Paris en ligne

Marmoth85

28 juillet 2017

Table des matières

Introduction

Dans ce document, nous allons proposer quelques définitions, notations, calculs et techniques de base pour les paris sportifs en ligne. Pour le moment, on ne s'intéresse qu'aux évènements sportifs à deux issues (victoire joueur 1 ou joueur 2, équipe gagnante au basket ball, set gagnant...). Cela dit, les calculs différent pour les systèmes ayant plus d'évènements possibles, sont légèrement plus complexes, mais les idées restent absolument les mêmes. Aussi, pour résumer les principales notations, nous les notons dans le tableau suivant :

Évènement	Joueur/Équipe 1	Joueur/Équipe 2
Côte	a	b
Mise	x	y
Gain brut	g_1	g_2
Gain net	G_1	G_2

Soit donc un évènement sportif présentant deux issues possibles, comme par exemple un match de tennis (vainqueur joueur 1 ou joueur 2). Chaque joueur possède un côte (a et b) fixée par le bookmaker. Le parieur va pouvoir miser une quantité d'argent x sur le premier évènement et/ou y sur le second suivant sa stratégie. Deux cas sont alors possibles :
— le joueur 1 s'impose dans son match : le gain g_1 est donné par ax,
— le joueur 2 s'impose dans son match : le gain g_2 est donné par by.
Ce qu'on appelle donc gain, ou gain brut pour expliciter la chose, c'est la somme avec laquelle on repars une fois le match terminé : c'est le terme "gain" utilisé par les bookmakers.

Nous autres, parieurs, aimons parler de gain en termes de bénéfices et ça n'est pas la même notion puisqu'il faut tenir compte des mises, $x + y$ ici. Les gains, ou gains net donc, sont alors données par $G_1 = ax - x - y$ et $G_2 = by - x - y$. A présent que les notions de base sont posées, on aborde la technique du pari sûr, le surebet, qui vise à gagner quoiqu'il se passe dans le match concerné.

1 La technique du surebet

1.1 Principe et stratégies de mises

Pour faire un pari certain, ou surebet, il faut rechercher plusieurs choses. Le premier volet est de se dire de chercher un couple $(x, y) \in (\Re^+ \times \Re^+)$ tel que G_1 et G_2 sont positifs ou nuls. Cela revient au final à rechercher les mises optimales pour un match donné, optimales par rapport à notre objectif. Le fait est que les calculs montrent que dans l'immense majorité des cas, ce couple n'existe pas et renvoie a des conditions sur les côtes elles-mêmes. Deux stratégies de mises sont possibles :

1. On pronostique que l'un des joueurs va gagner (le joueur 1 par exemple) et on veut donc que le gain soit maximal en cas de victoire de ce joueur et que si la malchance nous frappe, qu'au moins la deuxième mise permet de ne pas perdre d'argent, en remboursant le match. Soit on gagne une somme sympa, soit on repars avec notre mise : 100% cool.
2. On vise le même gain, quelque soit le vainqueur du match. Forcément, le gain sera plus faible qu'avec la stratégie précédente si notre pronostique est bon, mais là on est certain de l'emporter ce gain. La logique est la même, mais les calculs sont plus simples. Exemple : si on souhaite repartir avec 100 € de gain (brut), il faudra miser $\frac{100}{a}$ € sur le joueur 1 (en cas de victoire on aurait $g_1 = \frac{100}{a}a$) et $\frac{100}{b}$ € sur le joueur 2.

Au final, vouloir gagner à coup sûr, se traduit comme suit :

$$\begin{cases} ax - x - y \geq 0 \\ by - x - y \geq 0 \end{cases} \tag{1}$$

1.2 Détecter les surebets

Prenons, le cas le plus simple pour commencer, la stratégie 2 que l'on vient de présenter. Ainsi donc on obtient $G_1 = G_2 = G$. On en déduit facilement que $x = \frac{G}{a}$ et que $y = \frac{G}{b}$. Immédiatement, on voit que les deux inégalités de notre système deviennent :

$$G - \frac{G}{a} - \frac{G}{b} \geq 0 \tag{2}$$

et on en conclu à la condition suivante, comme condition de surebet :

$$\frac{1}{a} + \frac{1}{b} \leq 1 \tag{3}$$

Notez que cette condition ne porte que sur les côtes des deux joueurs/équipes : les mises ont complètement disparues de l'inégalité ! Cela signifie que lorsque cette condition est vérifiée, un surebet est possible en choisissant bien les mises x et y. Les combinaisons de x et y sont infinies, mais x et y sont liées entre elles : il ne faut donc pas les choisir complètement au hasard !

