

Probabilité et Statistique pour le DEA de Biosciences

Avner Bar-Hen

Université Aix-Marseille III
2000–2001

Table des matières

1	Introduction	3
2	Introduction à l'analyse statistique	5
1	Introduction	5
2	Planification	5
3	Observation des résultats	5
4	Statistiques descriptives	7
5	La moyenne	8
6	L'écart-type	8
7	Le coefficient de variation	9
3	Expérimentation	11
1	Introduction	11
2	Nécessité des répétitions	11
3	Disposition au hasard	12
4	Facteurs croisés et facteurs hiérarchisés	12
5	Dispositif complètement randomisé	13
6	Constitution de blocs	14
7	Les blocs complets randomisés	14
8	Plans factoriels	16
9	Le Split-plot : Facteurs contrôlés subsidiaires	17
10	Conclusion	19
4	Estimation et tests d'hypothèse	21
1	Introduction. Notions de probabilité	21
2	Population totale, échantillon, loi de distribution	21
3	Échantillon au hasard	22
4	Notion d'estimateur	23
5	Test d'une hypothèse	23
6	Risques	24
5	Tests de comparaison de moyennes	27
1	Fluctuations d'échantillonnage pour la moyenne arithmétique	27
2	Comparaison de deux moyennes	27
3	Tests de comparaison de variances	30

4	Test de Bartlett	30
6	Analyse de variance	33
1	Cas d'un seul facteur.	34
2	Cas de deux facteurs	35
3	Cas non orthogonal	36
7	Comparaisons entre les niveaux des facteurs	39
1	La plus petite différence significative (PPDS)	39
2	Les tests de Duncan et de Newman-Keuls	40
3	Autres méthodes de comparaisons multiples de moyennes	41
4	Exemple	41
5	La méthode des contrastes	43
8	Hypothèses de l'analyse de variance	47
1	Test de normalité	47
1.1	Test d'ajustement du χ^2	48
1.2	Test de Kolmogorov-Smirnov	49
1.3	Test de Lin et Mudholkar	49
1.4	Test de Shapiro-Wilks	50
2	Additivité	51
3	Non-indépendance des erreurs	52
4	Variance hétérogène	52
9	Transformation des données	55
1	Transformation logarithmique	55
2	Transformation racine carrée	55
3	Transformations Arc sinus	56
4	Partitionnement de l'erreur	56
10	Analyse de variance non paramétrique	59
1	Le test des rangs de KRUSKAL-WALLIS : cas d'un facteur	59
2	Le test de FRIEDMAN : cas de deux facteurs	61
11	Corrélation de rangs	65
1	Introduction	65
2	Coefficient de corrélation de rangs de Spearman	65
3	Coefficient de corrélation de rangs de Kendall	67
4	Coefficient de concordance de Kendall	68
12	Tests sur les résidus d'une régression	69
1	Test de Durbin-Watson	69
2	Test des suites	70

13 Tables statistiques	73
La loi normale	74
La loi du chi-deux	75
La loi de Student	76
La loi de Fisher (5%)	77
La loi de Fisher (1%)	78
Valeurs critiques pour le test de Kruskal-Wallis	79
Valeurs critiques pour le test de Friedman	80
Valeurs critiques pour le coefficient de corrélation r_s de Spearman	81
Valeurs critiques pour le coefficient de corrélation τ de Kendall	82
Valeurs critiques pour le coefficient de de concordance de Kendall	83
Valeurs critiques du test de Newman et Keuls au seuil 5%	84
Valeurs critiques du test de Duncan au seuil 5%	85
Valeurs critiques du test de Dunett au seuil 5%	86
Valeurs des scores normaux $a_n(i)$	87
Valeur critique du test de Shapiro-Wilks	88
Valeur critique du test de Durbin-Watson au seuil de 5%	89
Valeur critique du test du nombre de paires au seuil de 5%	90
Bibliographie	91

Chapitre 1

Introduction

Le chapitre 3 présente les principes de l'expérimentation. L'idée est de rappeler les principes fondamentaux contenus dans la construction des plans d'expérience.

Le chapitre 4 présente les concepts de base en statistiques. Il a pour but de définir le cadre de la théorie des tests. En effet, une mauvaise utilisation des termes de base conduit souvent à une conclusion erronée.

Les chapitres 5 à 10 sont centrés sur l'analyse de variance. Le chapitre 5 est une version simplifiée de l'analyse de variance dans le cas de deux moyennes. Conclure à une différence significative (à un seuil donné) entre les moyennes étudiées n'est généralement pas suffisant. Afin de rechercher l'origine de ces différences, les techniques de comparaisons multiples sont présentées dans le chapitre 7.

L'analyse de variance n'est valide que sous certaines hypothèses qu'il est important de vérifier. Le chapitre 8 présente des méthodes de vérification de ces hypothèses et les chapitres 9 et 10 indiquent des manières de procéder lorsque les hypothèses de base ne sont pas respectées.

Dans la même idée, nous supposons que les méthodes de calcul des coefficients de corrélation sont connues du lecteur. Le chapitre 11 présente des tests de corrélation non paramétriques utilisables lorsque les hypothèses des tests de corrélation ne sont pas remplies. Enfin le chapitre 12 présente des tests sur l'autocorrélation des données en régression.

Dans le chapitre 13, nous avons regroupé l'ensemble des tables statistiques nécessaires à l'utilisation des tests présentés tout au long du document.

En résumé, nous avons essayé de rappeler les notions de base dans les trois premiers chapitres et nous présentons un ensemble de techniques utilisables pour vérifier les hypothèses de base de l'analyse de variance et de la régression et pour analyser les données lorsque ces conditions ne sont pas remplies. Cette approche nous permet de présenter des techniques utilisables dans d'autres contextes (test du chi-deux ou de Kolmogorov-Smirnov, par exemple).

La présentation n'est pas exhaustive et un certain nombre de techniques très utilisées ne sont pas présentées (ACP, AFC, classification, discrimination, régression non linéaire...).

Chapitre 2

Introduction à l'analyse statistique

1 Introduction

L'expérimentateur qui opère se heurte à de nombreuses difficultés. La grande variabilité des caractères étudiés, les nombreuses et incontrôlables causes qui peuvent les influencer rendent suspect, a priori, tout résultat isolé. Entre la constatation du fait expérimental et la conclusion plus générale que l'on prétend en tirer, se situe une phase intermédiaire qui est celle de l'interprétation. Une interprétation incorrecte d'observations, inattaquables en elles-mêmes, peut conduire à des conclusions tout à fait erronées. Les méthodes statistiques permettent d'éprouver la validité des résultats, en fonction même de leur variabilité, avec la plus grande rigueur scientifique. Elles permettent une analyse, base de toute interprétation.

2 Planification

Toute expérience un peu complexe utilise la combinaison de principes de base. Il doit toujours y avoir un certain nombre de répétitions et les causes de variation sont soit contrôlées, soit réparties au hasard. Ces principes seront développés dans le chapitre 3. Mais la disposition adoptée n'est pas indifférente : il faut qu'elle soit pratiquement réalisable sur le terrain, sans que les difficultés d'exécution ne deviennent exagérées, et il est essentiel qu'elle permette une estimation correcte de l'erreur expérimentale, base indispensable pour juger de la signification des écarts observés. La mise au point du plan expérimental semble donc, dans chaque cas particulier, nécessiter la collaboration de la technique agricole ou forestière et de la technique statistique.

3 Observation des résultats

Avant de commencer une analyse, il est important de bien observer ses résultats, et donc de les présenter correctement. Il est nécessaire que les données soient lisibles et

compréhensibles tant pour l'expérimentateur que pour quelqu'un n'ayant pas une connaissance détaillée de l'expérience. Il est aussi utile de pouvoir les reprendre (et les comprendre) après une longue période.

Les données doivent être toujours accompagnées d'une explication détaillée de l'expérience et des remarques de celui qui les a relevées. L'expérimentateur doit les contrôler rapidement afin de pouvoir aller sur le terrain vérifier les anomalies éventuellement détectées. L'introduction des données à l'ordinateur est une source d'erreur importante. Si les données sont copiées à la main, il faut qu'une autre personne contrôle le travail.

Les données doivent être accompagnées du protocole complet de l'expérience afin de pouvoir resituer les données suspectes. Ceci permet de savoir si on peut corriger la donnée ou si on doit l'éliminer de l'analyse statistique. Cette dernière possibilité n'est acceptable que lorsqu'une cause indépendante du traitement est identifiée (exemple 2.1 et 2.2).

Exemple 2.1 Circonférence d'arbres (en cm)

BLOCS							
I		II		III		IV	
ESP	CIRC	ESP	CIRC	ESP	CIRC	ESP	CIRC
D	37.0	C	56.7	E	21.5	A	59.2
A	61.0	B	50.2	C	18.4	D	40.7
E	20.0	A	2.1	B	1.5	C	62.2
B	53.7	E	18.4	D	33.4	B	48.6
C	1.8	D	30.9	A	50.3	E	19.0

Les données correspondent aux circonférences de cinq espèces d'arbres. En observant les données, fournies avec le plan de la parcelle, nous remarquons trois valeurs particulièrement faibles pour les espèces C (CIRC=1.8), A (CIRC=2.1) et B (CIRC=1.5) dans les blocs I, II, III respectivement. Il est manifeste que ces faibles circonférences ne sont dues ni à des effets espèces ni à des effets blocs. Les notes prises au champ permettent de découvrir que ces arbres ont subi de fortes attaques d'insectes.

Exemple 2.2 Circonférence moyenne de quatre espèces d'arbres (en cm.)

BLOC	ESP	CIRC	ESP	CIRC	ESP	CIRC	ESP	CIRC
I	A	40.0	B	30.0	C	15.0	D	10.0
II	B	45.0	A	42.5	D	18.0	C	12.5
III	D	35.0	A	40.0	C	16.0	B	13.0

Le plan du dispositif indique que les blocs sont adjacents. Les faibles circonférences dans la partie droite de la parcelle sont dues à la présence d'une pente qui crée un gradient de fertilité.

Pour observer les données, il peut être utile de les réarranger par traitement afin de mettre en évidence des grandes différences entre des parcelles ayant reçu le même traitement.

Dans le cas de mesures répétées dans le temps, faire un graphe des données par rapport au temps où par rapport à la position est une excellente méthode pour détecter des valeurs aberrantes.

Les exemples 2.1, 2.2, 2.3, 2.4 représentent des cas typiques.

