

A propos des notions de Moyenne, de Pourcentage de réussite et de Ranking

J-Pol Guillement

Pour quantifier le niveau d'un joueur au billard la "moyenne" est le moyen le plus consensuel. Ce n'est pas le seul. Il y a aussi la "probabilité (ou pourcentage) de réussite" et le "Ranking". Nous allons rappeler les définitions et indiquer les relations entre ces notions. On trouvera (en fichier attaché) les détails et des explications sur le calcul du Ranking, et une "discussion finale" sur son utilisation.

1. La moyenne.

La moyenne au cours d'une partie est **par définition** le nombre de points effectués divisé par le nombre de reprises. La moyenne sur un ensemble de parties est le nombre de points effectués sur l'ensemble des parties divisé par le nombre total de reprises.

Le calcul de la moyenne est simplissime, à la portée de tout un chacun, aucune mauvaise interprétation rendant le résultat erroné n'est possible.

La définition donnée correspond à la moyenne probabiliste des séries de chacune des reprises :

Par exemple considérons une partie dont les dix reprises sont

s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
2	5	0	7	15	1	2	0	5	13

La moyenne est

$$\Sigma s_i / 10 = (2 + 5 + 0 + \dots + 13) / 10 = 50 / 10 = 5.00.$$

C'est la "moyenne probabiliste" de la variable aléatoire s_i représentant la série obtenue à la reprise i .

On entend parfois dire que la moyenne attribuée à un joueur est **fausse**, suivant en cela, paraît-il, le grand champion Belge du moment.

Une définition n'est ni juste ni fausse, elle peut être mal formulée, sans intérêt, son calcul peut être faux, son utilisation mal appropriée, mais le qualificatif "faux" ne peut s'appliquer à une définition.

2. La moyenne calculée sur un ensemble de parties n'est pas la moyenne des moyennes. (Ce serait vrai si toutes les parties avaient le même nombre de reprises).

Considérons l'exemple d'une deuxième partie dont les séries sont

s_1	s_2	s_3	s_4	s_5
20	30	10	30	10

La moyenne de la deuxième partie est $100/5 = 20$, la moyenne sur l'ensemble des deux parties est $(50 + 100)/(10 + 5) = 150/15 = 10$, la moyenne des moyennes est $(5 + 20)/2 = 12.5$

3. Probabilité (ou pourcentage) de réussite.

La notion de probabilité de réussite est intéressante. Elle ne fait intervenir ni la notion de reprise ni celle de série. Elle se mesure à chaque tentative. C'est la probabilité de faire le point à chaque coup de queue. A chaque essai on marque 0 ou 1 point. La probabilité de réussite est le nombre de points marqués divisé par le nombre de tentatives. C'est un nombre compris entre 0 et 1. (Le pourcentage de réussite est 100 fois la probabilité de réussite).

Nous allons voir qu'il y a une relation simple entre moyenne et probabilité de réussite, **à la condition de supposer que la probabilité de réussite reste constante** au cours de chaque tentative. Cette hypothèse n'est pas valable à la partie Libre où le rôle du rappel est essentiel.

En notant m la moyenne et p la probabilité de réussite, la relation est

$$p = \frac{m}{1+m}, \quad m = \frac{p}{1-p}$$

qui montre, avec l'hypothèse indiquée ci-dessus, que moyenne et probabilité de réussite sont "mécaniquement" liés.

Justifions ces relations en deux temps, d'abord avec l'exemple d'un joueur qui fait 1 de moyenne et 0.5 de probabilité de réussite, et ensuite de façon plus générale et plus mathématique.

- (a) i. Considérons un joueur qui fait 1 de moyenne :
 Sur 100 reprises il fait 100 points et il a 100 échecs (on ignore le dernier coup).
 Il a tenté 200 coups, sa probabilité de réussite est 1/2, c'est cohérent avec la formule

$$p = \frac{m}{1+m}$$

- ii. Inversement considérons un joueur qui a une probabilité **constante** 1/2 de faire le point à chaque fois qu'il joue.
 Sur 100 reprises, il retourne s'asseoir sans faire de point 50 fois, et 50 fois réussit le premier point.
 Quand il fait le premier point, il a une chance sur deux de faire le second. Il fait donc (théoriquement) 25 fois le second point. Et ainsi de suite ...
 Au total, il fait (théoriquement)

$$50 + 50/2 + 50/4 + \dots = 100 \text{ points,}$$

et donc 1 de moyenne.
 C'est cohérent avec la formule

$$m = \frac{p}{1-p}$$

- (b) Justification plus générale.
 Soit X le nombre de points marqués au cours d'une partie de r reprises. Le nombre de coups joués est (en ignorant la possibilité de série arrêtée à la dernière reprise)

