

## Équipe 06 – Préparation au café du savoir 1

### Question 1.

On prévoit catégoriser la force des équipes des données complètes en observant leur pointage total des 18 trous. Ensuite, on fera des distributions pour chaque trou et chaque force avec les 2000 données (nombre de forces X nombre de trous=nombre de distributions). Ces distributions sont celles a priori.

On veut aussi utiliser les données complètes pour tester nos futurs modèles. On pense utiliser une approche de MSE pour pouvoir les comparer.

Les données partielles seront utilisées pour appliquer le modèle, soit faire nos prédictions.

### Question 2.

Pour les données supérieures à 10 coups (données extrêmes) dans le fichier de parties complètes, on envisage deux solutions. La première est d'enlever la partie complète contenant une de ces données, car elles représentent uniquement 1,3% des parties complètes et il serait très peu probable qu'il ait réellement fallu exactement 30 ou 33 coups précisément pour réussir le trou (surtout en équipe de 4).

L'autre alternative serait de remplacer les 30, 33 et « NA », par la médiane (en faisant attention d'avoir uniquement des valeurs entières), car il y a moins que 50% de chances que la vraie valeur soit supérieure ou inférieure à la médiane. On aurait voulu prendre la moyenne, mais on doit utiliser des valeurs entières de coups pour être réalistes, donc on s'est tournés vers la médiane.

### Question 3.

$S_i$  = nombre de coups totaux pour l'équipe  $i$ ,  $i = 1: 2000$

$$P(\Theta = \theta) = \begin{cases} \frac{\text{nb de } S_i < 80}{2000}, \theta = \text{Bon} \\ \frac{\text{nb de } 80 \leq S_i < 90}{2000}, \theta = \text{Moyen} \\ \frac{\text{nb de } 90 \leq S_i}{2000}, \theta = \text{Mauvais} \\ \dots \end{cases}$$

$\Theta$  : talent du joueur exprimé selon une catégorie.

**\*\*NOTE** :Le nombre de catégories et les seuils peuvent varier (ce sera à tester).

$$P(\Theta = \theta | X_1 = x_1, X_2 = x_2, \dots) = \frac{\prod_{i=1}^n P(X_i = x_i | \Theta = \theta) P(\Theta = \theta)}{P(\Theta = \theta)},$$

$X_i$  : nombre de coups des trous précédents.