

Béatrice Degraeve

Statistiques en psychologie et neuropsychologie

**Méthodes, applications
et exercices corrigés**

DUNOD

Mise en page : Belle Page

© Dunod, 2022

11 rue Paul Bert - 92240 Malakoff

ISBN 978-2-10-083936-0

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

Table des matières

<i>Introduction</i>	9
Jeux de piste.....	14
CHAPITRE 1 – CONCEPTS DE BASE	15
1. Pourquoi, mais pourquoi des statistiques ?	17
2. Types de variables.....	18
3. Des statistiques dans quel cadre ?	25
4. Nature des variables.....	28
5. Population <i>versus</i> échantillon	29
Exercices.....	30
CHAPITRE 2 – STATISTIQUES DESCRIPTIVES	33
1. Réaliser un tri à plat	35
2. Résumer un ensemble de données : les indices de tendance centrale.....	37
3. Les données sont-elles homogènes ou hétérogènes : les indices de dispersion.....	48
4. Loi normale	65
Exercices.....	87
CHAPITRE 3 – UN PAS VERS L'INFÉRENCE : ESTIMER DES PARAMÈTRES À PARTIR D'UN ÉCHANTILLON.....	89
1. Distribution d'échantillonnage des moyennes et erreur-type (ou erreur standard).....	91
2. Estimer des paramètres : les intervalles de pari et de confiance	98
CHAPITRE 4 – STATISTIQUES INFÉRENTIELLES : LOGIQUE DES TESTS	105
1. Le principe des tests d'hypothèse.....	107
2. Concrètement comment ça se passe (presque à chaque fois) ?	124
3. Taille de l'effet, type d'erreurs et puissance.....	129
4. Critiques et contres-critiques.....	132

CHAPITRE 5 – COMPARER DES RÉPARTITIONS D’EFFECTIFS	135
1. Choisir le bon test	137
2. Comparer une répartition d’effectifs (N) observés à une norme : χ^2 d’ajustement ou de conformité.....	137
3. Comparer plusieurs répartitions d’effectifs (N) observés : le χ^2 d’homogénéité.....	148
4. Le χ^2 de Mac Nemar (échantillons appariés)	154
5. Correction de continuité ou correction de Yates.....	159
Exercices.....	162
 CHAPITRE 6 – COMPARER DES PROPORTIONS, DES POURCENTAGES	 163
1. Comparer un pourcentage observé (p) à une norme	166
2. Comparer deux pourcentages observés provenant de deux groupes indépendants	170
 CHAPITRE 7 – EXAMINER L’INDÉPENDANCE ENTRE DEUX VARIABLES NOMINALES.....	 181
1. Le χ^2 d’indépendance	183
2. Les limites d’utilisation et d’interprétation	187
3. Taille de l’effet	187
Exercices.....	188
 CHAPITRE 8 – COMPARER DEUX MOYENNES	 189
1. Vérification des conditions d’applications et choix du test	191
2. Comparer une moyenne à un standard : test de Student pour échantillon unique.....	191
3. Comparer deux moyennes quand la dispersion est homogène (homoscédasticité) : test de Student pour échantillons indépendants	198
4. Comparer deux moyennes quand la dispersion est hétérogène (hétéroscédaticité) : test de Welch pour échantillons indépendants	214
5. Test de Student pour échantillons appariés.....	220
6. Alternatives non paramétriques au test de Student.....	229
Exercices.....	254
 CHAPITRE 9 – COMPARER TROIS MOYENNES ET PLUS, EN FONCTION D’UNE SEULE VI.....	 257
1. Vérification des conditions d’applications et choix du test	259
2. L’analyse de variance (ANOVA) simple pour échantillons indépendants.....	259
3. L’alternative non paramétrique à l’ANOVA simple pour échantillons indépendants : le test de Kruskal-Wallis.....	277
4. L’ANOVA simple pour échantillons appariés	287

5. L'alternative non paramétrique à l'ANOVA simple pour échantillons appariés : le test de Friedman	299
Exercices	305
CHAPITRE 10 – ANALYSER LA FORCE D'ASSOCIATION ENTRE DEUX VARIABLES	309
1. Vérification des conditions d'applications et choix du test	311
2. ρ ou r de Bravais-Pearson	312
3. Régression linéaire simple	331
4. ρ ou r_s de Spearman	338
5. Les indices Phi (φ) ou V de Cramer	344
6. Association entre deux variables de natures différentes	345
Exercices	346
CHAPITRE 11 – ANALYSE DE PUISSANCE ET CALCUL DE LA TAILLE D'ÉCHANTILLON NÉCESSAIRE	349
1. Erreur de type I, erreur de type II et puissance statistique	351
2. Comment calculer la puissance d'un test ? Illustration avec le test z	354
3. Comment calculer la taille d'échantillon nécessaire ?	363
4. Autres possibilités pour déterminer la taille d'échantillon <i>a priori</i>	368
Exercices	368
CHAPITRE 12 – ÉVALUATION DES PERFORMANCES ET STATISTIQUES DU CAS UNIQUE	371
1. Les tests et l'étude de leurs qualités psychométriques	373
2. Évaluer les performances individuelles en référence à celles d'une norme établie	393
3. Comparer les performances individuelles à celles d'un groupe : statistiques du cas unique	402
Exercices	414
ANNEXES	415
Annexe 1 - Tester la normalité et l'homogénéité d'une distribution	417
Annexe 2 - Table de la loi normale centrée-réduite (bilatérale)	418
Annexe 3 - Table de la loi normale centrée-réduite (unilatérale)	420
Annexe 4 - Table du χ^2 (valeurs critiques)	422
Annexe 5 - Table de Student (valeurs critiques)	423
Annexe 6 - Table de Mann-Whitney (valeurs critiques)	425
Annexe 7 - Table Wilcoxon (valeurs critiques)	427

