Causal Inference and Fairness in Insurance Pricing

Insurance Data Science Conference 2023

Research by

Olivier Côté* **(Université Laval)**
Marie-Pier Côté **(Université Laval)**
Arthur Charpentier **(Université du Québec à Montréal)**

What is fairness?

Unfair discrimination for ratemaking

“**A rate is** reasonable and **not** excessive, inadequate, or **unfairly discriminatory if it is an actuarially sound estimate of the expected value of all future costs** associated with an individual risk transfer.”

– Casualty Actuarial Society (1988)

What is fairness?

Unfair discrimination for ratemaking

“**A rate is** reasonable and **not** excessive, inadequate, or **unfairly discriminatory if it is an actuarially sound estimate of the expected value of all future costs** associated with an individual risk transfer.”

– Casualty Actuarial Society (1988)

The debate regarding the formal definition of fairness never really settled (Frezal and Barry, 2020)

What is fairness?

Unfair discrimination for ratemaking

"**A rate is** reasonable and **not** excessive, inadequate, or **unfairly discriminatory if it is an actuarially sound estimate of the expected value of all future costs** associated with an individual risk transfer."

– Casualty Actuarial Society (1988)

The debate regarding the formal definition of fairness never really settled (Frezal and Barry, 2020) , and the rise of machine learning and big data increases the potential for harm due to unfairness (Embrechts and Wüthrich, 2022).

Why should we care about fairness?

1. Regulatory framework
   - "Financial institutions must ensure that the use of artificial intelligence systems **does not undermine fairness**." (Recommendation 9, Autorité des Marchés Financiers, 2021)

Why should we care about fairness?

1. Regulatory framework
   - “Financial institutions must ensure that the use of artificial intelligence systems **does not undermine fairness**.” (Recommendation 9, Autorité des Marchés Financiers, 2021)
2. Responsibility of the modeller
   - “The statistician **cannot evade the responsibility** for understanding the process he applies or recommends.” (Fisher, 1956)

Why should we care about fairness?

1. Regulatory framework
   - "Financial institutions must ensure that the use of artificial intelligence systems **does not undermine fairness**." (Recommendation 9, Autorité des Marchés Financiers, 2021)
2. Responsibility of the modeller
   - "The statistician **cannot evade the responsibility** for understanding the process he applies or recommends." (Fisher, 1956)
   - "A model's **blind spots** reflect the judgments and priorities of its creators." (O'Neil, 2016)

Why should we care about fairness?

1. Regulatory framework
   - "Financial institutions must ensure that the use of artificial intelligence systems **does not undermine fairness**." (Recommendation 9, Autorité des Marchés Financiers, 2021)
2. Responsibility of the modeller
   - "The statistician **cannot evade the responsibility** for understanding the process he applies or recommends." (Fisher, 1956)
   - "A model's **blind spots** reflect the judgments and priorities of its creators." (O'Neil, 2016)
3. Maintaining public trust

Causal inference explained

Why causal inference and fairness?

The overall goals of causal inference and fairness in insurance are congruent:

Causal inference:
Estimate a **target effect** while
avoiding **undesired** bias
from **irrelevant** confounders.

Why causal inference and fairness?

The overall goals of causal inference and fairness in insurance are congruent:

Causal inference:
Estimate a **target effect** while avoiding **undesired** bias from **irrelevant** confounders.

Fairness in insurance:
Estimate a **premium** while avoiding **unfair** bias from **prohibited** confounders.

How can causal inference be useful?

- Causal inference offers a toolbox to deal with various types of biases

How can causal inference be useful?

- Causal inference offers a toolbox to deal with various types of biases **(Unfair biases)**.

How can causal inference be useful?

- Causal inference offers a toolbox to deal with various types of biases **(Unfair biases)**.
- Causal inference offers a deeper understanding of relationships in a dataset

How can causal inference be useful?

- Causal inference offers a toolbox to deal with various types of biases **(Unfair biases)**.
- Causal inference offers a deeper understanding of relationships in a dataset **(True risk factors, Araiza Iturria et al., 2022)**.

How can causal inference be useful?

