Two Factorizations and a Density Ratio

On 'Parameterizing and Simulating from Causal Models' by Evans and Didelez

---

F. Richard Guo

RSS Discussion Meeting, 3 Oct, 2023

Statistical Laboratory, Cambridge

Two factorizations

Suppose we partition a random vector as (X, Y, Z). Any distribution over (X, Y, Z) can be factorized in two ways:

$$\left(P_{ZX}, P_{Y|Z,X}\right) \underset{C}{\overset{C^{-1}}{\leftrightarrows}} P_{ZXY} \underset{A}{\overset{A^{-1}}{\rightleftarrows}} \left(P_{ZX},\, P_{Y|X},\, \phi_{ZY|X}\right).$$

Two factorizations

Suppose we partition a random vector as (X, Y, Z). Any distribution over (X, Y, Z) can be factorized in two ways:

$$\left(P_{ZX}, P_{Y|Z,X}\right) \underset{C}{\overset{C^{-1}}{\leftrightarrows}} P_{ZXY} \underset{A}{\overset{A^{-1}}{\rightleftarrows}} \left(P_{ZX},\, P_{Y|X},\, \phi_{ZY|X}\right).$$

- C is the 'usual' factorization; A is the focus of Evans and Didelez.

Two factorizations

Suppose we partition a random vector as (X, Y, Z). Any distribution over (X, Y, Z) can be factorized in two ways:

$$\left(P_{ZX}, P_{Y|Z,X}\right) \underset{C}{\overset{C^{-1}}{\leftrightarrows}} P_{ZXY} \underset{A}{\overset{A^{-1}}{\rightleftarrows}} \left(P_{ZX}, P_{Y|X}, \phi_{ZY|X}\right).$$

- C is the 'usual' factorization; A is the focus of Evans and Didelez.

☞ Concretely, $\phi_{ZY|X}$ is a **conditional copula density**.

$$\forall x, \quad (z, y) \mapsto \phi_{ZY|X}(z, y \mid x) \text{ is a density over } [0, 1]^2 \text{ with uniform margins.}$$

Two factorizations

Suppose we partition a random vector as (X, Y, Z). Any distribution over (X, Y, Z) can be factorized in two ways:

$$\left(P_{ZX}, P_{Y|Z,X}\right) \underset{C}{\overset{C^{-1}}{\leftrightarrows}} P_{ZXY} \underset{A}{\overset{A^{-1}}{\rightleftarrows}} \left(P_{ZX},\, P_{Y|X},\, \phi_{ZY|X}\right).$$

- C is the 'usual' factorization; A is the focus of Evans and Didelez.

☞ Concretely, $\phi_{ZY|X}$ is a **conditional copula density**.

$$\forall x, \quad (z, y) \mapsto \phi_{ZY|X}(z, y \mid x) \text{ is a density over } [0,1]^2 \text{ with uniform margins.}$$

☞ Factorization A:

$$\begin{aligned} p_{ZXY}(z, x, y) &= p(x)\, p(z \mid x)\, p(y \mid x)\, \phi_{ZY|X}(F(z \mid x), F(y \mid x) \mid x) \\ &= \boxed{p_{ZX}(z, x)\, p_{Y|X}(y \mid x)\, \phi_{ZY|X}(F(z \mid x), F(y \mid x) \mid x).} \end{aligned}$$

Density ratio

Density ratio

But we are not directly interested in P_{ZXY}, but, rather a (causal) distribution P^*_{ZXY} related to P_{ZXY}.

Density ratio

But we are not directly interested in P_{ZXY}, but, rather a (causal) distribution P^*_{ZXY} related to P_{ZXY}.

- **Density ratio**

$$\frac{p^*(Z, X, Y)}{p(Z, X, Y)} = r(z, x; p)$$

such that

(1) $r(z, x; p) > 0$ strictly positive almost everywhere; ☞ This ensures $p/p^* = r^{-1}$.
(2) r does not depend on Y;
(3) r can be identified from P_{ZXY}.

Density ratio

But we are not directly interested in P_{ZXY}, but, rather a (causal) distribution P^*_{ZXY} related to P_{ZXY}.

- **Density ratio**

$$\frac{p^*(Z,X,Y)}{p(Z,X,Y)} = r(z,x;p)$$

such that

(1) $r(z,x;p) > 0$ strictly positive almost everywhere; ☞ This ensures $p/p^* = r^{-1}$.
(2) r does not depend on Y;
(3) r can be identified from P_{ZXY}.