Revenons à la stratégie 1, celle qui voulait maximiser un des deux gains au dépend du deuxième gain possible. Si notre favoris est l'équipe ou le joueur 1, on va vouloir maximiser G_1 et avoir G_2 égal à 0. Pour avoir $G_2 = 0$, il faut calculer la mise minimale à poser sur le joueur 2 juste pour se rembourser. Évidemment, la mise y va dépendre de x, logique.

$$G_2 = by - x - y = 0 \Leftrightarrow y(b-1) - x = 0$$
$$y = \frac{x}{b-1}$$

Quant à lui, x peut prendre la valeur qu'on veut bien lui donner. On peut la mettre à $\frac{g_1}{a}$ si on veut, n'importe quoi d'autre : le plus important est la relation entre x et y et qui permet à la fois de maintenir le surebet mais aussi de satisfaire à la stratégie.

En cas de victoire du joueur 2, on a :

$$by - x - y = b\frac{x}{b-1} - x - \frac{x}{b-1} = \frac{x}{b-1}(b-1) - x = 0$$

par construction de y. En cas de victoire du joueur 1, on a :

$$ax - x - y = ax - x - \frac{x}{b-1} = x(a - 1 - \frac{1}{b-1}) \geq 0$$

Le joueur 1 étant notre favoris dans ce scénario, la mise x n'est pas nulle donc on peut diviser par x de part et d'autre de l'inégalité :

$$a - 1 - \frac{1}{b-1} \geq 0 \Leftrightarrow \frac{a(b-1) - b + 1 - 1}{b-1} \geq 0 \Leftrightarrow \frac{a(b-1) - b}{b-1} \geq 0$$
$$\Leftrightarrow ab - a - b \geq 0 \Leftrightarrow 1 - \frac{a}{ab} - \frac{b}{ab} \geq 0$$

Au final, on obtient à nouveau

$$\frac{1}{a} + \frac{1}{b} \leq 1$$

C'est donc un invariant pour détecter un surebet. Ce sont les côtes qui sont essentielles pour en détecter un, pas les mises. Les mises sont simples à gérer : il ne faut pas faire n'importe quoi, mais c'est une formalité.

1.3 Un exemple concret

Supposons que Djokovic côté à 1,67 affronte Nadal, côté à 2,85. On a bien la somme des inverses de côte qui est inférieur à 1 (0,949 pour être précis) : ce match peut donc offrir un surebet en choisissant bien les mises. On part sur la stratégie la plus simple, la 2.
On veut repartir avec $g = 100$ €.
On a alors $x = \frac{100}{1{,}67} = 59,88$ € et $y = \frac{100}{2{,}85} = 35,09$ €.
Par construction des mises, quelle que soit l'issue du match, on a $g_1 = g_2 = 100$ €.
En effet, si Djokovic gagne, on repart avec $59,88 \times 1,67 \approx 100$ €. Si Nadal l'emporte, on repart avec $35,09 \times 2,85 \approx 100$ €.
Le total des mises est de 94,97 €.
Donc le gain total est de $100 - 94,97 = 5,03$ €.

Repassons à la première stratégie, consistant à maximiser les gains en cas de victoire de Djokovic sur Nadal (on peut faire le contraire sans problème...). Djokovic dans cet exemple part largement favoris et notre pronostique sera de le penser vainqueur. On va donc chercher à maximiser les gains pour ce cas en particulier. Si le pronostique n'est pas bon, on repartira quand même avec la mise de départ. Le gain (brut) visé est de 100 €. Alors on a $x = \frac{100}{1{,}67} = 59,88$ € comme dans le cas précédent.
La mise y est donnée par : $y = \frac{x}{b-1} = \frac{59{,}88}{1{,}85} = 32,37$ €.

Ainsi, si Djokovic remporte le match, on gagne $59,88 \times 1,67 = 100$ € soit donc notre objectif initial. La somme des mises se monte à 92,25 € soit un gain net de 7,75 €, ce qui est effectivement supérieur à la stratégie précédente.
En revanche, si Nadal s'impose, alors on gagne $32,37 \times 2,85 = 92,25$ € ce qui correspond exactement au montant des mises. Le gain net est alors nul, conformément à ce qu'on voulait. Entre ces deux stratégies, on peut en imaginer une infinité en se basant sur la proportion des gains désirés dans tel cas ou tel autre cas.