- Causal inference offers a toolbox to deal with various types of biases **(Unfair biases)**.
- Causal inference offers a deeper understanding of relationships in a dataset **(True risk factors, Araiza Iturria et al., 2022)**.
- Causal inference allows to answer questions that goes beyond the observable dataset

How can causal inference be useful?

- Causal inference offers a toolbox to deal with various types of biases **(Unfair biases)**.
- Causal inference offers a deeper understanding of relationships in a dataset **(True risk factors, Araiza Iturria et al., 2022)**.
- Causal inference allows to answer questions that goes beyond the observable dataset **("What would have been the premium, had there been no disparity?")**.

Reviewing the discrimination-free formula

Important definitions

Table 1: Key Definitions

Notation	Description	Example
X	Allowed variables	Vehicle model
D	Prohibited variables	Ethnic origin
Y	Response variable	Claim amount

Important definitions

Table 1: Key Definitions

Notation	Description	Example
X	Allowed variables	Vehicle model
D	Prohibited variables	Ethnic origin
Y	Response variable	Claim amount

$$
\begin{tikzcd}
D \arrow[d] \arrow[dr] & \\
X \arrow[r] & Y
\end{tikzcd}
$$

Figure 1: Typical directed acyclic graph (DAG) of fairness in insurance

Discrimination-free formula

Recall the discrimination-free premium proposed by Lindholm et al. (2022) with the real world measure $\mathbb{P}$:

$$\mu^{DF}(X) = \int_d \mathbb{E}\left[Y|X, D\right] \mathrm{d}\mathbb{P}\left(D = d\right).$$

Counterfactual identification

Counterfactual for a discrimination-free premium

What is a counterfactual (or potential outcome)?

Counterfactual for a discrimination-free premium

What is a counterfactual (or potential outcome)?

- Y is the response variable.

Counterfactual for a discrimination-free premium

What is a counterfactual (or potential outcome)?

- Y is the response variable.
- Y^x is the potential outcome, had X been equal to x.

Counterfactual for a discrimination-free premium

What is a counterfactual (or potential outcome)?

- Y is the response variable.
- Y^x is the potential outcome, had X been equal to x.

There is a relation between the **discrimination-free formula** and the expected counterfactual (Araiza Iturria et al., 2022) :

$$\mathbb{E}\left[Y^x\right]$$

Counterfactual for a discrimination-free premium

What is a counterfactual (or potential outcome)?

- Y is the response variable.
- Y^x is the potential outcome, had X been equal to x.

There is a relation between the **discrimination-free formula** and the expected counterfactual (Araiza Iturria et al., 2022) :

$$\mathbb{E}\left[Y^x\right] \overset{\text{Causal assumpt.}}{=} \underbrace{\mathbb{E}_D\left\{\mathbb{E}\left[Y|X,D\right]\right\} = \mu^{DF}(X)}_{\text{Discrimination-free formula}}$$

Counterfactual for a discrimination-free premium : remark

What are those causal assumptions?

Counterfactual for a discrimination-free premium : remark

What are those causal assumptions?

1. The prohibited attribute has to be a **confounder**.

$$\begin{array}{ccc} D & & \\ \downarrow & \searrow & \\ X & \longrightarrow & Y \end{array}$$

Figure 2: Desired DAG to satisfy assumptions

Counterfactual for a discrimination-free premium : remark

What are those causal assumptions?

1. The prohibited attribute has to be a **confounder**.
2. Positivity, exchangeability and consistency must be valid.

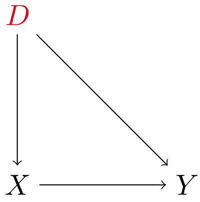


Figure 2: Desired DAG to satisfy assumptions

Inverse probability weighting

Towards a propensity score weighting

We start again with the discrimination-free formula, focusing on the weighting term:

$$\mu^{DF}(X) = \int_d \mathbb{E}\left[Y|X, D = d\right] \mathrm{d}\mathbb{P}(D = d).$$

