Detralytics X Allianz Allianz Trade

ADAPTER SA TARIFICATION :

Des arbres satisfaisant des propriétés d'équité populaires en assurance

Julien Trufin, Thomas Hames, Sophie Navarro, Harrison Verelst

Introduction

- L’assurance repose sur ***l’évaluation du risque*** à partir de variables observables.

- Cela crée des ***groupes d’assurés***.

- Mais certaines variables de segmentation sont corrélées à des ***caractéristiques protégées***.

Équité actuarielle

- ***Principe*** : des individus présentant le même niveau de risque payent la même prime.

- ***Problème*** : certaines variables sont corrélées à des critères discriminatoires.

- => Nécessité de ***redéfinir la notion d'équité*** en présence de données sensibles.

Limite de l’exclusion des variables sensibles

- Ignorer une variable comme le genre ne garantit pas l’équité.

- Des corrélations fortes avec d’autres variables permettent de le retrouver indirectement.

Notions d’équité

- ***Équité individuelle*** : des individus avec des variables non-discriminantes similaires sont traités de manière similaire.

- **Équité de groupe** : traitement équitable entre groupes démographiques.

- Compromis entre équité et équité actuarielle.

Trois approches pour corriger l’inéquité

- ***Pré-traitement*** (pre-processing) : transformer les variables pour les rendre indépendantes des attributs sensibles.

- ***En cours de traitement*** (in-processing) : modifier l’algorithme d’apprentissage.

- ***Post-traitement*** (post-processing) : ajuster les prédictions finales.

Équité de groupe

- Le modèle utilise des variables explicatives Z = (X,S) où S est la variable sensible.

- ***Fairness through unawareness*** (FTU) : ignorer S n'assure pas l'équité car S peut être corrélée à X.

Trois principes de l'équité de groupe

- 1**. *Indépendance*** : $\hat{Y} \perp S$ (ex: Parité démographique)
- 2. ***Séparation*** : $\hat{Y} \perp S \mid Y$ (ex: Égalité des chances)
- 3. ***Suffisance*** : $Y \perp S \mid \hat{Y}$ (ex: Test de calibration)

Parité démographique – En cours de traitement (in processing)

- ***Approche*** : construire des arbres de régression garantissant la parité démographique par rapport à une variable S protégée (in-processing).

- ***Typiquement*** : à chaque split d'un arbre de régression, la variable sélectionnée (ainsi que sa valeur) est celle qui minimise la déviance du nœud en question.

Déviance Poisson : $$D(y,\hat{y}) = 2\sum_{i=1}^{n}\left[y_i \log\left(\frac{y_i}{\hat{y}_i}\right) - (y_i - \hat{y}_i)\right]$$

- ***Idée*** : ne considérer que les splits pour lesquels la variable sensible conserve la même distribution que celle observée au niveau du portefeuille.*

*Référence: Simon, Denuit and Trufin (2025), DPTree and DPForest: Tree-based methods fulfilling demographic parity, submitted for publication.

Application sur un cas théorique – Description des cas générés

Nous considérons un exemple issu du domaine de l'assurance automobile.
Quatre variables explicatives $X = (X_1, X_2, X_3, X_4)$ sont supposées disponibles pour chaque souscripteur, à savoir :

- X_1 : **sexe** du souscripteur (valeurs possibles : *femme* ou *homme*) ;
- X_2 : **âge** du souscripteur (valeurs entières comprises entre **18** et **65** ans inclus) ;
- X_3 : indicateur de **paiement fractionné** de la prime annuelle (*oui* ou *non*) ;
- X_4 : indicateur précisant si le véhicule est une **voiture de sport** (*oui* ou *non*).

Les variables X_1, X_2, X_3, et X_4 sont supposées suivre les lois de probabilité suivantes :

- $P[X_1 = \text{femme}] = P[X_1 = \text{homme}] = 0{,}5$;
- $P[X_2 = k] = \frac{1}{48}$, pour $k \in \{18, 19, \ldots, 65\}$ (*loi uniforme*) ;
- $P[X_3 = \text{oui}] = P[X_3 = \text{non}] = 0{,}5$;
- $P[X_4 = \text{oui}] = P[X_4 = \text{non}] = 0{,}5$.

Par ailleurs, les variables explicatives sont supposées **indépendantes**, à l'exception de X_1 (sexe du souscripteur) et X_4 (voiture de sport), entre lesquelles une dépendance contrôlée par un paramètre $\rho \in [0{,}5; 1]$ est introduite. On a alors :

$$P[X_1 = \text{homme} \mid X_4 = \text{oui}] = \rho, \quad P[X_1 = \text{homme} \mid X_4 = \text{non}] = 1 - \rho.$$

Le paramètre ρ permet ainsi de moduler le niveau de dépendance entre le sexe du souscripteur et la possession d'un véhicule de sport.

Application sur un cas théorique – Description de la Target

Y représente le nombre de sinistres enregistrés par le souscripteur.

Conditionnellement aux variables explicatives $X = x$, la variable Y suit une loi de Poisson de paramètre $\mu(x)$, c'est-à-dire :

$$Y \mid X = x \sim \text{Poisson}(\mu(x))$$

La fonction $\mu(x)$, qui représente la fréquence attendue des sinistres, est définie par :

$$\mu(x) = 0{,}1 \times \left(1 + \frac{1}{\sqrt{x_2 - 17}}\right) \times (1 + 0{,}15 \times I[x_4 = \text{oui}])$$

où :

- x_2 est l'âge du souscripteur,
- $I[\cdot]$ est la fonction indicatrice qui vaut 1 si la condition est vraie, 0 sinon,
- x_4 indique si le véhicule est une voiture de sport.

On remarque que $\mu(x)$ ne dépend pas du sexe du souscripteur, mais varie en fonction de son âge et de la possession d'une voiture de sport.

Application sur un cas théorique – Description des cas

L'étude théorique va être menée en 2 cas principaux:

Cas 1 : cas pour lequel **le genre n'a pas d'effet direct sur le nombre de sinistre**

- Et nous ferons varier la corrélation entre le Genre et le fait que la voiture soit une voiture de sport pour induire un effet indirect du Genre sur le nombre de sinistre

Cas 2 : cas pour lequel **le genre a un effet direct sur le nombre de sinistre**

- Et nous ferons varier la corrélation entre le Genre et le fait que la voiture soit une voiture de sport pour induire **en plus** un effet indirect du Genre sur le nombre de sinistre

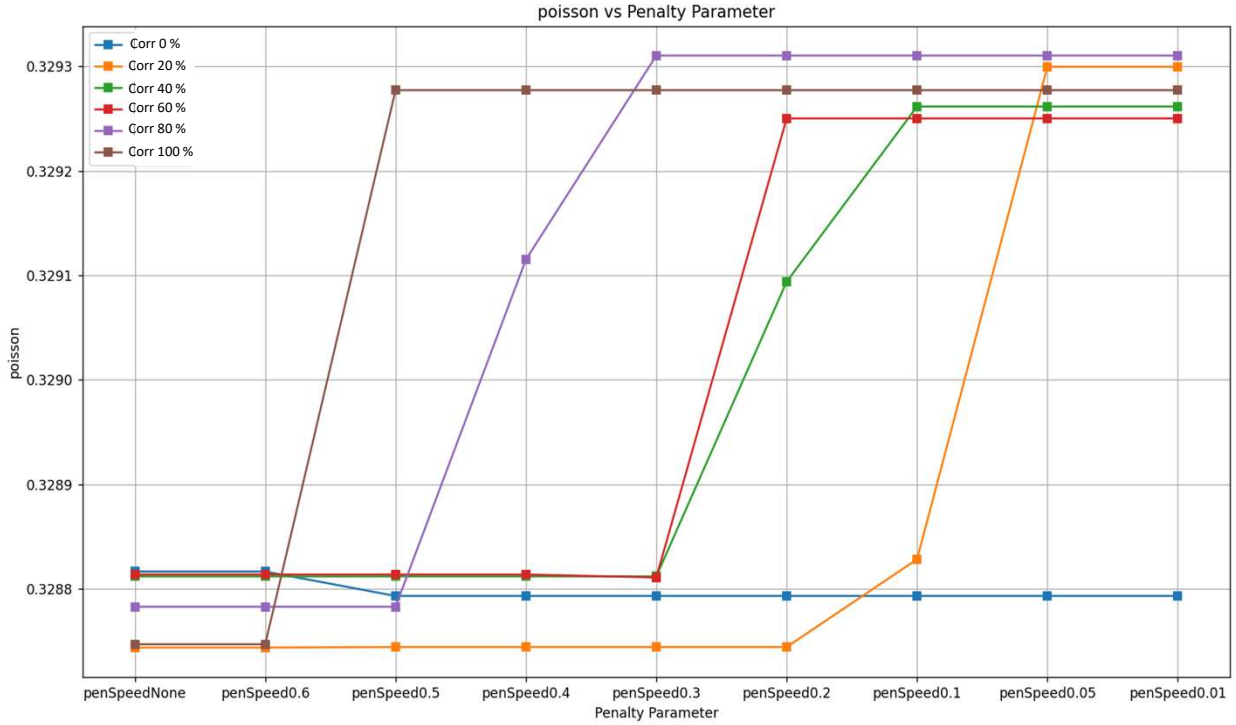
Application sur un cas théorique – Cas 1
Pas de lien direct Genre - Nombre de Sinistre (μ(x))

Résultat sur test set: Déviance de poisson

Comme il n'y a pas de lien direct entre mu(x) et le genre, les effets se font au travers de la corrélation entre le genre et le fait la voiture soit une voiture de sport ou pas.

Le fait qu'une voiture soit une voiture de sport a un impact sur mu(x).

Au plus la corrélation augmente entre le genre et le fait qu'une voiture soit une voiture de sport, au moins il est possible de permettre d'écart de proportion sans pénaliser la déviance.



