Sensibilité des indices de qualité d'un classifieur probabiliste

Ndeye Awa Dieye, Ndèye Niang, Giorgio Russolillo

Sensibilité des indices de qualité d'un classifieur probabiliste

Ndèye Awa DIEYE[a], Ndèye NIANG[a], Giorgio RUSSOLILLO [a]

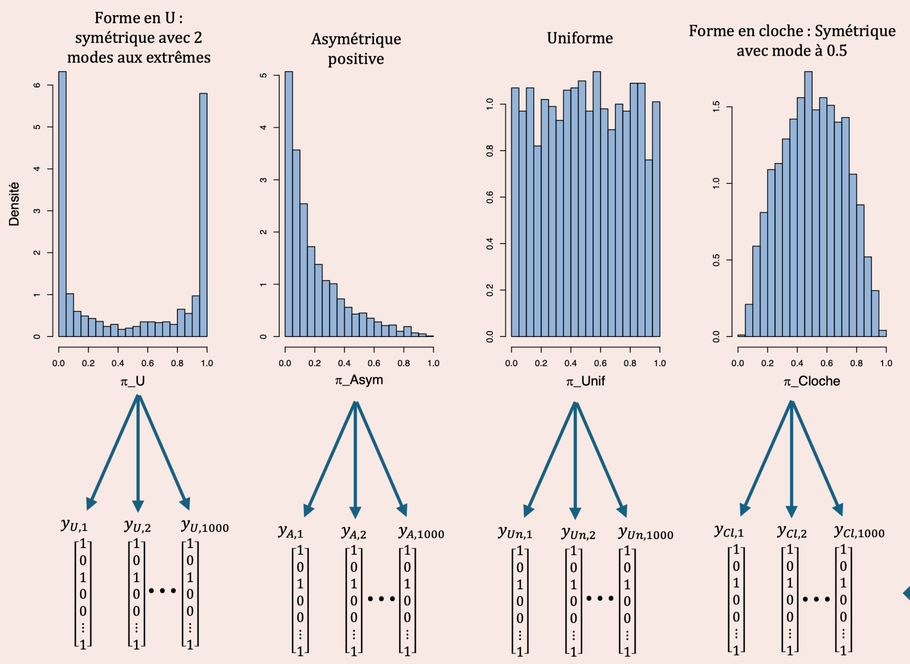
a. Laboratoire Cédric - MSDMA, CNAM, Paris, France

Contexte

- Classification supervisée binaire avec la variable de réponse $y \in \{0, 1\}$.
- Un classifieur probabiliste estime la probabilité π_i de l'évènement $y = 1$ pour l'observation i.
- **Objectif** : Étudier la sensibilité des indices de qualité d'un classifieur probabiliste à différents niveaux d'écarts à l'ajustement parfait selon la forme de distribution des probabilités π et de la distribution des écarts à ces probabilités.

Méthodologie

1. On simule 4 vecteurs de probabilités de taille 2000 issues du modèle logistique $\pi_i = \frac{e^{\beta x_i}}{1+e^{\beta x_i}}$ avec $\beta = 1$



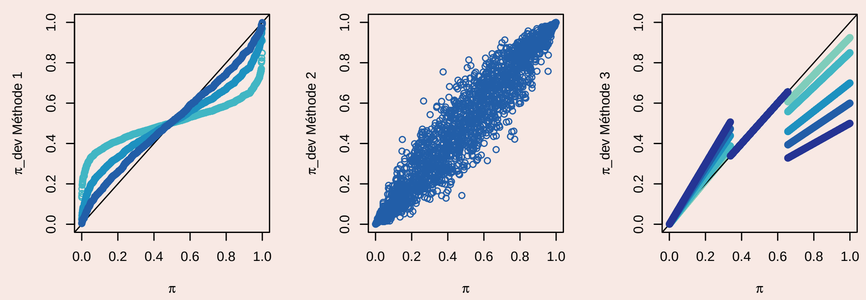
2. 1000 vecteurs y issus de lois de Bernoulli de paramètre π sont générés à partir de chaque vecteur de probabilités.

3. On dévie ces probabilités suivant 3 méthodes :

Méthode 1 : la valeur du coefficient β est remplacée par les valeurs 0.75, 0.5 et 0.25. Cette méthode donne des déviations **monotones, symétriques, plus accentuées sur les valeurs extrêmes**.

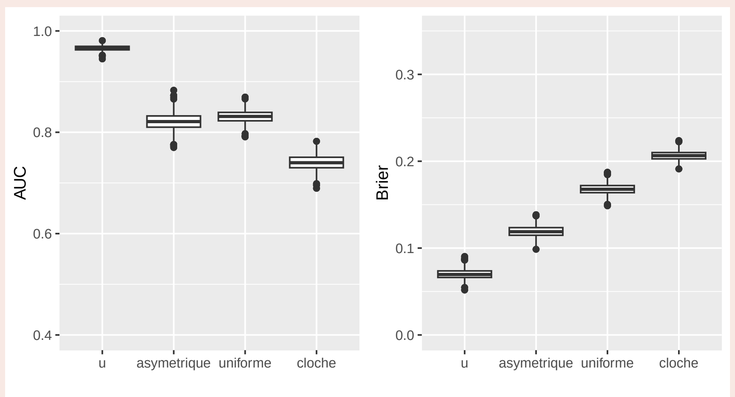
Méthode 2 : des bruits gaussiens de moyenne nulle et d'écart-type 0.2, 0.5, 1 et 2 sont ajoutés aux vecteurs π. Cette méthode donne des déviations **non monotones, symétriques, plus accentuées sur les valeurs centrales**.

