

Conf'Lunch Hasard et Football

Fabien Torre

INRIA Lille - EPI Mostrare

Vendredi 30 mars 2012

- 1 Humains et ordinateurs face au hasard
- 2 Estimer le championnat de football
- 3 Gagner (plus) au Loto

Championnat de football ou loto, lequel est le plus aléatoire ?

L'être humain et l'aléatoire

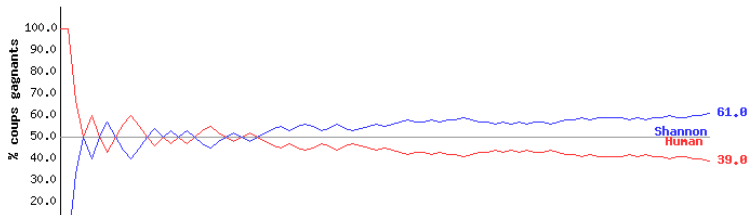
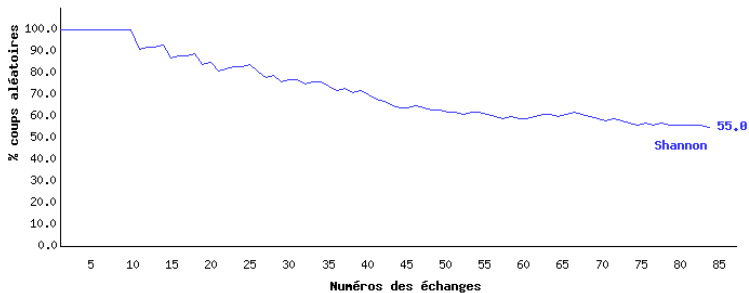
Difficulté à le reconnaître

10 02 06 05 18 11 19 08 15 03 07 18 10 12 10 20 14 06 11 14
07 01 11 01 19 20 14 08 14 03 11 06 01 12 19 20 01 16 09 01
01 13 11 17 06 08 16 04 18 14 09 05 19 03 15 10 20 02 12 07
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20

Difficulté à le produire

- Choisir pile ou face, mais différemment de la machine,
- *a priori*, il faut jouer au hasard pour obtenir le nul...
- *Mind Reading Machine* de [Claude Shannon, 1953] : mémorise 8 situations et les actions de l'humain associées...
- **une partie ?**

Pile/Face contre Shannon



Nécessaire à l'être humain

- Pour être imprévisible : bluff au poker, penalty au football, etc.
- problèmes (conjugaux) mal posés :
 - quelle teinte pour le nouveau canapé ? **rose !**
 - de quelle couleur peindre l'escalier ? **bleu !**
- éviter manipulation et pièges abscons
[Petit traité de manipulation à l'usage des honnêtes gens, Robert-Vincent Joule et Jean-Léon Beauvois, 1987, mai 2012],
- quand on manque d'information ou quand on ne comprend pas les conseils donnés
[La machine qui gagna la guerre, Isaac Asimov, 1961].

Et pour les informaticiens ? aléatoire facile à produire ? dans quel but ?

Interlude



Facile pour les informaticiens

Production

aléatoires	10 02 06 05 18 11 19 08 15 03 07 18 10 12 10
pseudo-aléatoires	07 01 11 01 19 20 14 08 14 03 11 06 01 12 19
quasi-aléatoires	01 13 11 17 06 08 16 04 18 14 09 05 19 03 15
réguliers	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15

Intérêts multiples

- Accélérer un algorithme en l'affaiblissant (**grippe ici ?**),
- se rapprocher d'une complexité en moyenne, éviter la complexité au pire (**tri rapide**),
- produire de la diversité (**chaises IKEA**),
- pallier un manque de connaissances... (**Asimov, football ?**)

Interlude 2



Chaises obtenues par algorithmes génétiques [EPI TAO & EZCT].

Championnat en cours : dernière journée jouée

ParisSG	1	1	Bordeaux
Nice	1	1	Marseille
Lyon	2	1	Sochaux
Montpellier	1	0	SaintEtienne
Evian	0	3	Lille
Ajaccio	1	1	Lorient
Toulouse	1	0	Auxerre
Valenciennes	1	0	Rennes
Brest	0	1	Nancy
Dijon	2	0	Caen

Derniers résultats connus,
reste 9 journées à jouer.

