_1 \qquad \textbf{(7.24)}$$

Clearly, any model M for which the causal diagram $G(M)$ is acyclic must be recursive.

page 382 *change* "Markov (assumptions underlying, 30)" to "Markov (assumption, 30, 69)"

page 382 *append* "69" after "causal, 30" in "Markov condition (causal, 30)"

page 384 *add* as subentry after "structural model, 27, 44, 202" "Markovian, 30, 69".

PRINTING CORRECTIONS TO BE IMPLEMENTED BY CAMBRIDGE

Updated 3/13/07

page 52 *replace* in line 17-18 ”protection” with ”projection”

page 73 *replace* equation between (3.11) and (3.12)

$$\begin{aligned} P(pa_i|do(x_i')) &= P(pa_i); \\ \frac{P(s_i, pa_i|do(x_i'))}{P(s_i', pa_i|do(x_i'))} &= \frac{P(s_i, pa_i)}{P(s_i', pa_i)}. \end{aligned}$$

should be:

$$\begin{aligned} P(pa_i|do(x_i')) &= P(pa_i); \\ \frac{P(s_i, pa_i, x_i'|do(x_i'))}{P(s_i', pa_i, x_i'|do(x_i'))} &= \frac{P(s_i, pa_i, x_i')}{P(s_i', pa_i, x_i')}. \end{aligned}$$

page 82 *replace* at end of paragraph 2: “since there is no back-door path from X to Z, we simply have” with “since there is no unblocked back-door path from X to Z in Figure 3.5, we simply have”

page 82 *replace* in Definition 3.3.3 (Front-Door): “(ii) there is no back-door path from X to Z; and” to “there is no unblocked back-door path from X to Z; and”

page 103 *replace* last paragraph on page 103 (including footnote 15) with:
To place this result in the context of our analysis in this chapter, we note that one class of semi-Markovian models satisfying assumption (3.62) corresponds to graphs in which all arrowheads pointing to X_k originate from observed variables. Indeed, in such models, the parents $PA_k = L_k, X_{k-1}$ of variable X_k satisfy the back-door condition of Definition 3.3.1,

$$(X_k \perp\!\!\!\perp Y|PA_k)_{G_{\underline{X}_k}},$$

which implies (3.62).[15] This class of models falls under Theorem 3.2.5, which states that all causal effects in this class are identifiable and are given by the truncated factorization formula of (3.14); the formula coincides with (3.63) after marginalizing over the uncontrolled covariates.

[15]Condition (3.62) is too restrictive and lacks intuitive basis; a graphical, more general condition leading to (3.63) is formulated in (4.5), Theorem 4.4.1, read: $P(y|g = x)$ is identifiable and is given by (3.63) if every action-avoiding back-door path from X_k to Y is blocked by some subset l_k of non-descendants of X_k. (by “action-avoiding” we mean a path containing no arrow entering an X variable later than X_k) see ⟨http://bayes.cs.ucla.edu/BOOK-2K/yudkowsky.html⟩.

page 174 *replace* in 4th line from end of page: “is not a statement about C being a positive causal factor for E, properly written” with “is not a statement about C having a positive influence on E, properly written”

page 195 *replace* ”Figure 6.1” with ”Figure 6.3” in 5th line before end of 3rd paragraph

ADDITIONAL PRINTING CORRECTIONS TO BE IMPLEMENTED BY CAMBRIDGE

Updated 1/10/08

page 24 Chapter 1, line 7

replace

(iii) $P_x(v_i|pa_i) = P(v_i|pa_i)$ for all $V_i \notin X$ whenever pa_i is consistent with $X = x$.

with:

(iii) $P_x(v_i|pa_i) = P(v_i|pa_i)$ for all $V_i \notin X$ whenever pa_i is consistent with $X = x$, i.e., each $P(v_i|pa_i)$ remains invariant to interventions not involving V_i.

page 39 Footnote 26, last sentence

replace

Causal assumptions of the type developed in Chapter 2 (see Definitions 2.4.1 and 2.7.4) must be invoked before applying such sequences in policy-related tasks.

With:

Such sequences are statistical in nature and, unless causal assumptions of the type developed in Chapter 2 (see Definitions 2.4.1 and 2.7.4) are invoked, they cannot be applied to policy-evaluation tasks.

replace first Remark

Remark: The distinction between probabilistic and statistical parameters is devised to exclude the construction of joint distributions that invoke hypothetical variables (e.g., counterfactual or theological). Such constructions, if permitted, would qualify any quantity as statistical and would obscure the distinction between causal and noncausal assumptions.