Rho = 0,5 = pas de corrélation entre Genre et Sport
Rho = 0,75 = corrélation de 50% entre Genre et Sport
Rho = 1 = corrélation de 100% entre Genre et Sport

On ne protège pas la variable

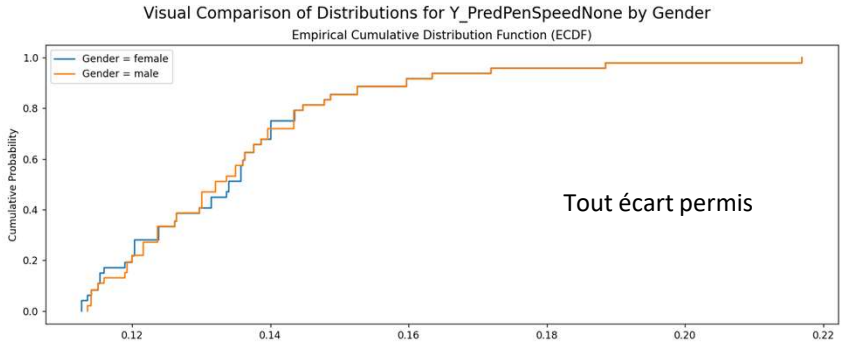
On permet 20% d'écart p/r à la prop. initiale

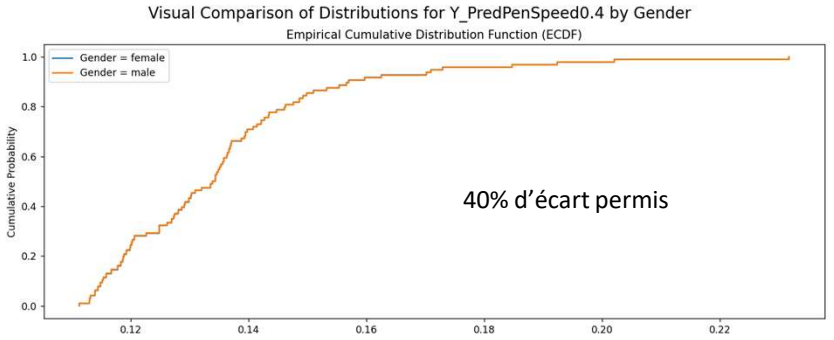
On permet 1% d'écart p/r à la prop. initiale

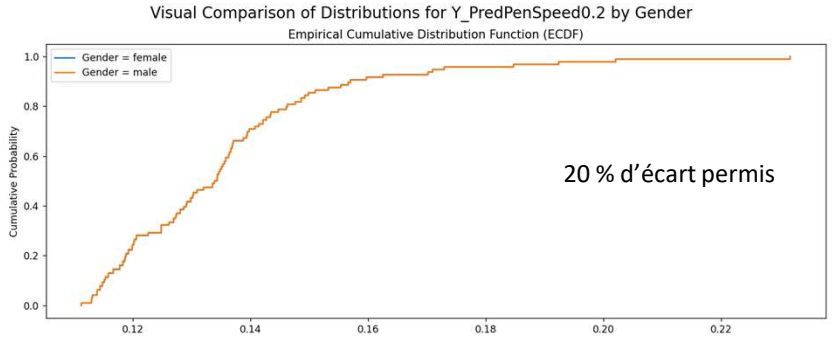
Application sur un cas théorique – Cas 1
Pas de lien direct Genre x Nombre de Sinistre (μ(x))

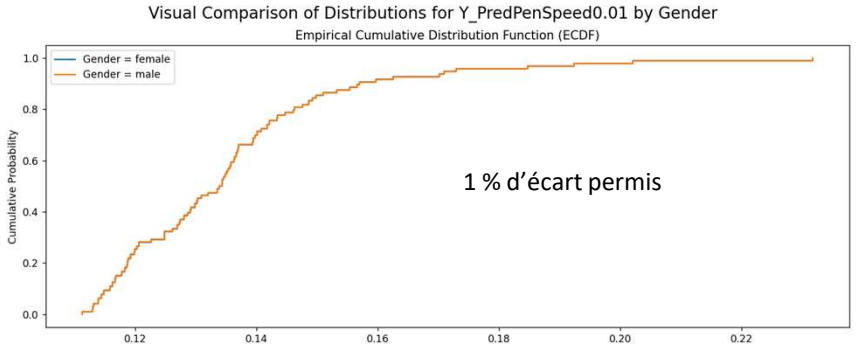
Pas de corrélation entre Genre x Sport

Comme il n'y a pas de lien direct ou indirect entre le genre et mu(x), pénaliser ou pas l'écart permis ne créé par de différence de distribution de prédiction en fonction du genre. L'algorithme n'a pas d'intérêt à aller chercher des effets en fonction du genre puisqu'il n'y a pas de lien entre le genre et mu(x).





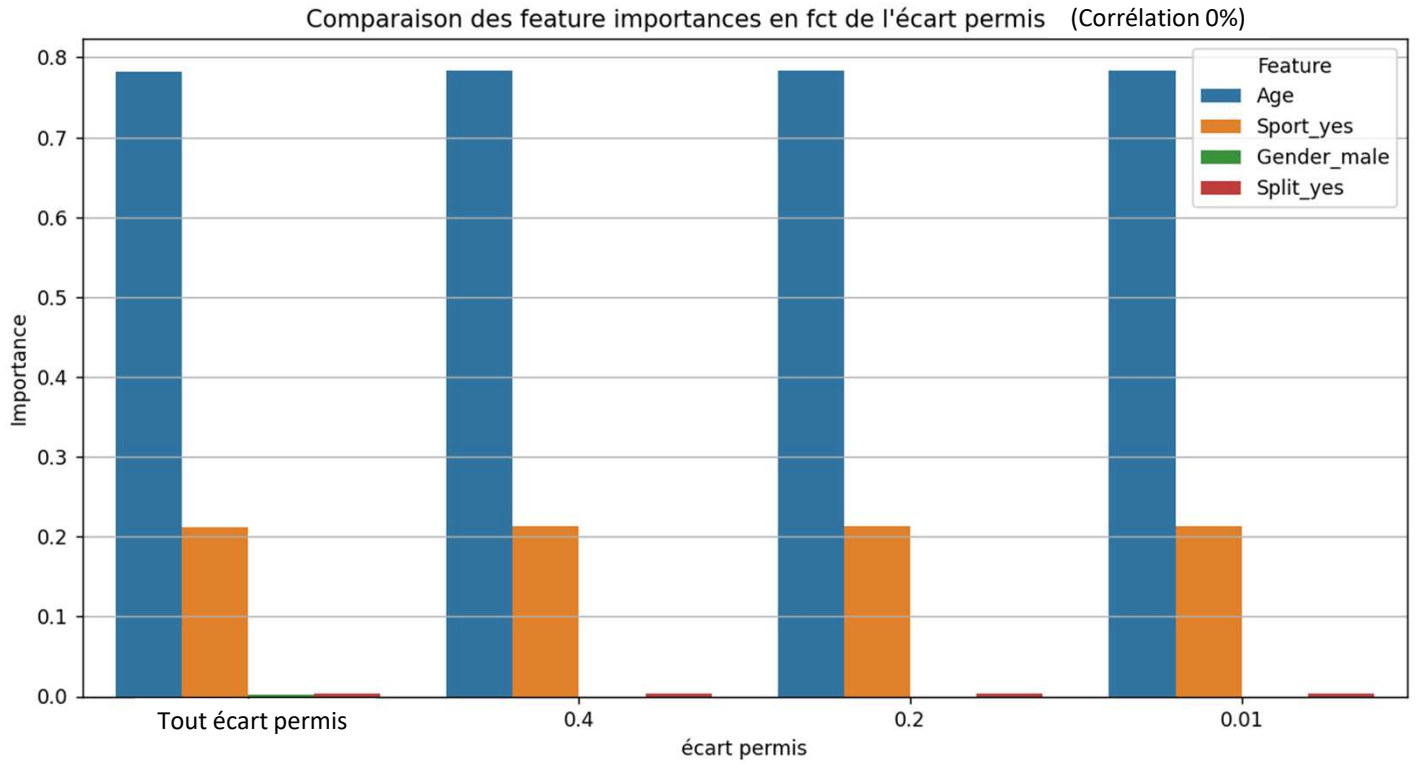




Application sur un cas théorique – Cas 1
Pas de lien direct Genre x Nombre de Sinistre (μ(x))

Pas de corrélation entre Genre x Sport

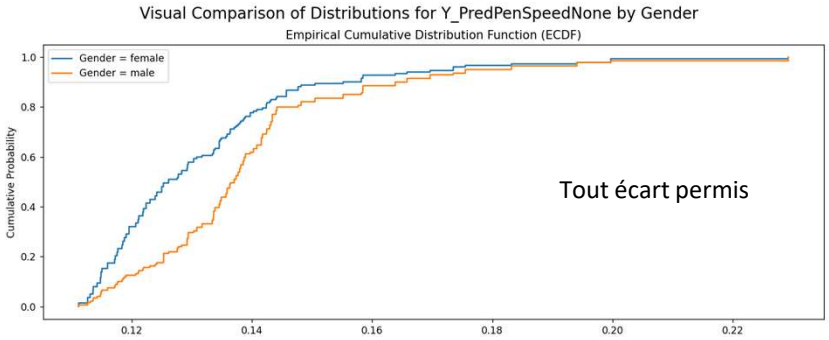
Les features importances ne bougent pas puisqu'il n'y a pas de réel lien entre le genre et le nombre de sinistre
Et puisqu'il n'y a pas de corrélation entre le genre et le sport

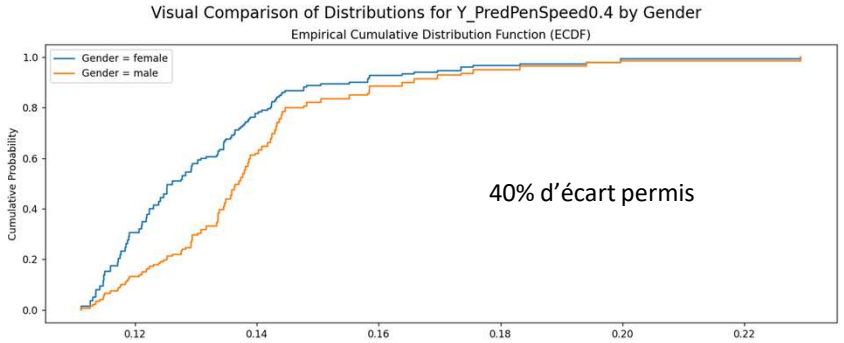


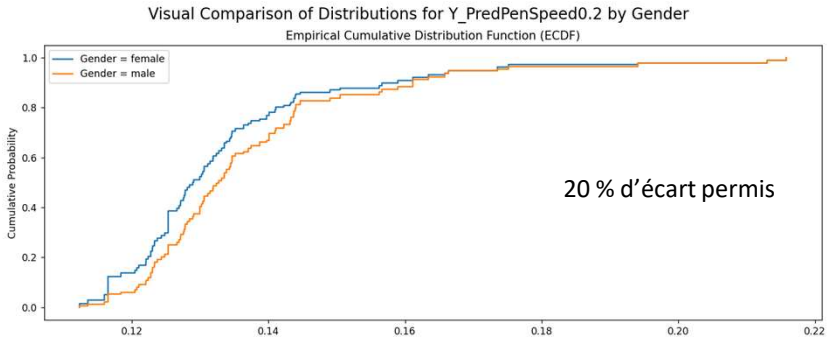
Application sur un cas théorique – Cas 1
Pas de lien direct Genre x Nombre de Sinistre (μ(x))

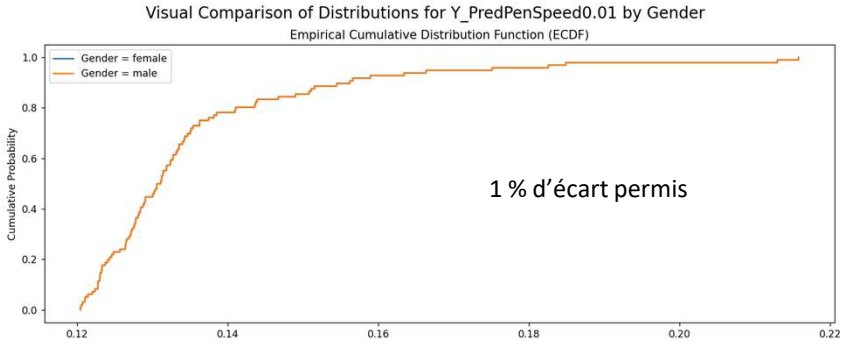
Corrélation Genre x Sport = 40 %

Pas de lien direct entre le genre et mu(x) mais il y a un lien indirect sur mu(x) via corrélation entre le Genre et le Sport, pénaliser ou pas l'écart permis créé une réelle différence de distribution de prédiction en fonction du genre.
L'écart entre les deux distribution dépend aussi de l'écart permis – au moins on permet d'écart, au plus les distributions se rapprochent.





