Predictive Modeling and Balance Property through Auto-calibration

Julien TRUFIN
Professor of Actuarial Science at Université Libre de Bruxelles
Scientific Advisor at Detralytics

January 21 2025
Petit-déjeuner de la **chaire d'excellence ACTIONS**

Table of contents

Part 1 - Global and local balances in insurance pricing

Setting

- Y = aggregate claim amount or its frequency component.
- $\boldsymbol{X} = (X_1, X_2, \ldots, X_p)$ = risk or rating factors.
- Pure premium= amount collected by the insurer to compensate for the claims, without loss nor profit.
- $\pi(\boldsymbol{X})$ candidate pure premium.
- $\mu(\boldsymbol{X}) = \mathrm{E}[Y|\boldsymbol{X}]$ true pure premium.

Global balance

- **Global balance** is the very **basic requirement** for a pure premium in insurance.
- Formally, it requires that

$$\mathrm{E}[Y] = \mathrm{E}[\pi(\boldsymbol{X})].$$

- In practice, this means that

$$\frac{1}{n}\sum_{i=1}^{n} Y_i \approx \frac{1}{n}\sum_{i=1}^{n} \pi(\boldsymbol{X}_i) \text{ for large } n$$

- In words, global balance for $\pi(\cdot)$ ensures that

  aggregate loss $\approx$ total pure premium income, for large n.

Method of marginal totals (MMT)

- **Early 1960s** in North America, still reflects actuarial thinking.
- With $x_{i,j} \in \{0,1\}$ for all j, the **pure premium** is written as

$$\pi(\boldsymbol{x}_i) = e_i \gamma_0 \prod_{j=1}^{p} \gamma_j^{x_{i,j}}$$

where

$$\begin{aligned} e_i &= \text{risk exposure for policy } i \\ \gamma_0 &= \textbf{base premium } \pi(0) \text{ per unit of exposure} \\ \gamma_j &= \textbf{relativity} \text{ associated with risk factor } j. \end{aligned}$$

- Considering $\boldsymbol{x}_{i_1} = \boldsymbol{x}_{i_2}$ except for $x_{i_1,j} = 1$, $x_{i_2,j} = 0$,

$$\gamma_j = \frac{\pi(\boldsymbol{x}_{i_1})}{\pi(\boldsymbol{x}_{i_2})}.$$

Method of marginal totals (MMT)

- MMT imposes

  - **global balance**

$$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \pi(\boldsymbol{x}_i)$$

  - and **local balance** over all contracts with $x_{i,j} = 1$

$$\sum_{i|x_{i,j}=1} y_i = \sum_{i|x_{i,j}=1} \pi(\boldsymbol{x}_i) \text{ for } j = 1, 2, \ldots, p.$$

- Referring to contingency tables, this type of local balance is called **marginal balance**.
- Base premium γ_0 and relativities γ_j solve the system of balance equations.

From MMT to GLMs

- Assume that given $\boldsymbol{X}$, the response Y obeys a distribution from the **ED family**: Binomial, Poisson, Gamma, Tweedie, etc.
- The **pure premium** rate for policy i is of the form

$$g^{-1}\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{i,j}\right)$$

where g is the link function.

- With canonical link function

Minimizing deviance $\Leftrightarrow$ Solving MMT equations.

$\Rightarrow$ **GLM estimation with canonical link is equivalent to the actuarial MMT** which predates GLMs.

$\Rightarrow$ MMT can be implemented with any tool developed for GLMs.

From MMT to GLMs

- If $x_{i,j} \in \{0,1\}$ for all j and $g = \log$ then

$$\pi(\boldsymbol{x}_i) = \exp\left(\ln e_i + \beta_0 + \sum_{j=1}^{p} \beta_j x_{i,j}\right) = e_i \exp(\beta_0) \prod_{j|x_{i,j}=1} \exp(\beta_j)$$

where

$$\begin{aligned} \exp(\beta_0) &= \text{base premium per unit of exposure for the reference} \\ &\quad \text{class, for which } x_{i,1} = x_{i,2} = \ldots = x_{i,p} = 0 \\ &= \gamma_0 \end{aligned}$$

$$\begin{aligned} \exp(\beta_j) &= \text{relativity of risk factor } j\text{, i.e. effect of switching} \\ &\quad \text{from } x_{i,j} = 0 \text{ to } x_{i,j} = 1 \\ &= \gamma_j. \end{aligned}$$

- We then **recover the multiplicative structure** of the premium.