Exemple 2.3 *Essai fertilisation (circonférence en cm.)*

FERT	CIRC	FERT	CIRC	FERT	CIRC
1	22.9	2	50.6	3	61.2
1	17.3	2	39.1	3	70.4
1	36.4	2	52.3	3	71.6
1	20.2	2	46.7	3	69.2
1	19.7	2	43.1	3	59.8
1	26.8	2	38.0	3	65.0
1	147.5	2	46.4	3	73.4
1	16.4	2	47.6	3	81.8

Exemple 2.4 *Mesure du poids de la masse foliaire de quatre espèces (tonnes/ha)*

PARC	ESP A	ESP B	ESP C	ESP D
1	1.69	2.72	1.25	8.70
2	1.77	3.16	3.25	7.60
3	1.31	2.48	1.50	19.00
4	1.99	2.50	2.75	12.25
5	2.03	2.66	3.00	8.71
6	1.87	2.18	1.50	7.95
7	1.75	4.00	2.32	6.03
8	1.91	3.54	3.71	7.19
9	1.88	1.72	2.63	9.86

La précision des données semble douteuse car tous les poids importants se terminent par .00 et .05. Il est possible que plusieurs machines aient été utilisées et que certaines ne pouvaient mesurer à 10 kg près. En faisant un graphe des rendements en fonction des parcelles, on note des différences importantes dans la variabilité des traitements.

4 Statistiques descriptives

Plus le nombre de données est grand et moins il est lisible. Il devient donc rapidement utile de regrouper les valeurs afin de faire ressortir les choses importantes. Les techniques graphiques sont particulièrement adaptées pour ce genre d'analyse. Néanmoins, pour des techniques plus avancées, nous renvoyons le lecteur à des livres plus spécialisés et nous concentrerons d'abord notre attention sur les notions de moyenne, de variance et de coefficient de variation.

5 La moyenne

C'est le paramètre de position le plus classique. Intuitivement, elle indique où se trouve le milieu des valeurs. Toute seule la moyenne a peu d'intérêt car elle ne donne aucune information sur la répartition des données ; de plus elle est très sensible aux valeurs extrêmes : quelques valeurs aberrantes peuvent lui enlever toute signification. La moyenne arithmétique d'un ensemble de nombres est égale à la somme des valeurs divisée par le nombre de valeurs.

Exemple 2.5 *La moyenne arithmétique des nombres 4, 20, 30, 54 vaut :*

$$\frac{4 + 20 + 30 + 54}{4} = 27$$

Remarque : *la moyenne arithmétique des nombres 1, 2, 5, 100 vaut aussi 27*

Pour ceux qui aiment les formules, la moyenne arithmétique des valeurs x_1, x_2, \dots, x_n vaut $\bar{x} = \frac{\sum x_i}{n}$

Il est à noter qu'il existe d'autres définitions de la moyenne :

La moyenne géométrique : $G = \sqrt[n]{x_1 \dots x_n}$. On peut noter que $\log G$ correspond à la moyenne arithmétique du logarithme des x_i .

La moyenne quadratique : $Q = \sqrt{\sum_i \frac{x_i^2}{n}}$

La moyenne harmonique : $H = \frac{1}{\sum_i \frac{1}{nx_i}}$

La moyenne géométrique est inférieure à la moyenne arithmétique et les deux moyennes ne sont égales que si toutes les valeurs x_i sont égales entre elles.

La moyenne arithmétique est toujours supérieure à la moyenne harmonique et toujours inférieure à la moyenne quadratique.

La moyenne arithmétique est le paramètre de positionnement le plus utilisé, notamment en raison de ses propriétés statistiques. La moyenne géométrique est utilisée dans l'étude des rapports, par exemple en matière économique dans la définition de certains nombres-indices. Il s'agit en effet de paramètres de position destinés à mesurer les variations relatives d'un ensemble de variables de même nature ou soumises à des influences communes (indice de prix ou de production). La moyenne quadratique intervient, de manière plus ou moins directe, dans la détermination du diamètre de l'arbre de section moyenne ou dans les calculs de surface terrière moyenne.

6 L'écart-type

Nous aimerions aussi savoir comment les valeurs se répartissent autour de la moyenne. La statistique la plus souvent utilisée pour cela est l'écart-type. C'est la racine carrée de la variance qui est donnée par la formule :

$$\hat{\sigma}^2 = s^2 = \frac{\sum_i (x_i - \bar{x})^2}{n - 1} \quad (2.1)$$

L'écart-type peut être vu comme la moyenne des distances à la moyenne. Nous divisons par $n - 1$ plutôt que par n pour des raisons qui seront abordées plus loin.

La quantité $\sum_i (x_i - \bar{x})^2$ s'appelle la somme de carrés des écarts et est notée SCE.

L'écart-type est particulièrement utile quand les données sont symétriques, c'est-à-dire réparties identiquement de chaque côté de leur moyenne, et qu'il n'y a pas de valeurs extrêmes. Dans ce cas, l'interprétation intuitive est que 66% des valeurs sont dans l'intervalle :

moyenne +/- écart-type

et 95% des valeurs sont dans l'intervalle :

moyenne +/- 2 × écart-type

Exemple 2.6 La variance des valeurs 4, 20, 30, 54 vaut :

$$\frac{(4 - 27)^2 + (20 - 27)^2 + (30 - 27)^2 + (54 - 27)^2}{4 - 1} = \frac{1316}{3}$$

Donc l'écart-type vaut 20.94. On peut noter que l'écart-type des valeurs 1, 2, 5, 100 vaut 84.34

7 Le coefficient de variation

Une statistique qui est souvent utilisée pour mesurer la variabilité d'un jeu de données est le coefficient de variation (CV). C'est l'écart-type exprimé comme un pourcentage de la moyenne.

$$CV = 100 \times \frac{\sigma}{\bar{x}}$$

L'idée est de rendre comparable la variabilité de plusieurs jeux de données quand les unités de mesure sont différentes. Il est quand même important de noter que le CV dépend de la moyenne ; il n'est donc utile que lorsque les jeux de données sont de même type et avec des moyennes positives.

Par exemple, les deux ensembles de nombres 1, 2, 3, 4 et 21, 22, 23, 24 ont la même variabilité mais ont des CV de 52% et 5% respectivement.

Le CV donne quelques idées sur la qualité de l'expérience si les niveaux de rendement et les conditions environnementales sont proches. Il n'est pas possible de donner de règle générale pour savoir à partir de quel niveau un CV est acceptable.

Il est intéressant de noter que, dans le cas de mesure avec calibration de l'instrument, les moyennes sont dépendantes de la calibration et donc que les coefficients de variation sont différents suivant la base de la calibration.

Chapitre 3

Expérimentation

1 Introduction

C'est un lieu commun de dire qu'avant d'entreprendre une expérience, ou une série d'expériences, il est indispensable d'établir un protocole, à la base duquel se trouve le plan d'expérience. Celui-ci ne doit pas se borner à dresser l'inventaire des facteurs dont on désire constater ou mesurer l'influence, et à choisir des méthodes d'observation, de mesure ou d'analyse. Si l'on ne prend pas certaines précautions, on risque d'obtenir des résultats confus d'où nulle conclusion précise ne peut être tirée ou, pire encore, d'attribuer à l'effet des facteurs étudiés des écarts apparents qui proviennent de l'imprécision des mesures, de l'intervention d'autres facteurs ou de la variabilité propre à la matière elle-même. Un bon plan expérimental doit permettre de séparer les deux sources de variation — variation contrôlée et variation aléatoire — et de distinguer les écarts significatifs de ceux qui ne le sont pas. D'une façon générale, on appelle facteurs contrôlés :

- les facteurs dont l'étude est l'objet même de l'expérience ;
- les autres facteurs connus susceptibles d'influencer les résultats, lorsqu'il auront été introduits dans le dispositif expérimental de telle façon qu'il soit possible de tenir compte de leur effet lors de leur interprétation.

Les facteurs non contrôlés seront généralement des facteurs à caractère aléatoire.

2 Nécessité des répétitions

Quel que soit le soin exercé dans le choix des parcelles et dans la conduite de l'essai, il est impossible de tirer une conclusion d'une expérience consistant à cultiver une seule parcelle pour un seul traitement. Les multiples facteurs de variation autres que ceux dont on cherche à mesurer l'effet ont pu exercer une action déterminante sur l'observation, et rien ne permet de faire le partage entre l'effet des traitements et celui de tous les autres facteurs. Les présomptions que l'on peut avoir ne donnent aucune garantie, et mieux vaut s'abstenir que d'expérimenter dans de telles conditions.

Si par contre l'expérience est répétée dans plusieurs parcelles unitaires, il devient possible de comparer le résultat des différents traitements et d'en déduire une estimation de la va-

riabilité propre au champ d'expérience, variabilité que l'on désigne généralement sous le terme d'erreur expérimentale. À partir de cette estimation, on peut comparer valablement les moyennes obtenues pour chaque traitement et prononcer un jugement sur leurs valeurs respectives. De plus, un résultat important des statistiques théoriques nous assure que la réponse moyenne a plus de chances d'être proche de la "vraie" moyenne, c'est-à-dire que l'on peut espérer "gommer" en partie les facteurs aléatoires si les répétitions sont placées correctement. Suivant la façon dont l'expérience est organisée, l'erreur expérimentale peut être plus ou moins élevée. Mais qu'elle soit faible et permette une discrimination entre des rendements voisins, ou qu'elle soit forte et conduise à juger non significatifs des écarts apparemment importants, la conclusion reste valable ; seule se trouve en cause la qualité de l'expérience, c'est-à-dire sa précision.

3 Disposition au hasard

C'est ce qu'on appelle plus communément *la randomisation*.

Les facteurs aléatoires existent toujours en agronomie (fertilité, climat, variabilité des plantes...). Chercher à en éliminer le plus possible risque de conduire à créer des conditions artificielles, et les conclusions risqueraient d'être inapplicables dans la pratique. D'autre part, un résultat n'est intéressant que s'il est possible de le généraliser à des conditions suffisamment étendues.

Du fait de la disposition au hasard, toute influence des multiples causes de variation groupées dans l'erreur a autant de chances de s'exercer sur chacun des traitements (gradient de fertilité, lumière...). La disposition au hasard est le seul moyen d'éviter que des liaisons entre parcelles voisines ne favorisent certains traitements au détriment des autres. Elle doit s'exercer sur tous les facteurs perturbateurs connus qui sont susceptibles d'influencer les résultats, afin d'éliminer toute action systématique d'un facteur aléatoire (gradient de fertilité, par exemple).