1.4 Résumé

On a donc établi, démontré et illustré combien $\frac{1}{a} + \frac{1}{b} \leq 1$ est intéressant pour définir une ou plusieurs stratégies plus ou moins voisines du surebet.
En pratique, sur des côtes pré-matches, cette condition n'est jamais réalisée sauf très rare exception en fouillant sur plein de bookmaker différents. Souvent, la somme des inverses de côte, sur un site français ARJEL est voisine de 1,15. En gros ça revient à dire que 15% de la somme des paris est reversés en taxes à l'état et aux marges du bookmaker concerné. A l'étranger, les marges sont nettement plus faibles, ce qui donne des côtes particulièrement intéressantes ! Malheureusement, c'est difficile d'y jouer car les français y sont persona non grata, en raison de la législation française...

2 Lien entre côte et probabilité

2.1 Une relation inverse

La côte et la probabilité d'un évènement sportif sont liés : plus un évènement est probable (plus il tend vers les 100%) plus sa côte baisse (et tend vers 1, ce qui signifie "remboursement du pari sans gain"). A l'inverse, plus les chances qu'un évènement se réalise sont faibles (tendant vers 0), plus sa côte augmente (et tend théoriquement vers l'infini même si d'un point de vue pragmatique le bookmaker stoppent l'augmentation à partir d'un certain seuil qui leur est propre, ce qui signifie tout de même de gros gains potentiels si jamais cet évènement rare se réalise).

Ainsi, la côte d'un évènement évolue de façon inverse à la probabilité qui lui est associée. On pourrait presque dire que $p_1 \approx \frac{1}{a}$ et $p_2 \approx \frac{1}{b}$ même si techniquement ce n'est pas le cas. Mais avec une grosse louche, on voit l'idée.

D'après le théorème des probabilités totales, on sait que $p_1 + p_2 = 1$ et c'est une évidence. Or, ici, dans le cas général, on a

$$\frac{1}{a} + \frac{1}{b} \geq 1$$

Donc clairement, on n'est pas dans un cadre de surebet et clairement cela montre que $\frac{1}{a}$ et $\frac{1}{b}$ ne sont pas des probabilités, même si on sent que ces grandeurs y sont fortement liées.

2.2 Estimer les probabilités à partir des côtes

2.2.1 Marges du bookmaker et taxes nationales

En revanche, on peut chercher à estimer p_1 et p_2.
Notons $\alpha = \frac{1}{a} + \frac{1}{b}$ et appelons cette grandeur "facteur de pénalité" ou encore "facteur de marge".
Une première idée toute simple est de repartir de l'équation précédente et de tout diviser par α, sorte de normalisation des grandeurs utilisées, de façon à bien avoir quelque chose égal à 1 et donc à obtenir une approximation des probabilités recherchées.
Ainsi donc, on a

$$\frac{1}{\alpha a} + \frac{1}{\alpha b} = 1$$

On peut alors voir les termes $\frac{1}{\alpha a}$ et $\frac{1}{\alpha b}$ comme des approximations de p_1 et p_2.

Explicitions p_1 et p_2 :

$$p_1 = \frac{1}{\alpha a} = \frac{1}{a} \times \frac{1}{\frac{1}{a} + \frac{1}{b}} = \frac{1}{a} \times \frac{1}{\frac{b+a}{ab}} = \frac{1}{a} \frac{ab}{b+a} = \frac{b}{b+a}$$

De la même façon,

$$p_2 = \frac{1}{\alpha b} = \frac{1}{b} \frac{1}{\frac{1}{a} + \frac{1}{b}} = \frac{a}{a+b}$$

Par cette technique on arrive à approcher les probabilités que le bookmaker a estimé pour l'évènement sportif concerné. D'une part, ce calcul peut être différent de quelques pourcents de l'estimation réelle (mais inconnue) faite par le bookmaker et d'autre part, peut-être que le bookmaker se trompe dans son estimation. Il ne faut pas donc prendre cette estimation pour argent comptant, mais en même temps, on voit mal comment obtenir quelque chose de plus précis et fiable avec des outils grand public. On n'aura pas d'autre choix que de s'en contenter.

2.2.2 Écart relatif des côtes

Une autre façon d'approcher les probabilités p_1 et p_2 est d'étudier l'écart relatif entre les deux côtes. Si le premier évènement est fort probable, alors le deuxième sera fortement improbable, ce qui signifie que la côte a sera basse et la côte b sera élevée. L'écart relatif de a et b renseigne donc sur l'écart relatif entre p_1 et p_2. En particulier, on voit que le rapport $\frac{b}{a}$ va renseigner sur p_1 et que le rapport $\frac{a}{b}$ sur p_2.