Modern learning tools

- **Linear score** $\beta_0 + \sum_{j=1}^{p} \beta_j x_{i,j}$ in GLMs **replaced with much more flexible ones**.
- Premiums still determined by **minimizing deviance**.
- However, deviance now

  - **penalizes departures** of $\pi(\boldsymbol{x}_i)$ from y_i

  - **maximizes correlation** between $\pi(\boldsymbol{x}_i)$ and y_i

  measured on different scales.
- Documented in Denuit, Sznajder and Trufin (2019), Denuit, Charpentier and Trufin (2021).
- **Balance is no more imposed** and optimizing deviance may favor correlation.
- Lack of balance is not **problematic** in many applications (like credit scoring or fraud detection) but well **in pure premium calculation**.

Local balance by autocalibration

- **Autocalibration** by Krüger and Ziegel (2021) **imposes**

$$\mathrm{E}\big[Y\big|\pi(\boldsymbol{X}) = s\big] = s \text{ for all } s.$$

  - Balance thus operates within each group of contracts charged the same premium, preventing any transfer.

- **Example**: two values for $\pi(\boldsymbol{X})$: π_1 and π_2. Then, $\pi(\boldsymbol{X})$ autocalibrated means that

$$\frac{1}{n_1}\sum_{i:\pi(\boldsymbol{X}_i)=\pi_1} Y_i = \pi_1 \quad \text{and} \quad \frac{1}{n_2}\sum_{i:\pi(\boldsymbol{X}_i)=\pi_2} Y_i = \pi_2.$$

  where $n_j = \sum_{i=1}^{n} \mathrm{I}[\pi(\boldsymbol{X}_i) = \pi_j]$, $j = 1, 2$.

Part 2 - Testing for local balances using the concept of auto-calibration

Lorenz and concentration curves

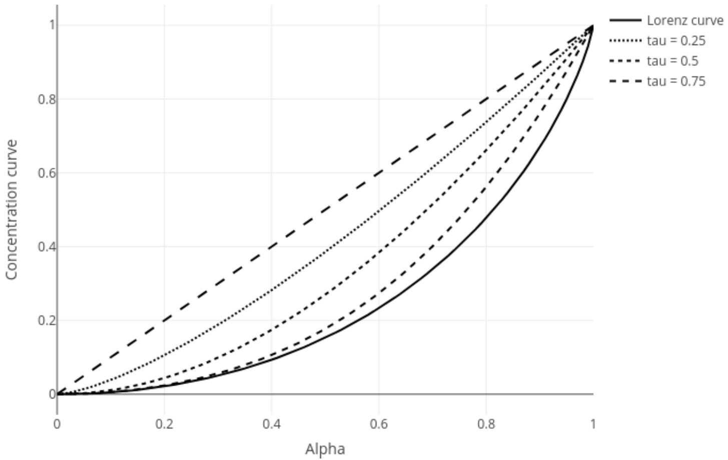
- The **Lorenz curve** associated to the predictor $\widehat{\pi}$ is defined by

$$\mathrm{LC}[\widehat{\pi}(\boldsymbol{X});\alpha] = \frac{\mathrm{E}\big[\widehat{\pi}(\boldsymbol{X})\mathrm{I}[\widehat{\pi}(\boldsymbol{X}) \leq F_{\widehat{\pi}}^{-1}(\alpha)]\big]}{\mathrm{E}[\widehat{\pi}(\boldsymbol{X})]}, \quad \alpha \in (0,1).$$

- The **concentration curve** of the response Y with respect to the predictor $\widehat{\pi}$ is defined as

$$\mathrm{CC}[Y,\widehat{\pi}(\boldsymbol{X});\alpha] = \frac{\mathrm{E}\big[Y\mathrm{I}[\widehat{\pi}(\boldsymbol{X}) \leq F_{\widehat{\pi}}^{-1}(\alpha)]\big]}{\mathrm{E}[Y]}, \quad \alpha \in (0,1).$$