Méthode 3 : le premier tercile de la distribution des π est multiplié par 1.075, 1.15, 1.3, 1.4, 1.5 et le dernier tercile par 0.925, 0.85, 0.7, 0.6, 0.5. Cette méthode donne des déviations **non monotones, asymétriques, plus accentuées sur les valeurs proches de 1**.

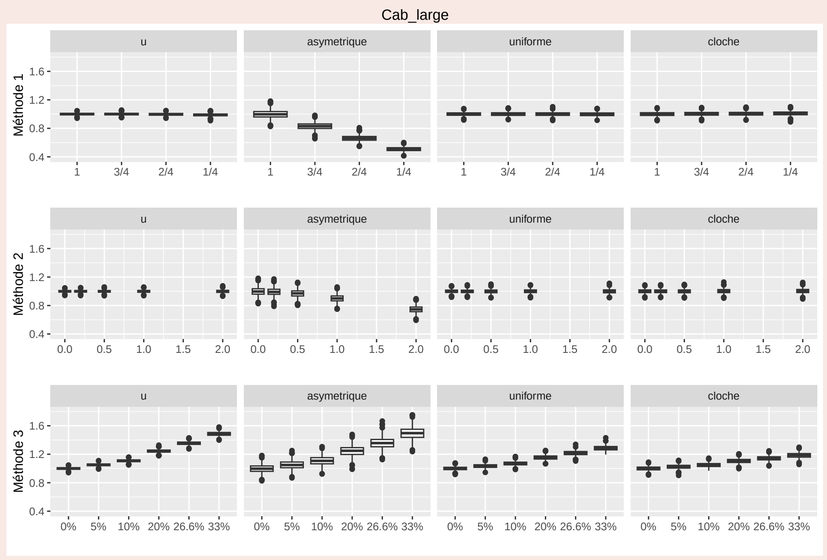


4. On calcule l'Aire sous la courbe ROC (AUC), le score de Brier, le rapport entre la probabilité estimée moyenne et le taux observé d'évènement (Calibration-in-the-large) et l'Expected Calibration Error (ECE) [1].

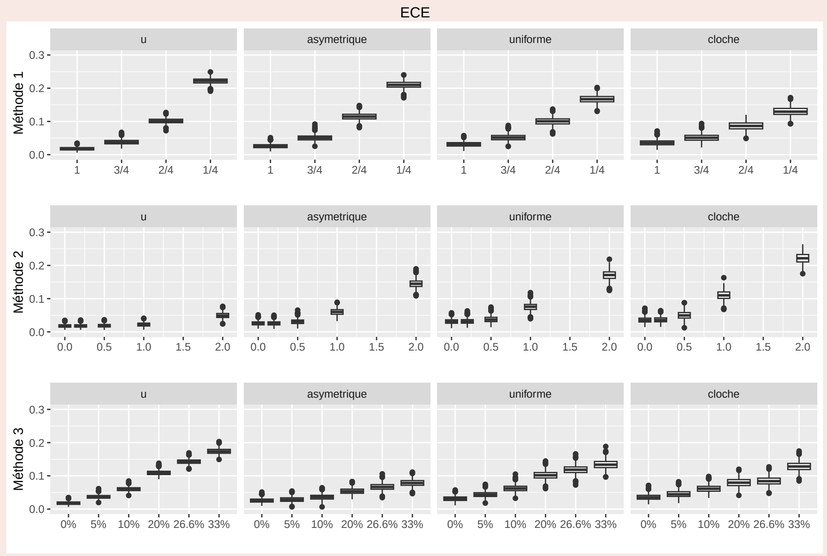
Résultats



- Les valeurs de l'AUC et du score de Brier, obtenues en comparant les réponses aux probabilités qui les ont générées, sont meilleures lorsque la distribution des probabilités présente plus de valeurs extrêmes.



- La mesure de la calibration-in-the-large est plus sensible lorsque on retrouve une asymétrie concernant soit la distribution des probabilités, soit l'ampleur des écarts.



- L'ECE se révèle plus sensible lorsque les plus larges écarts se situent au niveau des probabilités avec une densité plus élevée.

Conclusion

- Sensibilité des indices à la distribution de π, qui n'est pas connue dans la pratique.
- Prudence dans l'interprétation des indices des classifieurs probabilistes.

Références

[1] Y. Huang, W. Li, F. Macheret, R. A. Gabriel, and L. Ohno-Machado, "A tutorial on calibration measurements and calibration models for clinical prediction models," *Journal of the American Medical Informatics Association*, vol. 27, pp. 621–633, 2020. [Online]. Available: https://doi.org/10.1093/jamia/ocz228

mail : ndeye-awa.dieye@lecnam.net