Annexe 8 - Table de Fisher (valeurs critiques).....	428
Annexe 9 - Table de Kruskal-Wallis (valeurs critiques).....	432
Annexe 10 - Table de corrélation de Pearson (valeurs critiques).....	434
Annexe 11 - Table de corrélation de Spearman (valeurs critiques).....	436
 CORRECTION DES EXERCICES	 439
 <i>Bibliographie</i>	 461



Introduction

Cet ouvrage présente les tests paramétriques et non paramétriques classiquement enseignés en licence et master de psychologie, tout en reprenant certaines bases relatives aux statistiques descriptives indispensables à la réalisation de ces tests. Il vise à satisfaire deux exigences :

- disposer d'un ouvrage récent articulant les méthodes « classiques » en statistiques ainsi que des méthodes répondant à **des nouveaux besoins** (c'est-à-dire incluant également les analyses « non paramétriques », « de puissances » et de « cas uniques ») ;
- fournir aux étudiants ou aux professionnels des **outils de compréhension pratiques** en proposant des exercices d'applications ainsi que des exemples de réalisation de ces analyses au moyen de logiciels gratuits, open-sources et multi-plateformes comme Jamovi® (The Jamovi Project, 2021), en plus du détail de la réalisation des calculs « à la main ».

Disposer d'un ouvrage récent sur le programme de statistiques répond à plusieurs constats. À la suite d'un ensemble de réflexions sur les meilleures pratiques statistiques à adopter (e. g. Muller, Perugini, & Yzerbyt, 2016), un certain nombre de **changements ont pris progressivement place dans la pratique professionnelle et la recherche en psychologie** ces dernières années (pré-enregistrements, analyses non paramétriques et vérification des conditions d'application ; voir Delacre, Lakens, & Leys, 2017 ; Delacre, Leys, Mora, & Lakens, 2019). Ainsi, l'enseignement évoluant à son tour, de nouveaux besoins apparaissent dès la licence générale de psychologie. Néanmoins, si de nouveaux besoins ont émergé, il n'existe pas, à notre connaissance, de support adapté en français à un public étudiant ou professionnel qui puisse actuellement y répondre.

- Si les **méthodes non paramétriques** sont parfois abordées dans certains ouvrages, le détail par étape et l'explication des calculs « à la main » de ces tests sont rarement rapportés. Or les étudiants de licence sont généralement évalués sur ces aptitudes. Ces notions correspondent donc à un besoin pour lequel il n'existe, à l'heure actuelle, pas de réelle réponse.
- Les **analyses de puissances** correspondent à de nouvelles exigences en termes de recherches effectuées par les chercheurs mais aussi dans le cadre de la pratique clinique menée par les professionnels en (neuro)psychologie. En effet, à présent, de nombreuses publications en psychologie (telles que le *Journal of Neuropsychology* ou l'*International Review of Social Psychology*) offrent la possibilité (voire recommandent) que les études soient « pré-enregistrées ». Cet enregistrement implique (entre autres) de rapporter *a priori* une analyse de puissance justifiant la taille d'échantillon utilisée. Dans le cadre de la pratique clinique, la justification de la taille des échantillons est également requise par les autorités de santé (Légifrance, 2019) et doit être précisée dans le protocole clinique rédigé en amont de l'inclusion des patients par les neuropsychologues professionnels ou les chercheurs dans ce domaine. Savoir réaliser ce type d'analyse est donc indispensable pour répondre aux exigences actuelles de la discipline. Cette pratique a émergé récemment et s'est vite généralisée dans la recherche et la pratique professionnelle des (neuro)psychologues. Cependant, aucun ouvrage ne traite de la question à l'heure actuelle. Ici aussi, aucun

ouvrage en français ne propose des exemples de réalisation de ces analyses au moyen de logiciel gratuit, open-source comme G*Power (Erdfelder, Faul, & Buchner, 1996).

