

UCL 2024/25 : Simulation Monte Carlo de la Phase à Élimination Directe

Ethan Robertson

Résumé

En m'appuyant sur mon analyse précédente de la force des équipes utilisant un système de classement Elo, j'ai développé une simulation Monte Carlo pour prévoir les résultats de la phase à élimination directe de l'UEFA Champions League (UCL) 2024/25. Cette simulation exploite les estimations de probabilité dérivées des classements Elo pour modéliser la progression du tournoi selon différents scénarios possibles. L'objectif principal de cette étude est d'estimer la probabilité de chaque équipe de progresser à travers les tours à élimination directe et finalement de remporter la compétition. De plus, je développe une stratégie de paris optimale en utilisant la solution explicite du Critère de Kelly formulée par Smoczyński & Tomkins, permettant une allocation optimale d'un capital fixe sur plusieurs paris tout en maximisant la croissance logarithmique attendue de la richesse.

1 Introduction

En m'appuyant sur mon analyse précédente de la force des équipes utilisant un système de classement Elo, j'ai développé une simulation Monte Carlo pour prévoir les résultats de la phase à élimination directe de l'UEFA Champions League (UCL) 2024/25. Cette simulation exploite les estimations de probabilité dérivées des classements Elo pour modéliser la progression du tournoi selon différents scénarios possibles.

L'objectif principal de cette étude est d'estimer la probabilité de chaque équipe de progresser à travers les tours à élimination directe et finalement de remporter la compétition. En simulant la structure du tournoi de manière répétée, j'obtiens une distribution probabiliste des résultats, permettant une évaluation quantitative des chances de victoire de chaque équipe.

De plus, je vise à développer une stratégie de paris optimale en utilisant la solution explicite du Critère de Kelly formulée par Smoczyński & Tomkins.¹ Cette approche permet

1. Smoczyński, K., & Tomkins, J. (2004). An explicit solution to the multivariate Kelly problem.

une allocation optimale d'un capital fixe sur plusieurs paris, maximisant la croissance logarithmique attendue de la richesse.

Pour plus de détails sur le système de classement Elo et son application, voir *Classements Elo UCL & Modélisation Prédictive (Saison 2024/25)*.

2 Méthodologie

2.1 Modèle de Probabilité Basé sur Elo

La base de la simulation est un système de classement Elo attribuant à chaque équipe un score numérique. À la fin de la phase de groupes, les classements Elo mis à jour ont été utilisés pour estimer les probabilités face-à-face. La probabilité que l'Équipe A gagne contre l'Équipe B est calculée selon la formule suivante :

$$P_A = \frac{1}{1 + 10^{(R_B - R_A)/400}} \quad (1)$$

Dans cette équation, P_A représente la probabilité que l'Équipe A gagne, tandis que R_A et R_B désignent les classements Elo des Équipes A et B, respectivement. Le facteur d'échelle de 400 contrôle la sensibilité aux différences de classement, assurant des distributions de probabilité raisonnables à travers différents écarts de classement.

2.2 Cadre de Simulation Monte Carlo

Le processus de simulation a impliqué l'exécution de 1 000 000 de simulations Monte Carlo de la phase à élimination directe. Pour chaque simulation de match, un nombre aléatoire u a été tiré d'une distribution uniforme $\mathcal{U}(0, 1)$. Si $u < P_A$, l'Équipe A était considérée comme qualifiée ; sinon, l'Équipe B progressait. Ce processus a été répété pour chaque match du tournoi, des huitièmes de finale jusqu'à la finale. Les taux de progression de chaque équipe ont été soigneusement enregistrés et agrégés pour produire les estimations finales de probabilité.

