

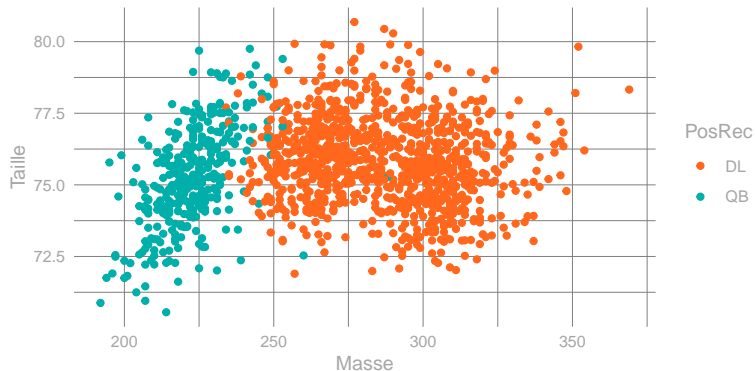
Introduction aux modèles prédictifs

Cas discret

Véronique Tremblay

Qu'est-ce qu'on cherche à faire?

On cherche à prédire G à partir de X .



Fonction de perte

On pose une matrice de coût \mathbf{L} qui contient le prix à payer pour avoir mal classifié une valeur. Par exemple, si la variable a $K = 2$ modalités:

$$\mathbf{L} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

La fonction de perte est $L(k, l)$, le prix à payer pour classer une observation qui appartient à la classe k dans la classe l .

La fonction de perte (0-1).

$$L(G, \hat{G}) = \mathbb{I}(G \neq \hat{G})$$

La fonction de risque

On cherche à minimiser

$$\text{EPE} = \mathbb{E}[L(G, \hat{G}(x))]$$

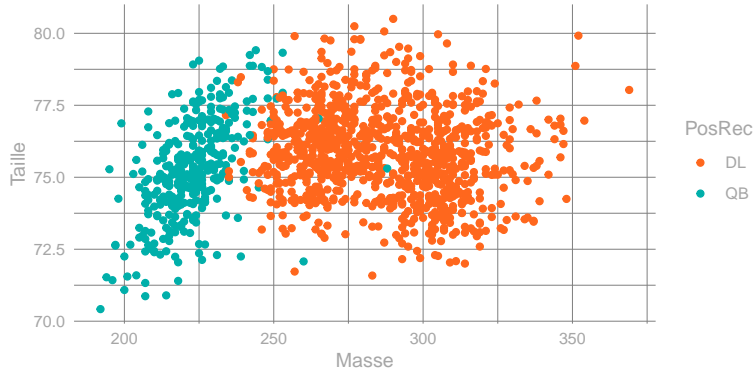
.

En conditionnant sur X ...

$$\hat{G}(x) = \operatorname{argmin}_{g \in \mathcal{S}} \sum_{k=1}^K L(\varsigma_k, g) \Pr(\varsigma_k | X = x)$$

Avec la fonction de perte 0-1, On choisit simplement la classe g pour laquelle $\Pr(g|X = x)$ est la plus élevée.

Les plus proches voisins



- On cherche une fonction qui modélise G à partir de X
- Plusieurs approches
- Dans les prochaines capsules
 - Autres approches
 - Compromis biais variance
 - Évaluation des modèles