

Analyse Factorielle Analyse (factorielle) des Correspondances (simples)

Thierry Dhorne

29 septembre 2015



Introduction

- Les méthodologies statistiques
- ❖ Statistique descriptive
- ❖ Statistique extractive - Objectif
- Méthodes d'analyse des données
- ♦ Objectifs du cours
- Prérequis
- Contenus
- ❖ Méthodes d'analyse factorielle

Mesure du lien

Couples de variables qualitatives

Introduction



Les méthodologies statistiques

Introduction

- Les méthodologies statistiques
- **❖** Statistique descriptive
- **❖** Statistique extractive - Objectif
- ❖ Méthodes d'analyse des données
- Objectifs du cours
- Prérequis
- Contenus
- Méthodes d'analyse factorielle

Mesure du lien

- régression, analyse de la variance,....
- méthodes prédictives
- séries chronologiques, méthodes spatiales
- méthodes de prévision (autoprojectives)
- synthèse (descriptive)
- méthodes d'analyse des données



Statistique descriptive

Introduction

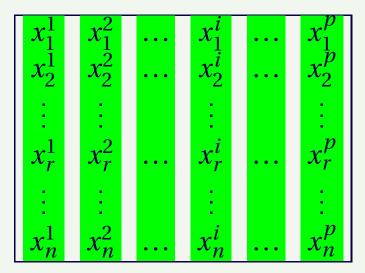
- Les méthodologies statistiques
- ❖ Statistique descriptive
- ❖ Statistique extractive Objectif
- Méthodes d'analyse des données
- ♦ Objectifs du cours
- Prérequis
- Contenus
- Méthodes d'analyse factorielle

Mesure du lien

Couples de variables qualitatives

- le tableau de données n'est pas structuré
- > toutes les variables sont jugées (a priori) aussi informatives

Tableau individus × variables non structuré





Statistique extractive - Objectif

Introduction

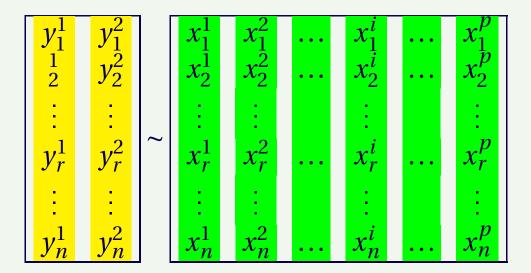
- Les méthodologies statistiques
- ❖ Statistique descriptive
- ❖ Statistique extractive - Objectif
- Méthodes d'analyse des données
- ❖ Objectifs du cours
- Prérequis
- Contenus
- Méthodes d'analyse factorielle

Mesure du lien

Couples de variables qualitatives

- mais on postule l'existence de variables potentiellement plus intéressantes
- synthétiques, structurantes,...

Tableau individus × variables extraction d'information



obtenues à partir des variables de départ



Méthodes d'analyse des données

Introduction

- Les méthodologies statistiques
- **♦** Statistique descriptive
- **❖** Statistique extractive - Objectif
- **❖** Méthodes d'analyse des données
- ❖ Objectifs du cours
- Prérequis
- Contenus
- Méthodes d'analyse factorielle

Mesure du lien

- synthèse qualitative (catégorielle) ⇒ groupes
- méthodes de classification automatique, classification non supervisée, clustering,
- synthèse quantitative ⇒ valeurs, indices (Ex : facteurs d'intelligence : QI)
- méthodes factorielles



Objectifs du cours

Introduction

- Les méthodologies statistiques
- **♦** Statistique descriptive
- **❖** Statistique extractive - Objectif
- ❖ Méthodes d'analyse des données

❖ Objectifs du cours

- Préreguis
- Contenus
- Méthodes d'analyse factorielle

Mesure du lien

Couples de variables qualitatives

- UE31 M3101
- Statistique
- Analyse des données
- Volume Horaire: 45h 15h CM, 20h TD, 10h TP
- Objectifs du module
- Comprendre, mettre en œuvre et interpréter les résultats des analyses exploratoires multivariées usuelles

2015-2016



Prérequis

Introduction

- Les méthodologies statistiques
- **❖** Statistique descriptive
- **❖** Statistique extractive - Objectif
- ❖ Méthodes d'analyse des données
- Objectifs du cours

Prérequis

- Contenus
- ❖ Méthodes d'analyse factorielle

Mesure du lien

- M1201 Statistique descriptive 2
- M2104 Mathématiques pour l'analyse des données



Contenus

Introduction

- Les méthodologies statistiques
- **❖** Statistique descriptive
- **❖** Statistique extractive - Objectif
- ❖ Méthodes d'analyse des données
- Objectifs du cours
- Préreguis

Contenus

Méthodes d'analyse factorielle

Mesure du lien

- Analyse en composantes principales (ACP)
- Analyse factorielle des correspondances (AFC)
- Classification Ascendante Hiérarchique (CAH)
- Classification par centres mobiles (k-means)
- ★ Prolongements possibles
- Analyse Factorielle des Correspondances Multiples (AFCM)
- Multidimensional scaling (MDS)



Méthodes d'analyse factorielle

<u>Introduction</u>

- Les méthodologies statistiques
- ❖ Statistique descriptive
- **❖** Statistique extractive - Objectif
- ❖ Méthodes d'analyse des données
- Objectifs du cours
- Prérequis
- Contenus
- **❖** Méthodes d'analyse factorielle