Pour effectuer une randomisation correcte, il est indispensable de faire appel à un dispositif qui ne laisse aucune place aux préférences personnelles, souvent inconscientes, de l'expérimentateur : tirage de numéro dans une urne ou liste de nombres aléatoires sont vivement recommandés.

Remarque : *Un bon dispositif doit permettre d'évaluer l'effet des traitements et la précision de cette estimation ; nous devons en outre être capables de comparer les différents traitements. Les répétitions, afin d'estimer l'erreur expérimentale, et la randomisation, afin que tous les facteurs aléatoires aient autant de chances d'agir (gradient de fertilité, action conjuguée de deux traitements...), sont deux techniques fondamentales pour atteindre ce but.*

4 Facteurs croisés et facteurs hiérarchisés

Un facteur est dit à effet aléatoire si les niveaux de ce facteur représentent un échantillon non exhaustif de l'ensemble des niveaux de facteurs possibles. Par exemple, en sélection,

des variétés données d'arbres représentent souvent un échantillon aléatoire parmi toutes les variétés possibles d'une espèce. C'est-à-dire que si l'on recommençait l'expérience, des clones différents seraient utilisés.

Si les niveaux du facteur représentent un échantillon exhaustif, on parlera d'un facteur à effet fixe. Par exemple, dans un essai de fertilisation NPK, l'expérimentateur n'est intéressé que par certains niveaux. Les facteurs sont donc à effet fixe.

Dans les modèles à effets croisés, les facteurs peuvent être considérés indépendamment les uns des autres (même si l'interaction peut être un objet de l'étude).

Dans le cas des modèles hiérarchisés, les critères de classification sont subordonnés l'un à l'autre. Par exemple si, pour étudier la progression d'une maladie, on choisit des feuilles au hasard : la feuille 1 de l'arbre 1 n'a rien à voir avec la feuille 1 de l'arbre 2 et l'on dit que le facteur feuille est hiérarchisé dans le facteur arbre. La moyenne des feuilles 1 n'a donc pas de sens et l'effet feuille ne peut être étudié qu'à l'intérieur de chacun des niveaux du facteur arbre.

Dans ce type de problème, le critère de classification subordonné est généralement aléatoire et le critère principal peut être indifféremment fixe ou aléatoire.

5 Dispositif complètement randomisé

Dans ces dispositifs, on utilise les répétitions et la randomisation indépendamment l'un de l'autre :

- chaque traitement est répété plusieurs fois, le nombre des répétitions pouvant varier d'un traitement à l'autre ;
- l'affectation des traitements aux parcelles unitaires est décidée par le hasard, sans aucune restriction (randomisation complète ou totale).

Exemple 3.1 Soient 4 traitements T_1 , T_2 , T_3 et T_4 répartis de façon aléatoire sur respectivement 9, 11, 9 et 7 parcelles. En faisant l'hypothèse que l'essai est réalisé sur un champ d'un seul tenant découpé en 36 parcelles, on peut obtenir le plan d'expérience suivant :

T_1	T_4	T_2	T_4	T_1	T_2
T_4	T_1	T_3	T_2	T_3	T_4
T_2	T_4	T_1	T_2	T_1	T_4
T_1	T_3	T_4	T_3	T_2	T_3
T_3	T_2	T_1	T_2	T_1	T_3
T_2	T_3	T_1	T_2	T_3	T_2

Avantages d'un tel dispositif :

- Souplesse totale : on peut utiliser n'importe quel nombre de traitements et n'importe quel nombre de répétitions. Le nombre de répétitions peut différer entre les traitements (même si cela est un peu plus compliqué à analyser). Tout le matériel expérimental disponible peut donc être utilisé ;

- analyse statistique facile, même si des parcelles sont détruites. La perte relative d'informations due aux parcelles manquantes est plus petite que dans tout autre dispositif.

Inconvénients :

L'inconvénient majeur concerne la précision. La randomisation n'étant soumise à aucune restriction, la totalité de la variation entre les parcelles est imputée au hasard et la précision des tests est faible si une disparité quelconque existe entre les parcelles.

Conclusion :

Un dispositif complètement randomisé pourra convenir si :

- les facteurs de variation autres que les traitements étudiés peuvent être regardés avec certitude comme constants dans toutes les parcelles ;
- une proportion importante de parcelles unitaires est susceptible d'être détruite ;
- le matériel expérimental est relativement homogène.

6 Constitution de blocs

La première condition de la conclusion précédente n'est pratiquement jamais observée dans les expériences mettant en cause la fertilité du sol. On cherche donc à grouper les parcelles unitaires en groupes de tailles à peu près égales, qu'on appelle blocs, de façon qu'à l'intérieur de chaque bloc les facteurs de variation soient homogènes. Les blocs doivent être suffisamment différents pour que la variabilité inter-blocs (d'un bloc à l'autre) soit supérieure à la variabilité intra-bloc (à l'intérieur d'un bloc). Cependant, les blocs ne doivent pas être trop différents si l'on ne veut pas que les traitements agissent de manière différente les uns par rapport aux autres, d'un bloc à un autre. Si le meilleur traitement du premier bloc se révèle être le plus mauvais dans le deuxième bloc, car les conditions d'application sont trop différentes, il sera impossible de comparer les deux traitements.

La randomisation se fait séparément dans chaque bloc de façon que si de légères disparités subsistent entre les parcelles d'un bloc, les traitements du bloc aient des chances égales d'en être affectés.

Selon que tous les traitements figurent ou non dans chaque bloc, les blocs seront complets ou incomplets.

7 Les blocs complets randomisés

Complet signifie que tous les traitements figurent dans chaque bloc. En général, les traitements sont répétés le même nombre de fois, si bien que l'on a autant de répétitions que de blocs, mais ceci n'est pas obligatoire. Si l'intérêt de l'expérience est de tester plus particulièrement un traitement, on lui consacra plus de parcelles unitaires qu'aux autres ; la méthode des blocs peut encore être utilisée si le traitement privilégié est répété le même nombre de fois dans chaque bloc. Il n'existe pas de règles absolues pour définir comment installer les blocs car tout dépend du terrain dont on dispose. Si on ne distingue aucun

gradient de fertilité, on ne peut compter que sur la notion de proximité pour estimer les variations du sol ; les parcelles seront contiguës et le bloc "le plus carré possible", toujours dans le but d'avoir un bloc le plus homogène possible. Les parcelles seront elles-mêmes "les plus carrées possible", afin d'avoir le moins de bordure possible, mais elles seront disposées de façon à avoir le plus de limites communes possibles. Si on distingue un gradient de variation (pente, ensoleillement ...), les blocs seront des rectangles allongés dans la direction perpendiculaire à ce gradient, et les parcelles des rectangles allongés dans le sens de ce gradient – plutôt que des carrés – afin que les parcelles extrêmes d'un bloc ne soient pas trop éloignées l'une de l'autre.

Exemple 3.2 gradient de fertilité —————>

<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>
<i>L</i>	<i>L</i>	<i>L</i>	<i>L</i>
<i>O</i>	<i>O</i>	<i>O</i>	<i>O</i>
<i>C</i>	<i>C</i>	<i>C</i>	<i>C</i>
<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>

gradient de fertilité —————>

Si une analyse de sol ou un autre critère (ensoleillement...) nous conduit à distinguer plusieurs types de sol, nous pourrions être amenés à rassembler dans un même bloc deux zones disjointes de faible surface. Un bloc de ce type est appelé bloc éclaté. Dans la mesure où l'installation des traitements peut prendre un certain temps, il est essentiel de travailler bloc par bloc plutôt que traitement par traitement.

Exemple 3.3 Soient 4 traitements T_1, T_2, T_3 et T_4 répétés 2 fois dans 3 blocs. En faisant l'hypothèse que les blocs sont des champs d'un seul tenant, un plan d'expérience possible est :

<i>Bloc 1</i>	<i>Bloc 2</i>	<i>Bloc 3</i>
T_2 T_4	T_1 T_3	T_1 T_3
T_3 T_3	T_3 T_1	T_4 T_2
T_1 T_4	T_2 T_2	T_2 T_1
T_2 T_1	T_4 T_4	T_3 T_4

Avantages de la méthode :

- À nombre de répétitions égal, un dispositif en blocs complètement randomisés est plus précis qu'un dispositif complètement randomisé, pourvu que le critère ayant servi à implanter les blocs corresponde effectivement à la réalité. Si le vrai gradient de variation est perpendiculaire à ce que l'on a supposé, la précision sera mauvaise et un dispositif complètement randomisé aurait sans doute été meilleur ;
- Cette méthode est très souple car elle n'impose aucune restriction aux nombres de traitements et de répétitions ;
- Elle est très robuste car si tout un bloc est détruit ou si toutes les parcelles soumises à un traitement doivent être écartées de l'analyse, le dispositif reste interprétable sans complication des calculs. En outre si quelques parcelles sont manquantes, les calculs supplémentaires pour les estimer ne sont pas compliqués. Évidemment, si l'on sait au départ que ces parcelles manquantes seront vraisemblablement nombreuses, il est préférable d'adopter un dispositif complètement randomisé.

Inconvénients :

- Dès que le nombre de traitements à comparer devient élevé, il est souvent impossible de trouver une superficie suffisamment homogène pour contenir une répétition complète. Dans ce cas, les blocs doivent être incomplets, c'est-à-dire ne contenir qu'une partie des traitements. La manière d'agencer ces traitements dépasse notre propos ;
- Nous ne pouvons contrôler qu'un seul critère d'hétérogénéité. Pour contrôler plusieurs gradients de variation (double pente, pente et ensoleillement...) nous devons faire appel à d'autres dispositifs, tels les carrés latins, gréco-latins ou lattices.