Towards a propensity score weighting

We start again with the discrimination-free formula, focusing on the weighting term:

$$\mu^{DF}(X) = \int_d \mathbb{E}\left[Y|X, D=d\right] \mathrm{d}\mathbb{P}(D=d).$$

Multiplying by a real fraction equal to 1, we obtain :

$$\mu^{IPW}(X) = \int_d \mathbb{E}\left[Y|X, D=d\right] \frac{\mathrm{d}\mathbb{P}\left(D=d|X\right)}{\mathrm{d}\mathbb{P}\left(D=d|X\right)} \mathrm{d}\mathbb{P}(D=d)$$

Towards a propensity score weighting

We start again with the discrimination-free formula, focusing on the weighting term:

$$\mu^{DF}(X) = \int_d \mathbb{E}\left[Y|X, D = d\right] \mathrm{d}\mathbb{P}(D = d).$$

Multiplying by a real fraction equal to 1, we obtain :

$$\begin{aligned}
\mu^{IPW}(X) &= \int_d \mathbb{E}\left[Y|X, D = d\right] \frac{\mathrm{d}\mathbb{P}\left(D = d|X\right)}{\mathrm{d}\mathbb{P}\left(D = d|X\right)} \mathrm{d}\mathbb{P}(D = d) \\
&= \int_d \mathbb{E}\left[Y \frac{\mathrm{d}\mathbb{P}(D = d)}{\mathrm{d}\mathbb{P}\left(D = d|X\right)} \middle| X, D = d\right] \mathrm{d}\mathbb{P}\left(D = d|X\right).
\end{aligned}$$

Weights for fairness

We get a weight w that introduces fairness in our discrimination-free formula :

$$\mu^{IPW}(X) = \mathbb{E}_D \left\{ \mathbb{E}\left[Y \frac{\mathrm{d}\mathbb{P}(D)}{\mathrm{d}\mathbb{P}\left(D|X\right)} \middle| X, D \right] \middle| X \right\}$$

Weights for fairness

We get a weight w that introduces fairness in our discrimination-free formula :

$$\mu^{IPW}(X) = \mathbb{E}_D \left\{ \mathbb{E}\left[Y \frac{\mathrm{d}\mathbb{P}(D)}{\mathrm{d}\mathbb{P}\left(D|X\right)} \middle| X, D \right] \middle| X \right\}$$

A variety of weights (see, e.g. Li and Li, 2019) and a variety of estimators (See Fong et al., 2018, for non-parametric estimator) exist.

Intuitive properties of weights

The weights have intuitive properties.

Intuitive properties of weights

The weights have intuitive properties.

They do not distort variables on the aggregate level :

$$\mathbb{E}\left[w\right] = 1$$
$$\mathbb{E}\left[D \cdot w\right] = \mathbb{E}\left[D\right]$$
$$\mathbb{E}\left[X \cdot w\right] = \mathbb{E}\left[X\right]$$
$$\mathbb{E}\left[Y \cdot w\right] = \mathbb{E}\left[Y\right]$$

Intuitive properties of weights

The weights have intuitive properties.

They do not distort variables on the aggregate level :

$$\mathbb{E}[w] = 1$$
$$\mathbb{E}[D \cdot w] = \mathbb{E}[D]$$
$$\mathbb{E}[X \cdot w] = \mathbb{E}[X]$$
$$\mathbb{E}[Y \cdot w] = \mathbb{E}[Y]$$

They attempt to remove the dependence between X and D :

$$\text{Cov}(X \cdot w, D) = 0$$

Intuitive properties of weights

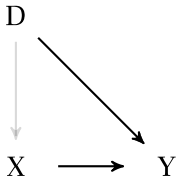
The weights have intuitive properties.