Team	QF Probability	SF Probability	Final Probability	Win Probability
FC Barcelona	62.15%	34.10%	19.49%	10.58%
FC Internazionale Milano	62.58%	34.69%	17.95%	9.24%
Arsenal FC	54.14%	27.53%	15.47%	8.44%
Club Atlético de Madrid	51.69%	26.83%	15.12%	8.26%
Borussia Dortmund	52.09%	26.45%	14.13%	7.12%
Paris Saint-Germain FC	51.44%	28.67%	13.75%	7.03%
Real Madrid CF	48.31%	24.19%	13.20%	7.00%
Liverpool FC	48.56%	26.43%	12.28%	6.13%
Bayer 04 Leverkusen	52.89%	26.78%	12.78%	6.06%
Lille OSC	47.91%	23.30%	11.97%	5.78%
PSV	45.86%	21.45%	11.21%	5.67%
FC Bayern München	47.11%	22.52%	10.10%	4.53%
Club Brugge KV	50.93%	23.09%	9.85%	4.51%
Sport Lisboa e Benfica	49.07%	21.81%	9.11%	4.11%
Aston Villa FC	37.85%	16.16%	7.28%	3.07%
Feyenoord Rotterdam	37.42%	16.00%	6.28%	2.46%

FIGURE 1 – Probabilités simulées pour chaque équipe d'atteindre les quarts de finale, demi-finales, finale, et de remporter l'UCL.

3 Résultats et Insights Clés

3.1 Probabilité de Progression par Phase

3.2 Analyse des Paris à Valeur

À travers une analyse comparative des probabilités dérivées du modèle et des probabilités implicites des bookmakers, plusieurs écarts de valeur notables ont été identifiés. La Figure ?? illustre ces différences pour les équipes sélectionnées.

L'analyse a révélé le FC Barcelone comme favori du tournoi, avec une probabilité de 10,58% de remporter la compétition selon nos simulations. De plus, des opportunités de valeur significatives ont été identifiées dans les cotes pour Dortmund, PSV et Lille OSC en comparaison avec les cotes de VegasInsider.

3.3 Gestion de Capital Basée sur le Critère de Kelly

Pour traduire les probabilités dérivées du modèle en décisions de paris exploitables, j'ai employé le Critère de Kelly, une approche mathématique d'optimisation du capital. Plus précisément, j'ai utilisé la solution explicite décrite par Smoczynski & Tomkins, qui étend le cadre Kelly classique aux paris multiples simultanés avec des résultats corrélés—un scénario plus réaliste dans les paris sportifs basés sur les tournois.

Stochastic Processes and Their Applications, 112(2), 197–213.

Team	Implied Probability	Elo Probability	Edge
Borussia Dortmund	1.96%	7.12%	5.16%
PSV	0.66%	5.67%	5.01%
Lille OSC	0.99%	5.78%	4.79%
Club Brugge KV	0.66%	4.51%	3.85%
Club Atlético de Madrid	4.76%	8.26%	3.50%
Bayer 04 Leverkusen	2.94%	6.06%	3.12%
Sport Lisboa e Benfica	1.49%	4.11%	2.62%
Feyenoord Rotterdam	0.50%	2.46%	1.96%
FC Internazionale Milano	7.69%	9.24%	1.55%
Aston Villa FC	2.94%	3.07%	0.13%
Paris Saint-Germain FC	7.69%	7.03%	-0.66%
Arsenal FC	14.29%	8.44%	-5.85%
FC Bayern München	12.50%	4.53%	-7.97%
Liverpool FC	15.38%	6.13%	-9.25%
FC Barcelona	20.00%	10.58%	-9.42%
Real Madrid CF	23.09%	7.00%	-16.09%

FIGURE 2 – Différences entre les probabilités impliquées par le modèle et les probabilités impliquées par les bookmakers pour les équipes sélectionnées.

3.3.1 Fondement Théorique

Le Critère de Kelly cherche à maximiser la croissance logarithmique attendue de la richesse. Pour un pari unique, la fraction optimale f^* du capital à parier est déterminée par :

$$f^* = bp - q \quad (2)$$

où f^* représente la fraction du capital à parier, b désigne les cotes décimales moins 1 (cotes nettes), p est la probabilité de gagner dérivée du modèle, et $q = 1 - p$ représente la probabilité de perdre.

Dans le contexte de paris multiples simultanés, comme c'est courant dans les tournois, le problème d'optimisation devient plus complexe. Smoczynski & Tomkins dérivent une solution sous forme fermée comme un problème d'optimisation sous contrainte qui maximise la richesse logarithmique attendue sous la contrainte que la somme de toutes les fractions de paris ne dépasse pas le capital total.

