

LE PSYCHOLOGUE

---

# L'analyse planifiée des données en psychologie

JEAN-MICHEL HOC

*Chargé de Recherche au CNRS*

*Préface de H. Rouanet et D. Léporel*



---

PRESSES UNIVERSITAIRES DE FRANCE

## SOMMAIRE

PRÉFACE par H. ROZANET et D. LAFINE.....	9
REMERCIEMENTS.....	13
INTRODUCTION .....	15
PREMIÈRE PARTIE	
<i>L'organisation des données</i> .....	21
CHAPITRE PREMIER. — <i>Variable, protocole et facteur</i> .....	21
1. Exemple illustratif .....	23
A / Objectif de l'étude .....	23
B / Méthode .....	24
a) Tâche expérimentale.....	24
b) Sujets .....	24
c) Conditions expérimentales .....	25
d) Variable observée .....	25
C / Données de base .....	25
2. Caractérisation préalable des données .....	27
A / Recherche de la variable .....	27
B / Représentation formelle de la variable .....	28
C / Protocole de base .....	28
3. Facteurs élémentaires .....	32
A / Recherche des facteurs élémentaires .....	32
B / Rôle des facteurs élémentaires dans l'analyse... ..	36
a) Facteurs de groupe et facteurs systématiques .....	36
b) Facteurs principaux et facteurs secondaires .....	38
CHAPITRE II. — <i>Relations de croisement et d'emboltement entre facteurs, plan d'analyse</i> .....	42
1. Facteur composé .....	43
2. Relation de croisement entre facteurs .....	46

6 L'ANALYSE PLANIFIÉE DES DONNÉES EN PSYCHOLOGIE

3. Relation d'emboîtement entre facteurs .....	50
4. Plan d'analyse .....	54
5. Protocoles dérivés .....	57
A / Sous-protocoles .....	58
a) Sous-protocoles séparés .....	58
b) Sous-protocoles appareillés .....	60
B / Protocoles dérivés par moyennage .....	62

DEUXIÈME PARTIE

<i>L'évaluation des effets observés</i> .....	71
---	----

CHAPITRE III. — *Analyse des effets liés à un facteur : les comparaisons* .....

1. Effet d'un facteur à deux modalités : l'écart entre deux moyennes .....	72
2. Effets liés à un facteur quelconque : contrastes entre moyennes .....	77
3. La notion de comparaison .....	80
A / Comparaison à un degré de liberté .....	80
B / Comparaison à plusieurs degrés de liberté ...	81
4. La somme des carrés et le carré moyen associés à une comparaison .....	84
A / Comparaison globale sur un facteur .....	85
B / Comparaison quelconque sur un facteur ....	89
a) Comparaison à un degré de liberté .....	89
b) Comparaison à plusieurs degrés de liberté..	92

Annexe au chapitre III : Formules « de définition » et formules « de calcul » .....	94
---	----

CHAPITRE IV. — *Analyse des effets liés à un emboîtement de facteurs* .....

1. Décomposition « inter » + « intra » .....	99
A / Comparaison et effet « inter » .....	100
B / Comparaison et effet « intra » .....	100
C / Equation de décomposition .....	104
2. Comparaisons entre parties d'un facteur : comparaison pondérée, comparaison équipondérée .....	108
A / Protocole pondéré par les effectifs .....	109
B / Moyennage pondéré par les effectifs ou équipondéré .....	110