Lorenz and concentration curves



A new characterization of auto-calibration

- **Proposition:** (Denuit et al., 2024) We have

$$\mathrm{CC}[\mu(\boldsymbol{X}), \widehat{\pi}(\boldsymbol{X}); \alpha] = \mathrm{LC}[\widehat{\pi}(\boldsymbol{X}); \alpha] \quad \text{for all } \alpha \in (0,1)$$

if, and only if $\widehat{\pi}_{\text{unbiased}}(\boldsymbol{X}) = \frac{\mathrm{E}[Y]}{\mathrm{E}[\widehat{\pi}(\boldsymbol{X})]}\widehat{\pi}(\boldsymbol{X})$ is auto-calibrated.

Statistical test for auto-calibration

- Denuit et al. (2024) introduced a **statistical test for auto-calibration**.
- The authors test the **null hypothesis**

$$\mathcal{H}_0 : \mathrm{CC}[\mu(\boldsymbol{X}), \widehat{\pi}(\boldsymbol{X}); \alpha] = \mathrm{LC}[\widehat{\pi}(\boldsymbol{X}); \alpha] \text{ for all } \alpha \in (0,1)$$

against the **alternative**

$$\mathcal{H}_1 : \mathrm{CC}[\mu(\boldsymbol{X}), \widehat{\pi}(\boldsymbol{X}); \alpha] \neq \mathrm{LC}[\widehat{\pi}(\boldsymbol{X}); \alpha] \text{ for some } \alpha \in (0,1).$$

Statistical test

- **Testing procedure**: based on the difference between sample versions

$$\widehat{\mathrm{CC}}[\mu(\boldsymbol{X}), \widehat{\pi}(\boldsymbol{X}); \alpha] = \frac{1}{n\bar{Y}} \sum_{i=1}^{n} Y_i \, \mathrm{I}\left[\widehat{\pi}(\boldsymbol{X}_i) \leq F_{\widehat{\pi}}^{-1}(\alpha)\right], \quad \alpha \in (0,1),$$

and

$$\widehat{\mathrm{LC}}[\widehat{\pi}(\boldsymbol{X}); \alpha] = \frac{1}{n\bar{\widehat{\pi}}} \sum_{i=1}^{n} \widehat{\pi}(\boldsymbol{X}_i) \, \mathrm{I}\left[\widehat{\pi}(\boldsymbol{X}_i) \leq F_{\widehat{\pi}}^{-1}(\alpha)\right], \quad \alpha \in (0,1).$$

Statistical test

- The **null hypothesis is rejected for large values of** the test statistic
$$\mathcal{T} = \sup_{\alpha\in(0,1)}|T_n(\alpha)|,$$
where
$$\begin{aligned} T_n(\alpha) &= \sqrt{n}\left(\widehat{\mathrm{CC}}[\mu(\boldsymbol{X}), \widehat{\pi}(\boldsymbol{X}); \alpha] - \widehat{\mathrm{LC}}[\widehat{\pi}(\boldsymbol{X}); \alpha]\right) \\ &= n^{-1/2}\sum_{i=1}^{n}\left(\frac{Y_i}{\bar{Y}} - \frac{\widehat{\pi}(\boldsymbol{X}_i)}{\bar{\pi}}\right)\mathrm{I}[\widehat{\pi}(\boldsymbol{X}_i) \leq F_{\widehat{\pi}}^{-1}(\alpha)], \quad \alpha \in (0,1). \end{aligned}$$

Statistical test

- **Proposition:** Assume that $(Y_i, \widehat{\pi}(\boldsymbol{X}_i))$, $i = 1, 2, \ldots, n$, are such that $\mathrm{E}[Y_i] \neq 0$, $\mathrm{E}[\widehat{\pi}(\boldsymbol{X}_i)] \neq 0$, $\mathrm{E}[\widehat{\pi}^2(\boldsymbol{X}_i)] < \infty$ and $\mathrm{E}[Y_i^2] < \infty$. Then, under the null hypothesis, $T_n(\alpha)$ converges weakly to a Gaussian process.