- Les traitements statistiques dédiés aux **analyses de cas uniques** correspondent aussi à des besoins nouveaux. Ces types d'analyses sont décrits dans deux articles scientifiques (Atzani, 2009 ; Jonin, 2013) mais restent peu accessibles à un public débutant. En effet, ces articles ne fournissent pas aux étudiants ou professionnels des outils de compréhension pratiques en proposant des exercices d'applications afin d'exploiter le plus justement possible ces données.
- Enfin, s'il existe certaines ressources numériques (voir Navarro, Foxcroft, & Meunier, 2020), aucun ouvrage français ne propose des exemples de réalisation de ces analyses au moyen de logiciels gratuits, open-sources et multi-plateformes comme Jamovi® (The Jamovi Project, 2021). Or les étudiants de Master ainsi que les professionnels en (neuro)psychologie ayant besoin de devenir autonomes sur la gestion d'un grand nombre de données, ont besoin de ce type de contenu.

Ainsi, le but de cet ouvrage est de proposer aux étudiants et professionnels en (neuro)psychologie de se familiariser avec les méthodes statistiques actuelles, au plus près des besoins actuels dans le domaine.

L'ouvrage, organisé autour de douze chapitres, reprend en premier lieu les bases essentielles à la bonne compréhension des tests présentés (comment calculer la variance, car nous en avons besoin pour le test de Student par exemple). Il aborde ensuite les tests « classiques » et leurs alternatives non paramétriques. Enfin, il présente certains développements plus récents en statistiques : les analyses de puissance et les analyses de cas uniques. Chaque chapitre suivra la même organisation :

- 1) description de l'analyse, son but et sa logique sous-jacente ;
- 2) étapes pour la réalisation des calculs « à la main » ;
- 3) exercices d'applications (avec correction en annexes) ;
- 4) exemples de réalisation de ces analyses sous logiciel libre (Jamovi®).

L'étape (2) (*réalisation des calculs « à la main »*) est elle aussi toujours présentée selon la même structure afin de faciliter l'explication des démarches de calculs :

- **Étape 1 : poser les hypothèses statistiques.** Cette étape est indispensable à la compréhension de l'objectif et le fonctionnement du test. C'est, par ailleurs, généralement un des aspects sur lesquels les étudiants sont évalués au moment de l'examen.
- **Étape 2 : calcul de la valeur de décision.** Cette étape correspond aux séries de calculs à déployer pour aboutir à la statistique de test (comment obtenir le χ^2 dans le test du χ^2 , comment obtenir le U de Mann-Whitney dans le test de Mann-Whitney, etc.). Nous souhaitons ici guider le lecteur *le plus didactiquement possible* pour comprendre comment aboutir aux statistiques de test et quelles sont les logiques sous-jacentes pour arriver à reproduire ces démarches par la suite sur d'autres données.

- **Étape 3 : lecture/calcul de la valeur critique.** Cette étape correspond aux démarches à opérer pour obtenir la « valeur critique » à laquelle nous allons confronter la statistique de test obtenue à l'étape précédente (utilisation des tables statistiques par exemple).
- **Étape 4 : prise de décision.** Cette étape correspond à la décision que nous prendrons à la suite de cette confrontation entre statistique de test et valeur critique. Concrètement, dans quels cas le test sera-t-il significatif ?

Cet ouvrage s'adresse donc **à des étudiants de Licence et Master**, en leur présentant les bases et développements actuels (nouvelles pratiques et méthodes statistiques) accompagnés d'exercices et d'illustrations sous logiciel. Il s'adresse aussi **à des professionnels en (neuro)psychologie** souhaitant accéder à un ouvrage synthétique et actualisé qui leur permettra de se documenter. Nous espérons que cet ouvrage permettra de compléter ou d'éclaircir certains aspects du programme de statistique qui pourraient rester obscurs, abstraits ou incohérents.

Le dernier mot est à l'intention des personnes qui *redoutent* le programme annoncé dans cette matière. Lorsque nous exigeons des résultats immédiats, des solutions instantanées, nous n'incarbons jamais cette phase (pourtant très importante) de *transition*, qui est pourtant celle où l'apprentissage et la progression se manifestent le plus. Ne vous demandez pas de tout comprendre d'un seul coup, du premier coup. Acceptez qu'il y ait des zones d'ombres à certains endroits, acceptez d'y revenir, **cela viendra avec le temps et le recul**. Parfois, s'y perdre un peu au cours de son apprentissage est même paradoxalement un signe de progression : aucun apprentissage n'est linéaire et sans embûches. Personne n'apprend en y voyant *de plus en plus clair*. Bien souvent (presque exclusivement ?) nous n'y comprenons d'abord plus rien avant d'y voir plus clair qu'avant. Et le phénomène ne se délimite pas à la vie étudiante mais s'étend bien au-delà (toute la vie ?). Bon nombre de chercheurs qui se respectent ont encore de belles zones d'ombres à venir dans des domaines qu'ils maîtrisent déjà bien et finiront par maîtriser encore un peu mieux (et c'est tant mieux !). Ne redoutez pas de ne plus comprendre, **redoutez plutôt de laisser les choses en cet état**. En vous tournant vers des ressources, c'est précisément ce que vous cherchez à combattre, alors bravo !

Trois personnages apparaîtront pour vous guider cet apprentissage :



M. Chavisé : apparaît dans des encarts pour préciser certains points, proposer des anecdotes ou fournir des explications supplémentaires plus avancées sur certains aspects.