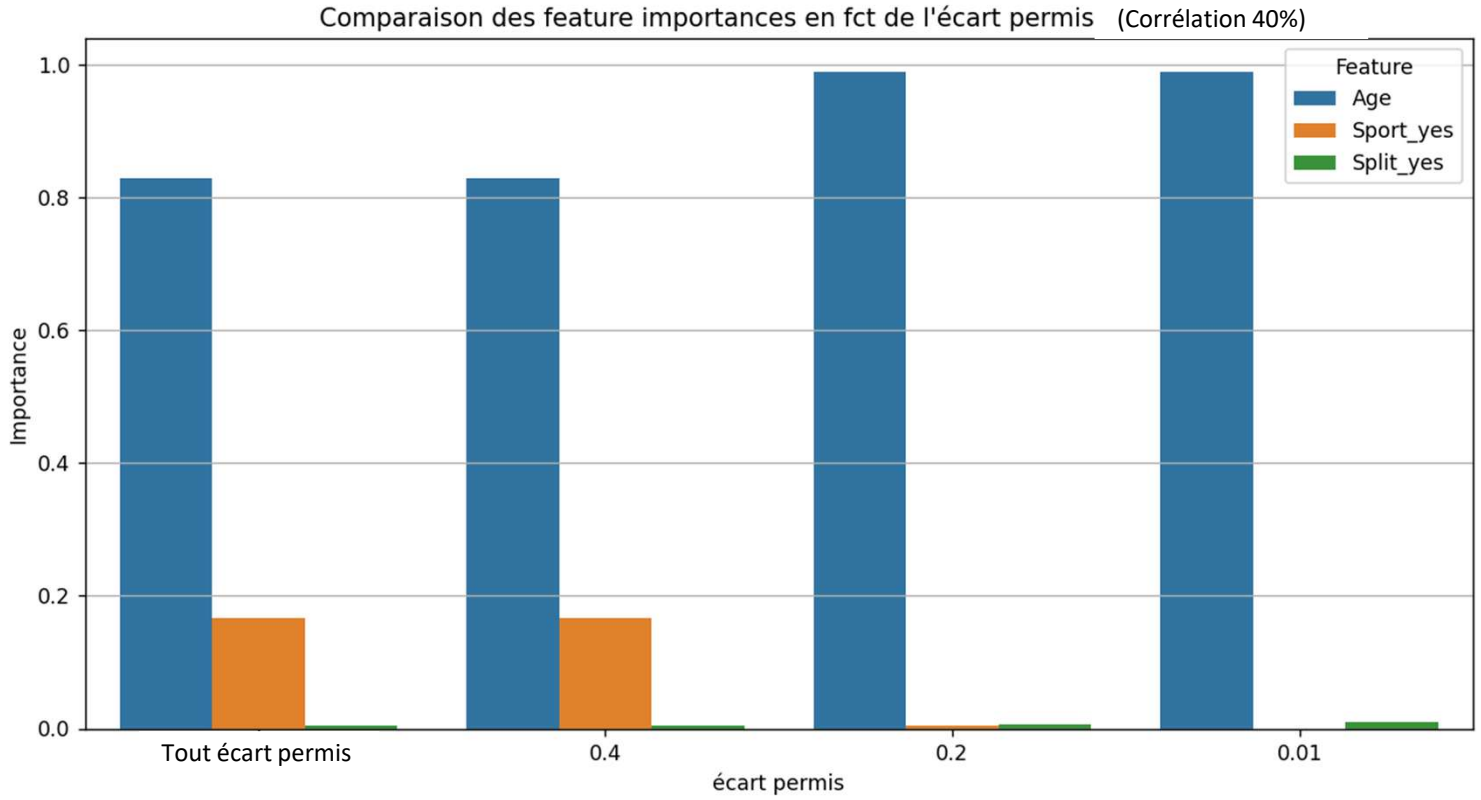




Application sur un cas théorique – Cas 1
Pas de lien direct Genre x Nombre de Sinistre (μ(x))

Corrélation Genre x Sport = 40 %

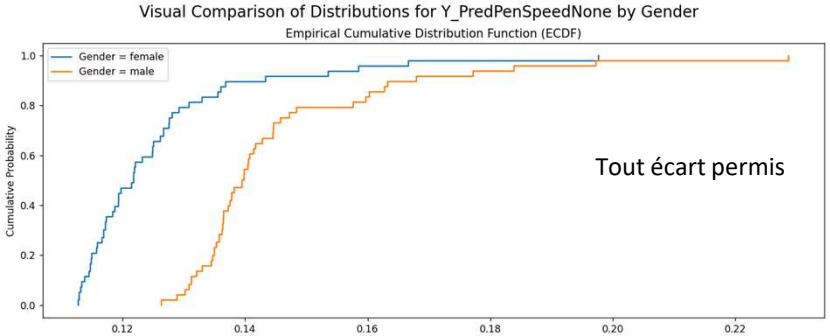
Les features importances bougent en fonction de l'écart que l'on permet par rapport à la proportion de genre.
Au début, on accepte encore la variable sport, mais plus la contrainte augmente, au plus celle-ci est rejetée de par sa corrélation avec le genre

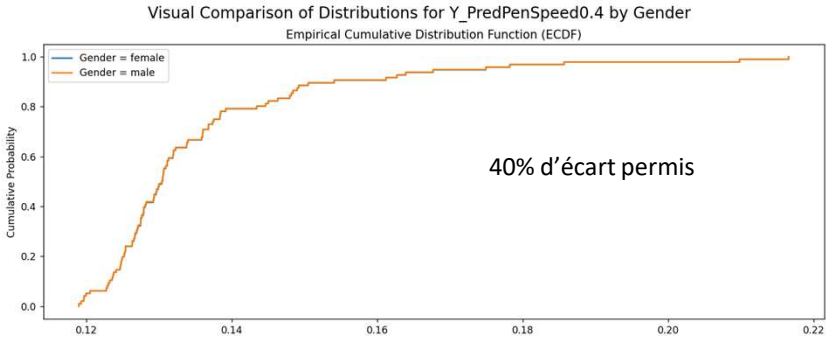


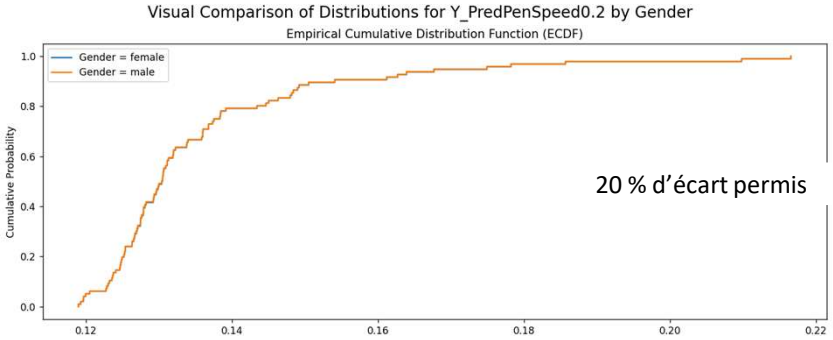
Application sur un cas théorique – Cas 1
Pas de lien direct Genre x Nombre de Sinistre (μ(x))

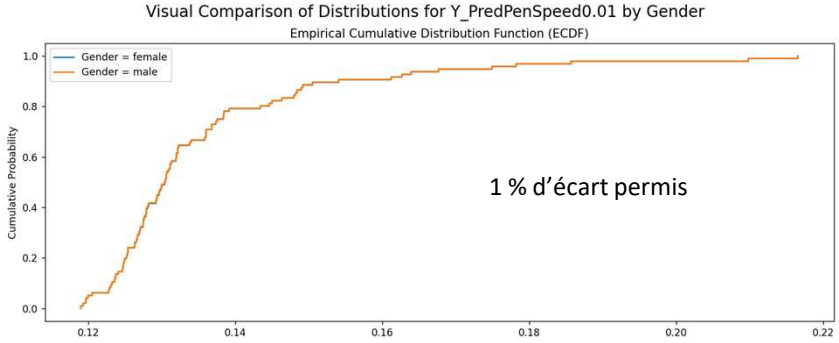
Corrélation Genre x Sport = 100 %

Pas de lien direct entre le genre et mu(x) mais il y a un lien indirect sur mu(x) via corrélation entre le Genre et le Sport.
Même conclusion que pour le cas précédent, sauf que cette fois-ci, le lien entre le Sport et le Genre est tellement important, qu'à la moindre pénalisation appliquée sur des écart de proportions de genre, les distributions de prédictions par genre deviennent identique – le modèle ne peut plus non plus utiliser la variable Sport.





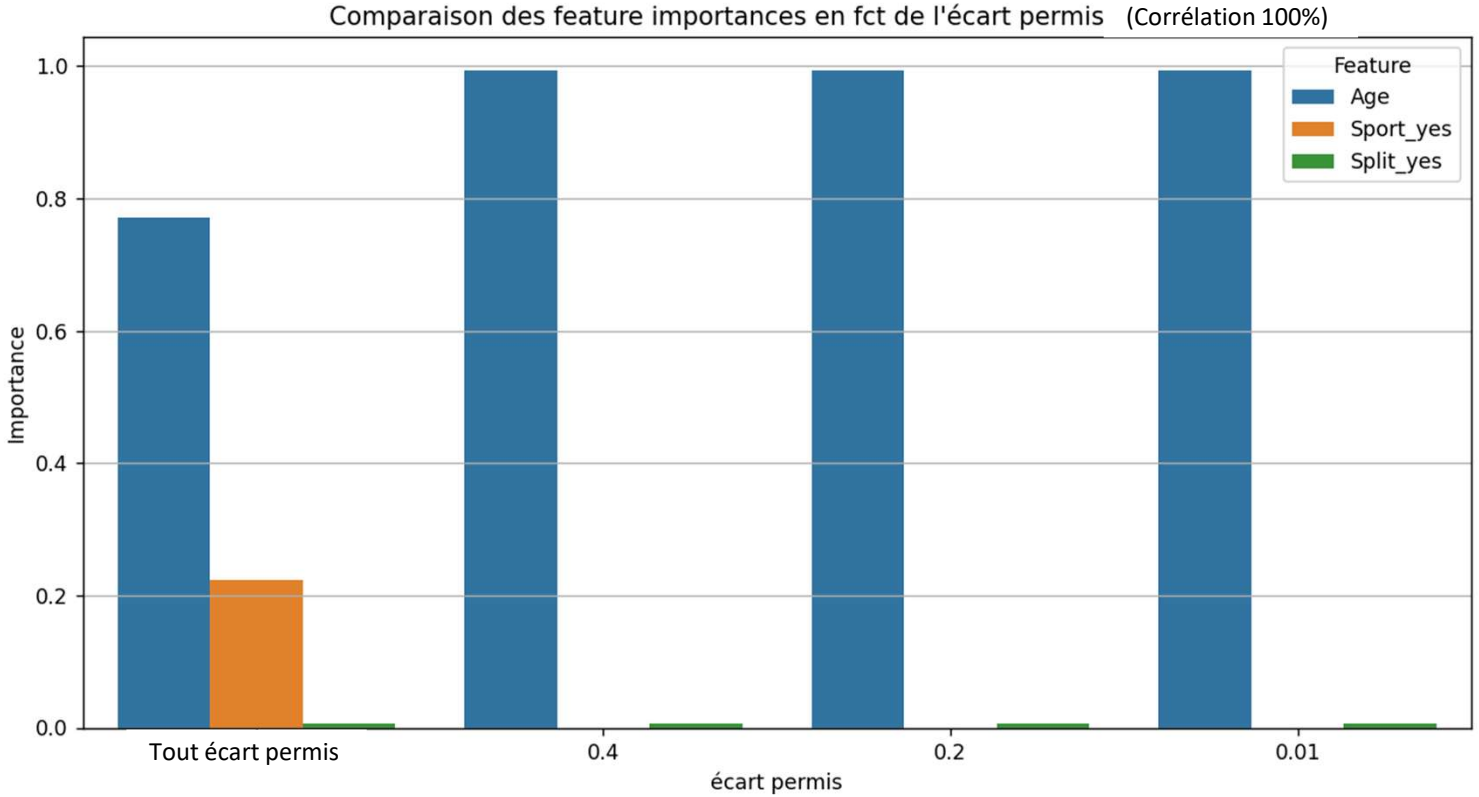




Application sur un cas théorique – Cas 1
Pas de lien direct Genre x Nombre de Sinistre (μ(x))

Corrélation Genre x Sport = 100 %

Comme la corrélation entre le genre et le sport est de 100% dans ce cas-ci, dès que nous pénalisons les moindres écarts par rapport aux déviations de proportion pour le genre, la variable sport est elle aussi rejetée des variables utilisées pour les splits.



