

Statistiques pour la psychologie — Licence

Régression linéaire simple

Nicolas Gauvrit
 Université de Metz
<http://adems.free.fr/>

10 octobre 2004

1 Un cas concret

Exercice 1 Des chercheurs¹ ont étudié les résultats en sciences (S) et en orthographe (O) d'élèves de primaire, pour les comparer selon le genre² G ³. Ils trouvent les résultats suivants

| G | S | O | G | S | O |
|-----|-----|-----|-----|-----|-----|
| 0 | 3 | 4 | 1 | 2 | 3 |
| 0 | 4 | 5 | 1 | 5 | 4 |
| 0 | 5 | 5 | 1 | 6 | 5 |
| 0 | 6 | 5 | 1 | 7 | 4 |
| 0 | 7 | 6 | 1 | 8 | 6 |
| 0 | 7 | 7 | 1 | 8 | 9 |
| 0 | 6 | 7 | 1 | 9 | 4 |
| 0 | 8 | 8 | 1 | 8 | 2 |

- Décrivez la situation statistique de deux manières différentes :
 - en supposant qu'il y a une variable "matière" avec pour modalité S et O
 - en supposant que les élèves sont les individus statistiques
- Représentez les données pour faire apparaître sur un unique graphique les deux groupes (garçons et filles), et les résultats en sciences et en orthographe.
- Les résultats de sciences sont-ils liés à ceux en orthographe :
 - chez les garçons ?
 - chez les filles ?
 - globalement ?
- Quelle méthode statistique utiliserait-on pour montrer que les résultats en sciences sont globalement meilleurs chez les garçons, et que ceux en orthographe sont globalement moins bons chez les garçons ?

2 Théorie

Exercice 2 Que dire de r_{XY} dans les cas suivants (réponse intuitive motivée) :

- X est une mesure de la timidité et Y une mesure de bien-être
- X mesure la vitesse de lecture et Y le niveau d'étude
- X est la taille et Y le QI chez des sujets adultes

Exercice 3 Que vaut r sur un échantillon de 2 individus ?

Exercice 4 Si on dispose de deux groupes A et B et d'une variable numérique X , comment utiliser la corrélation linéaire pour comparer X sur les groupes A et B ?

3 Pratique de la régression

Exercice 5 On mesure, sur un ensemble de volontaires, une note de stress (X) et le nombre de consultations Y au cours des trois mois suivants le passage du test de stress. On trouve :

| | | | | | | | | | | |
|-----|---|----|----|---|----|---|----|----|---|---|
| X | 0 | 10 | 15 | 5 | 20 | 3 | 12 | 25 | 2 | 1 |
| Y | 0 | 2 | 3 | 2 | 5 | 2 | 3 | 15 | 1 | 6 |

- Pourquoi ne pas prendre pour Y les consultations au cours des 3 mois précédents le passage du test de stress ?
- Faites une analyse de régression complète et concluez.

Exercice 6 On a relevé les notes de mathématiques (M) et de français (F) obtenues au baccalauréat par des étudiants de première année de psychologie. On trouve $r_{MF} = -.52$. Concluez.

¹Bacharach, V. R. et al. (2003). Racial and gender science achievement gaps in secondary education. *The Journal of Genetic Psychology*, 144(1).

²Les auteurs étudient aussi les différences blancs/noirs, que nous n'évoquons pas ici.

³On notera 0 pour "fille" et 1 pour "garçon".