

Biostatistiques

Introduction :

La statistique constitue, en biologie, l'outil permettant de répondre à de nombreuses questions qui se posent en permanence au biologiste :

1. Quelle est la valeur normale d'une grandeur biologique, taille, poids, glycémie ?
2. Quelle est la fiabilité d'un examen complémentaire ?
3. Quel est la portée d'un traitement ?
4. Le traitement A est-il plus efficace que le traitement B ?

1. La variabilité et l'incertain :

Toutes ces questions reflètent une propriété fondamentale des systèmes biologiques qui est leur variabilité. Cette variabilité est la somme d'une variabilité expérimentale (liée au protocole de mesure) et d'une variabilité proprement biologique. On peut ainsi décomposer la variabilité d'une grandeur mesurée en deux grandes composantes :

Variabilité totale = variabilité biologique + variabilité métrologique

- La variabilité biologique peut être elle-même décomposée en deux termes : d'une part la variabilité intra-individuelle, qui fait que la même grandeur mesurée chez un sujet donné peut être soumise à des variations aléatoires ; et d'autre part la variabilité inter-individuelle qui fait que cette même grandeur varie d'un individu à l'autre.

Variabilité biologique = variabilité intra-individuelle + variabilité inter-individuelle

La variabilité intra-individuelle peut être observée lors de la mesure de la performance d'un athlète qui n'est pas capable des mêmes performances à chaque essai, mais qui se différencie des autres athlètes (variabilité inter-individuelle). En général, la variabilité intra est moindre que la variabilité inter.

- La variabilité métrologique peut être elle aussi décomposée en deux termes : d'une part les conditions expérimentales dont les variations entraînent un facteur d'aléas ; et d'autre part les erreurs induites par l'appareil de mesure utilisé.

Variabilité métrologique = variabilité expérimentale + variabilité appareil de mesure

La mesure de la pression artérielle peut grandement varier sur un individu donné suivant les conditions de cette mesure ; il est ainsi recommandé de la mesurer après un repos d'au moins 15 minutes, allongé, en mettant le patient dans des conditions de calme maximal. Cette recommandation vise à minimiser la variabilité due aux conditions expérimentales. La précision de l'appareil de mesure est une donnée intrinsèque de l'appareil, et est fournie par le constructeur.

2. La mesure d'une grandeur :

2.1. Unités et équations aux dimensions :

Les grandeurs mesurées possèdent la plupart du temps une unité. La distance entre Annaba et Alger se mesurera par exemple en kilomètres, l'épaisseur d'un cheveu en microns, le poids d'une orange en grammes. Dans quelle mesure ces grandeurs peuvent-elles être comparées ? La distance entre Annaba - Alger et l'épaisseur d'un cheveu sont deux longueurs ; leur comparaison est possible si on les mesure avec une unité commune, par exemple

le mètre. En revanche, le poids de l'orange n'est pas comparable aux longueurs précédentes.

Deux grandeurs qui peuvent être comparées sont dites posséder la même dimension. Elles peuvent être caractérisées par leur dimension : on parlera par exemple de longueur. Les dimensions de toutes les grandeurs physiques peuvent s'exprimer en fonction de sept dimensions de base : la **longueur** notée **L**, la **masse M**, le **temps T**, l'**intensité électrique I**, la **température C**, l'**intensité lumineuse J**, et la **quantité de matière N**.

Par exemple une vitesse est une longueur divisée par un temps. On dira que sa dimension est LT^{-1} . Plus précisément, de l'équation donnant la vitesse v en fonction de la distance d parcourue pendant le temps t , $v=d/t$, on déduit la relation entre les dimensions (notées entre crochets) des deux membres de l'équation $[v] = [d]/[t] = LT^{-1}$. Cette relation est appelée équation aux dimensions. Une équation aux dimensions permet donc d'exprimer la dimension de n'importe quelle grandeur en fonction des dimensions élémentaires, à condition de connaître les relations entre elles. Elle permet aussi une première validation d'une relation entre grandeurs physiques : les dimensions de la partie gauche et de la partie droite de la relation doivent être identiques.

2.2. Erreurs de mesure :

La mesure d'une grandeur ne peut conduire à une valeur exacte. En premier lieu, l'instrument de mesure possède nécessairement une précision limitée : une règle graduée millimètre par millimètre ne peut donner une meilleure précision qu'un demi millimètre. En second lieu, la grandeur à mesurer peut être source de variabilité intra-individuelle : la répétition de la mesure avec le même instrument et dans des conditions identiques conduit alors à des résultats différents. Enfin, l'instrument de mesure peut être mal étalonné

ou mal adapté et conduire à un biais de mesure systématique : les valeurs mesurées seront systématiquement trop élevées, ou systématiquement trop basses.

Pour une grandeur X à mesurer, on note ΔX l'erreur de mesure. Cette erreur est généralement facilement connue si elle n'est due qu'à un problème de précision. S'il existe une variabilité intraindividuelle (raisonnablement faible), on fera intervenir l'écart-type des mesures. Si x est la valeur mesurée, la vraie valeur est donc comprise entre $x-\Delta X$ et $x+\Delta X$.

2.3. Expression de la précision d'une mesure :

Lorsqu'on mesure une quantité sur un individu, par exemple sa Tension Artérielle Systolique (Tas), le résultat obtenu est entaché de plusieurs erreurs provenant de *sources de variabilité* différentes.

La variabilité instrumentale

La TAS mesurée par un tensiomètre souffre d'une erreur de mesure attachée au tensiomètre utilisé : si on répète la mesure dans exactement les mêmes conditions (le mot « exactement » est à bien définir), on ne trouve pas en général la même valeur. La distribution des valeurs mesurées autour de la vraie valeur est Normale, de variance, et de moyenne 0 si l'appareil est exact, de moyenne μ_a si l'appareil présente un défaut de calage. μ_a est la tare d'une balance qu'on corrige pour tenir compte du poids de l'assiette dans laquelle se trouve la farine à peser par exemple.

La variabilité intra individuelle

Lorsqu'on répète la mesure de la pression artérielle chez le même individu, comme d'ailleurs c'est conseillé, on ne trouve pas exactement les mêmes valeurs, même si l'on suppose disposer d'un appareil de précision

parfaite, au point qu'il est considéré qu'un chiffre isolé, non répété, n'a pas de valeur. On parle alors de variation intra individuelle et on peut modéliser cette variation par une nouvelle loi normale, de moyenne 0 et de variance.

3. La décision dans l'incertain

L'objectif de ce cours est de vous fournir les bases indispensables permettant de comprendre les méthodes utilisées, d'interpréter correctement les résultats de nouvelles recherches, et d'adopter un mode de raisonnement qui soit à même d'aider à la décision dans l'exercice de la recherche en biologie.

Plus précisément nous étudierons successivement :

_ La statistique descriptive qui permet de représenter et de quantifier la variabilité d'une ou plusieurs grandeurs observées.

_ La statistique inductive qui inclura les tests statistiques permettant de retenir une hypothèse A plutôt qu'une hypothèse B à partir de données expérimentales (comme dans le cas de la comparaison de deux traitements, où l'hypothèse A est que les deux traitements sont équivalents et l'hypothèse B est qu'ils sont différents).