8 Plans factoriels

Avant de parler d'autres plans d'expérience, nous allons aborder les plans factoriels. Il est fréquent que l'on veuille étudier plusieurs variables, appelées facteurs, ayant plusieurs niveaux d'expérimentation. Par exemple, un chercheur désire connaître l'effet de l'azote et du phosphate sur le rendement d'une espèce d'arbre et il aimerait quantifier la différence de comportement entre les doses 30 kg et 90 kg par parcelle. Les deux facteurs, azote et phosphate, ont donc deux niveaux. Malheureusement il n'est pas évident que l'azote agisse de la même manière suivant que l'on applique 30 kg ou 90 kg de phosphate. Autrement dit, il est fort possible que le niveau de l'azote influence l'effet du phosphate. Cette relation entre les effets est appelée interaction. Il est de loin préférable d'étudier ensemble ces deux facteurs car chaque expérience nous donnera une information sur les facteurs (effets principaux) et sur leur interaction, au contraire de plusieurs expériences séparées. Chaque parcelle donnant une information sur plusieurs aspects, chaque effet est estimé à l'aide de plusieurs parcelles ; nous avons donc une forme de répétition. Si toutes les combinaisons de niveaux de facteurs sont présentes nous parlerons de plan factoriel complet.

Cette méthode est séduisante mais le nombre de combinaisons augmente très vite avec le nombre de facteurs et de niveaux de facteurs. Pour un essai N, P, K avec trois niveaux de facteurs, nous avons $3 \times 3 \times 3 = 27$ combinaisons distinctes ; 4 traitements à 3 niveaux donnent 81 combinaisons, etc.

Dans la méthode classique des blocs, toutes les combinaisons doivent être représentées par une parcelle dans chaque bloc ; ceux-ci occupent alors une surface importante sur le terrain et l'on peut difficilement réaliser de bonnes conditions d'homogénéité à l'intérieur des blocs (sauf peut-être dans les plaines de la Beauce ou du Middle-West...). L'erreur expérimentale, qui est liée à la dispersion des rendements des différentes parcelles d'un même bloc sous un même traitement risque d'être élevée, et de ne faire apparaître aucune différence significative. Le désir d'inclure dans une même expérience un grand nombre de facteurs de variation se trouve limité par l'imprécision dont sont entachés les résultats expérimentaux. La méthode du *confounding* consiste à subdiviser chaque bloc complet que l'on désigne souvent sous le terme de "*répétition complète*" en un certain nombre de blocs plus petits, ne contenant qu'une partie des traitements.

Ceux-ci seront répartis de façon appropriée, en vue de satisfaire aux conditions suivantes :

- abaisser l'erreur expérimentale, qui dépend des variations de fertilité à l'intérieur des blocs, variations plus faibles qu'à l'intérieur des répétitions complètes ;
- permettre le test des effets principaux et de certaines interactions, en sacrifiant une partie de ces dernières, généralement celles d'ordre plus élevé, qui se trouveront entièrement ou partiellement "confondues" avec les différences moyennes entre blocs.

Il est à noter que les plans factoriels peuvent être tout aussi bien mis en place au sein d'un plan complètement randomisé, si la surface est particulièrement homogène.

9 Le Split-plot : Facteurs contrôlés subsidiaires

Dans certaines expériences, on s'arrange pour que chaque résultat puisse être considéré comme appartenant à :

- un traitement principal ;
- une subdivision de ce traitement suivant un traitement secondaire ou subsidiaire.

L'installation des blocs suit les mêmes principes que précédemment. Supposons que le traitement principal (T) soit constitué par plusieurs méthodes de labour sur des parcelles réparties au hasard dans chaque bloc. Au lieu de faire porter l'essai sur une espèce unique, on partage chaque grande parcelle en plusieurs sous-parcelles sur lesquelles on cultive différentes espèces (V) : celles-ci constituent le traitement subsidiaire. Le plan expérimental, de type split-plot, est conforme au schéma ci-dessous.

Bloc 1	Bloc 2	
T_1 V_1 V_2 V_3 V_4	T_2 V_1 V_4 V_3 V_2	etc.
T_2 V_2 V_3 V_1 V_4	T_1 V_4 V_3 V_2 V_1	etc.
T_3 V_1 V_3 V_2 V_4	T_3 V_3 V_1 V_2 V_4	etc.

Si les combinaisons avaient été réparties au hasard dans chaque bloc, une même erreur expérimentale affecterait la comparaison des traitements et des espèces. Dans le dispositif en split-plot :

- l'erreur expérimentale affectant les comparaisons entre traitements provient des facteurs aléatoires à l'intérieur des grandes parcelles et n'est nullement modifiée par l'introduction du facteur subsidiaire Espèces ;
- l'erreur affectant les comparaisons entre espèces résulte des facteurs aléatoires à l'intérieur des sous-parcelles ;
- les comparaisons entre les moyennes Traitement \times Espèces n'ont pas la même précision suivant qu'elles impliquent des niveaux différents ou non du traitement. Ceci provient du fait que les comparaisons pour le même niveau du traitement se font à l'intérieur des grandes parcelles et ne font intervenir que les erreurs des sous-parcelles ; alors que les comparaisons entre différents niveaux de traitement se font entre grandes parcelles et font donc intervenir l'erreur des grandes parcelles et l'erreur des sous-parcelles.

Un inconvénient du split-plot est que les comparaisons n'ont pas la même précision suivant que l'on considère le traitement principal ou les traitements subsidiaires.

La variabilité des grandes parcelles est généralement plus élevée que la variabilité des sous-parcelles et elle est connue avec moins de précision. Il en résulte que les comparaisons ou les tests les plus précis sont effectués sur les traitements a priori secondaires et les interactions de ces traitements avec le traitement principal.

Le dispositif en split-plot présente des analogies avec une technique classique (non présentée ici) : la *confounding*. À l'intérieur d'un bloc les comparaisons entre les niveaux du traitement principal sont confondues avec les comparaisons entre grandes parcelles ; les tests du traitement principal supposent que les différences entre ces parcelles aient un caractère aléatoire (pas de gradient de variation). C'est pour cela que l'on dit souvent que le split-plot "sacrifie" le facteur principal.

Cette inégalité des rôles des facteurs pose le problème de l'utilisation d'un tel dispositif. Dans quelles conditions peut-on admettre de perdre une grande partie de l'information sur un facteur ?

- Si l'on connaît l'information apportée par un facteur ou qu'elle n'est pas intéressante, et que l'on se préoccupe surtout de l'interaction avec le deuxième facteur. On peut ima-

giner un essai Densité de semis \times Espèces d'arbre qui a pour but de tester de nouvelles espèces ainsi que de préciser leur peuplement optimal respectif. Le facteur "densité de semis" est volontairement sacrifié, parce que la comparaison globale faible-forte densité est déjà bien connue et présente peu d'intérêt ; par contre, les différences entre espèces et l'interaction densités \times espèces seront bien testées ;

- Quand un facteur doit être ajouté à une expérience en cours. Par exemple, dans un essai sur un défoliant chimique pour les pommiers, on découvre l'apparition d'une carence en potassium. L'étude de l'interaction défoliant \times carence en potassium grâce à un split-plot évitera le coût d'une nouvelle expérience ;
- Quand il est pratiquement impossible de faire autrement. Dans un essai de labourage ou d'irrigation, la grandeur des parcelles peut justifier l'emploi d'un split-plot, si l'intérêt de base n'est pas seulement le labourage ou l'irrigation mais son interaction avec un autre facteur. Cela permettra de réduire la taille et les coûts de l'expérience, mais il est important de garder à l'esprit que l'effet principal sera mal estimé. Dans les essais variétés \times date de semis, le problème est identique ; il est difficile d'avoir les dates de semis autrement qu'en grandes parcelles. Dans ce cas nous pourrions estimer l'effet des espèces et leur interaction avec les dates de semis correctement, mais nous sacrifierions la comparaison des différents niveaux du facteur date de semis.

10 Conclusion

Nous n'avons fait que soulever le coin du voile qui recouvre la théorie des plans d'expérience : celle-ci est en effet vaste et riche. Néanmoins, il doit être clair que mettre en place un plan d'expérience est une entreprise délicate qui nécessite que le chercheur sache précisément les questions qu'il se pose. Il faudra ensuite vérifier que le plan proposé répond effectivement aux problèmes étudiés.

Chapitre 4

Estimation et tests d'hypothèse

1 Introduction. Notions de probabilité

Le calcul des probabilités est un chapitre important des mathématiques, et il ne saurait être question d'en donner un aperçu, même sommaire. Nous nous contenterons de la définition la plus simple du terme "probabilité" :

La probabilité d'un événement est le rapport du nombre de cas favorables à l'événement sur le nombre de cas possibles, sous la condition que tous les cas soient également vraisemblables.

Par exemple, si l'on prend une carte au hasard dans un jeu bien battu de 32 cartes, la probabilité de tirer un trèfle est de $\frac{8}{32}$, car il y a huit éventualités favorables et 32 cas possibles.

Dans l'étude du sexe des nouveaux-nés sur une population suffisamment étendue, on a constaté qu'il y a à peu près autant de naissances masculines que de naissances féminines. A l'occasion d'une mise au monde prochaine, sur laquelle on ne possède aucun renseignement a priori, il est naturel d'attribuer autant de "chances" aux deux éventualités possibles : garçon ou fille ; en d'autres termes, on dira que la probabilité de naissance d'un garçon (ou d'une fille) est de $\frac{1}{2}$.

Nous passons donc d'une donnée expérimentale, la fréquence, à une donnée théorique, la probabilité. Il faut naturellement se demander dans quelle mesure ce "passage à la limite" est justifié. La réponse, dans les différents cas qui peuvent se présenter, est donnée par des règles déduites des probabilités théoriques. Les conclusions obtenues ne sont d'ailleurs jamais d'une certitude absolue, elles comportent seulement un degré plus ou moins élevé de vraisemblance.

Dans les problèmes pratiques, les probabilités exactes sont en général inconnues et doivent être estimées à partir des données de l'observation.

2 Population totale, échantillon, loi de distribution

Considérons deux séries d'observations relatives au même phénomène. Si les individus constituant l'une et l'autre série ont été pris strictement au hasard, les deux distribu-

tions présentent une allure analogue, et cette ressemblance est d'autant plus étroite que le nombre d'individus est élevé. Lorsque celui-ci augmente, la forme de la distribution a tendance à se stabiliser, et la limite vers laquelle elle tend définit une population hypothétique, dite population totale.

Toute loi de distribution observée sera considérée comme un échantillon issu d'une population totale. La distribution que nous aurions si nous pouvions observer tous les individus de la population qui nous intéresse s'appelle la loi de distribution théorique.