Application sur un cas théorique – Cas 2
Lien direct Genre x Nombre de Sinistre (μ(x))

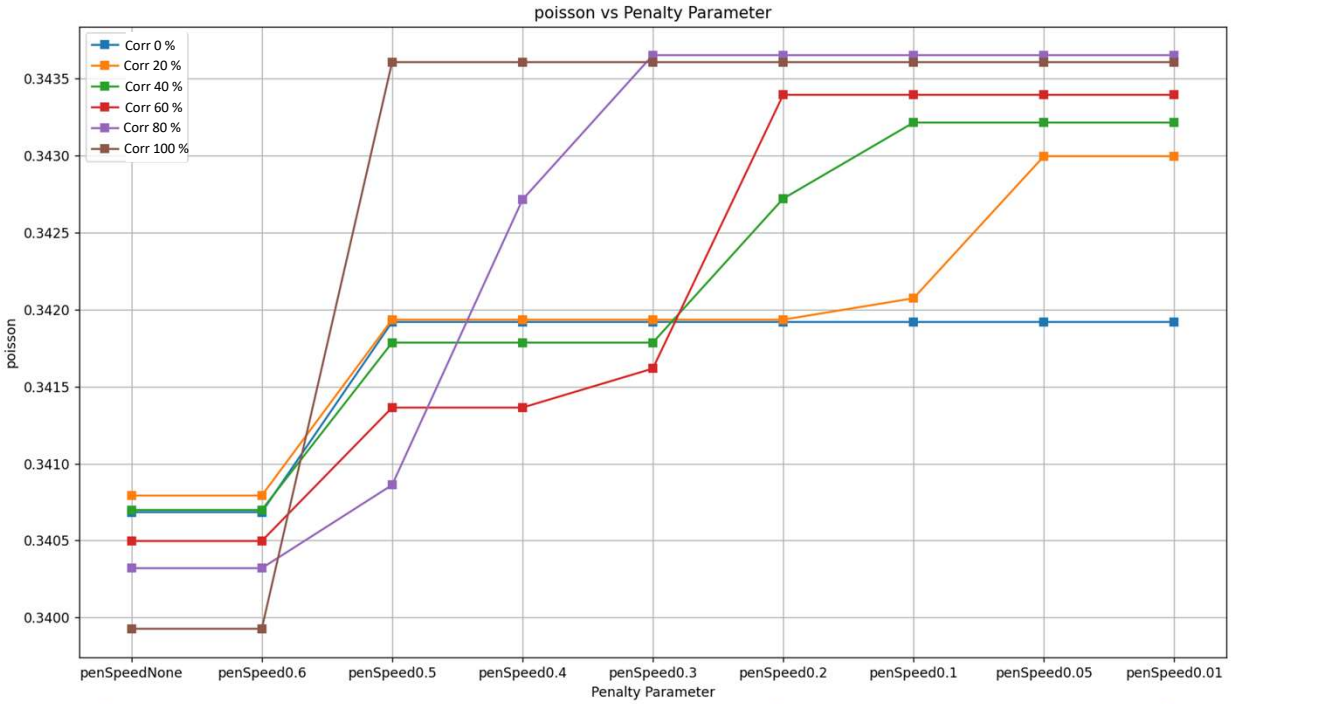
Comme il y a à présent un lien direct entre mu(x) et le genre, les effets ne se font plus seulement au travers de la corrélation entre le genre et le fait que la voiture soit une voiture de sport ou pas.
Dès que nous introduisons une pénalisation sur le genre, le score est négativement impacté.

De plus, au plus la corrélation augmente entre le genre et le fait qu'une voiture soit une voiture de sport, au plus la déviance est pénalisée tôt sur base de l'écart de proportion.

Le premier jump de déviance est lié à la perte de performance liée à la non utilisation du genre, la seconde perte de performance, plus graduelle, est liée à la non utilisation de la variable sport par effet de corrélation.

Rho = 0,5 = pas de corrélation entre Genre et Sport
Rho = 0,75 = corrélation de 50% entre Genre et Sport
Rho = 1 = corrélation de 100% entre Genre et Sport

Résultat sur test set: Déviance de poisson en fonction de l'écart permis



On ne pénalise pas

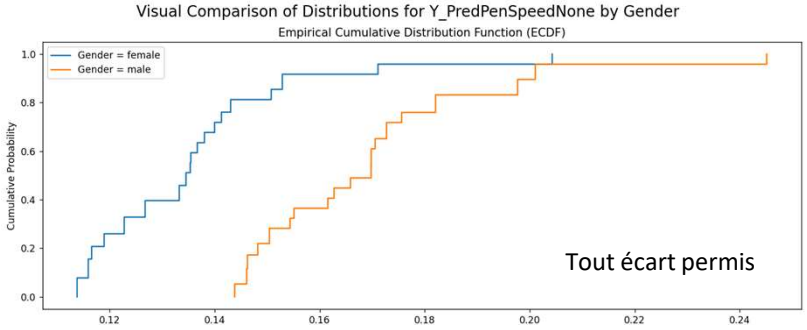
On permet 20% d'écart p/r à la prop. initiale

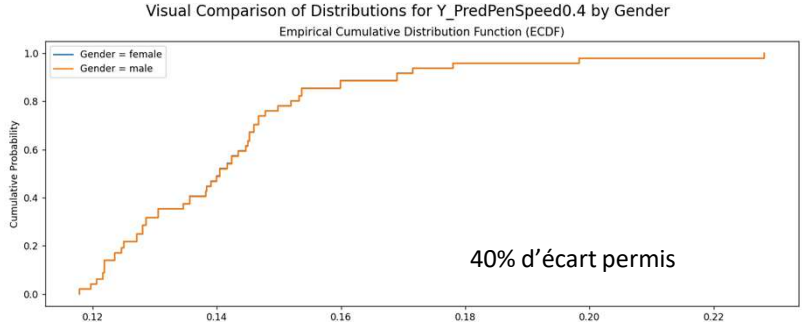
On permet 1% d'écart p/r à la prop. initiale

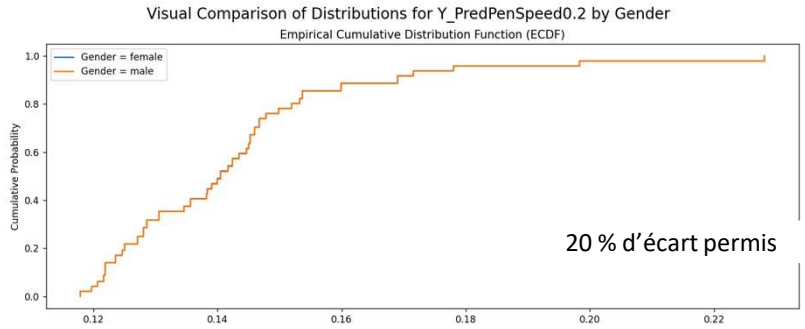
Application sur un cas théorique – Cas 2
Lien direct Genre x Nombre de Sinistre (μ(x))

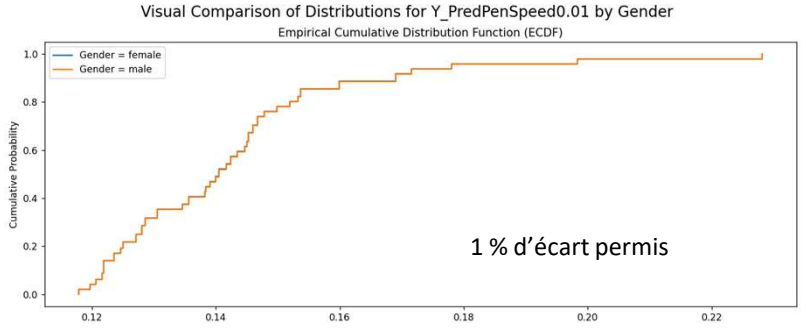
<u>Pas de corrélation entre Genre x Sport</u>

Comme il y a à présent un lien direct entre le genre et mu(x), pénaliser ou pas l'écart permis ramène à des distribution de prédiction équivalente par rapport au genre.





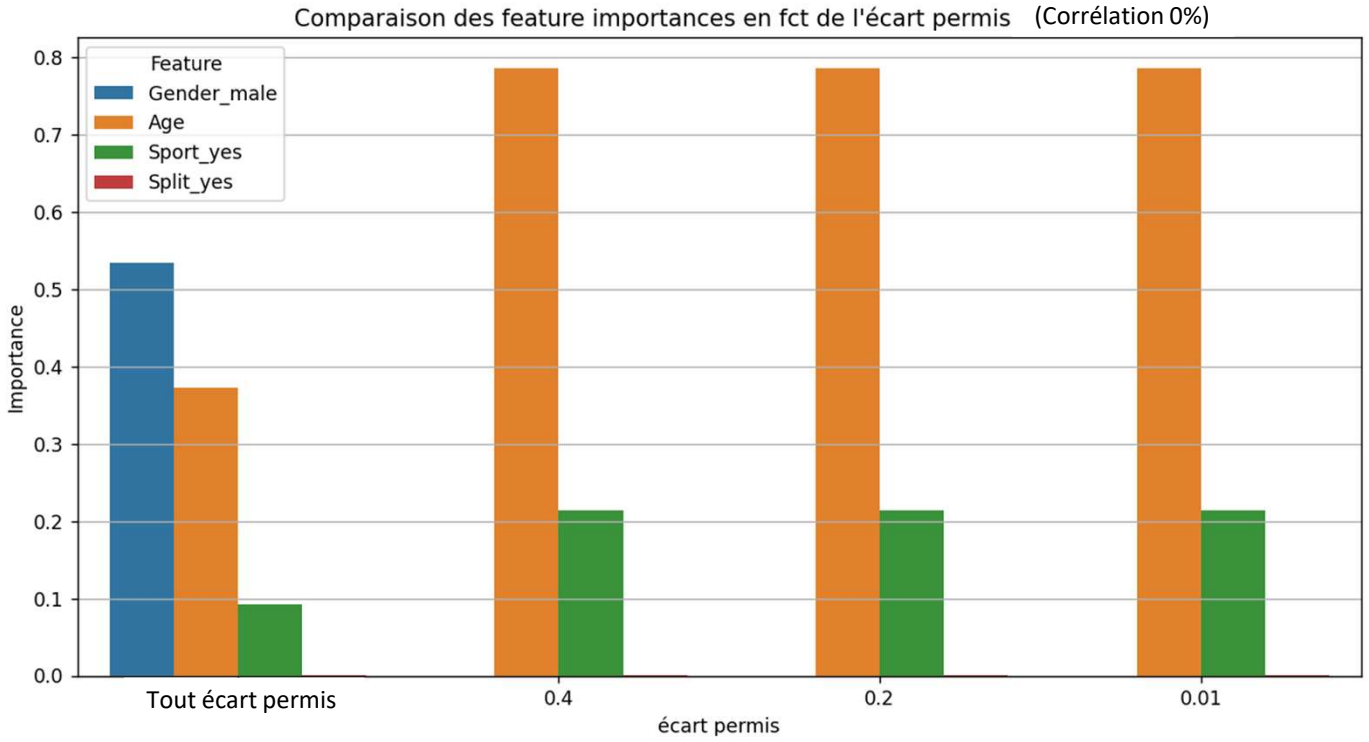




Application sur un cas théorique – Cas 2
Lien direct Genre x Nombre de Sinistre (μ(x))