Statistical test

- **The proposed test rejects the null hypothesis when**

$$\mathcal{T} = \sup_{\alpha\in(0,1)} |T_n(\alpha)| > c_\beta,$$

  for some critical value c_β such that
  $\mathrm{P}[\sup_{\alpha\in(0,1)} |T_n(\alpha)| > c_\beta] = \beta$.
- However, the analyst **cannot compute the critical value** c_β since the underlying distribution of $(Y_i, \widehat{\pi}(\boldsymbol{X}_i))$ is unknown. Given the asymptotic Gaussian behavior of the process, the **non-parametric Monte-Carlo methods** of Zhu et al. (2016) can be used.

Case study - Data set

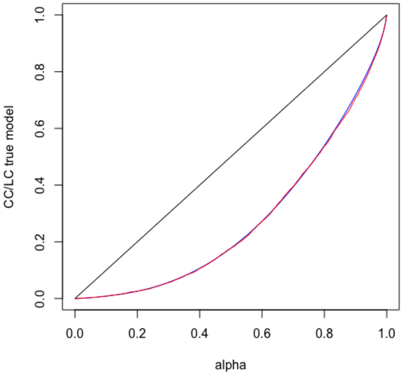
- **Swiss motor insurance database** with 500 000 insurance policies used in Wüthrich and Buser (2016) and in Wüthrich (2020).
- For **each policy** i :
  - **Numbers of claims** Y_i;
  - **Exposure-to-risk** $e_i \leq 1$ (i.e. the duration of observation expressed in years)
  - **Features** $\mathrm{X}_i = (X_{i1}, \ldots, X_{i8})$:
    - X_{i1} = the age of the driver (`age`);
    - X_{i2} = the age of the car (`ac`);
    - X_{i3} = the power of the car (`power`);
    - X_{i4} = the fuel type of the car (`fuel`);
    - X_{i5} = the vehicle brand (`vehicle brand`);
    - X_{i6} = the area code of the living place of the driver (`area`);
    - X_{i7} = the population density at the living place of the driver (`dens`);
    - X_{i8} = the Swiss canton of the license plate of the car (`ct`).

Case study - Data set

- **Important**: The values $e_i\mu(\boldsymbol{X}_i)$ of the **true model** are **also provided** (the Y_i have been generated from expected true model frequencies $e_i\mu(\boldsymbol{X}_i)$).
- We partition the data set into a **training set** $\mathcal{D}$ (80%) and a **validation set** $\overline{\mathcal{D}}$ (20%).

Case study - Data set

- Estimates of the CC (in red) and LC (in blue) on $\overline{\mathcal{D}}$ for the **true model**:



- Applying our **testing procedure** to the true model on $\overline{\mathcal{D}}$, **we get** $\widehat{p}_{\textbf{equal}} = 0.29$, meaning that the null hypothesis is not rejected at the level 5%.

Case study - Models under consideration

- Given $\boldsymbol{X} = \boldsymbol{x}$ and the exposure-to-risk e, Y is assumed to be **Poisson with mean** $e\mu(\boldsymbol{x})$.
- $\mathcal{D}$ is divided into $\mathcal{D}_1$ and $\mathcal{D}_2$, where $\mathcal{D}_1$ includes 80% of the observations of $\mathcal{D}$ and $\mathcal{D}_2$ gathers the remaining 20%.
- We first fit **two GAMs** on $\mathcal{D}_1$ using the `R` package `mgcv`:
  - $\widehat{\pi}^{\text{GAM1}}(\boldsymbol{x})$, with **only the feature** X_1 (`age`);
  - $\widehat{\pi}^{\text{GAM2}}(\boldsymbol{x})$, **using all 8 available features**.

  The covariates `age`, `ac`, `power` and `dens` are captured by **splines** and **we do not consider interaction terms**.