Les lois que l'on rencontre dans la pratique se rapprochent le plus souvent de l'un des types suivant :

1. La loi normale présente un grand caractère de généralité. Lorsqu'une variable se trouve soumise à un grand nombre d'influences aléatoires, toutes très petites et indépendantes les unes des autres, on démontre que sa distribution obéit à la loi normale. En fait, cette loi se rencontre dans l'étude d'un grand nombre de caractères, lorsque les valeurs possibles sont nombreuses ou infinies (variation continue). Dans la loi normale théorique, la variable peut prendre toutes les valeurs, ce qui n'arrive jamais dans la pratique ; ceci ne constitue pas un obstacle car plus on s'éloigne de la moyenne, moins l'on observe de valeurs ; en effet, 95% des valeurs sont contenues dans l'intervalle :

$$\text{moyenne } \pm 1.96 \times \text{écart-type.}$$

Il en résulte, que même dans un échantillon très abondant, les effectifs que l'on pourrait observer en dehors de l'intervalle des valeurs sont tout à fait négligeables. La loi normale est entièrement définie par sa moyenne et son écart-type.

2. La loi binomiale et la loi de Poisson apparaissent dans l'étude de séries où l'on observe l'existence ou l'absence d'un caractère déterminé. La loi binomiale ne dépend que d'un paramètre : la probabilité de l'événement favorable. La loi de Poisson ne dépend que de sa moyenne.

3 Échantillon au hasard

Aussi bien dans les problèmes d'estimation que dans l'application des tests statistiques, on supposera toujours que les échantillons ont été prélevés au hasard. Une méthode d'échantillonnage réalise cette condition lorsque tous les individus de la population ont des chances égales de figurer dans l'échantillon.

La manière pratique de réaliser un tel échantillonnage — qui seul permet une induction correcte de l'échantillon à la population totale — est affaire de cas particuliers. Théoriquement, on peut imaginer que tous les individus de la population portent des numéros, et qu'on tire dans une urne bien brassée les numéros qui désigneront les individus à choisir. Il existe d'ailleurs des tables toutes préparées de "nombres au hasard". On peut imaginer beaucoup d'autres procédés, mais on observera qu'une méthode peut être correcte ou incorrecte suivant l'objet auquel elle s'applique.

Il est très fortement conseillé de faire appel à des procédés dont l'impartialité est évidente : jeux de cartes bien battus, papiers dans l'urne, tables de nombres aléatoires...

4 Notion d'estimateur

Le phénomène étudié serait entièrement défini si l'on connaissait la loi de distribution de la population totale. Il arrive que cette loi soit donnée a priori et que l'on se propose de vérifier que les observations sont en accord avec elle. Mais le plus souvent, la forme de la loi est seulement suggérée par les observations ou par des considérations théoriques ; sa définition complète nécessite le calcul des paramètres figurant dans son expression analytique (c'est-à-dire son équation) : moyenne \bar{x} et écart-type σ , par exemple dans le cas de la loi normale. Le calcul exact est évidemment impossible, puisqu'on ne dispose que d'un échantillon limité. Le mieux que l'on puisse faire, c'est estimer, à partir des données, les meilleures valeurs à adopter pour ces paramètres.

Dépendant des quantités aléatoires x_1, x_2, \dots, x_N , l'estimateur est lui-même une variable aléatoire et suit donc aussi une loi de distribution. La valeur qu'il prend pour un groupe de valeurs (x_1, x_2, \dots, x_N) constitue l'estimation du paramètre considéré dans ce cas particulier. La condition indispensable à laquelle l'estimateur doit satisfaire est de tendre vers la vraie valeur, c'est-à-dire la valeur de la caractéristique pour la population totale, lorsque N augmente de plus en plus. Il existe en général de nombreuses fonctions qui satisfont à cette condition : elles conduisent, sur un échantillon déterminé, à des estimations numériquement distinctes. On choisit, de préférence, l'estimateur le plus précis, c'est-à-dire celui qui, appliqué à tous les échantillons imaginables de N (échantillon "au hasard" naturellement), conduit aux résultats les plus groupés ; cet estimateur est celui pour lequel les fluctuations d'échantillonnage sont les plus faibles.

Les problèmes d'estimation — dont les solutions mathématiques générales dépassent notre propos — comportent deux aspects essentiels :

- le choix de l'estimateur ;
- la précision procurée par l'estimateur, que l'on caractérise soit par son écart-type soit par un intervalle de confiance lié à cet estimateur ; c'est-à-dire un intervalle qui contient presque certainement la vraie valeur de la caractéristique.

Ces notions interviennent dans le chapitre 5.

5 Test d'une hypothèse

L'interprétation statistique utilise les observations pour en tirer des conclusions applicables à la population totale ou, compte tenu des fluctuations d'échantillonnage, à toute autre série issue de la même population. Ce problème peut se définir d'une façon très générale :

Une hypothèse H_0 est formulée. Soit elle se fonde sur des considérations théoriques que l'on se propose de vérifier, soit elle est suggérée par les données elles-mêmes. Sa validité résulte de la réponse à la question suivante :

Si l'hypothèse est exacte sur la population totale, est-il vraisemblable que les observations soient justement celles qui ont été constatées ?

La certitude absolue ne peut être atteinte, car elle nécessiterait la connaissance de la population entière, alors qu'on ne dispose que d'un échantillon. Mais si les données

expérimentales n'ont que très peu de chances d'apparaître par tirage au sort dans la population totale, il est logique de rejeter l'hypothèse ; dans le cas contraire, elle peut être adoptée, au moins provisoirement, et soumise éventuellement à d'autres vérifications.

Soumettre l'hypothèse à cette épreuve de validité s'appelle tester l'hypothèse. L'écart entre l'hypothèse et l'observation est dit significatif – l'hypothèse H_0 doit être abandonnée – lorsque la probabilité attachée à un tel écart est inférieure à un certain seuil.

En fait, la probabilité de voir apparaître un échantillon déterminé par tirage au sort au sein d'une population infinie – ou, dans les mêmes conditions, de voir telle caractéristique prendre une valeur déterminée – est toujours infiniment petite. La notion de seuil de signification s'applique au domaine hors duquel ne doit pas sortir l'échantillon – ou aux limites dont ne doivent pas s'écarter les valeurs expérimentales des caractéristiques – pour que l'hypothèse puisse être considérée comme valable.

Le seuil (ou niveau) le plus fréquemment utilisé en agronomie correspond à une probabilité de 0.05 (5%) ; il conduit à rejeter 5 fois sur 100 en moyenne l'hypothèse alors que celle-ci est vraie. Cette règle est particulièrement pratique pour éprouver la valeur expérimentale d'une caractéristique dont la distribution est normale : nous savons qu'un écart supérieur à 2 fois l'écart-type (écart mesuré de part et d'autre de la moyenne) correspond à une probabilité d'environ 0.05 (5%).

Ces notions interviennent dans le chapitre 5.

6 Risques

D'une façon générale, pour tester l'hypothèse H_0 , on utilise une fonction des observations (par exemple la moyenne ou la variance) dont on connaît la loi de distribution quand l'hypothèse H_0 est vraie. On calcule la probabilité de réalisation de l'événement observé si H_0 est vraie. La décision est soit de rejeter H_0 parce que trop peu vraisemblable (probabilité trop faible de réalisation) soit de ne pas rejeter H_0 car tout à fait vraisemblable. Si l'on rejette H_0 , implicitement, on admet la validité d'une hypothèse alternative H_1 .

Un test d'hypothèse constitue une sorte de démonstration par l'absurde en probabilité. On rejette éventuellement l'hypothèse émise au départ, non pas parce que l'on aboutit à une réelle absurdité ou une réelle impossibilité, mais bien parce que l'on aboutit à une quasi impossibilité.

L'ensemble des valeurs observées pour lesquelles l'hypothèse nulle est admissible forme le domaine d'acceptation ou la région d'acceptation. Les autres valeurs constituent le domaine de rejet, également appelé région de rejet ou région critique. Les valeurs limites sont appelées valeurs critiques ou seuils de signification. Cette dernière expression prête cependant à confusion, car elle est aussi utilisée fréquemment pour désigner le niveau de signification d'un test.

D'une façon générale, une distinction doit être faite entre deux types de tests : les tests unilatéraux et bilatéraux. Nous revenons sur ce point dans l'exemple 5.1 de la page 28.

Deux remarques s'imposent. D'abord, il y a une très grande différence entre ne pas rejeter H_0 et accepter H_0 . Il suffit par exemple de considérer que H_0 correspond à une hypothèse d'égalité des moyennes. Rejeter H_0 car sa probabilité d'apparition est faible

comporte le risque de rejeter H_0 alors que celle-ci est vraie. C'est ce qu'on appelle le risque de 1^{ère} espèce. Il est fixé par l'expérimentateur, généralement à 5% en agronomie. Ce seuil de signification convient parfaitement pour la vérification d'une hypothèse qu'on a des raisons a priori de considérer comme vraie. Mais il ne renseigne nullement sur le risque d'accepter cette hypothèse alors que celle-ci est fausse. C'est le risque de deuxième espèce.

La signification des deux risques apparaît clairement dans le problème de réception d'un lot d'objets à partir d'un échantillon prélevé au hasard : le risque de première espèce est le risque de refuser le lot alors que celui-ci a une qualité acceptable, le risque de deuxième espèce est d'accepter le lot alors que sa qualité est inacceptable.

Un test est d'autant plus puissant que le deuxième risque est faible : la quantité $1 - (\text{risque de 2^{ème} espèce})$ s'appelle la puissance de l'essai. Elle définit la "reproductibilité" d'un résultat ; c'est-à-dire la plus ou moins grande probabilité qu'il se reproduise.

En résumé, nous avons le tableau suivant :

		Réalité	
		H_0	H_1
Décision	H_0	OK	risque de 2 ^{ème} espèce
	H_1	risque de 1 ^{ère} espèce	OK

Chapitre 5

Tests de comparaison de moyennes

1 Fluctuations d'échantillonnage pour la moyenne arithmétique

La moyenne arithmétique \bar{x} d'un échantillon de taille N issu d'une loi normale de moyenne μ et d'écart-type σ est distribuée selon une loi normale de moyenne μ et d'écart-type $\frac{\sigma}{\sqrt{N}}$. Ayant estimé μ par la moyenne arithmétique des données, un intervalle de confiance à 95% de \bar{x} est donné par :

$$\bar{x} \pm 2 \times \frac{\sigma}{\sqrt{N}}$$

L'intervalle est symétrique par rapport à \bar{x} . L'intervalle de confiance à 95% est tel que sa limite supérieure n'a que 2.5 chances sur 100 d'être inférieure à μ et que sa limite inférieure n'a que 2.5 chances sur 100 d'être supérieure à μ . Dans le cas de la loi normale, l'intervalle symétrique est celui qui a la longueur minimum pour une probabilité inférieure donnée. De même, le coefficient 2 est une approximation de la vraie valeur qui peut être lue dans une table de la loi normale ; la valeur exacte vaut 1.96.