Pas de corrélation entre Genre x Sport

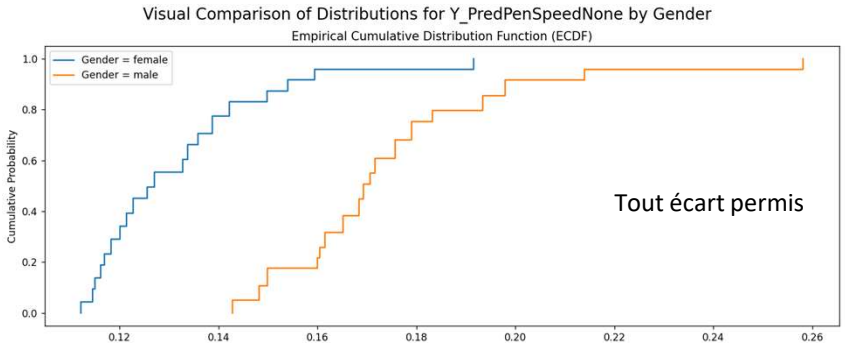
Par rapport aux features importances, on peut voir que pénaliser ou pas le genre ne change pas l'importances des autres variables étant donné qu'elles ne sont pas corrélée avec le genre. Dès que l'on pénalise sur base du genre, celui-ci n'est plus utilisé.

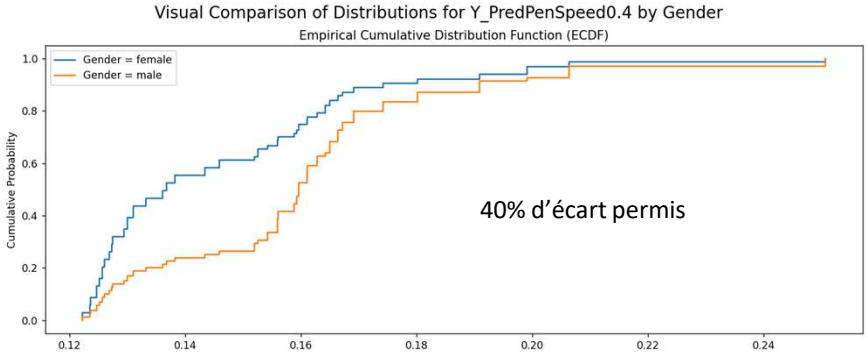


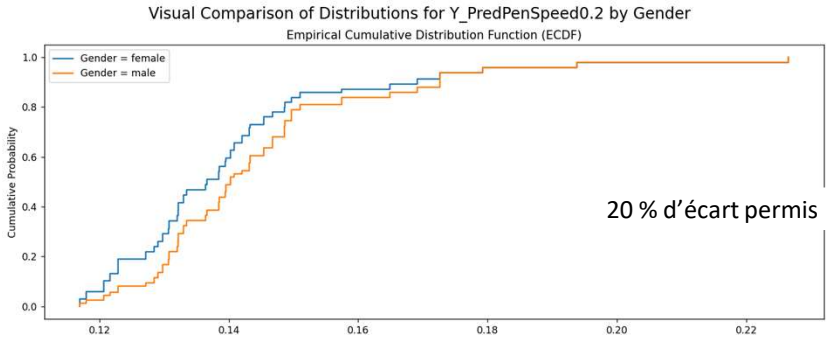
Application sur un cas théorique – Cas 2
Lien direct Genre x Nombre de Sinistre (μ(x))

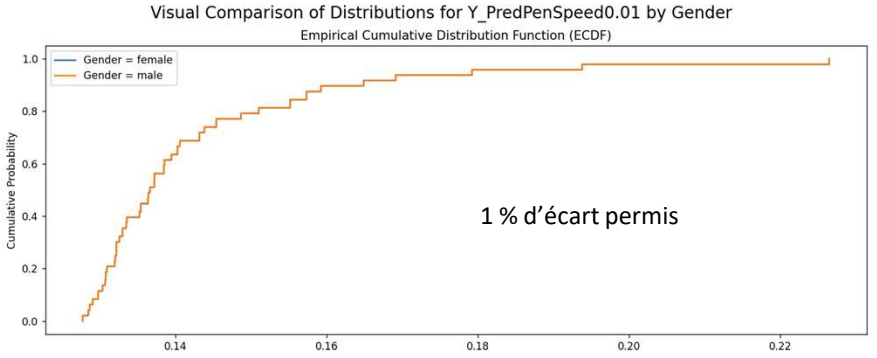
Corrélation Genre x Sport = 40 %

Lien direct entre le genre et mu(x) + lien indirect sur mu(x) via corrélation entre le Genre et le Sport, pénaliser ou pas l'écart permis créé une réelle différence de distribution de prédiction en fonction du genre.
L'écart entre les deux distribution dépend aussi de l'écart permis – au moins on permet d'écart, au plus les distributions se rapprochent.





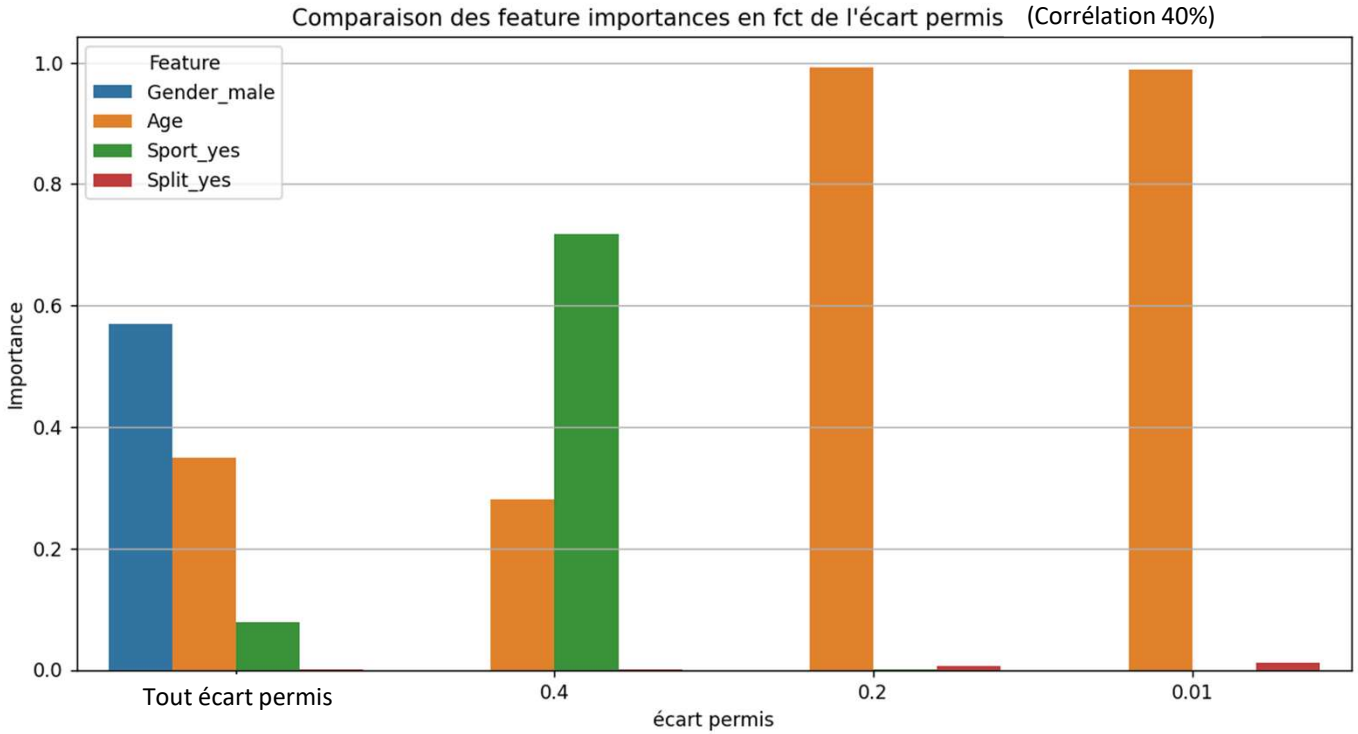




Application sur un cas théorique – Cas 2
Lien direct Genre x Nombre de Sinistre (μ(x))

Corrélation Genre x Sport = 40 %

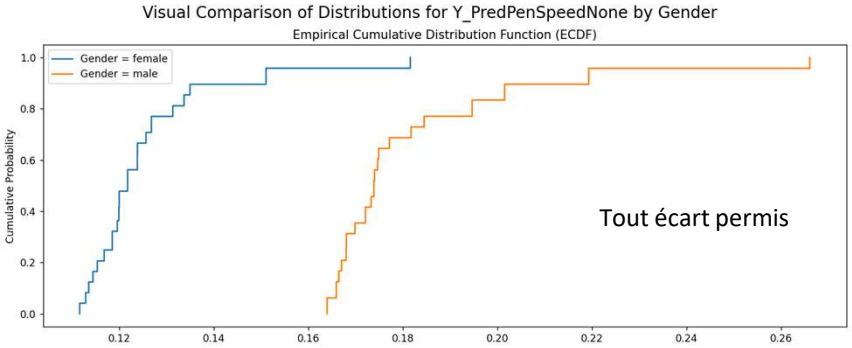
Les features importances bougent en fonction de l'écart que l'on permet par rapport à la proportion de genre.
Au début, on accepte encore la variable sport, mais plus la contrainte augmente, au plus celle-ci est rejetée de par sa corrélation avec le genre