They do not distort variables on the aggregate level :

$$\mathbb{E}[w] = 1$$
$$\mathbb{E}[D \cdot w] = \mathbb{E}[D]$$
$$\mathbb{E}[X \cdot w] = \mathbb{E}[X]$$
$$\mathbb{E}[Y \cdot w] = \mathbb{E}[Y]$$

They attempt to remove the dependence between X and D :

$$\mathrm{Cov}(X \cdot w, D) = 0$$



Example

Example

For every profile i :

- x_i : Occupation (Nursing ou Mechanic)
- d_i : Gender (**M**ale or **F**emale)
- e_i : Exposure to risk (vehicle year)
- Y_i : Observed pure premium ($)

Example

For every profile i :

- x_i : Occupation (Nursing 💉 ou Mechanic 🔧)
- d_i : Gender (**M**ale or **F**emale)
- e_i : Exposure to risk (vehicle year)
- Y_i : Observed pure premium ($)

Table 2: Dataset of motor vehicle claims

i	x_i	d_i	e_i	Y_i
1	💉	F	95	50
2	💉	M	5	200
3	🔧	F	10	75
4	🔧	M	90	250

Weight calculation

We use formula mentioned previously for weights :

$$w_1 = \frac{\hat{\mathrm{P}}(D = \mathtt{F})}{\hat{\mathrm{P}}(D = \mathtt{F} | X = \text{syringe})} = \frac{\left(\frac{95 + 10}{200}\right)}{\left(\frac{95}{95 + 5}\right)} \approx 0.5526$$

Table 3: Required information for weight calculation

i	x_i	d_i	e_i	★ w_i ★
1	syringe	F	95	
2	syringe	M	5	
3	wrench	F	10	
4	wrench	M	90	

Weight calculation

We use formula mentioned previously for weights :

$$w_1 = \frac{\hat{\mathrm{P}}(D = \mathtt{F})}{\hat{\mathrm{P}}(D = \mathtt{F} | X = \text{💉})} = \frac{\left(\frac{95 + 10}{200}\right)}{\left(\frac{95}{95 + 5}\right)} \approx 0.5526$$

Table 3: Required information for weight calculation

i	x_i	d_i	e_i	⭐ w_i ⭐
1	💉	F	95	0.5526
2	💉	M	5	
3	🔧	F	10	
4	🔧	M	90	

Weight calculation

We use formula mentioned previously for weights :

$$w_1 = \frac{\hat{\mathrm{P}}(D = \texttt{F})}{\hat{\mathrm{P}}(D = \texttt{F} | X = 💉)} = \frac{\left(\frac{95 + 10}{200}\right)}{\left(\frac{95}{95+5}\right)} \approx 0.5526$$

$$w_2 = \frac{\hat{\mathrm{P}}(D = \texttt{M})}{\hat{\mathrm{P}}(D = \texttt{M} | X = 💉)} = \frac{\left(\frac{90 + 5}{200}\right)}{\left(\frac{5}{95+5}\right)} = 9.5$$

Table 3: Required information for weight calculation

i	x_i	d_i	e_i	⭐ w_i ⭐
1	💉	F	95	0.5526
2	💉	M	5	
3	🔧	F	10	
4	🔧	M	90	

Weight calculation

We use formula mentioned previously for weights :

$$w_1 = \frac{\hat{\mathrm{P}}(D = \mathtt{F})}{\hat{\mathrm{P}}(D = \mathtt{F} | X = \text{💉})} = \frac{\left(\frac{95 + 10}{200}\right)}{\left(\frac{95}{95 + 5}\right)} \approx 0.5526$$

$$w_2 = \frac{\hat{\mathrm{P}}(D = \mathtt{M})}{\hat{\mathrm{P}}(D = \mathtt{M} | X = \text{💉})} = \frac{\left(\frac{90 + 5}{200}\right)}{\left(\frac{5}{95 + 5}\right)} = 9.5$$

Table 3: Required information for weight calculation

i	x_i	d_i	e_i	⭐ w_i ⭐
1	💉	F	95	0.5526
2	💉	M	5	9.5
3	🔧	F	10	
4	🔧	M	90	

Weight calculation

We use formula mentioned previously for weights :

$$w_1 = \frac{\hat{\mathrm{P}}(D = \mathtt{F})}{\hat{\mathrm{P}}(D = \mathtt{F} | X = \text{💉})} = \frac{\left(\frac{95+10}{200}\right)}{\left(\frac{95}{95+5}\right)} \approx 0.5526$$

