

# **Questionnaire Big Five**

## **Analyse des résultats**

# L'analyse factorielle

# 1. Définition

- L'**analyse factorielle** (AF) est une **méthode statistique** inventée par **Spearman** permettant d'analyser un ensemble de **corrélations**.

Le facteur général d'intelligence :

- ✓ est **commun** aux différentes épreuves d'intelligence
  - ✓ permet de **résumer** les corrélations observées
  - ✓ est **hypothétique**
- L'AF est une **méthode de réduction des données**
    - ✓ nombre de facteurs < nombre de variables observées
    - ✓ facteurs = dimensions fondamentales

(en psychologie, variables observées = comportements)

## 2. Fonctionnement de l'analyse factorielle

- On distingue **cinq étapes** :
  1. on part d'une matrice sujets-variables
  2. on calcule la matrice des **corrélations observées**
  3. on estime la matrice des saturations factorielles
  4. on calcule la matrice des **corrélations reproduites**
  5. on compare les corrélations reproduites aux corrélations observées

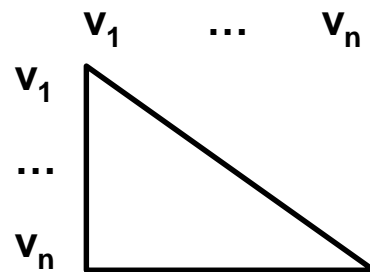
## 2. Fonctionnement de l'analyse factorielle

### 1. Matrice sujets variables

	$v_1$	...	$v_n$
1			
2			
3			
4			
...			
p			



### 2. Matrice des corrélations observées

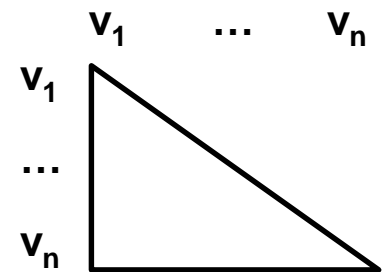


### 3. Matrice des saturations factorielles

	$F_1$	...	$F_m$
$v_1$			
...			
$v_n$			



### 4. Matrice des corrélations reproduites



- *nombre de facteurs*
- *interprétation des facteurs*



### 5. Différences ?

### 3. L'hypothèse fondamentale

*Les corrélations observées sont dues à la présence de **facteurs latents communs aux variables***

- Un facteur commun est :
  - ✓ **une variable latente** (une entité hypothétique)
  - ✓ **une variable causale** (les covariations des variables sont dues aux variations d'un facteur commun)
- L'hypothèse fondamentale de l'analyse factorielle se décline sous deux aspects :
  - ✓ le principe d'**indépendance locale**
  - ✓ la notion de **corrélation reproduite**

### 3. L'hypothèse fondamentale

#### *Le principe d'indépendance locale*

- Chez des enfants de 5 à 12 ans, on constate une corrélation élevée entre la richesse du vocabulaire et la longueur du pouce...  
En fait, cette corrélation est *artificielle* car elle due à la présence d'une **covariable** : l'âge
- La corrélation entre deux variables lorsqu'une covariable est contrôlée est appelée **corrélation partielle**.  
Chez des enfants du **même âge**, il n'y a plus de corrélation entre la richesse du vocabulaire et la longueur du pouce
- **Principe d'indépendance locale** : si l'on maintient constant un facteur, les corrélations entre les variables observées sont nulles

### 3. L'hypothèse fondamentale

#### *La notion de corrélation reproduite*

- La **saturation** ( $\lambda$ ) d'une variable observée dans un facteur est la **corrélation entre le facteur et la variable** ( $-1 \leq \lambda \leq +1$ )

Une saturation correspond donc au degré auquel le facteur influence la variable

- Si l'on connaît les saturations des variables dans les facteurs, on peut reproduire les corrélations entre les variables.

Ce sont les corrélations attendues sur la base des facteurs communs

- Si deux variables observées sont influencées par un facteur commun : **corrélation reproduite = produit de leurs saturations respectives dans ce facteur**



## 4. La matrice des saturations factorielles

- Cette matrice indique les **saturations des variables dans les différents facteurs.**

Matrice centrale dans l'AF : elle montre comment les variables sont reliées aux facteurs

	<b>Facteur 1</b>	<b>Facteur 2</b>	<b>...</b>
variable 1			
variable 2			
...			