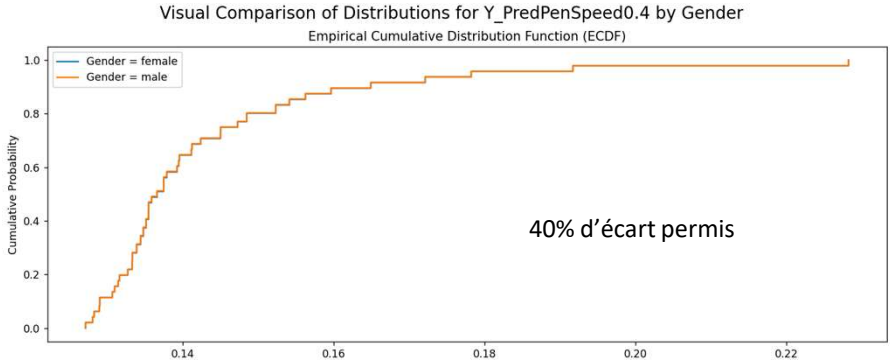


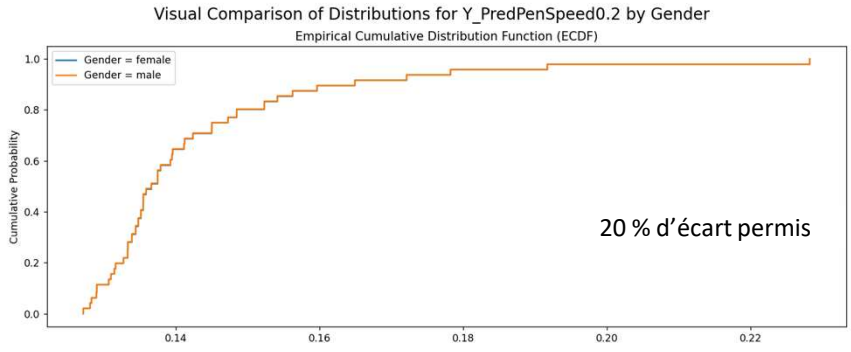
Application sur un cas théorique – Cas 2
Lien direct Genre x Nombre de Sinistre (μ(x))

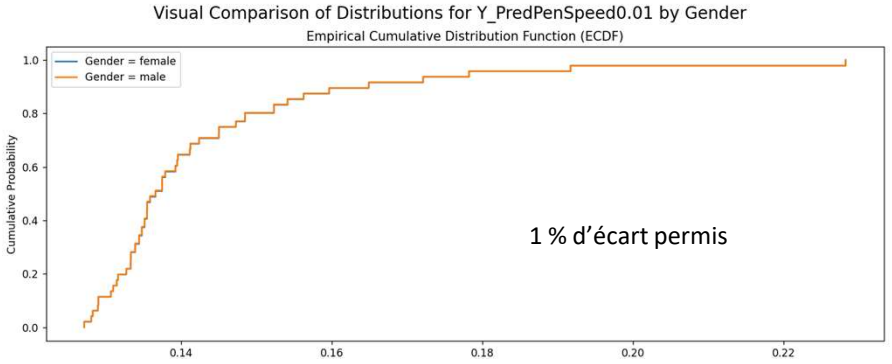
<u>Corrélation Genre x Sport = 100 %</u>

Lien direct entre le genre et mu(x). Et comme le Genre est 100% corrélé avec le Sport, dès que nous introduisons une pénalisation sur le genre, la variable sport est directement rejetée également.





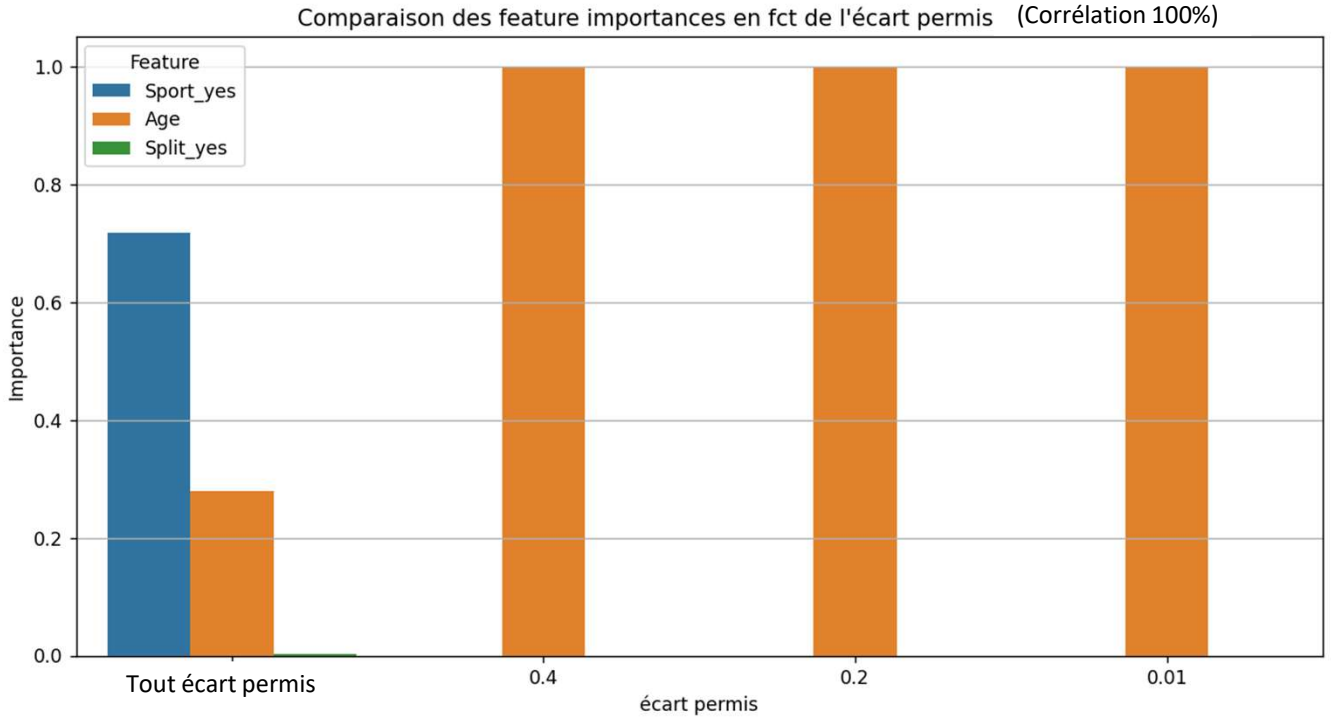




Application sur un cas théorique – Cas 2
Lien direct Genre x Nombre de Sinistre (μ(x))

Rho Genre x Sport = 100 %

Comme la corrélation entre le genre et le sport est de 100% dans ce cas-ci, dès que nous pénalisons les moindres écarts par rapport aux déviations de proportion pour le genre, la variable sport est elle aussi rejetée des variables utilisées pour les splits.



Application sur un cas théorique – Résumé Cas 1 & Cas 2

Impact sur les performances du modèle de protéger la variable Genre

Corrélation **Lien Genre – μ(x)**	***Sans*** Corrélation entre Genre et Voiture de Sport	***Avec*** Corrélation entre Genre et Voiture de Sport
Lien Indirect (Cas 1)	**Pas d'impact** sur les performances. Le Genre n'a pas de lien avec le nombre de sinistre.	**Impact** sur les performances lié à la perte d'utilisation de la **variable « Voiture de Sport »** qui à un lien avec le nombre de sinistre
Lien Direct (Cas 2)	**Impact** sur les performances lié à la perte d'utilisation de la **variable « Genre »** qui a un lien avec le nombre de sinistre	**Impact** sur les performances lié à la perte d'utilisation de la **variable « Voiture de Sport » ET de la variable « Genre »** qui ont toutes deux des liens avec le nombre de sinistre

RESULTATS CAS PRATIQUE Dataset FREQMTPL de Casdataset

Application sur freMTPL dataset de Casdataset
Description de la base de donnée

Le dataset freqMTPL est un **jeu de données** très utilisé en **actuariat non-vie** pour l'enseignement et la recherche, en particulier pour la tarification des contrats d'assurance automobile.

Variables explicatives:

VehiclePower : Puissance du véhicule assuré (catégorie discrète)
VehicleAge : Age du véhicule (en années)
DriverAge : Age du conducteur principal (en années)
Density : Densité de population de la zone de résidence (habitants/km²)
Region : Région géographique du contrat (variable catégorielle)
VehBrand : Marque du véhicule assuré
VehGas : Type de carburant utilisé par le véhicule

Target:

ClaimNb : Nombre de sinistres déclarés pendant la période d'exposition

Exposition au risque:

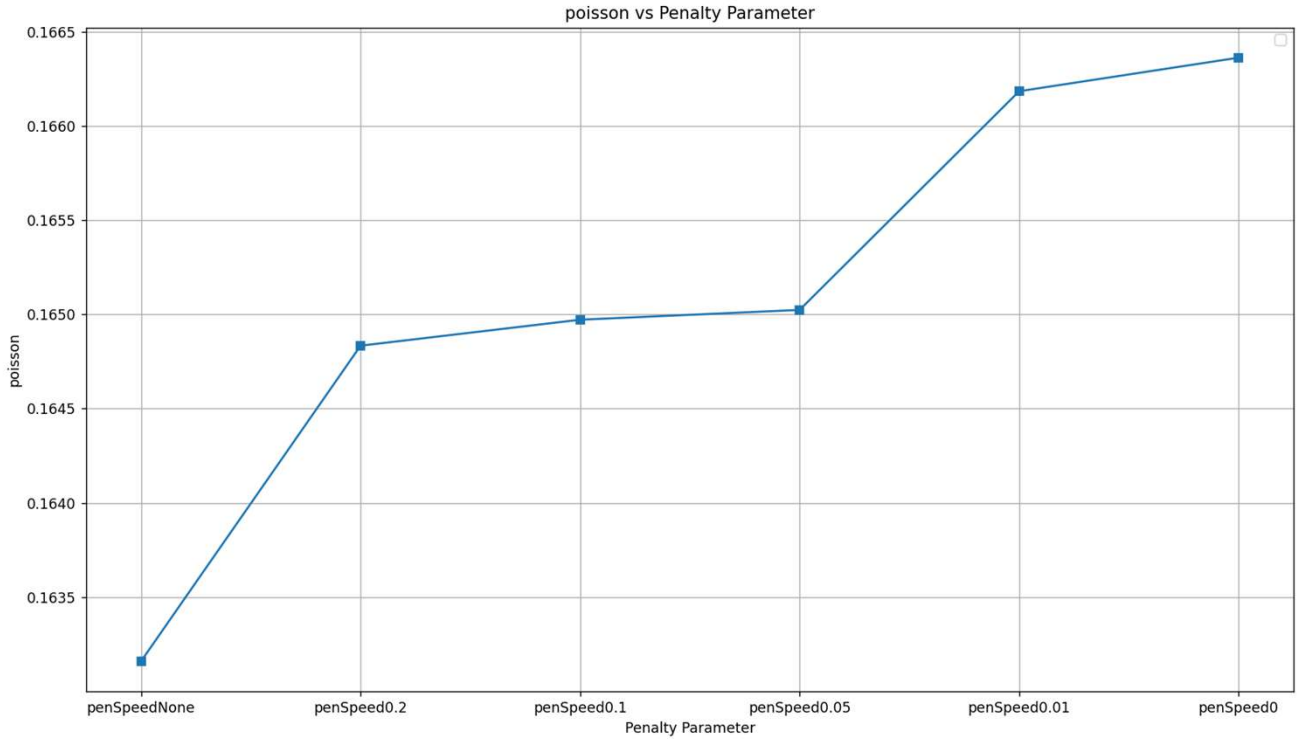
Exposure : Durée de couverture du contrat, en années (exemple : 0.5 pour 6 mois)

