Le caractère le plus général est le caractère quantitatif. En effet, il permet de mesurer et de comparer les données numériquement, ce qui rend possible l'utilisation de l'ensemble des outils statistiques.

- paramètres de position (moyenne, médiane, mode…)
- paramètres de dispersion (variance, écart-type, écart interquartile…)
- paramètres de forme (asymétrie, aplatissement…)

On voit que tous ces paramètres sont construits à partir de valeurs numériques, et donc que l'application des ces paramètres est uniquement possible si le caractère est quantitatif.

*"Les paramètres statistiques concernent **principalement** les variables quantitatives, et **ponctuellement** qualitatives."*

Les caractères quantitatifs **discrets** correspondent à des valeurs numériques qui sont exclusivement des nombres entiers, isolés à l'intérieur d'un intervalle de variation. Les caractères quantitatifs **continus**, au contraire, sont des valeurs pouvant prendre toutes les valeurs possibles à l'intérieur d'un intervalle de variation. Le plus souvent des nombres décimaux ou des %. On peut observer, en prenant l'exemple de la moyenne et de la médiane, que le calcul de ces paramètres change selon la nature de la variable. La distinction est importante entre caractères quantitatif discret et quantitatif continu car elle permet d'adapter correctement les formules et méthodes (calcul des paramètres statistiques) utilisées lors de l'analyse statistique.

Paramètre de position

Il existe plusieurs façons de calculer des moyennes en fonction de la nature (discrète ou continue) des variables. La moyenne arithmétique est la plus utilisée, mais elle ne suffit pas toujours dans la mesure où elle n'est pas toujours adaptée à la nature des données ou à la situation étudiée. Mais il existe d'autres moyennes qui vont être plus adaptés : la moyenne quadratique , la moyenne harmonique , la moyenne géométrique, et la moyenne mobile. Et donc plusieurs types de moyenne permettent de choisir la formule la plus adaptée au phénomène observé.

La médiane est présentée comme la valeur (m_e) qui coupe une série de données en deux parties d'effectif égal. On calcule une médiane pour obtenir une mesure de position qui ne soit pas influencée par des valeurs extrêmes. Contrairement à la moyenne arithmétique, la médiane dépend du classement des données et non de leur amplitude. Elle permet donc de représenter le centre d'une distribution même si la elle est fortement dissymétrique ou contient des valeurs aberrantes.

Il est possible de calculer un mode (m_o) lorsqu'une série statistique présente au moins une valeur qui apparaît plus souvent que les autres. Le mode correspond à la modalité, la valeur dont l'effectif (ou la densité) est maximal, la plus fréquente. Le mode existe seulement si une valeur se détache clairement, et elle n'est pas toujours unique : certaines distributions peuvent être bimodales ou plurimodales.Ainsi, on peut calculer un mode dès qu'une ou plusieurs valeurs dominantes se dégagent dans la distribution.

Paramètre de concentration

La médiale est utile parce qu'elle partage la valeur totale du caractère étudié en deux parts égales. Contrairement à la médiane, qui sépare les individus en deux groupes d'effectif identique, la médiale sépare selon l'importance totale du caractère (par exemple la masse salariale). Elle sert donc à analyser les situations où la répartition du total est plus significative que la simple répartition des individus. L'indice de concentration de C. Gini sert à mesurer le degré d'inégalité ou de concentration d'une distribution. Il est souvent utilisé pour analyser la répartition d'une variable (par exemple les revenus ou la population). L'indice se mesure entre 0 et 1, plus l'indice est élevé (proche de 1), plus la distribution est concentrée, et plus les inégalités sont fortes. Son intérêt est de donner un indicateur synthétique qui résume la manière dont un caractère (revenus, surfaces, ressources…) est réparti..Ainsi, la médiale et l'indice de Gini permettent d'aller au-delà des mesures de position en apportant une information sur la dispersion et la concentration des données.

