

Présidence du Concours National Commun 2024
École supérieure des industries du textile et de l'habillement - ESITH -

Fiche TIPE

Nom et Prénom : OUDRHIRI IDRISSE SAFWANE

CIN : BE934829

Titre : Modélisation statistique des matchs de football grâce à la loi de Poisson.

Motivation : Allant des simples spectateurs jusqu'aux scientifiques les plus aguerris, prédire les résultats des matchs de football a toujours été un défi que se sont lancés les amateurs de ce sport. N'échappant pas à cette règle, j'ai choisi de mettre mes connaissances à profit pour étudier son aspect probabiliste.

Ancrage : Mon Tipe établit un lien fort entre les mathématiques et le domaine des jeux et sports, en effet la loi de Poisson est efficace pour modéliser les événements rares, ce qui est le cas des buts dans les matchs de football qui est l'un des sports des plus populaires.

Encadrant : LACHHAB ABDENNACER

Positionnement thématique : - MATHÉMATIQUES (Autres) - INFORMATIQUE (Informatique pratique)

Travail en groupe : Non

Mots clés en français : Modelisation statistique, Football, Loi de Poisson, Maximum de vraisemblance, Variables aléatoires

Mots clés en anglais : Statistical modelisation , Football , Poisson law , Maximum likelihood, Random variables

Bibliographie commentée : Les prédictions et par extension les paris sportifs ont toujours joué un rôle central dans l'histoire du sport, allant des jeux olympiques Antiques jusqu'à nos jours, les méthodes de prédiction des événements sportifs ont bien évoluées: des prévisions se basant uniquement sur l'intuition jusqu'à l'essor de l'intelligence artificielle tout en passant par la découverte de la puissance des mathématiques dans ce domaine attisant la curiosité de tous de part sa nature imprévisible. Mais comment deux concepts aussi abstraits que sont les mathématiques et le hasard peuvent s'associer pour prévoir la réalité? L'idée intuitive qui vient quand on parle de prédire mathématiquement des résultats de matchs de football est de prédire le nombre de buts marqués et encaissés par chacune des deux équipes. Le modèle le plus simple pour faire cela est de découper le match en n périodes et d'estimer une probabilité p constante de marquer un but sur chaque période, la variable donnant le nombre de buts marqués suivra donc une loi binomiale de paramètre n et p . Naturellement, la première idée qui vient à l'esprit pour améliorer ce modèle est de faire tendre le nombre de périodes vers l'infini: c'est l'approximation de la loi binomiale par la loi de Poisson, dont la principale difficulté est d'en estimer le paramètre. Le premier modèle de poisson pertinent a été proposé par Maher en 1982, venu faire face à Reep et Benjamin qui ont réussi à modéliser avec succès la succession de passe lors d'un match grâce à la loi binomiale négative, en effet, il a été le premier à prendre en compte de nouveaux paramètres inhérents à un match de football tels que les forces d'attaque et de défense de chacune des équipes ainsi que le facteur de l'avantage du terrain (à domicile ou à l'extérieur). D'autres modèles alternatifs sont apparus suite à cette publication centrale de Maher: une modélisation par la loi de Poisson bivariée fut proposée par Dimitris Karlis et Ioannis Ntzoufras, elle permet de modéliser conjointement les variables aléatoires représentant le nombre de buts marqués par chacune des deux équipes et ce pour prendre en compte l'impact d'un but marqué ou encaissé sur chacun des deux adversaires, mais ce modèle peut s'avérer assez compliqué sans pour autant offrir des avantages conséquents, ce pourquoi ces deux même chercheurs ont eu l'idée de modéliser la différence de buts des deux équipes pour créer ce lien entre elles, la variable aléatoire qui en résulte, étant une différence de Poisson, suivra donc la loi de Skellam, cette simple petite différence confère les avantages de la loi de Poisson bivariée avec bien moins de désavantages et de difficultés. Le modèle de la loi de Poisson paraît cohérent, mais comment en estimer le paramètre? Pour ce faire plusieurs méthodes sont possibles: la régression logistique, la régression de Poisson, la méthode des moindres carrés, les chaînes de Markov... Mais l'idée principale réunissant ces différentes méthodes est de se baser sur les résultats des matchs passés pour maximiser la vraisemblance, ou en d'autres termes avoir l'estimation la plus précise possible des forces d'attaque et de défense de chaque équipe qui permettront par la suite d'estimer ce fameux paramètre.

Problématique : Comment modéliser mathématiquement les matchs de football grâce à la loi de Poisson et l'optimiser? Et comment concevoir un programme informatique capable de prédire les résultats des matchs de football grâce à cette loi?

Objectifs : -étudier avec rigueur mathématique la modélisation des matchs de football par la loi de Poisson. -vérifier la pertinence de celle-ci en simulant matchs et championnats via un informatique. -explorer des pistes d'amélioration et comparer les résultats finaux avec les résultats initiaux.

Liste de références bibliographiques

Publication		
Auteur 1, Auteur 2 ...	Nom du periodique	Titre de l'article
PAUL STEFFEN	Université de Bordeaux	Statistical modeling of event probabilities subject on a sports bet : Theory and applications to soccer, tennis and basketball
DIMITRIS KARLIS ET IOANNIS NTZOUFRAS	IMA Journal of Management Mathematics	Bayesian modelling of football outcomes: using the Skellam's distribution for the goal difference
MAHER M.J	Statistica Neerlandica	Modelling association football scores

Site internet	
URL avec informations complémentaires	Date de consultation
https://www.science4all.org/article/a-model-of-football-games/	16.12.2023