M. Chabasourdi : il surgit et pose des questions (parce qu'il est abasourdi). Il essaie de suivre et de s'accrocher mais a parfois besoin qu'on éclaircisse davantage les choses pour être sûr d'avancer dans le bon sens (et il a bien raison).



M. Chamusant : toujours là pour nous donner des astuces, combines, parades ou subterfuges. Grâce à lui, nous pourrions aller plus vite sur certaines étapes ou obtenir d'autres façons d'arriver au même résultat.

Jeux de piste

Trouvez toutes les lettres grecques de l'ouvrage !

Lettre	Se prononce	Signifie...	Vu pour la première fois page...
α	Alpha	La probabilité de commettre une erreur de type I.	
β	Béta	La probabilité de commettre une erreur de type II.	
γ	Gamma	Indice de taille d'effet.	
Δ, δ	Delta	Δ renvoie à la différence, à l'écart ; δ de Ferguson : indice de sensibilité.	
ε	Epsilon	ε renvoie à l'erreur. ε^2 : epsilon-carré, indice de taille d'effet.	
η	Êta	η^2 : éta-carré, indice de taille d'effet.	
κ	Kappa	Indice de fidélité.	
λ	Lambda	Valeur critique en bilatéral.	
μ	Mu	Moyenne d'une population.	
ν	Nu	Degrés de libertés (<i>ddl</i>).	
π	Pi	Proportion théorique.	
ρ	Rho	Corrélation d'une population.	
σ	Sigma	Écart-type d'une population.	
τ	Tau	Tau-b de Kendall, indice de taille d'effet.	
ϕ	Phi	Indice de taille d'effet.	
χ	Chi	χ^2 : Statistique du test du chi-deux.	
ψ	Psi	<i>Revised Standardized Difference Test</i> (RSDT).	
ω	Omega	ω^2 : oméga-carré, indice de taille d'effet.	

Chapitre 1

Concepts de base



Sommaire

1. Pourquoi, mais pourquoi des statistiques ?	17
2. Types de variables.....	18
3. Des statistiques dans quel cadre ?.....	25
4. Nature des variables.....	28
5. Population <i>versus</i> échantillon.....	29
Exercices	30

1. Pourquoi, mais pourquoi des statistiques ?

La psychologie est une science humaine qui prend pour objet d'étude des *personnes*. C'est bien là toute la difficulté : **il n'y en a pas deux pareilles !** Chaque caractéristique d'un individu (qu'il s'agisse de ses traits individuels, de son comportement, de ses opinions, valeurs...) résulte d'une complexe interaction entre ses prédispositions et son environnement. Avec une infinité de combinaisons possibles (autant que d'êtres humains ?), le constat est le suivant : l'espèce humaine est composée d'un cortège d'individus *tous différents*. Même s'ils font partie de la même espèce, même s'ils ont les mêmes parents, même s'ils sont jumeaux. Même ce qui pourrait nous sembler de plus identique : nous-mêmes à deux instants différents, allons changer suivant le moment de la journée, notre état émotionnel, de fatigue... Il est fort à parier que je n'agisse et ne pense pas de la même façon *avant* et *après* mon premier café du matin. Ainsi, si nous souhaitons **dégager les grands principes psychologiques de l'être humain** (les heuristiques, les biais, les phénomènes et processus qui nous touchent généralement et sur lesquels s'échafaudent les théories ; voir figure 1.1), nous ne pouvons donc pas nous appuyer sur *un seul exemplaire* pris au hasard. Puisque nous ne réagissons pas tous de la même façon, il faut prendre plus de recul et passer outre la variabilité de chacun pour énoncer des lois générales. Il faudra alors baser l'étude de ces phénomènes psychologiques sur **des échantillons de plusieurs individus**, de préférence suffisamment *conséquents* pour faire **émerger les invariants malgré toutes les potentielles sources de variabilités non investiguées**, d'autant plus que le processus psychologique est *fin* ou non universel (et qu'on n'a pas encore compris ce qui le modère vraiment...).

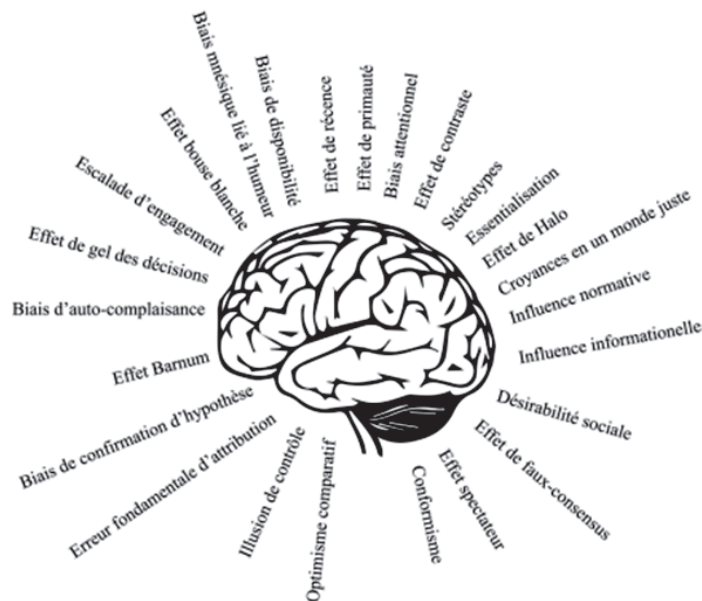


Figure 1.1 – Quelques exemples de principes psychologiques de l'être humain.

