

## Problème Pratique de Statistique - 7

# Typologie de golfeurs

Données proposées avec une remarquable introduction par Eric Madigou, INSA 4<sup>o</sup> année BIM, 2003. Question : existe-t-il plusieurs manières d'être un grand joueur de golf ?

*Etant moi-même joueur de golf, j'ai décidé de m'intéresser aux statistiques de l'année 2002 des 50 meilleurs joueurs du PGA tour. Le PGA tour se déroule aux Etats-Unis et est considéré comme le meilleur circuit au monde. Tous les champions actuels s'y côtoient, on peut ainsi y voir évoluer le désormais très célèbre Tiger Woods et d'autres champions bien connus des initiés comme Phil Mickelson ou Ernie Els.*

*Le site [www.pgatour.com](http://www.pgatour.com) regroupe toutes les statistiques de ces joueurs de façon très détaillée et a donc été ma source de renseignements.*

*Dans cette analyse, je vais essayer de « décrypter » toutes ces données diverses avec les outils que je possède. Ainsi je pourrais peut-être faire ressortir des faits intéressants pouvant me permettre de mieux comprendre le golf.*

*Pour les non-initiés, commençons par donner quelques définitions nécessaires à la compréhension de l'analyse (le schéma de la page suivante montre la localisation des différentes aires de jeu) :*

- *Par* : le par est le nombre de coups standard à effectuer sur un trou. Il existe des pars 3, 4 et 5 en fonction de la longueur du trou. Le par est aussi le nombre de coup standard à effectuer sur un parcours. La plupart des parcours sont des pars 72. Faire le par sur un parcours nécessite un niveau de golf très élevé.
- *Birdie* : un joueur fait un birdie quand il joue un trou 1 coup sous le par (par 4 en 3 coups par exemple).
- *Eagle* : un joueur fait un eagle quand il joue un trou 2 coups sous le par. Les eagles se réalisent en général sur les pars 5.
- *Drive* : le driver est le club qui permet de faire les coups les plus longs en distance. Les joueurs utilisent le driver pour jouer le premier coup des pars 5 et souvent celui des pars 4. Ce coup est appelé un drive. Ils ne drivent pas sur les pars 3 car alors ils dépasseraient le trou.
- *Fairway* : le fairway est la zone du terrain où il est préférable d'envoyer sa balle car elle est bien tondu.
- *Rough* : le rough est la zone qui entoure le fairway et qui est peu ou pas tondu. Elle sanctionne un mauvais drive car il est plus difficile de jouer dans le rough que sur le fairway.
- *Green* : le green est la zone du terrain qui entoure le drapeau. Elle est parfaitement tondu pour permettre au joueur de putter.
- *Putt* : un putt est l'action de jouer un coup sur le green. C'est le coup de golf qui nécessite le plus d'application et de précision car il faut étudier la pente du green et contrôler précisément la force du coup.
- *Bunker* : un bunker est une zone de sable censée gêner l'approche du green. Il existe aussi des bunkers de parcours.



Un tournoi du PGA tour se déroule sur 4 jours en 4 rounds (les joueurs jouent le parcours une fois par jour). Après les 2 premiers jours seuls les meilleurs se qualifient pour les 2 derniers rounds. Les autres sont éliminés. On dit qu'ils n'ont pas passé le cut.

J'ai choisi d'étudier 10 statistiques qui définissent d'une façon relativement complète les joueurs. Ces statistiques concernent toute la saison 2002.

- *Driving distance (dd)* : distance moyenne des drives.
- *Driving accuracy (da)* : pourcentage de coups joués au départ qui ont atteint le fairway.
- *Greens in regulation (gir)* : pourcentage de greens touchés en régulation, c'est à dire atteint au bout de 1 coup pour les pars 3, 2 pour les pars 4, 3 pour les pars 5.
- *Putting average (pa)* : nombre moyen de putts joués par trou.
- *Eagle (e)* : nombre de trou joués divisé par le nombre d'eagles.
- *Birdie average (ba)* : nombre de birdie divisé par le nombre de rounds joués.
- *Sand save (ss)* : pourcentage de par effectués alors que le joueur était dans le bunker.
- *Before cut (bc)* : score moyen sur les 2 premiers rounds des tournois.
- *Round 3 (r3)* : score moyen du round 3.
- *final round (fr)* : score moyen du dernier round.

Pour récupérer les données :

```
w=read.table("http://pbil.univ-lyon1.fr/R/pps/pps007.txt",h=T)
summary(w)
```

	dd	da	gir	pa	e
Min.	:264	:57.8	:60.9	:1.70	: 99.5
1st Qu.	:274	:66.3	:65.2	:1.74	:156.6
Median	:280	:68.5	:67.7	:1.76	:202.0
Mean	:280	:68.9	:67.2	:1.75	:250.9
3rd Qu.	:286	:70.6	:69.1	:1.77	:293.4
Max.	:294	:81.2	:74.0	:1.81	:837.0

D. Chessel - Biométrie et Biologie Évolutive - Université Lyon 1

	ba	ss	bc	r3	fr
Min.	:3.29	:29.7	:69.0	:69.3	:67.7
1st Qu.:	3.67	51.0	69.8	69.9	70.0
Median	:3.83	:54.0	:70.3	:70.4	:70.5
Mean	:3.83	:52.8	:70.3	:70.4	:70.5
3rd Qu.:	3.98	56.1	70.7	70.8	70.9
Max.	:4.47	:64.9	:71.6	:72.0	:72.4