Variable que nous allons protéger:

L'âge du conducteur est décomposé sur base des quartiles et nous protégeons les **4 catégories d'âge**

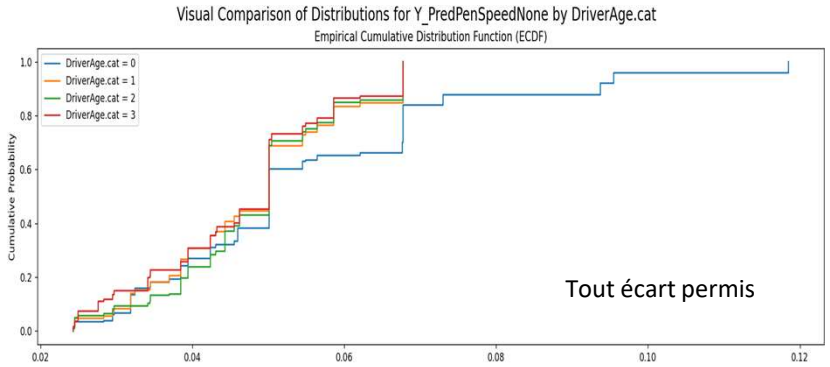
Application sur freMTPL dataset de Casdataset
Modéliser le nombre de Sinistre (μ(x))

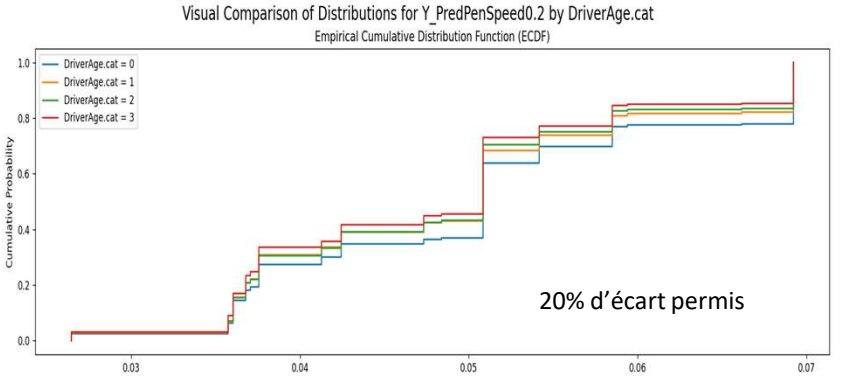
Modèles cross-validé 10 folds – Fine-tune Max-depth et nombre d'obs dans la feuille
Déviance en fonction de l'écart permis sur le **test set**

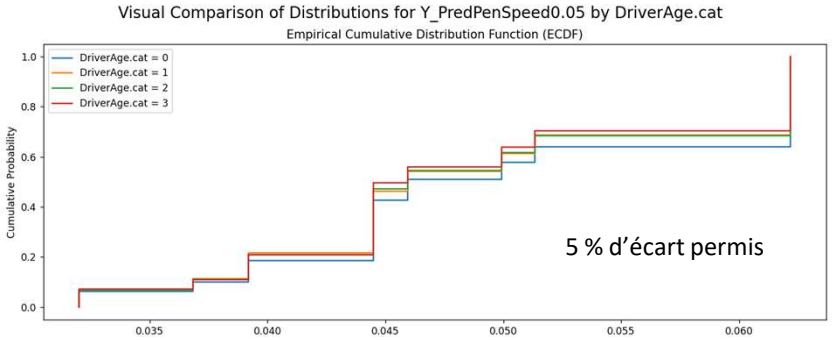


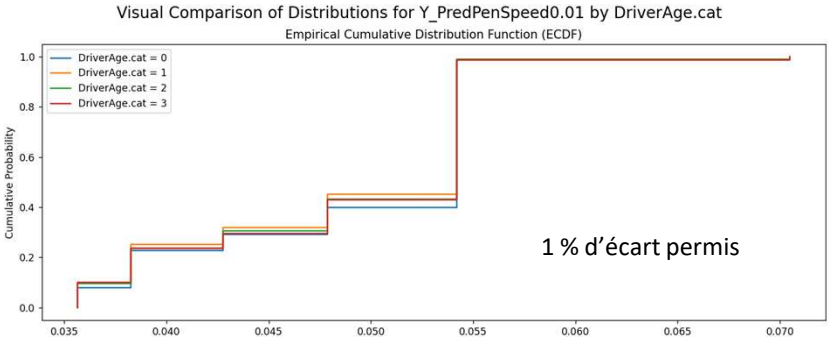
Application sur freMTPL dataset de Casdataset
Modéliser le nombre de Sinistre (μ(x))

Distribution de prédictions par catégorie d'age



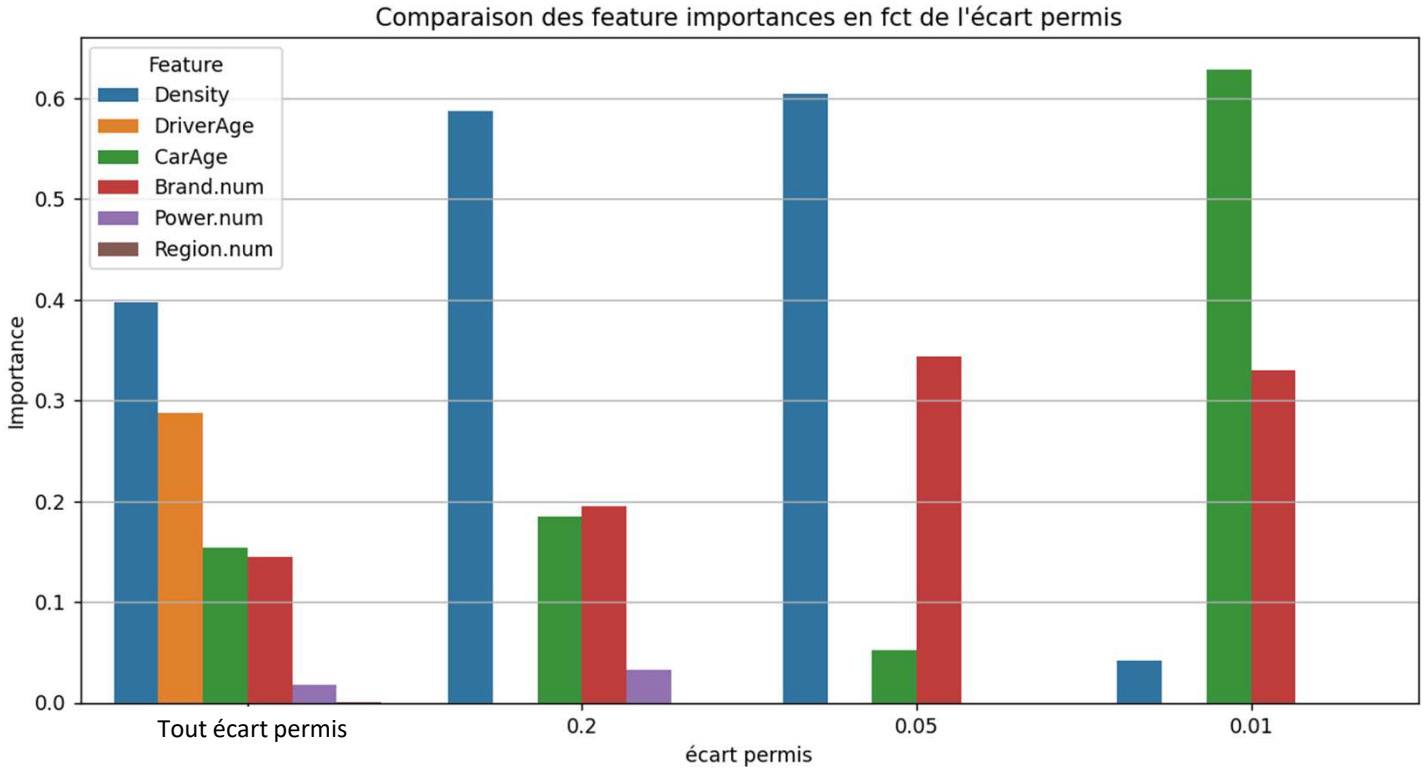






Application sur freMTPL dataset de Casdataset
Modéliser le nombre de Sinistre (μ(x))

Les features importances bougent en fonction de l'écart que l'on permet par rapport à la proportion d'âge.
Les variables qui sont corrélées avec l'âge disparaissent au fur et à mesure. On peut voir que la puissance du véhicule n'est plus utilisée lorsqu'on protège fortement les écart de proportion



RESULTATS CAS PRATIQUE
Dataset Allianz Trade

Application sur Allianz Trade dataset
Description de la base de donnée

Le dataset d'Allianz Trade est un **jeu de données** issu de l'**assurance crédit.**
La visée de ce dataset est d'étudier les acceptations de limite envers des clients qui font la demande pour en obtenir de nouvelles/des supplémentaires.

Variables explicatives:

+-10 variables financières : Visant à capter la santé financière de l'entreprise demandant la limite
+- 2 variables de Grading
Dont la variable de Grading indiquant si le grading est automatique ou manuel
+- 5 variables concernant la taille de l'entreprise

Target:

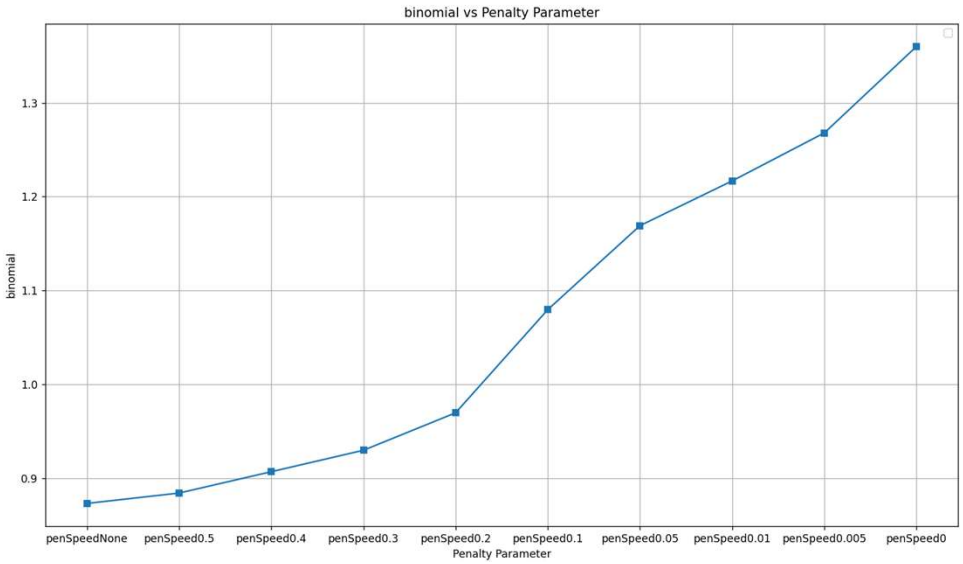
Une variable 0-1 : Indiquant si la limite a été accordée ou pas au client suite à sa demande