Évidemment, chaque échantillon sera aussi *différent*. Au final, lorsque nous comparerons deux groupes après une expérience, une observation, une enquête..., **nous observerons toujours une différence entre eux !** Une partie (voire la totalité !) des différences observées pourra être simplement la conséquence du hasard (plus globalement ce qu'on appellera *l'erreur de mesure*, nous y reviendrons). Ainsi, observer à « l'œil nu » une différence entre deux groupes ne nous permet en rien de savoir si cette différence traduit l'effet des variables ou du phénomène investigué, ou bien si elle traduit l'effet du hasard.

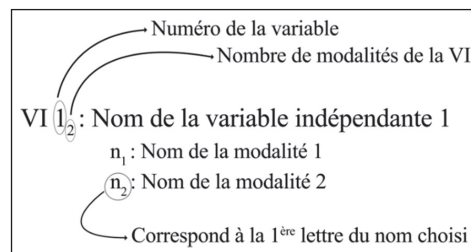
Les tests statistiques (du moins ceux abordés dans cet ouvrage) permettront de calculer la **probabilité qu'une telle différence puisse être observée simplement sous l'effet du hasard**. L'utilisation des statistiques nous permettra ainsi de savoir **dans quelle mesure et avec quel degré de confiance** nous pouvons tirer des conclusions à partir des résultats observés. Voilà pourquoi les étudiants en psychologie ont impérativement besoin de connaître au moins les bases des statistiques.

2. Types de variables

2.1 Variables indépendantes

Les variables *indépendantes* sont les variables que **l'expérimentateur manipule pour examiner leur influence sur les réponses ou le comportement du participant**. Elles ont au moins deux modalités : nous n'étudions jamais l'impact d'un facteur *dans l'absolu*, mais toujours en comparaison à une ligne de base, une condition contrôle ou un autre facteur. Pour évaluer l'effet d'un médicament par exemple, nous n'allons pas uniquement étudier son efficacité en tant que telle mais également la comparer avec une prise de placebo et/ou un autre médicament concurrent.

Nous pouvons noter les variables indépendantes (VI) telle que (voir figure 1.2) :



Nous obtenons donc ici :

- $VI1_2$: type de médicament administré ;
- t_1 : médicament expérimental ;
- t_2 : placebo.

Si nous manipulons une deuxième variable (à 3 modalités), nous noterons : $VI2_3$: ...

Figure 1.2 – Illustration de la notation des variables indépendantes (VI).

Nous pouvons distinguer les variables indépendantes **intra-sujets** des variables indépendantes **inter-sujets** :

- Les variables indépendantes **intra-sujets** soumettent les participants d'une recherche à *chacune* des modalités de la variable. Les réponses sont mesurées à plusieurs reprises, il s'agit d'un plan de recherche « à mesures répétées » ou « appariées ».
- Les variables indépendantes **inter-sujets** soumettent les participants d'une recherche à *une seule* modalité de la variable. Des groupes différents sont formés, il s'agit d'un plan de recherche « à groupes indépendants ».

Exemple

Si nous décidons de mettre au point une étude sur les effets de la musique sur la mémorisation (dans le but de savoir si elle facilite ou inhibe les performances mnésiques), notre **variable indépendante (VI) ici est la musique** : c'est bien le facteur que nous manipulons pour examiner l'influence qu'il peut avoir sur les réponses ou le comportement du participant. Puisque les VI ont au moins deux modalités (nous n'allons pas étudier l'impact d'un facteur *dans l'absolu*, mais nous allons à minima comparer son effet à une ligne de base), nous décidons ici de comparer la modalité « musique » à une modalité « silence » : les individus procèdent à un exercice de mémorisation cette fois-ci en silence et nous verrons quelle condition (« musique » ou « silence ») entraîne les performances mnésiques plus importantes en moyenne. Nous obtenons :

VI₁ : Contexte pendant la mémorisation :

c_1 : avec musique

c_2 : sans musique (dans le silence)

Nous pouvons opter pour deux designs expérimentaux différents, soit nous décidons que notre VI est intra-sujets, soit nous décidons que notre VI est inter-sujets :

VI intra-sujets	VI inter-sujets
Tous les participants effectuent la tâche de mémorisation (apprendre une liste de 10 mots) avec musique puis effectuent à nouveau cette tâche (avec une autre liste de 10 mots de difficulté similaire) mais sans musique (en silence, voir figure 1.3).	La moitié des participants effectue la tâche de mémorisation avec musique et l'autre moitié effectue à la même tâche mais sans musique (en silence, voir figure 1.4). Une seule liste de mots est nécessaire.