Si la cote a est deux fois plus grosse que b, cela signifie que la probabilité p_1 est deux fois plus faible que p_2.
De même, si on a $a = b$, on voit bien que les rapport de force est équilibré : chaque évènement aura donc 50% de chances d'arriver. Pourtant, a et b ne valent pas forcément 2... Parfois, on a $a = b = 1,7$, en particulier en live betting lorsqu'un évènement sportif est particulièrement indécis : le bookmaker ne prend pas de risque à surpayer le parier en créant des "value" intéressantes pour nous.

D'une façon plus générale, on voit bien que finalement, les valeurs de a et b prises seules peuvent donner un ordre d'idée, mais qu'en ayant ces deux informations simultanément, on peut en déduire la valeur de p_1 et p_2 en poussant cette idée dans ses retranchements. En particulier, si on note $\beta = \frac{a}{b}$, cela signifie que le joueur 2 a β fois plus de chances de gagner que le joueur 1. Et inversement, le joueur 1 a $\frac{1}{\beta}$ fois plus de chances de s'imposer dans son match contre le joueur 2.

Ainsi donc on obtient les deux égalités suivantes : $p_1 = \frac{b}{a}p_2$ et $p_2 = \frac{a}{b}p_1$.
Développons ces calculs :

$$p_1 + p_2 = 1 \Leftrightarrow \frac{b}{a}p_2 + p_2 = 1 \Leftrightarrow (\frac{b}{a} + 1)p_2 = 1 \Leftrightarrow p_2 = \frac{a}{a+b}$$

De même pour p_2 on obtient :

$$p_1 = \frac{b}{a}p_2 = \frac{b}{a}\frac{a}{a+b} = \frac{b}{a+b}$$

On constate que ces deux approches aboutissent à la même estimation des probabilités. On les gardera donc telles quelles comme point de repère et pour nous aider dans la prise de décision. A présent que nous disposons d'une estimation des probabilités de victoire de chacun des joueurs, on peut commencer à chasser des côtes intéressantes pour une stratégie long terme.

3 Long terme et espérance mathématique

3.1 Le cas général : paris fermés

Vouloir adopter une stratégie long terme est un bon choix en paris sportifs car c'est une stratégie qui est probablement la plus sûre : on peut avoir des périodes où l'on fait de gros gains, des périodes où ça va mal, mais si on dézoom et qu'on prend du recul, sur un an ou sur plusieurs années, en moyenne on fait des gains. Cette notion de "gain moyen" est en fait directement liée à la notion d'espérance mathématique issue du domaine des probabilités : c'est sa traduction pratique dans le cadre des paris. Notée $E[X]$ où X est le gain net moyen, l'espérance mathématique de nos gains est calculée comme suit :

$$E[X] = p_1 \times (ax - x - y) + p_2 \times (by - x - y) = axp_1 + byp_2 - (x + y)$$

Dans le cas général, on peut tout à fait envisager, selon notre stratégie, de miser à la fois sur le joueur 1 et 2. L'un des paris visera à espérer un gain et l'autre à diminuer la perte (ou de tenter un surebet en live betting pour quoi pas). En résumé, x et y peuvent être tous les deux strictement positifs. Par conséquent, il n'y a pas vraiment de simplifications à faire pour extirper des informations sur a et b dans le cas général à partir de cette expression de l'espérance mathématique. D'ailleurs, dans l'équation, j'ai noté x et y pour parler des mises, mais dans le cadre général, il faudrait parler de la somme totale des mises sur le joueur 1 et la somme de mises sur le joueur 2. Mais bon, ça viendrait compliquer la lisibilité de la chose pour une situation à peu près jamais rencontrée : on parie rarement 15 fois sur le même match, même si c'est évidemment possible.

Dans cette expression, on a $axp_1 + byp_2$ qui représente ce qu'on gagne en moyenne et $(x + y)$ nos mises. La différence des deux donne donc bien un bénéfice (ou une perte) moyenne pour le type d'évènement sportif considéré. Il faut se dire que si on joue ce match (ou le même type de match) 10 000 fois, on gagnera 10 000 fois l'espérance mathématique au bout du compte. Il faut donc éviter que celle-ci soit négative !