$$w_2 = \frac{\hat{\mathrm{P}}(D = \mathtt{M})}{\hat{\mathrm{P}}(D = \mathtt{M} | X = \text{💉})} = \frac{\left(\frac{90+5}{200}\right)}{\left(\frac{5}{95+5}\right)} = 9.5$$

$$w_3 = \frac{\hat{\mathrm{P}}(D = \mathtt{F})}{\hat{\mathrm{P}}(D = \mathtt{F} | X = \text{🔧})} = \frac{\left(\frac{95+10}{200}\right)}{\left(\frac{10}{10+90}\right)} = 5.25$$

$$w_4 = \frac{\hat{\mathrm{P}}(D = \mathtt{M})}{\hat{\mathrm{P}}(D = \mathtt{M} | X = \text{🔧})} = \frac{\left(\frac{90+5}{200}\right)}{\left(\frac{90}{10+90}\right)} \approx 0.5278$$

Table 3: Required information for weight calculation

i	x_i	d_i	e_i	⭐ w_i ⭐
1	💉	F	95	0.5526
2	💉	M	5	9.5
3	🔧	F	10	5.25
4	🔧	M	90	0.5278

New exposure

Table 4: Dataset of motor vehicle claims

i	x_i	d_i	e_i	w_i	e_i^*	Y_i
1		F	95	0.5526	52.5	50
2		M	5	9.5	47.5	200
3		F	10	5.25	52.5	75
4		M	90	0.5278	47.5	250

With

$$e_i \cdot w_i = e_i^*$$

Avoiding D using different approaches

Table 5: Aggregated profiles using e

x_j	e_j	$\mu^{U}(X)$ (Unaware)	$\mu^{IPW}(X)$ (Inverse pribability weighting)
	$95 + 5 =$ **100**		
	$10 + 90 =$ **100**		

Avoiding D using different approaches

Table 5: Aggregated profiles using e

x_j	e_j	$\mu^U(X)$ (Unaware)	$\mu^{IPW}(X)$ (Inverse pribability weighting)
	$95 + 5 =$ 100	$0.95 \cdot 50 +$ $0.05 \cdot 200 =$ 57.5	
	$10 + 90 =$ 100	$0.10 \cdot 75 +$ $0.90 \cdot 250 =$ 232.5	

Avoiding D using different approaches

Table 5: Aggregated profiles using e

x_j	e_j	$\mu^U(X)$ (Unaware)	$\mu^{IPW}(X)$ (Inverse pribability weighting)
	$95+5=$ 100	$0.95 \cdot 50 +$ $0.05 \cdot 200 =$ 57.5	$0.525 \cdot 50 +$ $0.475 \cdot 200 =$ 121.25
	$10+90=$ 100	$0.10 \cdot 75+$ $0.90 \cdot 250 =$ 232.5	$0.525 \cdot 75+$ $0.475 \cdot 250 =$ 158.125

Conclusion

Concluding on Causal Inference and Fairness in insurance pricing

The theoretical equivalence between the discrimination-free formula and causal tools goes beyond that.

Concluding on Causal Inference and Fairness in insurance pricing

The theoretical equivalence between the discrimination-free formula and causal tools goes beyond that.

Causal inference proposes many strategies to remove biases from confounders (Hernán and Robins, 2020; Moodie and Stephens, 2022):

Concluding on Causal Inference and Fairness in insurance pricing

The theoretical equivalence between the discrimination-free formula and causal tools goes beyond that.

Causal inference proposes many strategies to remove biases from confounders (Hernán and Robins, 2020; Moodie and Stephens, 2022):

- Standardization (g-formula)

Concluding on Causal Inference and Fairness in insurance pricing

The theoretical equivalence between the discrimination-free formula and causal tools goes beyond that.

Causal inference proposes many strategies to remove biases from confounders (Hernán and Robins, 2020; Moodie and Stephens, 2022):

- Standardization (g-formula)
- Inverse probability weighting (IPW)

Concluding on Causal Inference and Fairness in insurance pricing

The theoretical equivalence between the discrimination-free formula and causal tools goes beyond that.