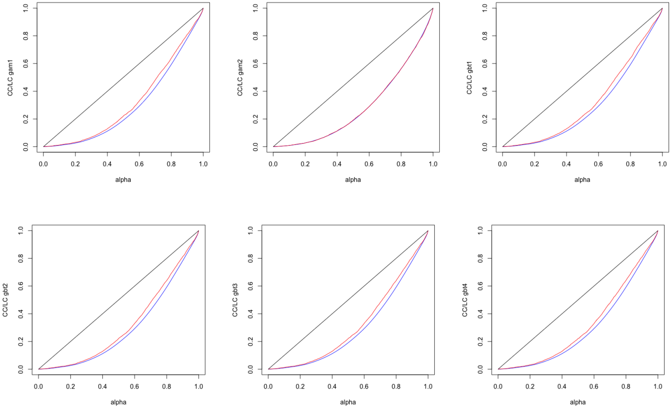
Case study - Models under consideration

- We fit **four GBTs** on $\mathcal{D}_1$ producing the following estimators:
  - $\widehat{\pi}^{\text{GBT1}}(\boldsymbol{x})$, with **only the feature** X_1 (age);
  - $\widehat{\pi}^{\text{GBT2}}(\boldsymbol{x})$, using all **8 available features** and $\texttt{ID} = 1$;
  - $\widehat{\pi}^{\text{GBT3}}(\boldsymbol{x})$, using all **8 available features** and $\texttt{ID} = 2$;
  - $\widehat{\pi}^{\text{GBT4}}(\boldsymbol{x})$, using all **8 available features** and $\texttt{ID} = 3$.

  The **optimal values for the number of trees** T are determined by minimizing the out-of-sample Poisson deviance loss computed on $\mathcal{D}_2$:
  - $T = 507$ for $\widehat{\pi}^{\text{GBT1}}$;
  - $T = 2977$ for $\widehat{\pi}^{\text{GBT2}}$;
  - $T = 2984$ for $\widehat{\pi}^{\text{GBT3}}$;
  - $T = 2943$ for $\widehat{\pi}^{\text{GBT4}}$.

Case study - Testing for equality of Concentration and Lorenz curves

- Estimates of the CC (in red) and LC (in blue) on $\overline{\mathcal{D}}$:



Case study - Testing for equality of Concentration and Lorenz curves

- **Values of** $\widehat{p}_{\textbf{equal}}$:

$\widehat{\pi}$	$\widehat{p}_{\text{equal}}$
$\widehat{\pi}^{\text{GAM1}}$	0.000
$\widehat{\pi}^{\text{GAM2}}$	0.744
$\widehat{\pi}^{\text{GBT1}}$	0.000
$\widehat{\pi}^{\text{GBT2}}$	0.000
$\widehat{\pi}^{\text{GBT3}}$	0.000
$\widehat{\pi}^{\text{GBT4}}$	0.000

Part 3 - Balance correction as a means for auto-calibration

Balance correction as a way to restore financial equilibrium

- Start from any predictor $\pi(\boldsymbol{X})$ **strongly correlated with the response** Y.

  $\Rightarrow$ The **ranks** produced by $\pi(\cdot)$ **are informative** to order contracts from the cheapest to the most expensive one.
- An accurate premium can then be obtained by averaging claim amounts over neighborhoods induced by $\pi(\cdot)$.
- The **balance-corrected premium** $\pi_{\text{bc}}(\boldsymbol{X})$ obtained from $\pi(\boldsymbol{X})$ by

$$\pi_{\text{bc}}(\boldsymbol{X}) = \mathrm{E}\left[Y \middle| \pi(\boldsymbol{X})\right]$$

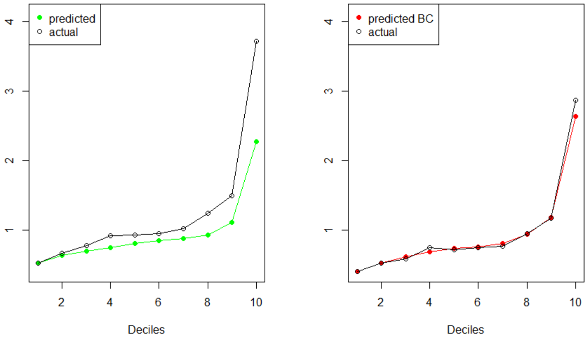
  satisfies autocalibration.
- In practice, **balance correction can be implemented** by

  - local regression (Denuit, Charpentier and Trufin, 2021, Ciatto et al. 2023);

  - isotonic regression (Wüthrich and Ziegel, 2023).