Figure 1.3 – Illustration d'un design expérimental intégrant une VI intra-sujets.



Figure 1.4 – Illustration d'un design expérimental intégrant une VI inter-sujets.

2.2 Variables dépendantes

La variable dépendante (VD) correspond à **ce que nous mesurons**. Il s'agit de la réponse ou du comportement du participant, qui sont donc « dépendants » de la situation expérimentale que nous avons créée. Cette mesure **ne varie pas** à travers les modalités de la variable indépendante, nous mesurons toujours la même chose (nous comparons simplement ces mesures entre les différentes modalités pour statuer sur la question qui nous intéresse, mais la VD est la même à chaque fois). Ainsi, **contrairement à la VI qui varie, la VD reste constante**. Dans notre expérience sur les effets de la musique sur la mémorisation, nous relevons des performances de mémorisation (liste de 10 mots) dans tous les cas.

2.3 Variables parasites

Les variables parasites correspondent à des variables dites « externes » que **l'expérimentateur ne veut pas étudier mais qu'il doit contrôler** car elles ont potentiellement un impact sur le phénomène étudié. Il s'agit de toute variable non investiguée qui peut influencer la VD et qui, de ce fait, peut compromettre l'établissement d'une relation claire entre la VI et la VD.

Si, par exemple, il y a d'importants travaux au moment où nous réalisons notre expérience sur les effets de la musique sur la mémorisation, le *bruit* pourrait nuire à la mémorisation (surtout pendant la condition « silence » !). Sans faire attention à cela, nous pourrions conclure que les performances sont moins bonnes dans la condition « silence » comparées à celles des individus lorsqu'ils écoutent de la musique, alors que peut-être cette chute des performances dans cette condition est, en fait, uniquement liée à la gêne occasionnée par les travaux... Il existe **différentes stratégies** pour contrôler une variable parasite :

- **La supprimer** : s'il y a du bruit par exemple, nous pouvons changer de lieu de passation, trouver un moyen de faire cesser le bruit ou attendre la fin des travaux.

- **La maintenir constante** : si nous pensons que l'heure de passation peut avoir un impact sur la mémorisation et que nous souhaitons contrôler cette variable parasite, nous ne pourrions pas l'éliminer comme le bruit... Nous pouvons toutefois choisir de faire les passations expérimentales uniquement entre 10 heures et 12 heures pour l'ensemble des participants. Ainsi, tout le monde sera dans les mêmes conditions et les choses seront donc comparables. Dans ce cas, nous avons maintenu la variable parasite à un niveau constant.
- **L'équilibrer** : si nous pensons que l'heure de passation peut avoir un impact sur la mémorisation et que nous avons deux créneaux de passation (fin de matinée et début d'après-midi), il faudra alterner entre nos deux conditions dans chacun des créneaux (plutôt que de faire uniquement la modalité « silence » en fin de matinée et uniquement la modalité « musique » en début d'après-midi). Il s'agit alors d'équilibrer pour rendre les choses comparables.
- **La répartir aléatoirement** (aléatorisation, randomisation) : nous partons du principe que si nous laissons jouer la variable parasite de manière aléatoire, le hasard réalisera un « équilibrage naturel ». Si nous pensons que l'heure de passation peut avoir un impact sur la mémorisation et que nous avons deux créneaux de passation (fin de matinée et début d'après-midi), nous pouvons choisir d'attribuer *aléatoirement* les conditions pour chaque créneau : le participant arrive au créneau fixé et tire au sort la condition expérimentale (« silence » ou « musique ») par exemple. Ici, nous laissons le hasard rendre les choses comparables. La randomisation est pertinente si la taille de l'échantillon est suffisamment importante : la probabilité d'arriver à un « équilibrage naturel » diminue si le nombre de participants ou d'essais est faible.
- **La contrebalancer** (supprimer les effets d'ordre) : lorsque l'expérience est réalisée en intra-sujet (tous les participants passent les *deux* conditions), nous pouvons imaginer que le fait de faire la première condition influence la performance obtenue à la deuxième (soit en l'améliorant, soit en l'altérant). Pour corriger cet effet potentiel de l'ordre, il suffit d'alterner (de contrebalancer) l'ordre pour la moitié des participants. Une partie des participants passera d'abord la condition « silence » puis « musique », et l'autre moitié passera d'abord la condition « musique » puis « silence ». Cette technique s'applique également dans le cadre d'un questionnaire. Si nous voulons nous assurer que le fait de répondre à une partie n'influencera pas les réponses à la partie suivante, nous pourrions contrebalancer l'ordre de présentation des parties en procédant à la création de plusieurs versions de ce questionnaire pour lesquelles les parties apparaîtront dans un ordre différent. Nous interrogeons ensuite un nombre de participants similaire par versions.
- **Appliquer le procédé de simple ou double aveugle**. Parfois, les participants et l'expérimentateur, par les attentes qu'ils peuvent nourrir, peuvent aussi influencer les résultats, même involontairement. L'expression « simple aveugle » signifie que **l'une des parties** (soit les participants, soit l'expérimentateur) ignore dans quelle condition le participant est placé. Quand il s'agit des participants, ils peuvent par exemple ignorer le traitement qui leur est