Championnat en cours : classement actuel

<i>position</i>	<i>équipe</i>	<i>points</i>
1	Montpellier	60
2	Paris Saint-Germain	60
3	Lille	53
4	Toulouse	50
...
7	Rennes	44
...
14	Lorient	32
15	Ajaccio	31
16	Brest	30
17	Caen	29
18	Nice	28
19	Sochaux	27
20	Auxerre	24

Classement établi
après la 29ème journée.

Questions sur le championnat en cours

Motivation

Donner à quelqu'un qui ne suit le championnat que de loin, une intuition sur des questions de classement.comme...

- Rennes peut gagner le championnat ?
- Rennes forcément devant Lorient ?
- Lorient relégué ? Lorient dernier ?
- avec quelle *probabilité* ?
- et combien de points pour avoir une chance ? pour être sûr ?

Comment ?

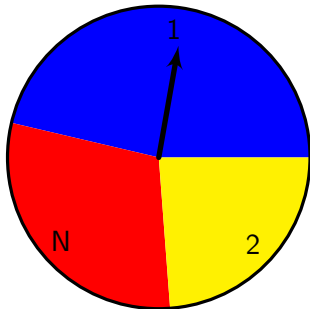
- Méthode de type *Monte Carlo*,
- décider de la fin des matches non joués... au hasard !
- comment terminer un match au hasard ?

Décider au hasard de la fin d'un match

Fin d'un match : décision 1N2

Observations des données passées

Résultats	Fréquences
1	46.32 %
N	29.87 %
2	23.81 %

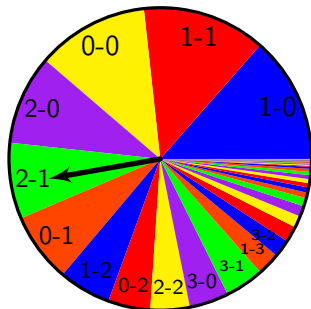


Décider au hasard de la fin d'un match

Fin d'un match : décision par scores

Observations des données passées

Résultats	Fréquences
1-0	13.53 %
1-1	13.20 %
0-0	11.97 %
2-0	9.58 %
2-1	8.13 %
0-1	7.43 %
1-2	5.58 %
0-2	4.51 %



Application à la prochaine journée

Rennes	-2	-0	Lyon
Marseille	-1	-0	Montpellier
Nancy	-0	-2	ParisSG
Bordeaux	-0	-0	Dijon
SaintEtienne	-1	-1	Nice
Lille	-1	-0	Toulouse
Lorient	-0	-0	Evian
Auxerre	-1	-0	Valenciennes
Caen	-1	-0	Ajaccio
Sochaux	-2	-0	Brest

Fins de matches
choisies au hasard.

Idem pour les autres
journées non jouées.

Une fin possible...

<i>position</i>	<i>équipe</i>	<i>points</i>
1	Rennes	66
2	Montpellier	65
3	Toulouse	64
4	Lille	62
5	Paris Saint-Germain	60
6	Lyon	60
...
14	Nancy	46
15	Caen	45
16	Nice	41
17	Sochaux	40
18	Auxerre	38
19	Brest	36
20	Lorient	35

Classement établi à partir
des matches déjà joués
et des matches aléatoires.

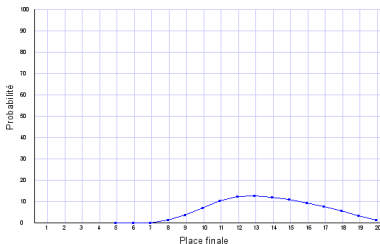
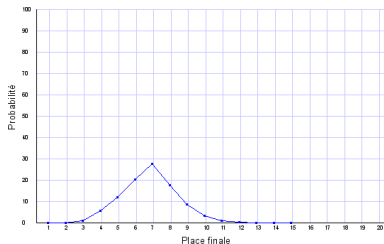
Absurde ! mais permet de
répondre en partie.

On va recommencer !
1 000 000 de fois ?