Il faut enfin noter que σ est supposé connu. Si N est assez grand (en pratique $N > 30$), σ peut être remplacé par son estimation : $\frac{\sum_i (x_i - \bar{x})^2}{N-1}$

2 Comparaison de deux moyennes

Il est cependant rare que la variance de la population dont est extraite l'échantillon soit connue. Si l'effectif est assez important, il est vraisemblable que l'estimation soit assez proche de la vraie valeur. Si tel est le cas et que le but est de comparer la moyenne observée à une moyenne théorique, il suffit de calculer :

$$T = \sqrt{N} \times \frac{(\bar{x} - \mu)}{\sigma}$$

et de comparer la valeur obtenue à la valeur lue dans la table de la loi normale. Par exemple, au seuil 5% la valeur lue est 1.96. Si la valeur calculée est plus grande que la valeur tabulée, l'hypothèse d'égalité des moyennes est rejetée au seuil décidé.

Si la variance est inconnue, la valeur est lue dans la table de Student. La table t est cependant un peu plus complexe puisqu'elle dépend de la taille de l'échantillon N . En fait, cette table est établie non pas en fonction de N mais en fonction du nombre de degrés de liberté dont dépend la variable. L'expression est formée à partir de l'estimation de σ qui utilise $N - 1$ comparaisons indépendantes (car la somme des écarts à la moyenne est nulle, ce qui explique le diviseur de l'estimateur de l'écart-type). Il faut donc lire la table t avec $N - 1$ degrés de liberté. Il est à noter que la table n'est donnée qu'avec 30 ddl au maximum ; au-delà, la loi normale est utilisée.

Exemple 5.1 *Supposons que la hauteur moyenne d'une espèce d'arbre soit de 5 mètres sous certaines conditions. Dans une expérience avec cinq parcelles d'une nouvelle espèce, la hauteur est estimée à 5.5 mètres avec un écart-type de 0.45 mètre. Est-ce que la nouvelle espèce a une hauteur plus grande que l'ancienne ou est-ce que la différence n'est due qu'au hasard ?*

À l'aide de la formule précédente, la valeur du test est calculée :

$$T = \sqrt{5} \times \frac{(5.5 - 5)}{0.45} = 2.5 \text{ avec 4 ddl}$$

Pour connaître la probabilité que cette différence ne soit due qu'au hasard, cette valeur est comparée à celle de la variable de Student avec $n - 1$ ddl. Pour cela il est important de bien comprendre une subtilité de la table t et des tests d'hypothèses. La question n'est pas de savoir si la hauteur de la nouvelle espèce est différente de l'ancienne mais si elle est plus grande, c'est-à-dire que seuls les hauteurs supérieures à 5.5 intéressent l'expérimentateur.

Si les hauteurs sont inférieures à 5 mètres, il n'y aura aucune évidence que la nouvelle espèce soit plus grande que l'ancienne. Ceci est un exemple de test unilatéral.

Si, par exemple, la question est de tester une balance, les écarts positifs ou négatifs entre le poids lu et le poids de l'étalon sont intéressants, le test est donc bilatéral.

Les valeurs de la table t sont le plus souvent données pour des tests bilatéraux. Pour effectuer, comme dans cet exemple, un test unilatéral, la probabilité lue doit donc être divisée par deux.

La valeur 2.5 se situe entre les colonnes 90 et 95% des valeurs de la table de Student avec 4 ddl. La probabilité que cette différence soit due au hasard est donc entre $\frac{(100-90)}{2}$ et $\frac{(100-95)}{2}$ %

Il arrive souvent que l'on veuille comparer deux moyennes \bar{x}_1 et \bar{x}_2 provenant de deux populations d'effectifs n_1 et n_2 et de variance σ_1^2 et σ_2^2 . La première étape consiste à estimer l'écart-type de la différence : néanmoins, si il y a de bonnes raisons de supposer les deux variance identiques, il est possible d'avoir une estimation conjointe de la variance :

$$\text{ETD}(\bar{x}_1 - \bar{x}_2) = \sqrt{\hat{\sigma}^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

qui possède $n_1 + n_2 - 2$ degrés de liberté. s correspond à la variance calculée en mélangeant les deux échantillons, c'est-à-dire $\hat{\sigma}^2 = \frac{\sigma_1^2 n_1 + \sigma_2^2 n_2}{n_1 + n_2}$

Exemple 5.2 *Considérons un essai où les accroissements en volume des six parcelles d'une nouvelle espèce d'arbre sont : 25, 21, 24, 20, 26 et 22 m³/parc. et les rendements des dix parcelles similaires de l'espèce standard sont : 22, 19, 18, 21, 21, 17, 23, 20, 17, 22 m³/parc. Pour tester si la nouvelle espèce a un meilleur rendement que l'ancienne, il suffit de calculer :*

$$\begin{aligned}\bar{x}_1 &= \frac{25 + 21 + \dots + 22}{6} = \frac{138}{6} = 23.0 \\ \bar{x}_2 &= \frac{22 + 19 + \dots + 22}{10} = \frac{200}{10} = 20.0 \\ \hat{\sigma}_1^2 &= \frac{(25 - 23)^2 + (21 - 23)^2 + \dots + (22 - 23)^2}{5} = \frac{28}{5} = 5.60 \\ \hat{\sigma}_2^2 &= \frac{(22 - 20)^2 + (19 - 20)^2 + \dots + (22 - 20)^2}{9} = \frac{42}{9} = 4.70\end{aligned}$$

Les deux variances 5.6 et 4.7 semblent assez proches, une estimation conjointe des deux variances est donc donnée par :

$$\hat{\sigma}^2 = \frac{28 + 42}{5 + 9} = 5.0 \text{ avec } 14 \text{ ddl et donc } T = \frac{23.0 - 20.0}{\sqrt{5.0 \left(\frac{1}{6} + \frac{1}{10}\right)}} = \frac{3.0}{1.155} = 2.60$$

En comparant 2.6 et les valeurs tabulées de la variable de Student avec 14 ddl, il apparaît que 2.6 est compris entre 95 et 99%. La différence est donc significative au seuil de 2.5% et non pas au seuil de 5% car le test est unilatéral.

Il n'est cependant pas toujours possible de considérer les variances égales. La conséquence est que l'estimation conjointe des variances n'est pas valide et que les degrés de liberté sont trop importants, ce qui a tendance à donner des résultats trop significatifs. Dans ce cas :

$$\text{ETD}(\bar{x}_1 - \bar{x}_2) = \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)^{\frac{1}{2}}$$

D'où l'intervalle de confiance :

$$\bar{x}_1 - \bar{x}_2 \pm t_{n-1}(5\%) \times \text{ETD}(\bar{x}_1 - \bar{x}_2)$$

où n représente le nombre de degrés de liberté d'ETD calculés à l'aide de la formule de Satterthwaite :

$$n = \frac{(w_1 + w_2)^2}{\frac{w_1^2}{n_1 - 1} + \frac{w_2^2}{n_2 - 1}}$$

avec $w_1 = \frac{s_1^2}{n_1}$ et $w_2 = \frac{s_2^2}{n_2}$

Remarque : Il est important de noter que dans l'exemple précédent les deux rendements sont indépendants. Si cela n'était pas le cas, par exemple dans le cas de mesure de poids avant et après traitement, la procédure consisterait alors à créer une nouvelle variable qui serait la différence des deux séries d'observations, ici une variable différence de poids, et à appliquer le test sur la moyenne arithmétique.

3 Tests de comparaison de variances

Dans un but pédagogique, commençons par comparer deux estimations de variance $\hat{\sigma}_1^2$ et $\hat{\sigma}_2^2$ avec $f_1 = n_1 - 1$ et $f_2 = n_2 - 1$ ddl respectivement. Pour tester l'hypothèse d'une différence significative, il faut calculer leur rapport plutôt que leur différence car si les variances sont indépendantes et proviennent d'échantillons suivant une loi normale, la distribution du rapport est connue pour suivre une loi F de Fisher avec f_1 et f_2 ddl.

Exemple 5.3 *Supposons que 1182158 et 1230043 soient les estimations avec 7 ddl de la variance de deux échantillons provenant de population suivant une loi normale. Pour les comparer il faut former leur rapport, la valeur la plus forte étant le numérateur de la fraction :*

$$F = \frac{1230043}{1182158} = 1.04$$

La table des valeurs tabulées de la loi F est à deux entrées. La ligne du haut correspond aux ddl du numérateur (f_1), la ligne du côté correspond aux ddl du dénominateur (f_2). f_1 correspond donc toujours aux ddl de la plus forte variance. Une erreur classique est de vouloir mettre le nombre de ddl le plus élevé au numérateur. Il est à noter que les valeurs théoriques correspondent à un test unilatéral. Si, comme dans l'exemple, l'intérêt est un test bilatéral (les deux variances sont-elles différentes ?), la table 5% correspond à un test au niveau 10%. Dans la table, la valeur $f_1 = 7$ n'est pas donnée. Pour $f_1 = 8$ et $f_2 = 7$, la valeur tabulée vaut 3.73. La valeur 1.04 étant bien inférieure, il n'y a aucune raison de rejeter l'hypothèse d'égalité des variances.

Il existe plusieurs tests permettant de comparer un groupe de variances. L'un des plus utilisés est le test de Bartlett.

4 Test de Bartlett

Supposons que l'on veuille comparer p estimations de variance $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_p^2$ ayant respectivement f_1, f_2, \dots, f_p degrés de liberté (voir équation 2.1).

Une estimation conjointe de la variance vaut :

$$\hat{\sigma}^2 = \frac{f_1 \hat{\sigma}_1^2 + f_2 \hat{\sigma}_2^2 + \dots + f_p \hat{\sigma}_p^2}{f}$$

avec $f = \sum_{i=1}^p f_i$

L'hypothèse nulle $H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_p^2$ peut être testée à l'aide de la statistique :

$$T = \frac{f \log_e \hat{\sigma}^2 - \sum_{i=1}^p f_i \log_e \hat{\sigma}_i^2}{1 + \frac{1}{3(p-1)} \left(\sum_{i=1}^p \frac{1}{f_i} - \frac{1}{f} \right)}$$

où \log_e correspond au logarithme népérien. La statistique T doit être comparée à la valeur d'une variable $\chi_{1-\alpha}^2$ avec $p - 1$ degrés de liberté.