- Une matrice de saturation est définie par deux éléments :
  - ✓ le **nombre** de facteurs
  - ✓ la **signification** des facteurs : on interprète chaque facteur en fonction des variables qui saturent dans le facteur ( $\geq 0.30$ )

## 4. La matrice des saturations factorielles

### *Le nombre de facteurs*

- Le nombre de facteurs retenus définit une **solution factorielle**

**solution à 1 facteur :**

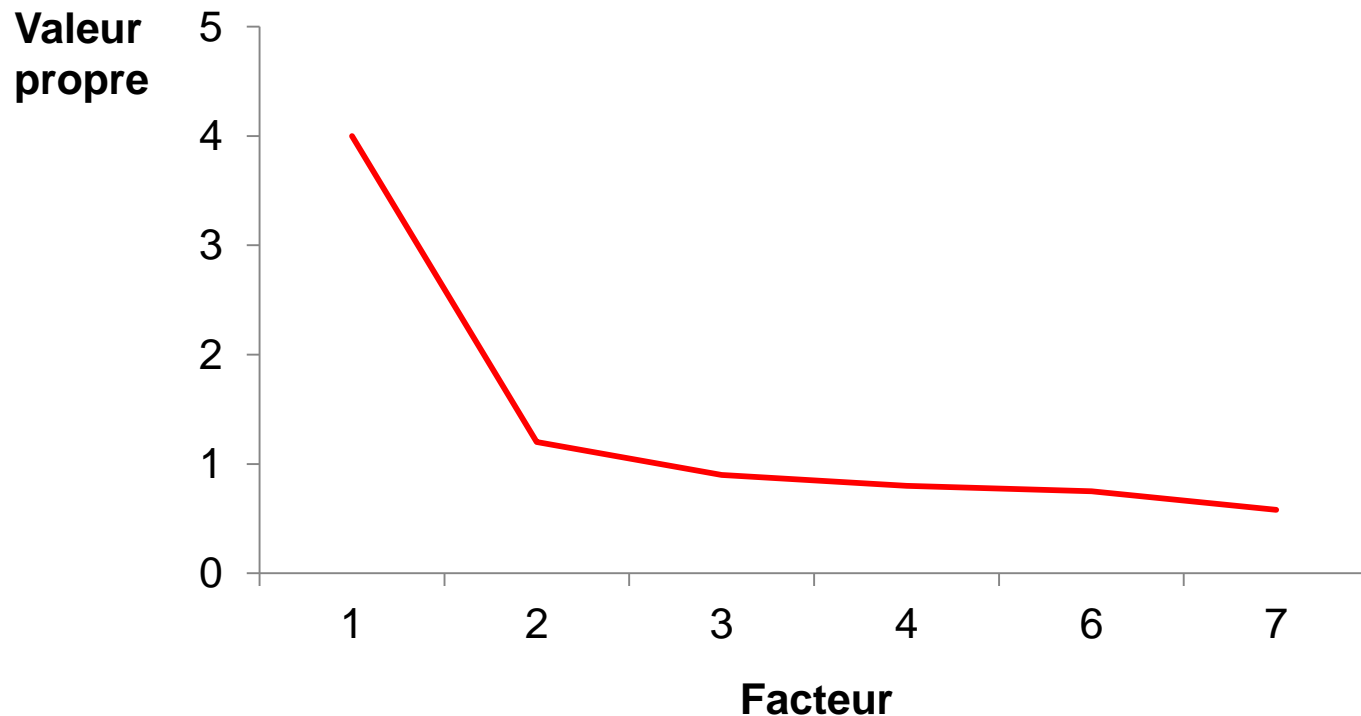
	<b>Facteur 1</b>
variable 1	
variable 2	
...	
variable $n$	

**solution à 2 facteurs :**

	<b>Facteur 1</b>	<b>Facteur 2</b>
variable 1		
variable 2		
...		
variable $n$		

## 4. La matrice des saturations factorielles

La **chute des valeurs propres** (Cattell) : ne pas retenir les facteurs dont les valeurs propres se situent sur une droite



## 4. La matrice des saturations factorielles

### *L'interprétation des facteurs*

- Après avoir décidé du nombre de facteurs retenus, il faut les interpréter : leur attribuer une **signification psychologique**.  
Un facteur est une entité statistique qui n'a pas de signification en soi
- Chaque facteur doit être interprété en fonction du **contenu des variables qui saturent fortement dans celui-ci** (en valeur absolue)

## 4. La matrice des saturations factorielles

- Pour que les facteurs puissent être facilement interprétés, la matrice des saturations doit présenter une **structure simple** : chaque facteur doit avoir des saturations proches de 0 ou de 1.

C'est le **principe de structure simple** (Thurstone, 1947)

### Structure simple

	Facteur 1	Facteur 2
v1	-0.08	0.81
v2	0.10	0.92
v3	0.09	-0.86
v4	0.92	-0.11
v5	-0.89	0.19
v6	0.94	-0.21

### Pas de structure simple

	Facteur 1	Facteur 2
v1	0.54	0.71
v2	0.69	0.74
v3	0.81	0.67
v4	-0.52	0.33
v5	0.36	-0.43
v6	0.75	0.67

# Analyse données questionnaire

- N = 68
- Analyse en composantes principales (ACP)
- Analyse de la matrice des saturations

*Pour chaque facteur, quels items saturent à  $\pm 0.30$  dans le facteur ?*