appliqué (médicament ou placebo). Cela n'est pas toujours possible : dans notre expérience sur la musique, les participants connaissent forcément la condition dans laquelle ils sont. Néanmoins, si nous ne pouvons pas toujours contrôler l'effet des attentes des participants, nous pouvons contrôler les effets de celles des expérimentateurs, qui ont envie de voir leur expérience fonctionner (et leurs hypothèses vérifiées...). Ici, par exemple, nous pourrions mettre un casque sur les oreilles des participants dans tous les cas (même en condition « silence »). Si le participant saura s'il y a eu de la musique ou non, l'expérimentateur peut, quant à lui, rester *aveugle*. L'expression « double aveugle » signifie que **les deux parties** (les participants et l'expérimentateur) ignorent dans quelle condition le participant est placé. Une variable parasite dont les effets nuisibles ont été neutralisés devient une variable *contrôlée*.

2.4 Variables confondues

Les variables confondues correspondent à **un type particulier de variable externe qui influence à la fois la VI et la VD**. Ces variables sont également susceptibles d'induire un biais dans l'analyse du lien que nous cherchons à mettre en évidence en nous amenant à produire de fausses associations. Prenons quelques exemples pour illustrer ce que sont ces variables confondues (et en quoi elles diffèrent des variables parasites).

Exemple 1

Un individu interrogé par un journaliste dit : « Il faut arrêter de créer des jeux vidéo violents, ceux qui y jouent sont plus violents que ceux qui jouent à des jeux non violents ! » Cet individu décrit une situation comportant une variable confondue ce qui l'amène à produire de fausses associations. Pourquoi ?

Les personnes jouant uniquement à des jeux vidéo *violents* en général sont aussi ceux qui ont un attrait plus marqué pour la violence par rapport à ceux qui ne jouent qu'à des jeux *non violents*. Il n'est donc pas surprenant de constater qu'au final, ces personnes-là ont aussi des comportements plus violents que ceux qui n'optent que pour des jeux sans violence. Ainsi, la variable « personnalité de l'individu » (appelons ce facteur ainsi très schématiquement pour les besoins de notre illustration) va jouer sur ses comportements mais va jouer aussi sur le choix de jeux vidéo qu'il va faire (voir figure 1.5). Cette variable individuelle de personnalité constitue ici une variable *confondue*. Si nous ne la prenons pas en compte, nous aboutissons à une conclusion erronée sur le plan statistique selon laquelle « jouer à des jeux vidéo violents rend violents ».

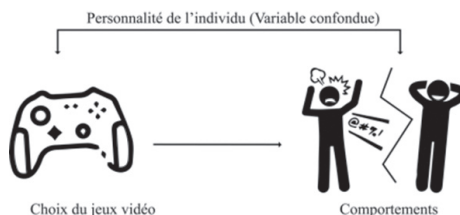


Figure 1.5 – Présence d’une variable confondue (facteurs de personnalité).

Exemple 2

Pour montrer que les variables confondues peuvent se cacher un peu partout, prenons un dernier exemple. Imaginons que nous voulons étudier le lien entre *la consommation d’alcool* et *la prise de risque* (la consommation d’alcool est-elle associée à une plus forte adoption de comportements à risque, comme prendre des substances psychoactives, faire des paris d’argents, etc. ?). Nous demandons à un échantillon d’individus d’évaluer leur consommation d’alcool et de préciser les différents comportements à risque qu’ils ont adopté pendant un mois. Une fois les résultats obtenus, admettons que nous décidions de réaliser un graphique représentant **le nombre de comportements à risque en fonction de la consommation d’alcool** et que nous obtenions ceci (voir figure 1.6) :

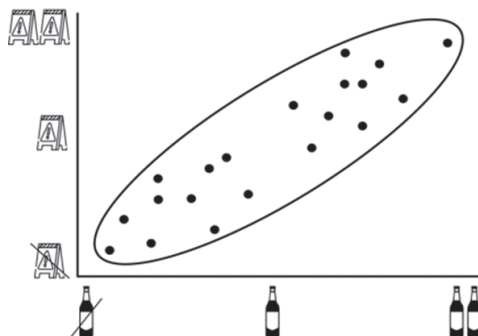


Figure 1.6 – Nombre de comportements à risque en fonction de la consommation d’alcool.

Sur la figure 1.6, chaque point représente une des personnes interrogées et quantifie sa consommation d’alcool et son nombre de comportements à risque.

Nous constatons une belle corrélation montrant que *plus les individus boivent d’alcool, plus ils adoptent des comportements à risque*.