C / Comparaison pondérée par les effectifs ou équi-pondérée .....	113
CHAPITRE V. — <i>Analyse des effets liés à un croisement de fac- teurs</i> .....	121
1. Décomposition « intra + inter » .....	122
A / Comparaisons « intra » partielles .....	125
B / Comparaison « inter » .....	126
C / Equation de décomposition .....	128
2. Décomposition en effets principaux et en effet d'in- teraction .....	130
A / Comparaison d'interaction .....	130
B / Equation de décomposition et condition d'obten- tion .....	133
3. Non-orthogonalité du croisement : un choix crucial .....	138
4. Comparaisons conditionnelles .....	140
 TROISIÈME PARTIE	
<i>L'analyse inférentielle des effets</i> .....	147
CHAPITRE VI. — <i>Principes de l'inférence statistique : du test de signification à l'inférence fiduciaire</i> .....	147
1. La méthode des tests de signification .....	149
A / Le vénérable exemple de Student .....	149
B / Test de signification .....	151
a) Cas d'école .....	151
b) Test du 't' de Student .....	154
2. La méthode d'inférence fiduciaire .....	155
A / Parabole du « démon de Fisher » .....	156
B / Inférence fiduciaire sur l'exemple de Student ..	159
C / Pratique de la méthode d'inférence fiduciaire ..	161
a) Procédure pour conclure à un effet notable .....	165
b) Procédure pour conclure à un effet néglig- geable .....	169
c) Absence de conclusion .....	173
CHAPITRE VII. — <i>L'inférence sur une comparaison</i> .....	177
1. La notion de facteur aléatoire .....	178
2. Plans du type « groupes indépendants » : $\underline{S} < \underline{G}$ > .....	180
3. Plans du type « groupes appariés » : $\underline{S} * \underline{T}$ .....	187

4. Plans du type « combinaison entre groupes indépendants et groupes appariés » : $S \times G \times T$ .....	190
A / Les comparaisons sur $G$ (type $V$ ) .....	191
B / Les comparaisons sur $T$ (type $W$ ).....	192
C / Les comparaisons d'interaction entre $G$ et $T$ (type $V, W$ ) .....	194
D / Les comparaisons sur $G$ conditionnellement à $T$ (type $V/T$ ) .....	193
E / Les comparaisons sur $T$ conditionnellement à $G$ (type $W/G$ ) .....	196
5. Inférence fiduciale sur un contraste entre moyennes dans le cadre $S \times G \times T$ .....	197
CONCLUSION .....	203
BIBLIOGRAPHIE .....	207
CORRIGÉS DES EXERCICES .....	209
TABLES STATISTIQUES .....	219

## Préface

*Avec l'usage d'instruments de calcul de plus en plus puissants, diversifiés et accessibles, s'est répandu, dans les domaines de recherche les plus divers, le goût des analyses statistiques plus poussées, qu'il s'agisse de dégager les lignes de force d'un processus à partir d'une masse de données, ou de procéder à des analyses fines répondant à des questions spécifiques que peut se poser le chercheur.*

*Parallèlement, on a vu se développer, dès les années cinquante, chez les « Anglo-Saxons », un nouveau courant de la statistique, dénommé par eux « analyse des données » (data analysis) : terme ne renvoyant bien sûr ni à une doctrine sectaire ni à des pratiques laxistes, mais à un certain état d'esprit à l'égard de la statistique, selon lequel il est parfaitement légitime d'examiner de près des données et de les interpréter, quelles que soient les intentions, les hypothèses qui ont pu présider à leur recueil. Quelques maximes favorites du statisticien Tukey, qui a tant œuvré pour répandre cet état d'esprit de la data analysis, sont à cet égard tout à fait significatives : « C'est au statisticien de s'adapter aux besoins de l'usager, et non l'inverse ! » ; ou encore : « Ce qui importe, ce sont les questions, bien plus que les modèles ! ». Pour un psychologue, surtout s'il est de tempérament un peu clinicien, ces idées auront peut-être l'air d'aller de soi, et de définir le cadre le plus sage qui soit pour que le fameux dialogue entre « le mathématicien » et « le psychologue » se déroule avec fruit. En réalité, elles ont dû s'affirmer contre l'idéologie statistique régnante (cette histoire reste à écrire), alors même qu'elles se trouvaient tout à fait en harmonie avec, d'une part, la tradition de la statistique dans l'expérimentation, d'autre part, l'utilisation de procédures inférentielles « non décisionnistes » (1).*

