

Éditorial du numéro spécial “Sport et Statistique”

Title: Editorial of the Special Issue “Sport and Statistics”

Gilles Celeux¹

La place de la statistique dans l'évaluation des performances sportives est devenu très importante. Ce phénomène esst né aux USA où il a été longtemps confiné. Depuis une dizaine d'années, il s'est étendu et a notamment atteint la France. Je pourrais multiplier les preuves de cette avancée de la statistique dans le domaine du sport de compétition. Mais je me contenterai d'en signaler un exemple, le dossier “L'équipe explore” *La Data révolution* (Delfosse, 2014) qui décrit de manière approfondie l'influence de l'analyse de données pour mieux comprendre et améliorer les performances des équipes de football. (Je remercie Charles Bouveyron d'avoir attiré mon attention sur ce dossier web.) Si l'évocation de données statistiques devient monnaie courante dans les émissions et les articles grand public consacrés au sport, on voit également se développer les articles scientifiques consacrés à l'analyse des performances sportives dans des revues académiques de statistique comme *Journal of the American Statistical Association* et la revue académique spécialisée sur le sport *Journal of Quantitative Analysis in Sports* fait la part belle à la statistique. Ainsi les bases de données sportives représentent aussi un champ d'investigation intéressant pour des méthodes et des modèles innovants de statistique (cf. par exemple Komar et al. 2014 qui construisent une classification fondée sur un modèle de mélange régularisé pour caractériser des stratégies d'apprentissage en natation).

Cet engouement pour l'usage des données statistiques dans le sport avait incité Jean-Louis Foulley à organiser une après-midi d'étude sur le thème "Sport et Statistique" sous l'égide de la SFdS en janvier 2014. Ce numéro spécial est l'émanation de cette demi-journée. Il comprend quatre articles dont l'objectif est avant tout de proposer des méthodologies statistiques pertinentes pour tirer des informations utiles de données sportives.

Dans l'article "Autour des procédures de classement : exemples du tennis, du tennis de table et du golf" Loïc Champagne et Léo Gerville-Réache proposent de voie de juste établissement des classements dans les sports individuels par des modèles probabilistes simples ajustés par des fréquences observées. Leur approche tient en particulier bien compte des caractéristiques propres aux sports envisagés.

Dans l'article "Prédiction des résultats de matchs de tennis professionnel par un modèle binomial avec des lois a priori informatives" Pierre Colin et Aurélien Bechler proposent une approche bayésienne informative tirant parti des performances passées et où ils distinguent judicieusement le rôle différent des moyennes et des variances des lois a priori.

Dans l'article "Un procédé bayésien simple de prédiction des résultats des matches de la ligue

¹ Inria-Saclay-Île-de-France, Orsay, F-91405, France. E-mail : Gilles.Celeux@inria.fr

des champions UEFA" Jean-Louis Foulley propose une utilisation efficace de l'actualisation annuelle des classements des équipes.

Dans l'article "Vers une évaluation objective de la performance d'une Équipe en basketball", Gilles Celeux et Valérie Robert proposent un score d'évaluation des équipes et des joueurs qui tient compte de l'issue des matches.

Ainsi tous ces articles ont en commun d'affiner des scores d'évaluation ou de prédiction de performances par un usage raisonné de données statistiques précises et maintenant facilement disponibles.

Je remercie les auteurs de leurs contributions et Jean-Louis Foulley de son activité d'animation pour l'exploitation statistique des données sportives dont ce numéro spécial est l'un des fruits.

Références

- Delfosse, A. (2014). La data révolution. In *L'équipe Explore*. <http://www.lequipe.fr/explore/la-data-revolution/#INTRO>.
Komar, J., Hérault, R., and Seifert, L. (2014). Key point selection and clustering of swimmer coordination through sparse Fisher-EM. <http://arxiv.org/pdf/1401.1489v1.pdf>.