Les réponses

Rennes champion : 0.01 % des cas.
espoir 65 points, pas de certitude.

Lorient relégué : 10.60 % des cas.
espoir 35 points, certitude 46 points.



Rennes devant Lorient : 98.97 % des cas.

Rennes : espoir 44 points, certitude 59 points.

Lorient : espoir 44 points, pas de certitude possible.

(voir les chiffres et courbes pour votre équipe)

Discussions

Faire mieux ?

- Mais pourquoi ? impossibilités/certitudes/points sont justes,
- seule question : les *pourcentages de cas observés* sont-ils les *probabilités réelles* ? sans doute pas... mais proches ?
- quel effort nécessaire ? pour parvenir à quoi ? évaluation ?

Justifications du modèle naïf

- L'issue d'un match de football est souvent aléatoire !
[*La glorieuse incertitude du football* par John Wesson,
Pour la Science, novembre 2002],
- 70 % des matches finissent avec une différence de 0 ou 1 but,
- sensible à la suppression d'une équipe... (*essayez !*)

L1 en 2008-2009	1. Bordeaux 80	2. Marseille 77
sans Le Mans (16ème)	1. Marseille 75	2. Bordeaux 74

Au Loto

Il y a parfois une part d'incertitude dans les jeux de hasard, Joe.

[William Dalton dans « Dalton City »]

Caractéristiques du jeu

- Tirage aléatoire de 5 numéros parmi 49, donc imprévisible ;
- pari mutuel : vous gagnez ce que les autres perdent.

Questions

- Est-ce les gens jouent au hasard ?
- utilisent-ils le *Système Flash* ?
- s'ils ne jouent pas au hasard, comment font-ils leurs paris ?
- quels numéros jouer au loto pour maximiser ses gains ?
(ou minimiser ses pertes...)

Apprentissage automatique du comportement des joueurs

À partir de 350 tirages/gains de la Française des Jeux :

tirages à petits gains

4	5	6	7	11
3	9	10	24	33
1	7	11	26	49
3	13	22	26	48
2	8	9	42	49
10	12	13	23	37
7	19	24	35	42
5	12	21	29	42

tirages à gros gains

2	4	37	40	44
16	31	36	39	41
30	36	38	43	47
4	14	22	35	38
6	29	32	44	45
13	18	32	43	46
3	28	37	38	42
9	31	35	40	48

Une théorie apprise pour le loto

Règles portant sur la présence/absence/moyenne des numéros.

règles identifiant les tirages rapportant plus

7	11	13	19	[25.4; 42.2]
---	----	----	----	--------------

6	7	9	10	11	12	22	27	29	[21.6; 39.0]
---	---	---	----	----	----	----	----	----	--------------

3	5	8	9	12	13	19	20	23	25	47	[22.6; 37.8]
---	---	---	---	----	----	----	----	----	----	----	--------------

7	12	21	22	24	26	27	49	[22.0; 34.2]
---	----	----	----	----	----	----	----	--------------

règles identifiant les tirages rapportant moins

35	37	[6.6; 23.0]
----	----	-------------

17	33	34	38	39	41	45	46	[6.6; 25.2]
----	----	----	----	----	----	----	----	-------------

1	2	4	16	20	32	34	36	41	[17.0; 27.2]
---	---	---	----	----	----	----	----	----	--------------

1	2	11	14	23	29	32	33	34	36	39	45	46	[16.0; 28.4]
---	---	----	----	----	----	----	----	----	----	----	----	----	--------------

Huit règles compréhensibles et prédictives (90% de réussite).

Apprentissage de combinaisons

Un résultat théorique

Plusieurs règles qui votent prédiront mieux qu'une seule règle...
à condition que :

- chacune des règles prédise un peu mieux que le hasard,
- les règles soient différentes les unes des autres.

Un double-objectif en *apprentissage automatique*

- Utiliser le hasard pour produire des règles diverses...
- en contrôlant que chacune soit un peu meilleur que le hasard !

7053 règles distinctes
qui votent :
99% de bonnes prédictions.

Conclusion

Dans la vie comme en informatique,
le hasard est souvent un allié précieux...



... mais pas toujours : jamais de tirage aléatoire dans un isolement !