Paramètre de dispersion

On ne peut pas utiliser directement l'écart à la moyenne car, lorsqu'on additionne tous les écarts, le résultat est toujours égal à zéro : les valeurs au-dessus et en dessous de la moyenne s'annulent. Pour éviter ce problème, on calcule la variance en élevant les écarts à la moyenne au carré, ce qui permet de mesurer correctement la dispersion des données. Cependant, la variance est exprimée dans une unité au carré, ce qui la rend difficile à interpréter. C'est pourquoi on utilise généralement l'écart type, qui correspond à la racine carrée de la variance. L'écart type est plus parlant, car il s'exprime dans la même unité que la variable étudiée. Ducoup, la variance sert surtout au calcul, tandis que l'écart type est plus simple à comprendre et à utiliser pour décrire la dispersion des données.

On calcule l'étendue car il s'agit d'une mesure simple et facile à calculer pour avoir une première idée de la dispersion des données. Elle correspond à la différence entre la valeur maximale et la valeur minimale d'une série statistique. Cela permet de voir rapidement sur quelle plage les valeurs sont réparties.

Créer un quantile permet de découper une série statistique ordonnée en plusieurs parts contenant le même nombre d'observations. Cela sert à mieux comprendre la répartition des données et à repérer leur position dans la distribution. Les quantiles sont utiles pour décrire la structure d'une série sans être trop influencés par les valeurs extrêmes. Les quantiles les plus utilisées sont les quartiles: le premier quartile (Q1), la médiane (Q2) et le troisième quartile (Q3) sont très courants. Ils permettent de décrire la position centrale des données et leur dispersion.

On construit une boîte de dispersion (boite à moustache) pour résumer une série statistique de manière simple et visuelle. Elle permet de décrire une distribution statistique en montrant son étendue statistique, les valeurs min et max,la médiane (quartile 2), le premier (Q1) et le dernier (Q3) quartile. C'est un outil statistique descriptif pratique pour comparer la distribution statistique de différentes variables. Pour l'interpréter : la boîte (le rectangle) représente l'intervalle entre le premier (Q_1) et le troisième (Q_3) quartile. Le trait à l'intérieur de la boîte correspond à la médiane. Les moustaches indiquent les valeurs minimale et maximale. Plus la boîte est grande, plus la dispersion est forte. La position de la médiane dans la boîte permet aussi de repérer une éventuelle dissymétrie de la distribution.

Paramètre de forme

- Les moments permettent de caractériser une distribution. La différence entre les moments centrés et les moments absolus est la référence sur lequel se base leur calcul. On va avoir les moments absolus (*k)* calculés directement à partir des valeurs de la variable. Ils dépendent donc de l'origine choisie et servent surtout à décrire la position générale des données. Et les moments centrés *(r)* qui sont calculés à partir des écarts des valeurs par rapport à la moyenne. Les moments centrés ne dépendent donc plus de l'origine des variables et permettent de mieux décrire la dispersion et la forme de la distribution, comme la variance pour la dispersion ou l'asymétrie et l'aplatissement pour la forme. On les utilise parce qu'ils apportent des informations différentes mais complémentaires : les moments absolus donnent une idée du niveau des données, tandis que les moments centrés permettent d'analyser comment les valeurs sont réparties autour de la moyenne.

On vérifie la symétrie d'une distribution pour comprendre comment les valeurs se répartissent. Dans une distribution symétrique, le mode, la médiane et la moyenne arithmétique sont égaux, ce qui rend l'interprétation des données plus facile et permet de comparer la distribution à des modèles comme la loi normale. Pour mesurer cette symétrie, on utilise le coefficient d'asymétrie β_1, basé sur le moment centré d'ordre 3 et l'écart-type :

- Si $\beta_1 = 0$, la distribution est symétrique.
- Si $\beta_1 > 0$, elle est étalée vers la droite (asymétrie positive).
- Si $\beta_1 < 0$, elle est étalée vers la gauche (asymétrie négative).

En résumé, calculer β_1 permet de **quantifier la dissymétrie** et de mieux comprendre la forme de la distribution avant d'appliquer d'autres analyses statistiques.