Typos in “Elements of Causal Inference: Foundations and Learning Algorithms”

Below, you find a collection of all typos and mistakes from our book that we know of. The part in blue is correct (hopefully!). We thank all readers who kindly sent us comments to any of these typos.

København, July 23, 2018
Jonas Peters, Dominik Janzing, and Bernhard Schölkopf
The correct definition for the inverse cdf is
\[ F_{Y|x}^{-1}(n_Y) := \inf \{ y \in \mathbb{R} : F_{Y|x}(y) \geq n_Y \}. \]

In 4.2.2., the first inequality on page 67 should read
\[ H(X) \geq H(Y). \]

In Problem 4.16, part (a) should read: Prove that \( f(x) = \mathbb{E}[Y \mid X = x] - \mu_{N_Y} \).

In Definition 6.1, “neither \( i_k \) nor any of its descendants is in \( S \)” should be understood as
\[ (\{i_k\} \cup \text{DE}_{i_k}) \cap S = \emptyset, \]
which is important for the case \( \text{DE}_{i_k} = \emptyset \).

We write that an SCM defines a unique distribution over \((X_1, \ldots, X_d)\) such that \( X_j = f_j(\text{PA}_j, N_j) \) in distribution. This is, admittedly, a confusing formulation. More formally, we can define an SCM as a pair of structural equations and a \( d \)-dimensional noise distribution. We then call \((X_1, \ldots, X_d, N_1, \ldots, N_d)\) a solution to the SCM if the noise variables have the correct distribution and the structural equations are satisfied almost surely. One can then show that all solutions induce the same distribution over the \( X \)'s, which we can call the entailed distribution. See [Bongers et al., 2016] for more details.

References