

Table des matières

1 Terminologie	2
2 Description des données	3
2.1 Paramètres numériques	3
2.1.1 Paramètres de position	3
2.1.2 Paramètres de dispersion	4
2.2 Et les moments dans tous ça ?	6
2.2.1 Coefficient d'asymétrie γ_1 : skewness	7
2.2.2 Coefficient d'aplatissement γ_2 : kurtosis	9

MAJ : 25 Janvier 2021... [ici](#)

J'ai ajouté dans cette mise à jour le détail des calculs pour les variances des exemples de votre cours.

Suite à une précision de votre professeur à la fac, j'ai également modifié les calculs des kurtosis.

Cette précision me permet de vous dire que les calculs de γ_1 et γ_2 sont effectués avec les variances des échantillons et non pas des variances estimées.

J'ai également ajouté des graphiques pour illustrer le signe du coefficient d'asymétrie, le skewness.

Tout semble cohérent désormais, si j'ai d'autres informations, vous serez rapidement tenus au courant grâce à la grande réactivité de Céline, merci à elle ;-)

*Comme toujours pour toutes remarques
 (orthographe(s), erreur(s) possible(s) de calculs ou suggestions),
 n'hésitez pas à contacter M. MERCIER à Exactétudes.*

À bientôt !



Dans ce cours de statistiques descriptives, vous avez quasiment tout vu au secondaire. La notion de moments est introduite cette année pour mener des calculs d'estimations pour certains paramètres. Ce petit cours ne se veut, de loin pas, exhaustif. Il sera très certainement amené à s'enrichir d'informations, exemples d'explication au fur et à mesure que vous en aurez besoin.

Accrochez-vous, on y va!

1 Terminologie

Considérons le tableau suivant :

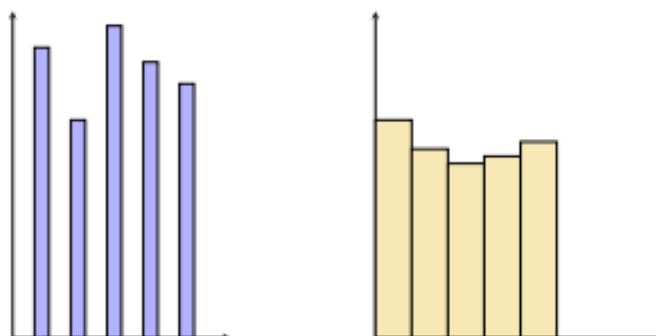
n° de patient	Sexe (0 H/1 F)	Âge	Poids
1	0	43	57
2	1	78	49
...
50	0	68	57

- ✗ Chaque cellule est une donnée.
- ✗ En colonne on peut voir différents caractères. S'ils sont mesurés, on les appelle des **variables**, comme l'âge par exemple.

PLUSIEURS TYPES DE VARIABLES

- ✗ Quantitatives : ce sont des variables qui peuvent être traitées comme des nombres. On peut les mesurer, les ajouter entre elles, etc. Elles peuvent être **discrètes** ou **continues**. Si une variable est discrète on peut alors compter les valeurs prises par cette variable. On peut représenter ses valeurs à l'aide d'un diagramme en bâtons par exemple. Vous avez peut-être entendu parler très récemment (18 janvier 2021) de la possibilité pour les smartwatch de détecter l'apparition de la COVID-19 avant même d'avoir les premiers symptômes? Il « suffirait » (information à vérifier, je ne suis pas médecin 😊) d'avoir en continue la pulsation du coeur en bpm (battements par minutes) et d'observer ses variations. Si les mesures se font toutes les minutes pendant 1 heure, on obtiendra alors 60 valeurs, différentes ou non. On aura alors 60 « bâtons » et la lecture du diagramme serait assez facile.

Si on mesure la pulsation durant une très grande période disons pendant 1 semaines toutes les minutes, il y aura trop de valeurs pour pouvoir utiliser un diagramme en bâtons qui soit pratique à lire. On pourra alors rassembler toutes ces valeurs en ce qu'on appelle des classes. Dans ce petit exemple d'application, on pourrait rassembler ces valeurs par tranche temporelle de 12 h par exemple et en faire la moyenne : on peut de ce fait utiliser un histogramme.



Exemple de diagramme en bâtons (à gauche) et d'histogramme (à droite)

- ✗ Qualitatives : ce sont celles qui ne peuvent pas être mesurées. Elles désignent alors plutôt un aspect visuelle du style la couleur des yeux. Elles peuvent être soit **ordinales** (elles peuvent être classées, ordonnées, comme les différents stades d'une maladie), soit **nominales** (sans hiérarchie, comme les groupes sanguins A, B, O, AB) ou **binaires** (seules deux modalités sont possibles, Garçon/fille, oui/non