☞ By integrating out y on both sides of $p^*(z,x,y) = r(z,x;p)p(z,x,y)$,

$$\boxed{r(z,x;p) = \frac{p^*(z,x)}{p(z,x)}, \quad p^*(y \mid z,x) = p(y \mid z,x).}$$

Choosing parametrization

$$
\begin{array}{ccccc}
\left(P_{ZX}, P_{Y|Z,X}\right) & \underset{C}{\overset{C^{-1}}{\leftrightarrows}} & P_{ZXY} & \underset{A}{\overset{A^{-1}}{\rightleftarrows}} & \left(P_{ZX},\, P_{Y|X},\, \phi_{ZY|X}\right) \\
 & & r^{-1} \Big\uparrow \Big\downarrow r & & \\
\left(P^*_{ZX}, P^*_{Y|Z,X}\right) & \underset{C}{\overset{C^{-1}}{\leftrightarrows}} & P^*_{ZXY} & \underset{A}{\overset{A^{-1}}{\rightleftarrows}} & \left(P^*_{ZX},\, P^*_{Y|X},\, \phi^*_{ZY|X}\right)
\end{array}
$$

☞ We parametrize p_{ZXY} (and hence P^*_{ZXY}) with **components** in this diagram.

Frugal parametrization

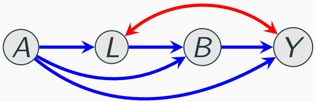
$$
\begin{array}{ccccc}
\left(P_{ZX}, P_{Y|Z,X}\right) & \underset{C}{\overset{C^{-1}}{\leftrightarrows}} & P_{ZXY} & \underset{A}{\overset{A^{-1}}{\rightleftarrows}} & \left(\boxed{P_{ZX}}, P_{Y|X}, \phi_{ZY|X}\right) \\
 & & r^{-1} \uparrow \downarrow r & & \\
\left(P^*_{ZX}, P^*_{Y|Z,X}\right) & \underset{C}{\overset{C^{-1}}{\leftrightarrows}} & P^*_{ZXY} & \underset{A}{\overset{A^{-1}}{\rightleftarrows}} & \left(P^*_{ZX}, \boxed{P^*_{Y|X}}, \boxed{\phi^*_{ZY|X}}\right)
\end{array}
$$

$$
\begin{array}{ccccc}
\left(P_{ZX}, P_{Y|Z,X}\right) & \underset{C}{\overset{C^{-1}}{\leftrightarrows}} & P_{ZXY} & \underset{A}{\overset{A^{-1}}{\rightleftarrows}} & \left(\boxed{P_{ZX}}, P_{Y|X}, \phi_{ZY|X}\right) \\
 & & r^{-1} \Big\uparrow \Big\downarrow r & & \\
\left(P^*_{ZX}, P^*_{Y|Z,X}\right) & \underset{C}{\overset{C^{-1}}{\leftrightarrows}} & P^*_{ZXY} & \underset{A}{\overset{A^{-1}}{\rightleftarrows}} & \left(P^*_{ZX}, \boxed{P^*_{Y|X}}, \boxed{\phi^*_{ZY|X}}\right)
\end{array}
$$

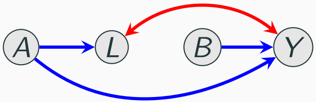
☞ Using $r(z,x;p) = p^*(z,x)/p(z,x)$, we get

$$
\begin{aligned}
p(z,x,y) = \frac{p^*(z,x,y)}{r(z,x)} &= \frac{A\left(p^*(z,x), p^*(y \mid x), \phi^*(z,y \mid x)\right)}{r(z,x;p)} \\
&= \frac{A\left(p(z,x)r(z,x;p), p^*(y \mid x), \phi^*(z,y \mid x)\right)}{r(z,x;p)}.
\end{aligned}
$$

Example: Verma graph

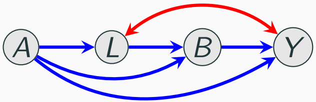


$p(A, B, L, Y)$

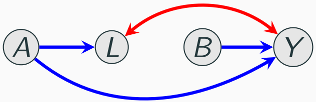


$p^*(A, B, L, Y)$

Example: Verma graph



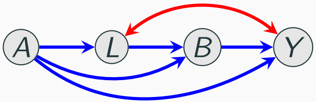
$p(A, B, L, Y)$



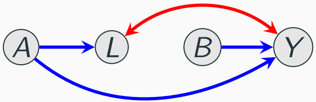
$p^*(A, B, L, Y)$

We choose $X = (A, B)$, $Z = L$ and $Y = Y$.