Causal inference proposes many strategies to remove biases from confounders (Hernán and Robins, 2020; Moodie and Stephens, 2022):

- Standardization (g-formula)
- Inverse probability weighting (IPW)
- Matching (optimal transport)

Concluding on Causal Inference and Fairness in insurance pricing

The theoretical equivalence between the discrimination-free formula and causal tools goes beyond that.

Causal inference proposes many strategies to remove biases from confounders (Hernán and Robins, 2020; Moodie and Stephens, 2022):

- Standardization (g-formula) **(Pope and Sydnor, 2011; Aseervatham et al., 2016; Lindholm et al., 2022; Araiza Iturria et al., 2022)**
- Inverse probability weighting (IPW) **(Lindholm et al., 2023)**
- Matching (optimal transport) **(Charpentier et al., 2023; Lindholm et al., 2023).**

Concluding on Causal Inference and Fairness in insurance pricing

The theoretical equivalence between the discrimination-free formula and causal tools goes beyond that.

Causal inference proposes many strategies to remove biases from confounders (Hernán and Robins, 2020; Moodie and Stephens, 2022):

- Standardization (g-formula) **(Pope and Sydnor, 2011; Aseervatham et al., 2016; Lindholm et al., 2022; Araiza Iturria et al., 2022)**
- Inverse probability weighting (IPW) **(Lindholm et al., 2023)**
- Matching (optimal transport) **(Charpentier et al., 2023; Lindholm et al., 2023).**

There is still some work to apply causal inference with high-dimensional X and D (Li and Li, 2019).

Thank you

Bibliography i

Araiza Iturria, C. A., Hardy, M., and Marriott, P. (2022). A discrimination-free premium under a causal framework. *Available at SSRN 4079068*.

Aseervatham, V., Lex, C., and Spindler, M. (2016). How do unisex rating regulations affect gender differences in insurance premiums? *The Geneva Papers on Risk and Insurance-Issues and Practice*, 41(1):128–160.

Autorité des Marchés Financiers (2021). L'intelligence artificielle en finance : Recommandations pour une utilisation responsable.

Casualty Actuarial Society (1988). Statement of Principles Regarding Property and Casualty Insurance Ratemaking. Adopté par le conseil d'administration de la CAS en mai 1988. Consulté le 13 février 2022.

Bibliography ii

Charpentier, A., Flachaire, E., and Gallic, E. (2023). Optimal transport for counterfactual estimation: A method for causal inference. *arXiv preprint arXiv:2301.07755*.

Embrechts, P. and Wüthrich, M. V. (2022). Recent challenges in actuarial science. *Annual Review of Statistics and Its Application*, 9:119–140.

Fisher, R. A. (1956). Statistical methods and scientific inference.

Fong, C., Hazlett, C., and Imai, K. (2018). Covariate balancing propensity score for a continuous treatment: Application to the efficacy of political advertisements. *The Annals of Applied Statistics*, 12(1):156–177.

Frezal, S. and Barry, L. (2020). Fairness in uncertainty: Some limits and misinterpretations of actuarial fairness. *Journal of Business Ethics*, 167(1):127–136.

Hernán, M. A. and Robins, J. M. (2020). Causal inference: What if.

Bibliography iii

Li, F. and Li, F. (2019). Propensity score weighting for causal inference with multiple treatments.

Lindholm, M., Richman, R., Tsanakas, A., and Wüthrich, M. V. (2022). Discrimination-free insurance pricing. *ASTIN Bulletin: The Journal of the IAA*, 52(1):55–89.

Lindholm, M., Richman, R., Tsanakas, A., and Wuthrich, M. V. (2023). What is fair? proxy discrimination vs. demographic disparities in insurance pricing. *Proxy Discrimination vs. Demographic Disparities in Insurance Pricing (May 2, 2023)*.

Moodie, E. E. and Stephens, D. A. (2022). Causal inference: Critical developments, past and future. *Canadian Journal of Statistics*.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.

Pope, D. G. and Sydnor, J. R. (2011). Implementing anti-discrimination policies in statistical profiling models. *American Economic Journal: Economic Policy*, 3(3):206–31.