Il s'agit d'un test unilatéral car T est nulle lorsque tous les $\hat{\sigma}_i^2$ sont égaux et donc égaux à $\hat{\sigma}^2$. T est positif dès qu'au moins deux $\hat{\sigma}_i^2$ sont différents. Le rejet de l'hypothèse nulle ne doit donc se faire que dans un sens : lorsque T est trop élevé.

Le fait que T soit positif provient du fait que la moyenne géométrique est inférieure à la moyenne arithmétique si toutes les valeurs ne sont pas égales entre elles.

Comme le test F d'égalité des variances, le test de Bartlett est très sensible à la non-normalité des populations-parents, quels que soient les effectifs des échantillons. De plus, il s'agit d'une méthode approximative qui n'est satisfaisante que si les effectifs sont suffisamment grands et si le nombre d'échantillons p n'est pas trop élevé par rapport à f_i .

Enfin, signalons que, pour deux populations, le test de Bartlett n'est strictement équivalent au test F que si les deux échantillons sont de même effectif ($n_1 = n_2$).

Exemple 5.4 *Soit trois écart-types : 1.359, 1.773, 2.441 estimé avec 13, 14, 10 observations respectivement. La statistique de test vaut 3.50 et la différence entre les trois écarts-types n'est donc pas significative au seuil de 5% (alors que les variances varient du simple au double...).*

Chapitre 6

Analyse de variance

Nous venons de voir comment tester des hypothèses concernant les moyennes de deux populations en utilisant les moyennes et les variances d'un échantillon. Néanmoins dans la pratique nous sommes souvent confrontés à des groupes de plusieurs moyennes. Par exemple nous aimerions tester une hypothèse d'égalité des moyennes sur les rendements de trois variétés de soja. Il ne serait pas très futé d'effectuer tous les tests de comparaison deux à deux possibles car non seulement la méthode risque d'être longue et laborieuse mais elle risque de conduire à des conclusions erronées car toutes ces comparaisons ne sont pas indépendantes. La technique utilisée pour effectuer ces tests s'appelle l'analyse de variance et son principe est très simple. Si nos variétés de soja ont la même moyenne et la même variance, toutes les parcelles présenteront la même variabilité, due à l'erreur expérimentale (hétérogénéité du sol, du matériel végétal, des pratiques culturales...). Si d'autre part les différentes variétés de soja ont la même variance mais des moyennes de rendement différentes, la variabilité entre deux parcelles plantées avec des variétés différentes sera certainement plus élevée que celle observée entre deux parcelles plantées avec une même variété. Le principe de base de l'analyse de variance est donc de chercher à voir si la variabilité des observations peut être en partie expliquée par les différences entre traitements. Pour cela nous aurons naturellement besoin de deux outils développés dans le chapitre précédent :

- comment comparer les estimations de deux variances à l'aide d'un test F ;
- comment obtenir une estimation de l'erreur expérimentale.

Néanmoins l'utilisation de l'analyse de variance présuppose quelques hypothèses. Même si elles sont en général raisonnables, il est important de les vérifier. Les hypothèses de base sont :

- Les échantillons de chaque population sont statistiquement indépendants ;
- Les populations ont toutes la même variance ;
- Les populations suivent une loi normale.

1 Cas d'un seul facteur.

La situation la plus simple est lorsque nous avons un échantillon provenant de plusieurs populations caractérisées par un seul facteur. C'est le cas par exemple lorsque des traitements ou des variétés ont été testés avec un dispositif complètement randomisé.

Exemple 4.2 : Supposons que nous ayons installé une expérience sur trois variétés avec quatre répétitions. Les rendements des douze parcelles sont : 3.8, 4.5, 3.8, 3.9, 4.3, 5.0, 4.5, 5.0, 4.1, 4.3, 4.7, 4.9 kg/parc.

La moyenne des rendements vaut 4.4 kg/parc. La somme des carrés des écarts vaut :

$$SCE = (3.8 - 4.4)^2 + (4.5 - 4.4)^2 + (3.8 - 4.4)^2 + \dots + (4.9 - 4.4)^2 = 2.16$$

Il y a 11 degrés de liberté (ddl) et la variance théorique que l'on a estimée d'après l'échantillon vaut : $\frac{2.16}{11} = 0.197$

Si nous avons utilisé la même variété pour les douze parcelles, la variabilité serait due aux différences entre parcelles. Comme nous avons utilisé trois variétés nous pouvons nous demander si les différences entre variétés n'ont pas introduit une nouvelle source de variabilité.

Les quatre premiers rendements correspondent à la variété A, les quatre suivants à la variété B et les quatre derniers à la variété C. la variance théorique que l'on a estimée d'après l'échantillon s'appelle le carré moyen (CM)

Rendement variété A : 3.8, 4.5, 3.8, 3.9 ; moy= 4.0 ; SCE= 0.34 ; DDL= 3 et CM= 0.113

Rendement variété B : 4.3, 5.0, 4.5, 5.0 ; moy= 4.7 ; SCE= 0.38 ; DDL= 3 et CM= 0.127

Rendement variété C : 4.1, 4.3, 4.7, 4.9 ; moy= 4.5 ; SCE= 0.40 ; DDL= 3 et CM= 0.133

Les rendements totaux de chaque variété sont de 16.0, 18.8 et 18.0 kg/parcelle. et la moyenne de ces rendements totaux vaut : $\frac{16.0+18.8+18.0}{3} = 17.6$ La somme de carrés des écarts inter-variétés vaut :

$$SCE = \frac{(16.0 - 17.6)^2 + (18.8 - 17.6)^2 + (18.0 - 17.6)^2}{4}$$

La quantité au numérateur correspond à la formule classique pour calculer la SCE mais comme chaque nombre est le total de quatre parcelles nous divisons par quatre pour avoir la SCE inter-variétés par parcelle. La SCE vaut donc 1.04 avec 2 ddl ce qui nous donne un CM de 0.52.

Nous pouvons remarquer que si nous ajoutons les SCE des trois variétés à la SCE inter-variétés nous retrouvons la SCE totale calculée au début (2.16) avec le même nombre de ddl (11). Nous avons donc divisé la SCE totale en quatre parties. Une des hypothèses de base étant l'égalité des variances pour les trois variétés, les trois carrés moyens sont une estimation de la même variance et nous pouvons donc obtenir une estimation globale :

$$\begin{aligned} \text{SCE} &= 0.34 + 0.38 + 0.40 = 1.12 \\ \text{DDL} &= 3 + 3 + 3 = 9 \\ \text{CM} &= \frac{1.12}{9} = 0.124 \end{aligned}$$

Cette partition de la SCE totale est généralement présentée sous la forme d'un tableau, appelé tableau d'analyse de variance :

Source	DDL	SCE	CM	F
Inter-variétés	2	1.04	0.520	4.18
Intra-variété	9	1.12	0.124	
Total	11	2.16		

Notre hypothèse nulle était qu'il n'y avait pas de différence entre les rendements moyens des trois variétés. Nous devons maintenant savoir quelle est la probabilité d'observer un CM inter-variétés au moins aussi élevé si notre hypothèse nulle est vraie.

Ceci nous conduit à effectuer un test F , $F = \frac{0.520}{0.124} = 4.18$ avec 2 et 9 ddl. Cette valeur n'est pas plus grande que la valeur tabulée de la distribution F avec 2 et 9 ddl au seuil 5%. Nous ne pouvons donc pas rejeter l'hypothèse d'égalité des rendements des trois variétés.

2 Cas de deux facteurs

Un autre utilisation classique de l'analyse de variance est le cas où nous avons deux facteurs, par exemple différentes variétés dans plusieurs blocs.

Exemple 4.3 : Supposons que nous ayons testé trois variétés sur douze parcelles réparties en quatre blocs complets randomisés

REP I			REP II			REP III			REP IV		
PARC	VAR	RDT	PARC	VAR	RDT	PARC	VAR	RDT	PARC	VAR	RDT
1	B	4.3	6	C	4.3	7	B	4.5	12	A	3.9
2	A	3.8	5	B	5.0	8	C	4.7	11	C	4.9
3	C	4.1	4	A	4.5	9	A	3.8	10	B	5.0

Commençons par former un tableau des totaux

VARIETE	REP I	REP II	REP III	REP IV	TOTAL
A	3.8	4.5	3.8	3.9	16.0
B	4.3	5.0	4.5	5.0	18.8
C	4.1	4.3	4.7	4.9	18.0
TOTAL	12.2	13.8	13.0	13.8	52.8

Nous pouvons vérifier que la somme des lignes est égale à la somme des colonnes. Comme précédemment la SCE totale vaut 2.160 avec 11 ddl, et la SCE entre variétés vaut 1.040 avec 2 ddl.

Une nouvelle source de variabilité est due aux blocs, nous avons donc une nouvelle partition de la SCE totale. La moyenne des quatre blocs vaut $\frac{12.2+13.8+13.0+13.8}{3} = 13.2$

$$\begin{aligned} \text{SCE} &= \frac{(12.2 - 13.2)^2 + (13.8 - 13.2)^2 + (13.0 - 13.2)^2 + (13.8 - 13.2)^2}{3} \\ &= 0.587 \text{ avec 3 ddl} \end{aligned}$$

Les totaux étant la somme de trois parcelles, le numérateur est divisé par trois pour ramener le résultat à l'échelle d'une parcelle.

La variation résiduelle est calculée en soustrayant les variations dues aux facteurs (ici variété et bloc) à la variation totale :

$$\text{SCE} = (2.160 - 0.587 - 1.040) = 0.533 \text{ avec } 11 - 3 - 2 = 6 \text{ ddl}$$

Nous obtenons donc le tableau d'analyse de variance :

Source	DDL	SCE	CM	F
Blocs	3	0.587	0.196	
Variétés	2	1.040	0.520	5.84*
Résiduelle	6	0.533	0.089	
Total	11	2.160		

Tester un effet bloc ne sert qu'à vérifier que les blocs ont été bien formés. Ici nous avons testé seulement le CM variété par rapport au CM résiduel. La valeur 5.84 est largement supérieure à la valeur tabulée au niveau 5% d'une distribution F avec 2 et 6 ddl. Nous en concluons que la variation entre les parcelles ayant reçu une même variété est plus faible qu'entre celles ayant reçu des variétés différentes (une fois la variabilité due au bloc prise en compte). Nous voyons comment le bloc a accru la précision de l'expérience car dans l'exemple précédent nous avons trop de variabilité entre parcelles ayant reçu la même variété pour mettre en évidence quoi que ce soit. On peut noter que la somme des SCE dues aux blocs et à la résiduelle est égale à la SCE intra-variété de l'exemple précédent.

