

Pour mesurer le poids d'un cochon, cherchez une planche solide et posez-la en équilibre sur une barrière. Attachez le cochon à une extrémité de la planche et courez à l'autre bout déposer une pierre sur l'extrémité opposée. Essayez différentes pierres jusqu'à trouver celle qui s'équilibre exactement avec le cochon. Tout ce qui vous reste alors à faire est de déterminer le poids de cette pierre...

Tiré de la *Collection de plaisanteries pour statisticiens* de Gary C. Ramseyer.

- 2.1 La validité d'une mesure
- 2.2 Les erreurs de mesure
- 2.3 Autres considérations sur les mesures
- 2.4 Les échelles de mesure
- 2.5 Organisation des mesures dans un tableau
- 2.6 Exercices

DÉFINITION 2.1

Mesurer, c'est associer des valeurs, numériques ou non, à des observations faites sur des objets, des individus, des groupes...

On parle d'*unité statistique* pour désigner l'objet, l'individu, le groupe... qui est mesuré.

Mesurer, c'est mettre en place une méthode, un protocole, pour appréhender la fréquence cardiaque maximum de Luis Fernandez, le nombre de licenciés d'un club sportif, la catégorie socioprofessionnelle des usagers d'une piscine municipale, le stress d'un compétiteur, la « glisse » d'un ski...

Dans ce chapitre, nous traiterons cinq questions concernant la mesure.

- Est-ce que la mesure représente bien la caractéristique qui est visée (question de validité) ?
- Qu'en est-il des erreurs de mesure (question de fiabilité) ?
- À part la fiabilité, quelles autres qualités peut-on attendre d'une mesure (question d'accessibilité, de « généralisabilité »...) ?
- Est-ce que tous les types de mesure se valent (question d'échelle de mesure) ?
- Comment organiser des mesures dans un tableau ?

2.1 La validité d'une mesure

2.1.1 Les formes de validité

DÉFINITION 2.2

Une mesure est dite *valide* si elle mesure effectivement... l'information que l'on souhaite mesurer.

La validité de mesure de certaines grandeurs physiques comme le poids, la taille, le temps... est assurée dès lors qu'on dispose des instruments de mesure adaptés : balance, toise, ruban gradué, pied à coulisse, plate-forme de force, chronomètre, cadran solaire...

D'autres caractéristiques, pourtant à première vue de mesure évidente, demandent que la méthode de mesure soit clairement précisée. Ainsi, pour estimer le salaire d'un individu, il faut déjà définir s'il est salarié... puis si ce salaire est défini en euros (ou en francs), brut ou net, mensuel ou annuel, avec ou sans primes...

Mais beaucoup de caractéristiques étudiées dans les sciences du sport sont en fait plus difficiles à appréhender : la forme physique, le stress, la performance... C'est pourquoi il est nécessaire d'envisager la validité d'une mesure sous d'autres points de vue.

Relier des mesures numériques

8

Les statistiques sont pourtant formelles, sur dix ailiers, il y en a neuf qui détestent se battre. Il faut que je tombe sur le dernier, dis donc !

Philippe Guillard, *Petits bruits de couloirs*, Paris, La Table ronde, 2001.

8.1 Associer deux variables quantitatives (beaux bébés !)

L'association entre la taille et le poids d'un individu a fait l'objet de nombreux travaux qui ont permis de déterminer des modèles idéaux de relation (formules de Lorentz, de Devine, de Peck...). Nous allons étudier cette association par le biais d'un échantillon très particulier, réunissant les médaillés masculins (OR) français des jeux olympiques de Sydney 2000⁽¹⁾. Pour les épreuves par équipes, on a retenu un seul médaillé, le premier dans l'ordre alphabétique⁽²⁾. Le jeu de données

Tableau 8.1

Taille et poids des médaillés (or) français aux JO de Sydney 2000

Noms	Taille	Poids
Andrieux	192	97
Asloum	165	63
Bette	186	70
Douillet	196	125
Dumoulin	171	64
Estanguet	182	75
Ferrari	187	83
Gané	176	79
Martinez	164	50
Rousseau	182	85

(tableau 8.1) provient du site web du journal *L'Équipe* (www.lequipe.fr).

On a cette fois affaire à deux variables clairement quantitatives. La relation va être décrite graphiquement grâce au nuage de points, quantifiée par une statistique de liaison — le coefficient de corrélation linéaire — ou résumée par une fonction mathématique, la droite de régression, qui permettra en outre de formuler des prédictions.

8.2 Un graphique statistique de liaison : le nuage de points

L'étude de la relation entre deux variables quantitatives passe nécessairement par la réalisation d'un *nuage de points*.

8.2.1 La construction du nuage de points

DÉFINITION 8.1

Le *nuage de points* décrit la relation entre deux variables quantitatives X et Y , mesurées sur les mêmes unités statistiques. Chaque unité apparaît comme un point dont la coordonnée horizontale est sa valeur en X et la coordonnée verticale sa valeur en Y ⁽³⁾.

- 8.1 Associer deux variables quantitatives (beaux bébés !)
- 8.2 Un graphique statistique de liaison : le nuage de points
- 8.3 La corrélation linéaire
- 8.4 Juger l'importance de la corrélation linéaire : le coefficient de détermination
- 8.5 Cas pratique : âge et fréquence cardiaque
- 8.6 Start-istiques (6)
- 8.7 La droite de régression
- 8.8 Prédications et résidus en régression linéaire
- 8.9 La droite de régression aux moindres carrés
- 8.10 Une nouvelle interprétation du coefficient de détermination
- 8.11 Cas pratique : lactates et vitesse du sprinter
- 8.12 Start-istiques (7)
- 8.13 Les liaisons dangereuses
- 8.14 La non-linéarité
- 8.15 Les points extrêmes
- 8.16 Start-istiques (8)
- 8.17 La régression linéaire multiple
- 8.18 Les lisseurs
- 8.19 La régression linéaire robuste
- 8.20 Exercices

⁽¹⁾ Il ne faut donc pas s'attendre à obtenir des résultats généralisables à une population masculine quelconque. Ces médaillés sont jeunes, et pourvus d'une masse musculaire peu commune...

⁽²⁾ Andrieux (Andrieux-Rolland), Bette (Bette, Dorfman, Hocde, Porchier), Gané (Gané, Rousseau, Tournant) et Ferrari (Ferrari, Guyart, Lhotellier, Plumenail).

⁽³⁾ On parle également d'abscisse pour désigner la coordonnée horizontale d'un point et d'ordonnée pour sa coordonnée verticale.