With:

Remark: The exclusion of unmeasured variables from the definition of statistical parameters is devised to rule out the construction of joint distributions that invoke counterfactual or metaphysical variables. Such constructions, if permitted, would qualify any quantity as statistical and would thus obscure the important distinction between quantities that can be estimated from statistical data alone, and those that require additional assumptions beyond the data.

insert (after "falsifiable")

Often, though not always, causal assumptions can be falsified from experimental studies, in which case we say that they are "experimentally testable." For example, the assumption that X has no effect on $E(Y)$ in model 2 of Figure 1.6 is empirically testable, but the assumption that X may cure a given subject in the population is not.

replace second Remark

Remark: The distinction between causal and statistical parameters is crisp and fundamental. Causal parameters can be discerned from joint distributions only when special assumptions are made, and such assumptions must have causal components to them. The formulation and simplification of these assumptions

will occupy a major part of this book.
With:
Remark: The distinction between causal and statistical parameters is crisp and fundamental – the two do not mix. Causal parameters cannot be discerned from statistical parameters unless causal assumptions are invoked. The formulation and simplification of these assumptions will occupy a major part of this book.

page 40 Addition to end of Chapter 1

Two Mental Barriers to Causal Analysis

The sharp distinction between statistical and causal concepts can be translated into a useful principle: one cannot substantiate causal claims from associations alone — behind every causal conclusion there must lie some causal assumption that is not testable in observational studies. Such assumptions are usually provided by humans, resting on expert judgment. The ramification of this is that psychological considerations of how humans organize causal knowledge are no longer a luxury; they become imperative and demand careful attention to the way assumptions are phrased.

Another ramification of this causal-statistical distinction is that any mathematical approach to causal analysis must acquire new notation for expressing causal assumptions and causal claims. The vocabulary of probability calculus, with its powerful operators of expectation, conditionalization and marginalization, is insufficient for expressing causal information. To illustrate, the syntax of probability calculus does not permit us to express the simple fact that "symptoms do not cause diseases," let alone draw mathematical conclusions from such facts. All we can say is that two events are dependent—meaning that if we find one, we can expect to encounter the other, but we cannot distinguish statistical dependence, quantified by the conditional probability $P(disease|symptom)$ from causal dependence, for which we have no expression in standard probability calculus.

The preceding two requirements: (1) to commence causal analysis with untested, judgmentally based assumptions, and (2) to extend the syntax of probability calculus, constitute the two main obstacles to the acceptance of causal analysis among professionals with traditional training in statistics. [Pearl, 2003]. [Cox and Wermuth, 1996] This book helps overcome the two barriers through an effective and friendly notational systems based on symbiosis of graphical and algebraic approaches.

page 280 Section 8.5.4, 2nd paragraph, 6th line
replace
the following four compliance-response populations: $\{(r_x = 0, r_y = 1),\ (r_x = 0, r_y = 2),\ (r_x = 1, r_y = 1),\ (r_x = 1, r_y = 2)\}$. Joe would have improved had he taken cholestyramine if his response behavior is either helped ($r_y = 1$) or always-recover ($r_y = 3$).
With

the four compliance-response populations:

$$\{(r_x = 0, r_y = 0),\ (r_x = 0, r_y = 1),\ (r_x = 1, r_y = 0),\ (r_x = 1, r_y = 1)\}.$$

Joe would have improved had he taken cholestyramine if and only if his response behavior is helped ($r_y = 1$).

page 298 after (9.38), within the next 5 lines; replace $y_{x'}$ with $y'_{x'}$ (4 places) **and** correction in last line of (9.39).
replace
In order to compute PNS, we must evaluate the probability of the joint event $y_{x'} \wedge y_x$. Given that these two events are jointly true only when $U = \mathrm{true}$, we have

$$\begin{aligned} \mathrm{PNS} &= P(y_x, y_{x'}) \\ &= P(y_x, y_{x'}|u)P(u) + P(y_x, y_{x'}|u')P(u') \\ &= \frac{1}{2}(1+0) = \frac{1}{2}. \end{aligned} \tag{39}$$

With
In order to compute PNS, we must evaluate the probability of the joint event $y'_{x'} \wedge y_x$. Given that these two events are jointly true only when $U = \mathrm{true}$, we have

$$\begin{aligned} \mathrm{PNS} &= P(y_x, y'_{x'}) \\ &= P(y_x, y'_{x'}|u)P(u) + P(y_x, y'_{x'}|u')P(u') \\ &= \frac{1}{2}(0+1) = \frac{1}{2}. \end{aligned} \tag{39}$$