3.2 Le cas du pari simple

Dans ce contexte, on se contente d'émettre un pronostique et de miser en accord avec celui-ci (en live betting ou en pré-match, peu importe). Dans un premier temps, on considère que notre pronostique se porte sur le joueur 1 (mais les calculs seront identiques pour le joueur 2). Comme on se porte sur le joueur 1, notre mise x et strictement positive et y est nulle. Donc on a :

$$E[X] = p_1(ax - x - y) + p_2(by - x - y) = p_1(ax - x) - p_2x = p_1ax - x = x(p_1a - 1)$$

Par conséquent, quelle que soit notre mise x, on a :

$$E[X] > 0 \Leftrightarrow a > \frac{1}{p_1}$$

De la même façon, si notre stratégie consiste à miser sur le joueur 2, on obtient :

$$E[X] > 0 \Leftrightarrow b > \frac{1}{p_2}$$

Dans le cas du paris simple comme c'est le cas dans le cadre de ce paragraphe, on a donc $E[X] = x(p_1a - 1)$ sur une stratégie de mise sur le joueur 1 et $E[X] = y(p_2b - 1)$ pour une stratégie de pari sur le joueur 2. Par ces deux expressions, on peut voir qu'une variable intéressante serait de considérer $\frac{E[X]}{x}$ pour le joueur 1 et $\frac{E[X]}{y}$ pour le joueur 2 : ces deux grandeurs expriment alors le pourcentage moyen de gain par paris... Pratique pour modéliser la trésorerie à long terme...

On a donc présenté les côtes minimales pour a et b qui nous permettent théoriquement d'assurer des gains moyens et une stratégie pérenne de paris. Théoriquement.

3.3 Et après ?

Après toutes ces informations mathématiques, on dispose de plein d'outils supplémentaires pour nous aider dans nos décisions. En revanche, ces faits et estimations scientifiques ne sont pas à prendre au pied de la lettre. Un pari sportif reste un paris sportif : si on peut avoir des critères scientifiques pour nous aider dans nos choix et nos stratégies c'est une excellente chose, mais il ne faut surtout pas négliger le côté sportif. Oui, c'est la base, j'enfonce une porte ouverte, mais il me semble important de le dire très explicitement.

Alors que faire après toutes ces étapes ? Une fois que nous avons établis quelques indicateurs de base pour nous aider dans nos pronostiques (probabilité, calcul de côte intéressante pour avoir une certaine rentabilité moyenne...), tout reste encore à faire :

— Bien connaitre les règles du sport auquel on s'intéresse... évident, oui, mais ça montre qu'une équation peut dire des choses mais si les règles du sport ou du site de paris en dit d'autres... il vaut mieux prendre du recul !

— Toujours regarder les matches en temps-réel dans le cas du live-betting (ne pas se fier uniquement aux côtes et probas),
— Construire une stratégie (un ou plusieurs paris, une ou plusieurs mises, fermer un paris, renverser le pronostique si un évènement précis survient...),
— Anticiper : définir des scénarios établis en avance, sniping d'évènements nous étant favorables,
— Faire attention aux variables cachées et imprévus (adversaire "bête noire", retour de blessure, fatigue, motivation, envie pressante - si si, ça s'est vu...),
— Définir des règles autour de sa stratégie et les respecter toujours,
— Gérer les émotions et le stress, c'est-à-dire s'en tenir aux règles !
— Tenir son cahier de paris pour analyser ses résultats : voir si on a gagné avec de la chance ou si le succès a été bien construit. De même pour les pertes : aléa ou erreur de stratégie, entorse aux règles établies... bref, comprendre ses erreurs et progresser en ne les renouvelant pas.

Conclusion

On a donc vu dans ce document ce qu'était une côte, en quoi elle était reliée aux probabilités des évènements lui correspondant, comment calculer la rentabilité moyenne d'un paris dans une stratégie long-terme et construire des surebet en live betting par exemple. Nous avons donc pas mal balayé mais ce ne sont là que des bases de travail. Il y aurait bien des choses à dire, notamment autour du paradoxe de Monty-Hall, en particulier voir si celui-ci s'applique aux paris sportifs ou non, auquel cas il y aurait un gros filon à suivre ! On peut également tenter de construire des martingales en se disant qu'un joueur précis n'allait pas toujours gagner et toujours perdre. L'idée est intéressante, mais il faut garder à l'esprit que c'est une martingale, donc c'est très risqué ! Enfin, tout cela pour dire qu'avec ces données là, on peut construire bien des choses : seules vos erreurs ou votre créativité seront vos freins !