Fondements de l'Analyse Multivariée

⇒ l'essentiel

Fondements de l'analyse multivariée

Variable dépendante et variable(s) indépendantes(s)

Une **variable dépendante** est une caractéristique pouvant prendre **plusieurs valeurs** différentes **lors d'une expérience**, dont la variation est **causée par la variation** d'une ou de plusieurs autres variables, à savoir : **les variables indépendantes**.

Les **variables dépendantes** sont **observées, mesurées** ou **enregistrées** tandis que les **variables indépendantes** sont celles qui sont **manipulées**, voire **contrôlées**, à valeurs fixées.

Variable : Grandeur susceptible de varier dans un ensemble donné, et telle qu'à chaque valeur prise par cette grandeur puisse correspondre, au moins théoriquement, un effectif d'individus (au sens large) ou une fréquence en pourcentage.

Fondements de l'analyse multivariée

Variable dépendante et variable(s) indépendantes(s)

Cette terminologie est également employée dans des études où l'on ne manipule pas vraiment des variables indépendantes mais où l'on affecte des individus à des "groupes expérimentaux" basés sur des propriétés préexistantes de ces individus.

Par exemple, si l'on cherche à comparer le taux d'une substance dans le sang chez les hommes et chez les femmes, le taux de cette substance est alors la variable dépendante (quantitative) et le sexe la variable indépendante (qualitative).

Fondements de l'analyse multivariée

		VD	
		QUALI	QUANTI
V	QUALI	Chi-deux	ANOVA
	QUANTI	Régression Logistique RR,OR	Régression Linéaire Tests Corrélation

Fondements de l'analyse multivariée

A/ Etude des relations entre deux variables qualitatives

L'une d'entre elle peut être la **variable dépendante** (*à expliquer par l'autre appelée alors variable indépendante*) mais ce n'est pas une nécessité.

Tests mis en œuvre :

test du **Chi-deux** d'indépendance,

test exact de Fisher,

RR (Risk Ratio) , OR (Odd Ratio)

et intervalles de confiance du RR et du OR.

Fondements de l'analyse multivariée

B/ Etude des relations entre une variable quantitative (dépendante) et une (ou plusieurs) variable(s) qualitative(s)

La variable quantitative est la variable dépendante (également appelée variable à expliquer) et la variable qualitative est la variable indépendante

> Fondée sur la **comparaison de moyennes** :

Test de **Student** (paramétrique), Mann-Whitney (non paramétrique), **ANOVA** (analyse de variance), Kruskal-Wallis (non paramétrique)

> Fondée sur la **comparaison de variance** :

Test de **Fisher**, test de Hartley

Fondements de l'analyse multivariée

C/ Etude des relations entre deux variables quantitatives

- Si les 2 variables sont toutes deux des variables aléatoires :

Tests de Corrélation de Pearson (linéaire, avec contrainte) ou test de Corrélation de Spearman (non nécessairement linéaire, non paramétrique)

- Si l'une des variables est indépendante, fixée, (variable explicative, encore appelée régresseur) :

Régression linéaire

(simple ou multiple) > « modèle linéaire généralisé »

La variable dépendante doit suivre une distribution normale et la variance être homogène sur tout le nuage de points (homoscédasticité).

Fondements de l'analyse multivariée

B/ Etude des relations entre une variable qualitative (dépendante) et une variable quantitative)

La variable qualitative est, dans ce cas, la variable dépendante et la variable quantitative est la variable indépendante.

Classiquement fondé sur le modèle du logit (il existe d'autres modèles) :

Régression logistique

(simple ou multiple)