$$n = X + r$$

La moyenne est

$$m = \frac{X}{r}$$

Quand on a une probabilité de réussite **constante** p , à chaque coup joué on marque p points. Sur l'ensemble des n coups on marque $X = n \times p$ points. Il en résulte que la moyenne est

$$m = \frac{X}{r} = \frac{pn}{r} = p \frac{X+r}{r} = p(m+1)$$

et donc

$$p = \frac{m}{1+m}, \quad m = \frac{p}{1-p}$$

4. Probabilité de réussite sur une position laissée par l'adversaire (adverse) et sur une position obtenue à la suite d'un point réalisé (propre).

Dans le paragraphe précédent on a fait l'hypothèse que la probabilité de réussite serait constante. Mais, en réalité, elle varie avec la qualité de regroupement et avec le souci de ne pas faciliter la tâche de l'adversaire ...

À la partie Libre la probabilité de réussite évolue de façon complexe au cours d'une série, mais au 3-Bandes, où les séries sont faibles, on peut limiter l'étude de la probabilité de réussite au cas de la position adverse et de la position propre.

Pour déterminer ces deux valeurs, en plus du nombre de points et du nombre de reprises, une donnée

supplémentaire est nécessaire, celle du nombre de traits. On dit qu'un joueur fait un trait, quand il fait une "série de 0 point" (le marqueur fait un trait sur la feuille de match).

Notons pts le nombre de points marqués, r le nombre de reprises, t le nombre de traits, p_a et p_p les probabilités de réussite sur position adverse et sur position propre. On a

$$p_a = \frac{r - t}{r}, \quad p_p = \frac{pts - (r - t)}{pts}$$

Justification

(a) Position adverse :

On joue r fois sur position adverse (y compris le point d'entrée), on marque 1 point quand on ne fait pas de trait, c'est-à-dire que l'on marque au total $(r - t)$ points sur position adverse. D'où la formule

$$p_a = \frac{r - t}{r}$$

(b) Position propre :

On joue sur position propre quand on a marqué un point, cela arrive pts fois. Les points marqués dans ces conditions sont pts moins ceux qu'on a marqué sur position adverse, c'est-à-dire $pts - (r - t)$. D'où la formule

$$p_p = \frac{pts - (r - t)}{pts}$$

Il est intéressant de regarder ces deux valeurs pour évaluer la qualité de repositionnement des joueurs de 3-Bandes.

5. Le Ranking.

Le Ranking (classement) est un concept statistique très élégant destiné à "prédire" le résultat d'une rencontre entre deux adversaires.

Le calcul du ranking d'un joueur fait intervenir ses résultats au cours des 18 mois précédents, mais aussi ceux de ses adversaires. Il tient compte des moyennes, des victoires et des défaites. Il n'est possible que par ordinateur et par accès à la base des résultats de la Fédération.

Voir ci-joints les pages du Code Sportif FFB relatives et une explication des formules utilisées.

Il est à noter que les auteurs des formules de ranking ont fait en sorte que le ranking (le coefficient) soit voisin (à 15% près) de 1000 fois la moyenne, aussi bien pour le 3-Bandes que pour la partie Libre. On peut le constater dans les colonnes "Ranking" et "MG" des tableaux suivants :

RANG	NOM	Ranking	Matches	MG	MG Adv	Esp Gain (%)	Gain (%)	POS	Ranking 31/2/2015	Rang 31/2/2015
42	LELIEVRE Vincent	824	44	0,822	0,836	47,4	47,7	1	825	43
131	SWITALA Joël	667	48	0,689	0,709	45,7	40,6	2	676	118
137	MARZIALE Geoffrey	656	12	0,602	0,638	41,1	54,2	3	655	140
142	STEIN Andreas	653	29	0,625	0,659	41,9	48,3	4	642	155
145	DERRADJI Samir	650	34	0,662	0,655	51,5	48,5	5	657	136
198	DAVID Nicolas	613	21	0,615	0,585	57,6	57,1	6	590	234
213	BENESTEAU Jean-Michel	605	9	0,662	0,626	58,6	44,4	7	602	216
229	THIERY Philippe	591	8	0,523	0,431	77,0	87,5	8	591	232
239	LE MAY Eric	584	20	0,558	0,519	61,2	67,5	9	585	239
271	MORALES Luis	565	26	0,537	0,495	62,3	69,2	10	565	270
280	BELLO François	559	39	0,553	0,602	37,0	38,5	11	556	284
308	BOUREL Loïc	545	6	0,477	0,544	30,5	50,0	12	545	312
329	DUPUIS Antoine	536	49	0,533	0,491	62,5	63,3	13	537	329
339	TELOT Patrick	532	31	0,514	0,494	56,2	61,3	14	532	340