Balance correction: Example (Ciatto et al. (2023)

- `freMTPL2freq` data set (in the `CASdatasets` package in `R`).
- 678 013 observations of the number of claims (response Y) in a French MTPL portfolio, with 9 features $\boldsymbol{X} = (X_1, \ldots, X_9)$.
- $\widehat{\pi}$: boosted Poisson model (training set = 60% of the data).
- $\widehat{\pi}_{\text{bc}}$: obtained by local regressions (on 20% of the data= validation set).
- Lift chart (on the test set = 20% of the data) :



Impact of balance correction on Bregman divergence

- For a convex function $\ell(\cdot)$, define the error measure for the mean as

$$L(y, m) = \ell(y) - \ell(m) - \ell'(m)(y - m).$$

  Functions $L(\cdot, \cdot)$ are called **Bregman loss functions**.
- Bregman loss functions are **important in insurance ratemaking** since any Bregman loss function is a strictly consistent loss function for the mean functional.
- Then,

$$\mathsf{E}\big[L\big(Y, \widehat{\pi}_{\mathrm{bc}}(\boldsymbol{X})\big)\big] \le \mathsf{E}\big[L\big(Y, \widehat{\pi}(\boldsymbol{X})\big)\big].$$

Impact of balance correction on concentration curves

- **Proposition:** (Denuit and Trufin, 2024)
  We have

  $$\mathrm{CC}[Y, \widehat{\pi}_{\mathrm{bc}}(\boldsymbol{X}); \alpha] \leq \mathrm{CC}[Y, \widehat{\pi}(\boldsymbol{X}); \alpha]$$

  for any probability level α.

Impact of balance correction on Lorenz curves

- **Proposition:** (Denuit and Trufin, 2024)
  Let $t \to m(t) = \mathsf{E}[Y|\widehat{\pi}(\boldsymbol{X}) = t]$. Assume that $m(\cdot)$ is continuous and strictly increasing. Then,

  $$t \mapsto \frac{m(t)}{t} \text{ non-decreasing} \Rightarrow \mathrm{LC}[\widehat{\pi}_{\mathrm{bc}}(\boldsymbol{X}); \alpha] \leq \mathrm{LC}[\widehat{\pi}(\boldsymbol{X}); \alpha]$$

  while

  $$t \mapsto \frac{m(t)}{t} \text{ non-increasing} \Rightarrow \mathrm{LC}[\widehat{\pi}_{\mathrm{bc}}(\boldsymbol{X}); \alpha] \geq \mathrm{LC}[\widehat{\pi}(\boldsymbol{X}); \alpha].$$

Part 4 - Performance criteria for auto-calibrated predictors

Context

- There are **many tools for model selection in machine learning** : deviance criteria, Gini index, concentration and Lorenz curves, correlation coefficients, . . .
- Interesting facts:
  - **Deviance is a consistent scoring rule** for the mean;
  - **Not the case in general for** the **other measures** listed before.

  $\Rightarrow$ Working with these tools for model selection **may thus lead to a wrong model choice**.
- Restricting the **Gini index** to the class of **autocalibrated regression models** makes it a **consistent scoring rule** (Wüthrich, 2023).

Pearson's linear correlation

- **Correlation coefficients** are often used to assess the strength of dependence within the pair $(Y, \widehat{\pi}(\boldsymbol{X}))$.
- The most elementary correlation coefficient is Pearson's linear one:
$$r(Y, \widehat{\pi}(\boldsymbol{X})) = \frac{\mathbb{C}\mathrm{ov}[Y, \widehat{\pi}(\boldsymbol{X})]}{\sqrt{\mathrm{Var}[Y]\mathrm{Var}[\widehat{\pi}(\boldsymbol{X})]}}.$$
- **Proposition 4.1:** (Denuit and Trufin, 2023)
Let $\widehat{\pi}(\boldsymbol{X})$ be an autocalibrated predictor. Then, $r(Y, \widehat{\pi}(\boldsymbol{X}))$ is maximum if, and only if, $\widehat{\pi}(\boldsymbol{X}) = \mu(\boldsymbol{X})$.