Admettons maintenant que quelqu’un nous demande quels sont les résultats si on sépare *les jeunes adultes* et *les adultes* ? Ah... Nous reprenons nos données et puis nous retraçons notre nuage de points en coloriant différemment les 18-25 ans et les 30-38 ans. Voilà ce que nous obtenons (voir figure 1.7) :

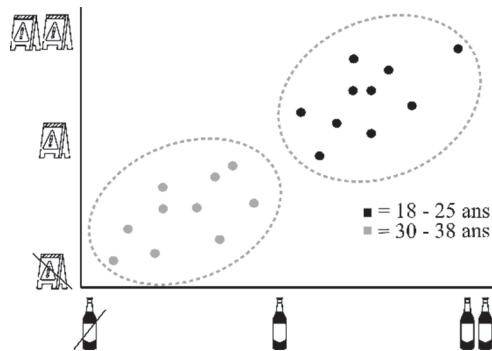


Figure 1.7 – Nombre de comportements à risque en fonction de la consommation d'alcool par tranche d'âge.

Si nous différencions par tranche d'âge, nous constatons qu'en réalité, il n'y a plus réellement de corrélation entre la consommation d'alcool et les comportements à risque.

Nous sommes à nouveau dans une situation comportant une variable confondue, qui aurait pu nous amener à produire une fausse association. En effet, nous essayons d'étudier le lien entre la consommation d'alcool et la prise de risque, et nous avons une variable confondue, l'âge des participants, qui **joue à la fois sur la consommation d'alcool** (les jeunes adultes ont tendance à boire plus d'alcool que les 30-38 ans) **et sur les comportements à risque** (les jeunes adultes prennent plus de risque par rapport aux 30-38 ans). Si nous ne la prenons pas en compte, nous aboutissons à une conclusion erronée sur le plan statistique selon laquelle « plus les individus boivent d'alcool, plus ils adoptent des comportements à risque » (alors que notre conclusion devrait plutôt être : les jeunes adultes boivent plus d'alcool et adoptent plus de comportements à risque que les adultes).

Quelles sont les possibles stratégies de contrôle ? Avec les variables parasites, il faut essayer de penser aux variables non désirables qui pourraient influencer la VD (le bruit, l'heure de la journée, les attentes...) et mettre en place des stratégies pour tenter de contrôler leurs effets mais avec les variables confondues, c'est parfois plus compliqué... Si nous connaissons la variable confondue, il faudra construire l'expérience ou analyser les données en conséquence. Si toutefois nous ne les connaissons pas (ou ne sommes pas sûrs de *toutes* les connaître), il faudra :

- s'entourer des personnes qui connaissent bien la littérature scientifique sur le domaine, ou réaliser soi-même une revue de la littérature poussée sur le sujet (se documenter) ;
- réaliser une recherche expérimentale plutôt qu'une enquête (recherche corrélacionnelle), ou opter pour une enquête à mesures répétées au cours du temps, (c'est-à-dire une enquête longitudinale).

En tout état de cause, il faut garder en tête que c'est **l'accumulation des études** (plutôt qu'une étude isolée) qui nous renseigne véritablement sur le phénomène étudié.

3. Des statistiques dans quel cadre ?

3.1 Un cadre expérimental

Le cadre expérimental consiste à **reproduire un phénomène dans des conditions contrôlées** pour tester une hypothèse. La démarche expérimentale permet de démontrer des **relations de cause à effet**. Si la variable étudiée a bien un impact sur le phénomène en question, les résultats devront logiquement être *différents* dans le groupe expérimental par rapport au groupe de contrôle. En nous assurant de la **constance des autres paramètres** (les variables parasites et confondues sont contrôlées), seule la variable « traitement » (la VI) peut expliquer les différences de résultats observées entre les groupes. Le contrôle de tous les autres paramètres minimise le risque de pouvoir attribuer les différences de résultats observées à d'autres facteurs que ceux qui sont investigués et manipulés. Ainsi, la méthode expérimentale consiste à reproduire un phénomène **en ne faisant varier que les paramètres d'intérêts** (la ou les VI) et observer les conséquences de cette variation sur un phénomène.

Exemple

Si nous décidons de mettre au point une étude sur les effets de la musique sur la mémorisation (dans le but de savoir si elle facilite ou inhibe les performances mnésiques), et souhaitons utiliser un cadre expérimental, nous pouvons :

- Avoir recours à un exercice de mémorisation d'une liste de 10 mots pour refléter les « performances mnésiques ». L'individu a 30 secondes pour retenir le plus de mots possibles à partir d'une liste de 10 mots. À l'issue des 30 secondes, la liste est retirée et il doit restituer le plus de mots possibles de la liste dans l'ordre qu'il souhaite. Le nombre de mots correctement restitués constituera donc notre indicateur de « performances mnésiques » (notre VD).
- Faire réaliser cet exercice à un individu en faisant varier l'environnement pendant la phase d'apprentissage. Nous introduisons de la musique pendant la phase de mémorisation, et puisque nous n'allons pas étudier l'impact de ce facteur *dans l'absolu*, nous allons comparer son effet à autre chose ou à une ligne de base. Nous pouvons décider de comparer la modalité « musique » à une modalité « silence » : les individus procèdent à l'exercice de mémorisation cette fois-ci en silence et nous verrons quelle condition ou *modalité* de notre VI (« musique » ou « silence ») entraîne des performances mnésiques plus importantes en moyenne.