3.3.2 Procédure d'Implémentation

Le processus d'implémentation a commencé par l'identification des paris à valeur en comparant la probabilité dérivée du modèle de gagner l'UCL (p_i) avec la probabilité

implicite basée sur les cotes des bookmakers ($\hat{p}_i = \frac{1}{\text{cotes}_i}$) pour chaque équipe i . Un pari était considéré comme ayant une valeur espérée positive lorsque $p_i > \hat{p}_i$.

Les fractions optimales $\{f_i^*\}$ ont été déterminées en résolvant le problème d'optimisation :

$$\max_f \mathbb{E} \left[\log \left(1 + \sum_i f_i X_i \right) \right] \quad (3)$$

sous la contrainte :

$$\sum_i f_i \leq 1, \quad f_i \geq 0 \quad \forall i \quad (4)$$

où X_i représente le gain net du pari i , prenant la valeur b_i si le pari i gagne et -1 sinon. L'optimisation a été implémentée en utilisant des méthodes numériques en Python.

TABLE 1 – Fractions de Paris Optimales Basées sur le Critère de Kelly et les Probabilités du Modèle

Équipe	Prob. Impl.	Prob. Elo	Eri	Paris Opt.
PSV	0,66%	5,67%	9,07	12,46%
Club Brugge KV	0,66%	4,51%	7,01	12,74%
Lille OSC	0,99%	5,78%	5,86	12,43%
Feyenoord Rotterdam	0,50%	2,46%	4,81	13,25%
Borussia Dortmund	1,96%	7,12%	3,27	12,10%
Sport Lisboa e Benfica	1,49%	4,11%	2,24	12,84%
Bayer 04 Leverkusen	2,94%	6,06%	1,42	12,36%
Club Atlético de Madrid	4,76%	8,26%	1,04	11,82%

3.4 Tests de Robustesse

Pour évaluer la sensibilité du modèle aux biais potentiels de classement, j'ai effectué une série de simulations Monte Carlo sous divers scénarios :

- Cas de base sans ajustements de classement
- Équipes de Premier League surévaluées de 40 points Elo
- Équipes de La Liga sous-évaluées de 30 points Elo
- Équipes de Bundesliga surévaluées de 50 points Elo

Les résultats ont révélé une remarquable cohérence à travers les scénarios, avec des ROI allant de 223% à 280%. Le scénario de Premier League surévaluée a donné les meilleurs résultats (279,68% ROI), suggérant un biais conservateur potentiel dans nos évaluations initiales des équipes de Premier League. Le scénario de La Liga sous-évaluée a montré

les rendements les plus faibles mais toujours substantiels (223,07% ROI), indiquant une robustesse contre une sous-estimation potentielle des équipes espagnoles.

3.5 Considérations de Gestion des Risques

Bien que le Critère de Kelly fournisse un dimensionnement théoriquement optimal des paris, sa nature agressive nécessite souvent des ajustements pratiques. Une approche plus conservatrice utilisant des paris Kelly fractionnaires est recommandée, où les positions sont réduites à une proportion fixe (par exemple, 50%) des montants optimaux. Cette modification aide à atténuer la volatilité tout en maintenant des rendements attendus significatifs sur le long terme.

Les résultats des tests de robustesse suggèrent que la stratégie maintient sa rentabilité même sous des biais de classement significatifs. Cependant, pour améliorer davantage la gestion des risques, je propose une approche diversifiée :

- 50% du capital alloué aux paris traditionnels sur les équipes sous-évaluées
- 40% alloué aux positions synthétiques (lay betting sur les équipes surévaluées)
- 10% réservé aux opportunités de couverture dynamique

Cette stratégie d'allocation vise à capturer de la valeur des deux côtés du marché tout en maintenant une réserve pour la gestion des risques. Les positions synthétiques sont particulièrement précieuses dans les cas où les équipes sont significativement surévaluées par le marché, car elles nous permettent de profiter à la fois des scénarios de surévaluation et de sous-évaluation.

