

Master : Marketing & Marketing Touristique

Semestre III

Module : Logiciels statistique 2

Chapitre I : Introduction aux Logiciels Statistiques et leur Mécanisme

I. Introduction

Les logiciels statistiques sont des outils informatiques permettant d'analyser des données de manière systématique et scientifique. Ils sont essentiels dans de nombreux domaines, tels que la recherche scientifique, les études de marché, la finance, la médecine, etc., pour traiter de grands volumes de données et en extraire des conclusions pertinentes.

II. Principaux logiciels statistiques

1. SPSS (Statistical Package for the Social Sciences) :

- Utilisé principalement dans les sciences sociales.
- Interface conviviale, basée sur des menus et des commandes prédéfinies.
- Outils pour l'analyse descriptive, l'analyse multivariée, les tests paramétriques et non paramétriques.
- Avantages : simplicité d'utilisation et large gamme de tests disponibles.

2. R :

- Logiciel open-source puissant, particulièrement apprécié dans les milieux académiques et de recherche.
- Basé sur un langage de programmation spécialisé pour l'analyse statistique.
- Flexibilité extrême, possibilité de créer des scripts personnalisés et d'installer des packages supplémentaires pour des analyses spécifiques.
- Avantages : gratuit, vaste communauté, extrêmement puissant pour des analyses complexes.

3. SAS (Statistical Analysis System) :

- Utilisé dans l'industrie, particulièrement dans les secteurs comme la finance, la santé, et le marketing.
- Interface basée à la fois sur des menus et des scripts de programmation.
- Outils robustes pour la gestion de grandes bases de données et l'analyse statistique.
- Avantages : stabilité, adapté pour des ensembles de données massifs.

4. STATA :

- Utilisé dans les domaines des sciences sociales, économiques, et politiques.
- Propose des analyses statistiques, des économétriques, et des graphiques.
- Interface conviviale, combinant la programmation et l'interaction graphique.
- Avantages : interface intuitive, large éventail de tests disponibles.

III. Mécanisme d'utilisation

1. Chargement des données :

- Les logiciels permettent d'importer des données depuis différents formats : fichiers CSV, Excel, SQL, ou encore des bases de données plus complexes. Dans R, par exemple, cela se fait via la commande `read.csv()`, tandis que SPSS et SAS offrent des interfaces d'importation simplifiées.

2. Préparation des données :

- Avant toute analyse, il est souvent nécessaire de nettoyer les données, par exemple en gérant les valeurs manquantes, en transformant les variables (e.g., normalisation, logarithmes), ou en créant de nouvelles variables dérivées. Cette étape est cruciale pour assurer la validité des résultats.

3. Analyse statistique :

- Une fois les données préparées, on peut effectuer différentes analyses : statistiques descriptives (moyenne, médiane, écart-type), tests d'hypothèse (tests t, ANOVA), régressions linéaires, analyses factorielles, etc.
- Exemples :
 - Dans SPSS, ces analyses se font via des menus déroulants.
 - Dans R, cela peut se faire par des commandes comme `summary()` pour les descriptifs ou `lm()` pour les régressions linéaires.

4. Visualisation des données :

- Les logiciels statistiques permettent également de visualiser les résultats via des graphiques (histogrammes, boîtes à moustaches, nuages de points, etc.).
- R est particulièrement réputé pour ses capacités graphiques, notamment via le package `ggplot2`.

5. Interprétation et présentation des résultats :

- Les logiciels fournissent des sorties détaillées qui incluent les résultats des analyses (valeurs p, intervalles de confiance, coefficients de régression). Ces résultats peuvent ensuite être exportés vers d'autres logiciels, tels que Word ou Excel, pour être présentés sous forme de rapports.