(1) Comme on sait, les principes de base de l'inférence fiduciaire ont été posés par le créateur même de l'analyse de la variance, R. A. Fisher,

C'est dans ce courant d'« analyse de données », au sens qui vient d'être évoqué, qu'il convient de situer les recherches de H. Rouanet et D. Lépine, étayées par les travaux informatiques de M.-O. Lebeaux, qui, entreprises vers la fin des années soixante, aboutirent à ce que nous appelons maintenant l'analyse des comparaisons (parce que la notion formalisée de comparaison y joue un rôle central). À l'origine, nous n'avions pas prémédité de remettre en cause le cadre traditionnel de l'analyse de variance, mais simplement d'utiliser toutes les ressources des mathématiques abstraites, alliées aux immenses possibilités ouvertes par le calcul automatique, pour aborder de front des situations techniquement complexes (mais usuelles dans la pratique), telles que les plans d'expérience « à mesures répétées », ou non orthogonaux, etc. Or, au fur et à mesure que se développait notre entreprise de mieux adapter les méthodes d'analyse des données expérimentales aux besoins des chercheurs, nous fûmes amenés à remodeler ce cadre traditionnel, à l'élargir dans des directions diverses, bref à mettre en place une nouvelle problématique, bien différenciée par rapport à l'analyse de variance classique. En 1975, on peut dire que l'analyse des comparaisons était constituée autour de trois idées-forces (structuration des données, analyse spécifique, inférence fiduciaire), qui devaient se retrouver dans tous les développements ultérieurs, et d'un programme informatique d'une grande généralité (le programme VAR 3), permettant la mise en œuvre pratique de ces idées-forces.

Les méthodes d'analyse des comparaisons, du moins dans leurs aspects les plus élémentaires (langage des plans quasi complets, langage des demandes d'analyse, voire inférences fiduciaires à un degré de liberté), se répandent progressivement dans la communauté des chercheurs qui utilisent la méthode expérimentale, aussi bien au laboratoire que sur le terrain ; le programme VAR 3 de Lebeaux, Lépine et Rouanet, implanté

dans les années trente, dans le but d'exprimer « ce que les données ont à dire sur les paramètres », en laissant de côté toute information ou opinion *a priori*.

dans un grand nombre de centres de calcul, a constitué, le plus souvent, l'instrument efficace de cette diffusion. Parallèlement, l'enseignement de la statistique en psychologie, au niveau licence et maîtrise, s'est trouvé renouvelé par les apports des nouvelles méthodes.

Le présent manuel répond à une demande, maintes fois émise par nos collègues, d'un texte élémentaire, capable de servir de support pour l'enseignement, et également de texte introductif pour les utilisateurs. Ce texte vient à son heure, ayant pu bénéficier de la longue pratique de son auteur, Jean-Michel Hoc, ainsi que de son expérience d'enseignant de la statistique en psychologie. Laisant à l'arrière-plan d'une part les considérations abstraites, d'autre part les aspects qui nécessitent un appareil plus technique, J.-M. Hoc a choisi de traiter en détail, à propos d'exemples concrets, les points suivants :

- l'organisation des données en vue de la planification des analyses ;
- l'évaluation descriptive des effets des facteurs ;
- l'analyse inférentielle des effets : tests de signification, et inférence fiduciaire pour les effets à un degré de liberté.

La présentation adoptée par J.-M. Hoc accorde une large place aux considérations méthodologiques, en insistant sur les critères de choix parmi les diverses procédures, plutôt que sur les justifications proprement mathématiques. De nombreux exercices, avec leurs corrigés, viennent opportunément compléter l'exposé. Tout permet de prédire que ce manuel connaîtra, auprès des chercheurs, enseignants et étudiants, le plein succès qu'il mérite.

H. ROUANET,  
Groupe Mathématique  
et Psychologie  
UMR MLF  
Université René-Descartes.

D. LÉPINE,  
Laboratoire de Psychologie  
expérimentale  
EPHE et Université  
René-Descartes.