Example: Verma graph



$p(A, B, L, Y)$



$p^*(A, B, L, Y)$

We choose $X = (A, B)$, $Z = L$ and $Y = Y$.

With density ratio $r(a, l, b; p) = p(b)/p(b \mid a, l)$, we can parametrize

$$P(A, L, B, Y) \underset{A}{\overset{A^{-1}}{\rightleftarrows}} \left(\boxed{P_{ALB}}, P_{Y|AB}, \phi_{YL|AB}\right)$$

$$P^*(A, L, B, Y) \underset{A}{\overset{A^{-1}}{\rightleftarrows}} \left(P^*_{ALB}, \boxed{P^*_{Y|AB}}, \boxed{\phi^*_{YL|AB}}\right).$$

Example: Structural nested model

With $r = 1$, this is a more direct parametrization

$$
\begin{array}{ccccc}
\left(\boxed{P_{ZX}}, P_{Y|Z,X}\right) & \overset{C^{-1}}{\underset{C}{\leftrightarrows}} & P_{ZXY} & \overset{A^{-1}}{\underset{A}{\rightleftarrows}} & \left(P_{ZX}, P_{Y|X}, \phi_{ZY|X}\right) \\
 & & r^{-1} \uparrow\downarrow r & & \\
\left(P^*_{ZX}, \boxed{P^*_{Y|Z,X}}\right) & \overset{C^{-1}}{\underset{C}{\leftrightarrows}} & P^*_{ZXY} & \overset{A^{-1}}{\underset{A}{\rightleftarrows}} & \left(P^*_{ZX}, P^*_{Y|X}, \phi^*_{ZY|X}\right)
\end{array}
$$

Example: "Partially" marginal model

For baseline covariates $Z = (Z_0, C)$, suppose that we want to study how C **modified** the effect of X on Y. ▶ i.e., interested in modeling $p(Y \mid C, \text{do}(X))$.

Example: "Partially" marginal model

For baseline covariates $Z = (Z_0, C)$, suppose that we want to study how C **modified** the effect of X on Y. ▶ i.e., interested in modeling $p(Y \mid C, \mathrm{do}(X))$.

☞ Yet, meanwhile, we want to control for the **confounding** induced by C (in addition to Z_0).

Example: "Partially" marginal model

For baseline covariates $Z = (Z_0, C)$, suppose that we want to study how C **modified** the effect of X on Y. ▶ i.e., interested in modeling $p(Y \mid C, \mathrm{do}(X))$.

☞ Yet, meanwhile, we want to control for the **confounding** induced by C (in addition to Z_0).

☞ Choosing density ratio $r(z_0, c, x, y; p) = p(x)/p(x \mid z_0, c)$, then we can look at

$$p^*(y \mid c, x) = p(y \mid c, \mathrm{do}(x))$$

for effect modification.

Example: "Partially" marginal model

For baseline covariates $Z = (Z_0, C)$, suppose that we want to study how C **modified** the effect of X on Y. ▶ i.e., interested in modeling $p(Y \mid C, \mathrm{do}(X))$.

☞ Yet, meanwhile, we want to control for the **confounding** induced by C (in addition to Z_0).

☞ Choosing density ratio $r(z_0, c, x, y; p) = p(x)/p(x \mid z_0, c)$, then we can look at

$$p^*(y \mid c, x) = p(y \mid c, \mathrm{do}(x))$$

for effect modification.

$$\begin{array}{ccccc}
\left(P_{Z_0CX}, P_{Y|Z_0,C,X}\right) & \underset{C}{\overset{C^{-1}}{\rightleftarrows}} & P_{Z_0CXY} & \underset{A}{\overset{A^{-1}}{\rightleftarrows}} & \left(\boxed{P_{Z_0CX}}, P_{Y|C,X}, \phi_{YZ_0|C,X}\right) \\
 & & r^{-1}\Big\uparrow\Big\downarrow r & & \\
\left(P^*_{Z_0CX}, P^*_{Y|Z_0,C,X}\right) & \underset{C}{\overset{C^{-1}}{\rightleftarrows}} & P^*_{Z_0CXY} & \underset{A}{\overset{A^{-1}}{\rightleftarrows}} & \left(P^*_{Z_0CX}, \boxed{P^*_{Y|C,X}}, \boxed{\phi^*_{YZ_0|C,X}}\right)
\end{array}$$

Cognate?

- The cognate definition essentially requires choosing

$$r(z, x; p) = \frac{w(z \mid x)}{p(z \mid x)}$$

for some kernel $w(z \mid x)$.

☞ But formulating it in terms of density ratio is perhaps more **general**?

Congrats & Thanks