3.6 Analyse de Distribution et Considérations de Risque

Un examen critique de la distribution des gains révèle un défi fondamental dans la stratégie de paris initiale. Comme le montre la Figure ??, la distribution présente une asymétrie significative vers la droite, indiquant que bien que l'espérance de gain reste positive, elle est principalement due à un petit nombre de gains extrêmement élevés se produisant avec une faible probabilité. Cette caractéristique, bien que mathématiquement valide, présente des défis pratiques pour la gestion des risques et l'allocation du capital.

L'analyse de la Valeur à Risque (VaR) au niveau de confiance de 95% quantifie ce risque, montrant que la stratégie pourrait subir des drawdowns substantiels dans la majorité des scénarios. Cette perspective a conduit au développement d'une approche plus équilibrée qui considère à la fois l'espérance de gain et la probabilité de succès, tout en maintenant la capacité d'exploiter les erreurs de prix du marché.

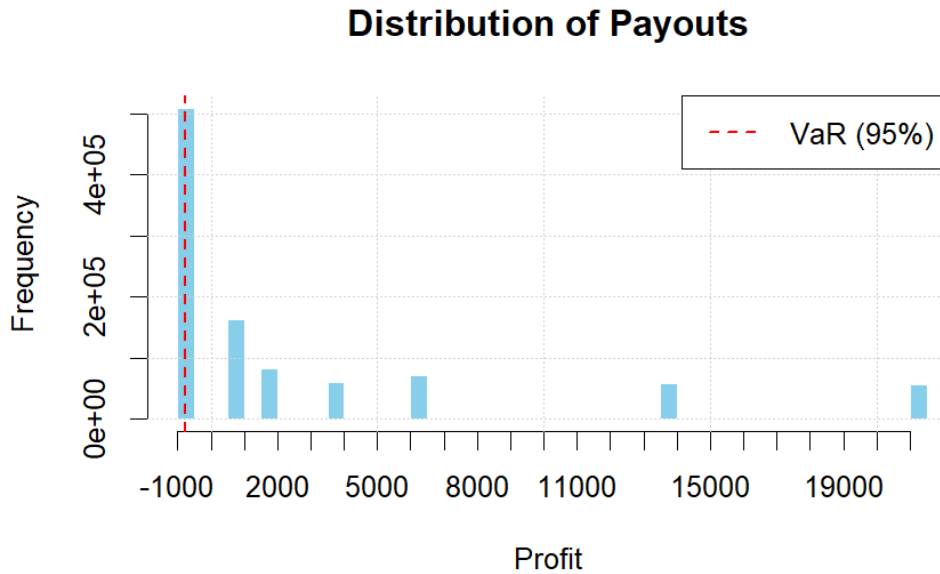


FIGURE 3 – Distribution des gains de paris montrant la nature asymétrique des rendements. La ligne rouge indique la Valeur à Risque (VaR) au niveau de confiance de 95%.

3.7 Stratégie de Paris Affinée

Pour répondre à ces défis tout en préservant les avantages fondamentaux de la stratégie, j'ai développé une approche à trois composantes qui équilibre risque et rendement :

3.7.1 Paris Traditionnels (50% du capital)

Cette composante se concentre sur les équipes qui satisfont deux critères clés :

- Valeur espérée positive basée sur les probabilités dérivées du modèle
- Seuil de probabilité minimum (par exemple, 5% de chances de gagner)

En implémentant ces filtres, nous assurons une distribution plus équilibrée des résultats potentiels tout en maintenant une exposition aux équipes sous-évaluées. Le seuil de probabilité aide à éviter une dépendance excessive aux scénarios de faible probabilité et de gain élevé.

3.7.2 Positions Synthétiques (40% du capital)

La composante synthétique est implémentée à travers deux mécanismes principaux :

1. **Paris Lay** : Lay betting direct sur les équipes surévaluées sur les bourses de paris, particulièrement celles avec des probabilités implicites significativement plus élevées que les estimations du modèle. Cela nous permet de profiter à la fois des scénarios de surévaluation et de sous-évaluation.

