

Questions de cours de la séance 6

1. Qu'est-ce qu'une statistique ordinale ? À quelle autre statistique catégorielle s'oppose-t-elle ? Quel type de variables utilise-t-elle ? En quoi cela peut matérialiser une hiérarchie spatiale ?
2. Quel ordre est à privilégier dans les classifications ?
3. Quelle est la différence entre une corrélation des rangs et une concordance de classements ?
4. Quelle est la différence entre les tests de Spearman et de Kendal ?
5. À quoi servent les coefficients de Goodman-Kruskal et de Yule ?

1. La statistique ordinale est une statistique qui repose sur le classement des objets étudiés. Elle ne cherche pas d'abord à mesurer une grandeur, mais à ordonner des entités les unes par rapport aux autres. Le cours souligne que cette approche est au cœur de la géographie humaine, discipline qui manipule constamment des classements : villes, régions, États, territoires.

Elle s'oppose à la statistique nominale, qui se limite à distinguer des catégories sans relation d'ordre entre elles. Là où la statistique nominale classe sans hiérarchiser, la statistique ordinale introduit une relation de rang.

La statistique ordinale utilise donc des variables ordinales, c'est-à-dire des variables qualitatives ordonnées. Ces variables permettent d'exprimer des positions relatives, sans supposer une distance mesurable entre les rangs.

En géographie, cette ordination matérialise directement une hiérarchie spatiale : elle rend visible quels territoires dominant, lesquels stagnent et lesquels déclinent. Le classement devient alors un outil de lecture des rapports de force spatiaux, qu'ils soient sociaux, économiques ou territoriaux.

2. Le cours indique clairement que l'ordre à privilégier est l'ordre croissant, aussi appelé ordre naturel. Cet ordre permet une lecture cohérente des séries statistiques et facilite l'identification des valeurs extrêmes, qu'elles soient exceptionnellement faibles ou élevées. Il existe toutefois des exceptions en géographie, comme la loi rang-taille, mais celles-ci restent marginales. L'ordre croissant constitue la norme, car il structure la statistique d'ordre et rend possible l'analyse des valeurs aberrantes et de la plus grande valeur d'une série.

3. La corrélation des rangs vise à mesurer le degré de liaison entre deux classements portant sur les mêmes objets. Elle permet de déterminer si deux variables ordinales évoluent de manière liée, indépendamment de la valeur absolue des rangs. La concordance de classements, en revanche, cherche à évaluer dans quelle mesure plusieurs classements sont similaires ou non. Elle s'intéresse davantage à l'accord global entre classements qu'à une relation linéaire entre deux séries. Ainsi, la corrélation des rangs compare deux ordinations, tandis que la concordance peut s'étendre à plusieurs classements, comme le montre la généralisation proposée par Kendall avec le coefficient W .

4. Le test de Spearman repose sur le calcul d'un coefficient de corrélation des rangs fondé sur les écarts quadratiques entre rangs. Il suppose que les rangs sont des permutations équiprobables et ne tolère pas, dans sa forme simple, les rangs ex æquo (une correction est alors nécessaire). Il est particulièrement adapté lorsque l'on compare deux classements.

Le test de Kendall, quant à lui, se fonde sur le nombre de paires concordantes et discordantes. Il adopte une logique plus combinatoire et s'interprète directement en termes de concordance. Son avantage majeur réside dans sa capacité à se généraliser à plusieurs classements, ce que ne permet pas le test de Spearman.

Les deux tests sont non paramétriques et adaptés aux variables ordinales, mais Kendall est souvent privilégié lorsque l'on cherche à analyser des accords multiples entre classements.

5. Les coefficients de Goodman-Kruskal et de Yule servent à mesurer l'association entre variables ordinales, à partir des paires concordantes et discordantes. Le coefficient de Goodman-Kruskal (γ) mesure le surplus de concordances par rapport aux discordances. Il fournit une indication sur la force et le sens de l'association, mais peut être nul même en l'absence d'indépendance stricte, ce qui invite à la prudence dans son interprétation. Le coefficient Q de Yule est un cas particulier du coefficient de Goodman-Kruskal, appliqué aux tableaux de contingence 2×2 . Il permet d'évaluer une association binaire de manière symétrique et s'interprète comme une mesure normalisée comprise entre -1 et $+1$. Dans le cadre de la géographie, ces coefficients offrent des outils robustes pour analyser les relations ordinales lorsque les hypothèses des tests paramétriques ne sont pas vérifiées.