Mesure du lien

- ACP : analyse en composantes principales
- synthèse descriptive d'un tableau individus × variables quantitatives
- synthèse de p variables (statistique ou aléatoire) quantitatives
- AFC : analyse factorielle des correspondances
- analyse d'une table de contingence bidimensionnelle
- étude du lien entre deux variables qualitatives
- ACM: analyse des correspondances multiples
- synthèse descriptive d'un tableau individus × variables qualitatives
- synthèse descriptive de p variables (statistique ou aléatoire) quantitatives



Introduction

Mesure du lien

 La quantification du lien (linéaire) entre variables statistique

Couples de variables qualitatives

Mesure du lien



La quantification du lien (linéaire) entre variables statistique

Introduction

Mesure du lien

La quantification
du lien (linéaire)
entre variables
statistique

- entre deux variables quantitatives
- le coefficient de corrélation linéaire
- entre une variable quantitative et une variable qualitative
- le rapport de corrélation
- entre deux variables qualitatives
- le coefficient de Cramer (obtenu à partir du χ^2 d'indépendance)



Introduction

Mesure du lien

Couples de variables qualitatives

- ❖ Tableaux individus x variables à deux variables qualitatives
- ❖ Tableau individus
- × variables
- ❖ Tableau disjonctif complet
- Manipulations algébriques
- ❖ Tableau de contingence à 2 x 2 modalités
- ❖ Tableau de contingence -Rappels



Tableaux individus x variables à deux variables qualitatives

Introduction

Mesure du lien

- ❖ Tableaux individus x variables à deux variables qualitatives
- ❖ Tableau individus × variables
- ❖ Tableau disjonctif complet
- ❖ Manipulations algébriques
- ❖ Tableau de contingence à 2 x 2 modalités
- ❖ Tableau de contingence -Rappels

- on considère deux variables statistiques qualitatives
- exemple
- le sexe
- la ponctualité
- l'objectif est d'étudier le lien (éventuel) entre les deux variables



Tableau individus × variables

Introduction

Mesure du lien

Couples de variables qualitatives

- ❖ Tableaux individus x variables à deux variables qualitatives
- ❖ Tableau individus× variables
- **❖** Tableau disjonctif complet
- Manipulations algébriques
- ❖ Tableau de contingence à 2 x 2 modalités
- ❖ Tableau de contingence -Rappels

c'est le tableau de données brutes

nom	sexe	présent-absent		
dupont	M	Α		
durand	F	Р		
:		:		
dubois	F	Α		
:				
dujnou	М	Α		



Tableau disjonctif complet

Introduction

Mesure du lien

Couples de variables qualitatives

- ❖ Tableaux individus x variables à deux variables qualitatives
- ❖ Tableau individus × variables
- Tableau disjonctif complet
- Manipulations algébriques
- ❖ Tableau de contingence à 2 x 2 modalités
- ❖ Tableau de contingence -Rappels

par simple recodage disjonctif :

par simple recodable ansjoinerin .							
nom	sexe		présent-absent				
	F	M	Р	Α			
dupont	0	1	0	1			
durand	1	0	1	0			
	:	:	:				
dubois	1	0	0	1			
	:	:	:	:			
dujnou	0	1	0	1			

* il y a maintenant une colonne par modalité de chaque variable



Manipulations algébriques

Introduction

Mesure du lien

Couples de variables qualitatives

- ❖ Tableaux individus x variables à deux variables qualitatives
- Tableau individus
- × variables
- ❖ Tableau disjonctif complet
- Manipulations algébriques
- ❖ Tableau de contingence à 2 x 2 modalités
- ❖ Tableau de contingence -Rappels

transposition et produit matriciel :

nom	F	М	Р	Α
dupont	0	1	0	1
durand	1	0	1	0
•	:	1		•
dubois	1	0	0	1
•		•		
dujnou	1	1	0	1

dupont	durand		dubois		dujnou	
F	0	1		1		0
M	1	0		0		1
Р	0	1		0		0
Α	1	0		1		1

	F	М	Р	Α
F	62	0	31	31
M	0	49	9	40
Р	31	9	40	0
Α	31	40	0	71

- matriciellement ceci s'écrit : ${}^{t}XX$ où X est le tableau disjonctif
- le tableau résultant s'appelle tableau de Burt du nom de son promoteur (C. Burt 1883-1971)



Tableau de contingence à 2 x 2 modalités

Introduction

Mesure du lien

- ❖ Tableaux individus x variables à deux variables qualitatives
- Tableau individus
- × variables
- ❖ Tableau disjonctif complet
- Manipulations algébriques
- ❖ Tableau de contingence à 2 x 2 modalités
- ❖ Tableau de contingence -Rappels

- la partie extra-diagonale du tableau de Burt est la table de contingence
- Valeurs observables (et observées)

		Colonnes			
		1	2	Totaux	
Lignes	1	n_{11}	n_{12}	$\overline{n_{1+}}$	
	2	n_{21}	n_{22}	n_{2+}	
	Totaux	n_{+1}	n_{+2}	$n_{++} = n$	



Tableau de contingence - Rappels

Introduction

Mesure du lien

Couples de variables qualitatives

- ❖ Tableaux individus x variables à deux variables qualitatives
- ❖ Tableau individus
- × variables
- ❖ Tableau disjonctif complet
- Manipulations algébriques
- ❖ Tableau de contingence à 2 x 2 modalités
- ❖ Tableau de contingence -Rappels

Tableau de contingence

Un tableau de contingence est obtenu

- à partir de deux variables qualitatives
- dont on croise chaque modalités une à une
- pour chaque combinaison
- on dénombre l'effectif des individus concernés