2. **Paris Combinés** : Création de positions synthétiques à travers des combinaisons de paris traditionnels. Par exemple, parier contre une équipe peut être réalisé en :
 - Soutenant toutes les autres équipes du tournoi
 - Utilisant des paris accumulateurs avec des combinaisons soigneusement sélectionnées
 - Implémentant des stratégies de spread betting lorsque disponibles

Les positions synthétiques sont particulièrement précieuses dans les cas où les équipes sont significativement surévaluées par le marché, car elles nous permettent de profiter à la fois des scénarios de surévaluation et de sous-évaluation.

3.7.3 Implémentation du Lay Betting Synthétique

En s'appuyant sur le cadre des positions synthétiques, nous avons implémenté une stratégie spécifique se concentrant sur les équipes qui sont surévaluées par le marché. La stratégie identifie les équipes où la probabilité implicite du marché dépasse notre probabilité modélisée d'au moins 5 points de pourcentage, créant des positions lay synthétiques contre ces équipes.

Pour l'UCL 2024/25, notre analyse a identifié cinq équipes significativement surévaluées :

TABLE 2 – Équipes Surévaluées et Positions Lay Synthétiques

Équipe	Prob. Modèle	Prob. Marché	Suréval.	Mise Lay
Real Madrid CF	12,00%	23,09%	+11,09%	100\$
FC Barcelona	10,58%	20,00%	+9,42%	100\$
Liverpool FC	7,50%	15,38%	+7,88%	100\$
Arsenal FC	8,00%	14,29%	+6,29%	100\$
FC Bayern München	6,50%	12,50%	+6,00%	100\$

La stratégie crée des positions lay synthétiques contre ces équipes, avec chaque position dimensionnée à 100\$. Étant donné la nature mutuellement exclusive des résultats du tournoi (une seule équipe peut gagner), la perte potentielle maximale est limitée à la plus grande responsabilité individuelle (700\$ dans ce cas), plutôt qu'à la somme de toutes les responsabilités. Cela crée un profil risque-rendement favorable :

- Mise Totale : 500\$ (5 positions \times 100\$)
- Responsabilité Maximale : 700\$ (si Bayern gagne)
- Meilleur Cas : Gain de 500\$ (si aucune de ces équipes ne gagne)
- Probabilité d'Aucun Gagnant : 55,92% (1 - 44,08%)

Cette approche fournit un profil risque-rendement équilibré tout en maintenant la capacité de profiter des erreurs de prix du marché. La stratégie est particulièrement efficace car elle :

- Se concentre sur les écarts de probabilité significatifs ($\geq 5\%$)
- Limite l'exposition à travers le dimensionnement des positions
- Bénéficie de la nature mutuellement exclusive des résultats du tournoi
- Maintient une probabilité favorable de succès (55,92%)

3.7.4 Couverture Dynamique (10% du capital)

La composante finale fournit de la flexibilité pour gérer l'exposition au risque tout au long du tournoi :

- Ajustement des positions en temps réel basé sur les résultats des matches
- Couverture opportuniste lorsque des cotes favorables se présentent
- Planification de contingence pour divers scénarios du tournoi

Cette approche affinée maintient la rigueur mathématique de la stratégie originale tout en abordant ses limitations pratiques. En incorporant des positions synthétiques, nous pouvons capturer de la valeur des équipes surévaluées sans compter uniquement sur des scénarios de faible probabilité et de gain élevé. La composante de couverture dynamique améliore davantage les capacités de gestion des risques.

4 Conclusion

La combinaison du cadre de simulation Monte Carlo et d'optimisation Kelly fournit une méthodologie robuste pour prévoir les résultats des phases à élimination directe de l'UCL et développer des stratégies de paris rentables. L'analyse révèle des opportunités significatives de paris à valeur, particulièrement dans les cas où les cotes du marché divergent substantiellement des probabilités dérivées du modèle. Cependant, l'implémentation de telles stratégies nécessite une considération attentive des principes de gestion des risques, étant donné les incertitudes inhérentes à la modélisation probabiliste et le potentiel de variance substantielle dans les résultats.