Ranking 3-Bandes Pays de la Loire mars 2015

RANG	NOM	Ranking	Matches	MG	MG Adv	Esp Gain (%)	Gain (%)	POS	Ranking 31/2/2015	Rang 31/2/2015
2	LEGROS Benoit	11 029	13	113,60	79,95	78,8	76,9	1	11 029	2
41	PELOUIN Jean-François	1 493	29	16,57	9,79	87,4	82,8	2	1 493	41
77	BLUSSEAU Nicolas	1 075	27	9,69	10,96	38,6	48,1	3	1 069	76
83	BOUYER Daniel	1 034	17	8,99	11,13	30,9	47,1	4	1 034	83
94	FERNANDES Antonio	979	26	9,71	10,63	41,6	42,3	5	975	92
97	PRIOU Denis	961	18	9,63	9,89	47,5	47,2	6	961	97
98	AGNAN Joël	960	38	9,55	6,43	81,3	81,6	7	960	98
109	MERCIER Jean-Philippe	897	22	7,80	6,25	69,6	81,8	8	910	105
121	ARNAUDEAU Fabrice	853	42	8,29	7,09	64,4	66,7	9	852	120
131	NORAIS Jean	833	16	7,52	6,75	60,1	68,8	10	830	128
160	HERVE Jacky	757	15	8,48	7,04	66,8	56,7	11	757	159
166	VERA Gilles	747	33	7,95	6,54	67,5	62,1	12	732	168
175	PERRINEL Pascal	728	29	6,33	6,30	50,4	69,0	13	728	174

Ranking Libre Pays de la Loire mars 2015

6. Discussion : moyenne vs ranking et FFB Sportif..

Prédire le résultat d'une rencontre avec la seule indication des moyennes est illusoire. D'abord à cause du caractère très aléatoire du billard, mais aussi parce que la moyenne ne permet pas de faire la différence par exemple entre un joueur qui pratique la défense, ce qui tend à faire chuter la moyenne et à favoriser la victoire, et un joueur qui ne se préoccupe que du rappel, au bénéfice de la moyenne et au détriment de la victoire.

Le ranking, qui prend en compte victoires et défaites, niveau des adversaires, semble beaucoup mieux armé. Mais s'il affine le critère de la moyenne, il ne peut pas, lui non plus, prendre en compte toute la subtilité des résultats. Comme par exemple le fait de ne perdre que par un ou deux points.

Il y a dans les tableaux de ranking des anomalies qui sautent aux yeux : ranking anormalement élevé pour les joueurs ayant participé à peu de matchs, majoritairement gagnés. (Augmenter la durée de prise en compte des résultats n'aurait pas de sens, ce qui compte c'est le niveau actuel d'un joueur, pas celui d'il y a 3 ans)

D'autre part les chiffres du ranking ne sont pas parlants : on voit mal dans un club deux joueurs ne se connaissant pas, échanger, avant une partie amicale, leur ranking plutôt que leur moyenne !

Le ranking est un superbe concept mathématique, complexe, mais je ne suis pas convaincu qu'il supplante la moyenne de l'année en cours si simple à appréhender et à calculer, sauf peut-être pour le haut niveau.

Par ailleurs le FFB **Sportif** contient beaucoup d'erreurs de saisie. Les formulaires de saisie devraient refuser les données erronées du genre

- des R4 qui jouent sur 3m10,
- des séries supérieures au nombre de points,
- des écarts de reprises entre deux adversaires (on ne parle pas de la différence d'une unité, mais du vainqueur qui fait une moyenne inférieure à celle du perdant ...),
- la date absente dans le label de la poule,

Et il devrait y avoir des procédures de vérification a posteriori.

D'autre part pour justifier l'utilisation du ranking plutôt que l'utilisation de la moyenne en cours, il serait intéressant d'avoir une procédure qui compare en permanence les prévisions qu'on peut en déduire, match par match. Cette procédure fournissant un compte rendu accessible sur le FFB Sportif du genre

- Telle semaine, tel mois, telle année, il y a eu tant de matchs de Libre, de 3-Bandes ... Les prévisions basées sur la moyenne en cours ont eu tel taux de succès, celles basées sur le ranking tel autre...