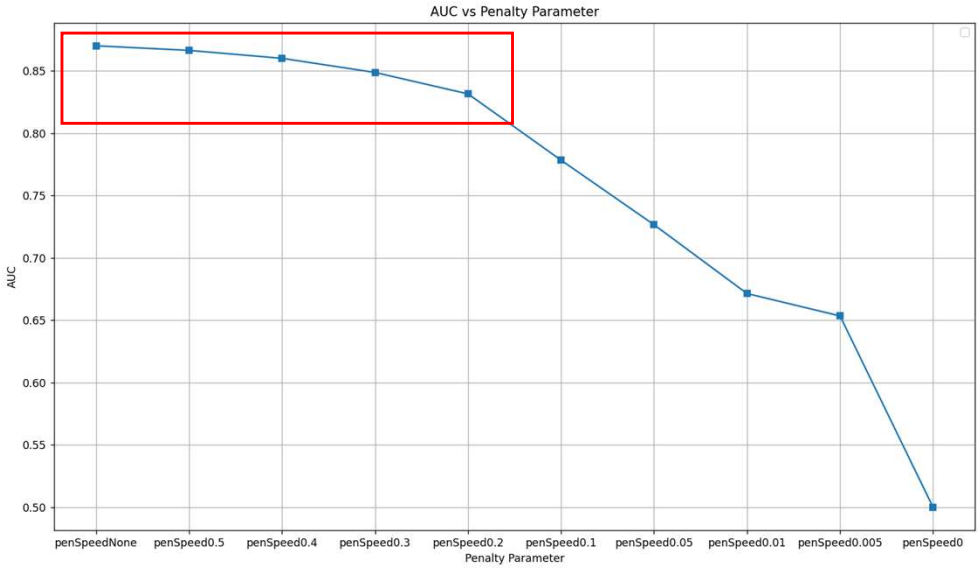
Variable que nous allons protéger:

Le dataset contient une variable qui indique si le Grading de l'entreprise a été calculé de manière automatique ou bien si celui-ci a été remplacé par une valeur manuelle.

Application sur Allianz Trade dataset
Modéliser l'acceptation ou le refus d'une limite

Dans ce cas-ci, la variable que nous protégeons est une variable indiquant si le grade est un calculé de manière automatique ou bien s'il a été remplacé par jugement d'expert.
Les graphiques de gauche et droite montrent respectivement la déviance et l'AUC
en fonction de l'écart de proportion permis sur la proportion de grade manuelle ou automatique.

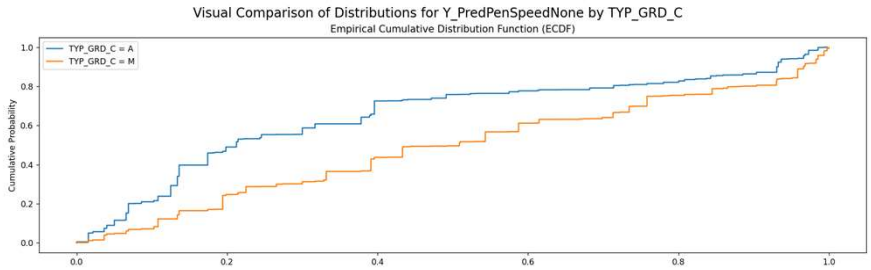


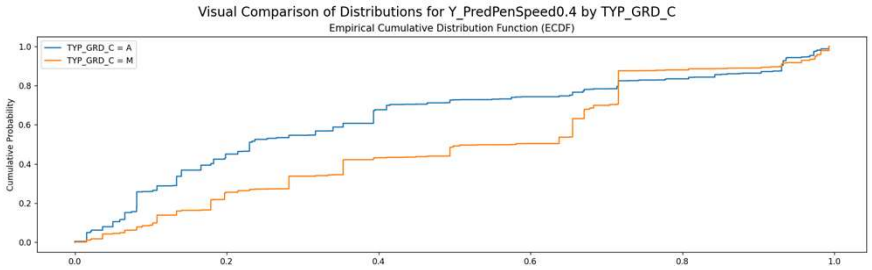


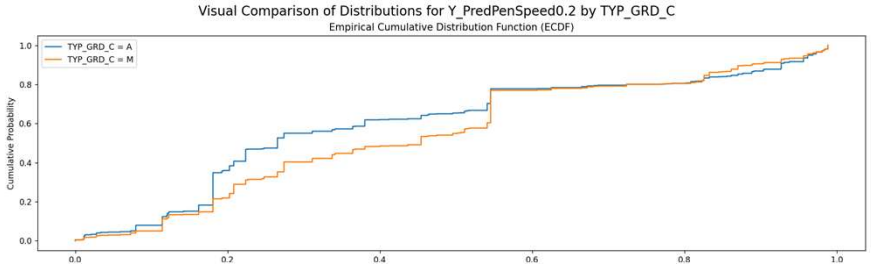
A titre indicatif, estimer les modèles sans utiliser la variable que nous protégeons actuellement permet d'arriver à un AUC de 87% sur le test set.

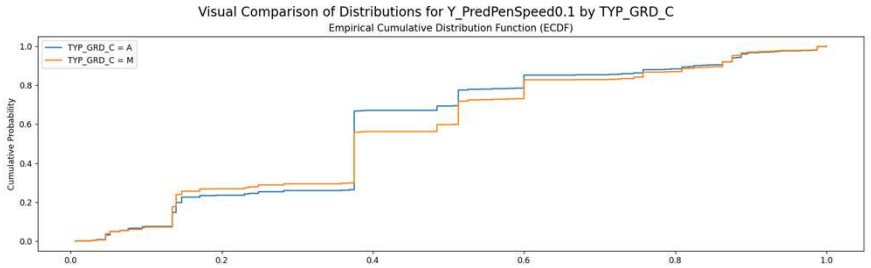
Application sur Allianz Trade dataset
Modéliser l'acceptation ou le refus d'une limite

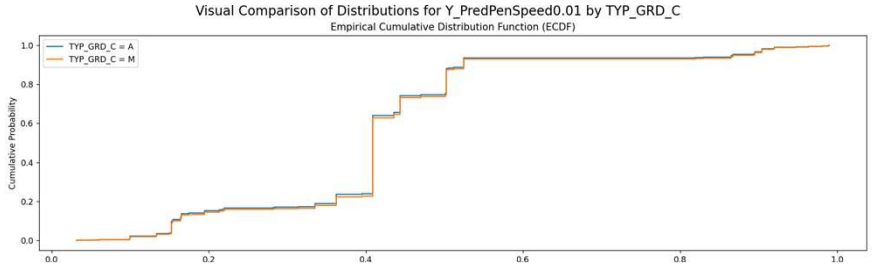
Distribution de prédictions par type de grading : Automatique ou Manuel

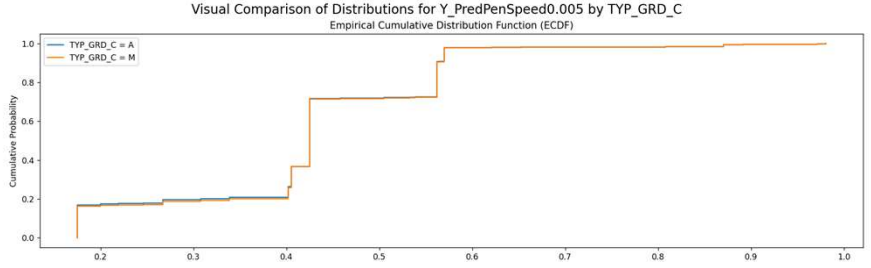




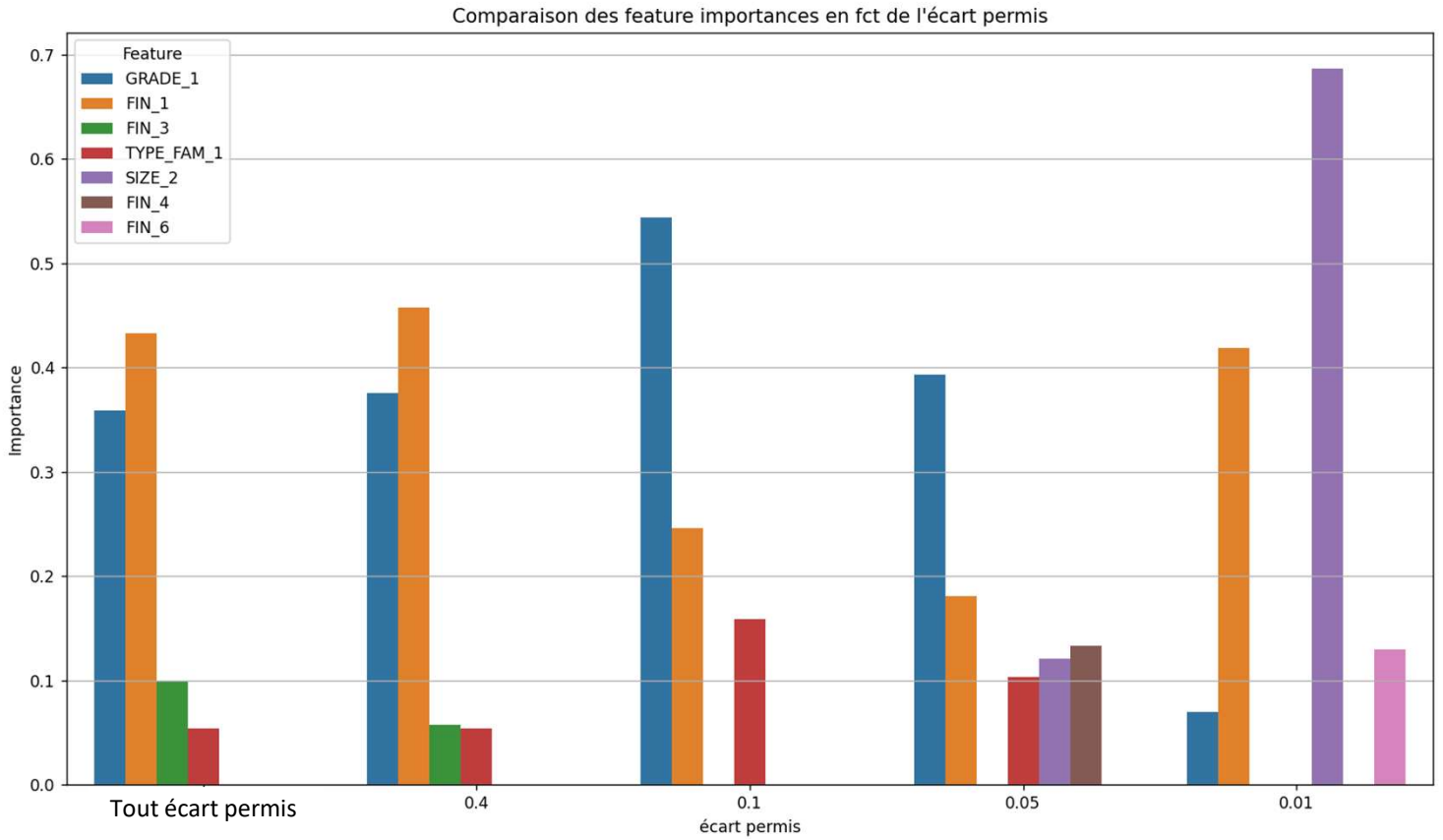








Application sur Allianz Trade dataset
Modéliser l'acceptation ou le refus d'une limite



Merci pour votre Attention !

Julien Trufin, Thomas Hames, Sophie Navarro, Harrison Verelst