Gini index and ICC

- **The integral of the concentration curve** (ICC) over the whole interval $[0, 1]$ is then given by

$$\begin{aligned}\mathrm{ICC}[Y, \widehat{\pi}(\boldsymbol{X})] &= \int_0^1 \mathrm{CC}[Y, \widehat{\pi}(\boldsymbol{X}); \alpha]\mathrm{d}\alpha \\ &= \frac{\mathsf{E}\left[Y \int_0^1 \mathsf{I}[\widehat{\pi}(\boldsymbol{X}) \leq F_{\widehat{\pi}}^{-1}(\alpha)]\mathrm{d}\alpha\right]}{\mathsf{E}[Y]}.\end{aligned}$$

- The **Gini index** can be defined as

$$\mathrm{Gini}[Y, \widehat{\pi}(\boldsymbol{X})] = \frac{\frac{1}{2} - \mathrm{ICC}[Y, \widehat{\pi}(\boldsymbol{X})]}{\frac{1}{2} - \int_0^1 \mathrm{CC}[Y, Y; \alpha]\mathrm{d}\alpha}.$$

Gini index and ICC

- The **Gini index is not in general a consistent scoring rule** for the mean, so neither is the ICC (Wüthrich, 2023).
- BUT the Gini index gives a **consistent scoring rule within the class of autocalibrated regression functions** (Wüthrich, 2023).

Concentration curve

- **Proposition 4.3**: (Denuit and Trufin, 2023)
  For any autocalibrated predictor $\widehat{\pi}(\boldsymbol{X})$, we have

  $$\mathrm{CC}[Y, \widehat{\pi}(\boldsymbol{X}); \alpha] \geq \mathrm{CC}[Y, \mu(\boldsymbol{X}); \alpha] \quad \text{for any } \alpha \in (0,1),$$

  with an identity if, and only if, $\widehat{\pi}(\boldsymbol{X}) = \mu(\boldsymbol{X})$.

Lorenz curve

- Since the LC coincides with the CC for autocalibrated predictors, Proposition 4.3 shows that **the true model also has the smallest Lorenz curve**.
- This **legitimates model assessment based on Lorenz curve under autocalibration**, as it is often performed in practice!

References

References

- Ciatto, N., Verelst, H., Trufin, J., Denuit, M. (2023). Does autocalibration improve goodness of lift? European Actuarial Journal 13, 479-486.
- Denuit, M., Charpentier, A., Trufin, J. (2021). Autocalibration and Tweedie-dominance for insurance pricing with machine learning. Insurance: Mathematics and Economics 101, 485-497.
- Denuit, M., Huyghe, J., Trufin, J., Verdebout, T. (2024). Testing for auto-calibration with Lorenz and Concentration curves. Insurance: Mathematics and Economics 117, 130-139.
- Denuit, M., Sznajder, D., Trufin, J. (2019). Model selection based on Lorenz and Concentration curves, Gini indices and convex order. Insurance: Mathematics and Economics 89, 128-139.
- Denuit, M., Trufin, J. (2023). Model selection with Pearson's correlation, concentration and Lorenz curves under autocalibration. European Actuarial Journal, 13, 871-878.
- Denuit, M., Trufin, J. (2024). Convex and Lorenz orders under balance correction in nonlife insurance pricing: review and new developments. Insurance: Mathematics and Economics 118, 123-128.

References

- Krüger, F., Ziegel, J.F. (2021). Generic conditions for forecast dominance. Journal of Business & Economic Statistics 39, 972-983.
- Wüthrich, M.V. (2020). Bias regularization in neural network models for general insurance pricing. European Actuarial Journal 10, 179-202.
- Wüthrich, M. V. (2023). Model selection with Gini indices under auto-calibration. European Actuarial Journal 13, 469-477.
- Wüthrich, M.V., Buser, C. (2016). Data analytics for non-life insurance pricing. SSRN Manuscript ID 2870308, Version of June 4, 2019.
- Wüthrich, M. V., Ziegel, J. (2023). Isotonic recalibration under a low signal-to-noise ratio. Scandinavian Actuarial Journal, 2024(3), 279–299.
- Zhu, X., Guo, X., Lin, L., Zhu, L. (2016). Testing for positive expectation dependence. Annals of the Institute of Statistical Mathematics 68, 135-153.