3 Cas non orthogonal

Lorsque les effectifs ne sont pas égaux, l'analyse de variance à deux facteurs se complique sensiblement, tant du point de vue pratique que du point de vue théorique. Ceci peut arriver par exemple lorsque le matériel étudié ne permet pas d'avoir des échantillons de même importance pour les différents facteurs, ou lorsque de nombreuses données sont manquantes.

La complication essentielle provient du fait que les différentes sommes des carrés des écarts ne sont plus indépendantes les unes des autres. Elles ne sont donc plus additives. Concrètement, ceci signifie que les valeurs du test F sont différentes si l'ordre d'étude des facteurs est changé.

Sans développer l'étude des dispositifs non orthogonaux, nous allons présenter succinctement deux approches de ce problème :

- La première méthode consiste à calculer l'apport des facteurs en fonction d'un ordre pré-établi (en général, leur ordre d'introduction dans le modèle). Dans ce cas, la somme des SCE associée aux facteurs plus la SCE résiduelle vaut bien la SCE totale. Il est par contre difficile de tirer des conclusions générales au niveau d'un facteur précis car les résultats dépendent de l'ordre. Imaginons par exemple, que l'effet de deux facteurs très dépendants soit mesuré. Le facteur interprété en second a peu de chance d'être significatif car la variabilité qu'il provoque sur les observations aura déjà été prise en compte par le premier facteur. Une inversion de ces deux facteurs aurait conduit aux résultats inverses. Les SCE calculées selon le principe d'un ordre des facteurs sont en général appelées SCE type I.
- La seconde méthode classique est de considérer l'effet de chaque facteur une fois que la variabilité associée à tous les autres a déjà été soustraite. L'ordre d'entrée dans le modèle n'a donc plus d'importance car l'on considère que chaque facteur est étudié en dernier. Ceci donne une bonne idée de l'apport spécifique de chacun des traitements. Il y a néanmoins deux inconvénients : tout d'abord la somme des SCE associée à chaque facteur et de la SCE résiduelle ne vaut plus la SCE totale (car les SCE ne sont pas indépendantes). Il est donc impossible de connaître la part de la variation totale attribuable à chaque facteur. De plus, si deux facteurs sont très corrélés, il y a de forts risques qu'aucun des deux facteurs ne soit significatif car la variabilité associée à un facteur aura déjà été prise en compte dans le modèle. Les SCE calculées selon ce principe sont en général appelées SCE type III.

Quelle que soit la méthode utilisée, l'interprétation des résultats dans le cas non orthogonal est souvent délicate et nécessite un examen approfondi des résultats.

Chapitre 7

Comparaisons entre les niveaux des facteurs

Lorsqu'à l'issue d'une analyse de variance, on est amené à rejeter une hypothèse d'égalité des niveaux de facteurs, la question se pose généralement de savoir quelles sont, parmi les moyennes considérées, celles qui diffèrent significativement. Cette question est en général résolue par les différentes méthodes de comparaisons multiples. Dans l'exposé de ces différentes méthodes, nous supposons toujours que les conditions de base de l'analyse de variance sont respectées.

1 La plus petite différence significative (PPDS)

Les méthodes de comparaison de moyennes ont été étudiées précédemment. Dans la mesure où l'analyse de variance n'est applicable que si les variances des populations sont égales, il est possible d'utiliser le carré moyen résiduel (CMR) comme estimation de la variance. Si de plus les moyennes sont estimées avec le même nombre n d'observations, la différence entre deux moyennes μ_1 et μ_2 est significative au seuil de 5% si cette différence est supérieure à

$$PPDS = \sqrt{\frac{2 * CMR}{n}} \times t(5\% \text{ et avec le nombre de ddl du CMR.}) \quad (7.1)$$

La valeur $\frac{2}{n}$ correspond à $\frac{1}{n} + \frac{1}{n}$ dans l'expression de la différence de deux moyennes, et le t se lit dans la table de Student.

Bien que fort populaire, cette méthode est peu correcte. En effet, les comparaisons ne sont pas indépendantes et, vu le nombre élevé de comparaisons, on est quasiment certain de conclure à des différences qui n'existent pas.

En réalité, la définition qui vient d'être donnée pour la plus petite différence significative n'est valable que pour l'analyse de variance à un facteur. D'une manière générale, le carré moyen résiduel doit être remplacé par le carré moyen qui a servi de base de comparaison pour les moyennes considérées (souvent le carré moyen de l'interaction), et l'effectif n est le nombre total d'observations qui interviennent dans chacune des moyennes. En

outre, le nombre de degrés de liberté de la distribution t est toujours le nombre de degrés de liberté du carré moyen qui a servi de base de comparaison.

2 Les tests de Duncan et de Newman-Keuls

Ces tests sont préférables à la PPDS si l'expérimentateur souhaite comparer tous les niveaux des facteurs. Ils sont basés sur la comparaison des amplitudes observées pour des groupes de 2, 3, ... moyennes avec une amplitude maximale théorique correspondant à une probabilité donnée de réalisation.

Concrètement, ceci signifie que l'on commence par calculer la plus petite amplitude significative relative à des groupes de deux moyennes, de trois moyennes, etc.

La plus petite amplitude significative est donc, pour un niveau de signification α , pour un nombre de moyennes et de degrés de liberté donnés :

$$q(1 - \alpha) \sqrt{\frac{2 * CMR}{n}}$$

Ce calcul nécessite l'emploi de tables particulières fournissant les valeurs $q(1 - \alpha)$ qui remplacent les quantités $\sqrt{2}t(1 - \alpha/2)$ de la PPDS (voir équation 7.1).

Les valeurs $q(1 - \alpha)$ sont calculées de telle sorte que, pour p populations normales de même moyenne et de même variance, la probabilité de dépasser la plus petite amplitude significative soit égale à α .

Après avoir déterminé de cette manière les amplitudes maximales admissibles relatives aux différents nombres de moyennes, on compare l'amplitude observée des p moyennes considérées à la plus petite amplitude significative correspondante, puis on traite de la même manière l'amplitude des différents groupes de $p - 1$ moyennes, de $p - 2$ moyennes, etc. On continue jusqu'au moment où l'amplitude de certains groupes ne dépasse plus la limite admissible correspondante. Les moyennes de ces groupes sont alors considérées comme non significativement différentes.

Il est bien évident, comme pour la PPDS, que le carré moyen résiduel doit éventuellement être remplacé par le carré moyen qui sert réellement de base de comparaison, et l'effectif n par le nombre d'observations qui interviennent dans le calcul de chacune des moyennes. De même, le nombre de degrés de liberté est, d'une façon générale, celui du carré moyen servant de base de comparaison.

Les différences entre les deux tests proviennent de l'utilisation de tables statistiques différentes. De manière générale, les résultats du test de Duncan sont intermédiaires entre ceux obtenus par la PPDS et ceux obtenus par Newman-Keuls.

Fort logiquement, les résultats sont identiques pour le cas de deux moyennes.

Il est cependant fort rare que l'expérimentateur soit intéressé par toutes les comparaisons, l'information étant divisée pour tester toutes les comparaisons possibles, il y a fort à craindre que les tests vraiment utiles soient peu puissants.

3 Autres méthodes de comparaisons multiples de moyennes

Il est peu puissant d'effectuer toutes les comparaisons de moyennes deux à deux et il faut mieux se limiter à quelques comparaisons particulières. Suivant les cas, différents tests de comparaisons multiples sont possibles et il n'est pas raisonnable de faire un recensement exhaustif des différents tests de comparaisons multiples proposés dans la littérature. Deux cas particuliers sont cependant assez fréquents :

Test de Dunnett : l'expérimentateur peut vouloir comparer un ensemble de traitements à un témoin. On effectue dans ce cas, pour p populations, $p - 1$ comparaisons. L'erreur globale de première espèce propre à la méthode de la plus petite différence significative doit en conséquence être réduite dans une moindre mesure que dans le cas général, où le nombre de comparaisons à effectuer est égal à $p(p - 1)/2$.

Comme précédemment, la plus petite amplitude significative est donc, pour un niveau de signification α , pour un nombre de moyennes et de degrés de liberté donnés :

$$d(1 - \frac{\alpha}{2})\sqrt{\frac{2 * CMR}{n}}$$

où n et CMR sont définis comme précédemment.

Les valeurs $d(1 - \alpha)$ sont calculées de telle sorte que, pour p populations normales de même moyenne et de même variance, la probabilité de dépasser la plus petite amplitude significative soit égale à α .

Test de Gupta : un autre problème particulier, qui se pose notamment en matière de sélection est la recherche de la ou des moyennes les plus élevées. D'une manière concrète, le but est de déterminer le plus petit ensemble de moyennes observées \bar{x}_i qui ait une probabilité donnée, par exemple $1 - \alpha$, d'englober la population de moyenne théorique maximum. Pour cela, on réunit dans un même groupe les valeurs \bar{x}_i telles que :

$$\bar{x}_i \geq \bar{x}_{\max} - d(1 - \alpha)\sqrt{\frac{2 * CMR}{n}}$$

où n et CMR sont définis comme précédemment, \bar{x}_{\max} désigne la moyenne observée la plus élevée, et $d(1 - \alpha)$ représentent les valeurs du test de Dunnett pour un test unilatéral. La méthode revient donc à effectuer un test unilatéral de comparaison de $p - 1$ moyennes observées avec une moyenne témoin, la valeur maximum servant de témoin.

4 Exemple

Quarante-deux pots, contenant chacun quatorze plants, ont été préparés en vue de comparer sept fumures différentes. Ils ont été répartis au hasard en six lots de sept pots et, ensuite, les sept traitements ont eux-mêmes été répartis au hasard à l'intérieur de chacun des six lots. A l'issue d'une période donnée, les accroissements individuels en hauteur ont

été mesurés, en centimètres, et ont conduit à la réalisation d'une analyse de variance qui a mis en évidence l'existence de différences très hautement significatives entre les sept fumures considérées. Le carré moyen de l'interaction bloc*fumure vaut 5.619, avec 30 degrés de liberté. De plus, il y a 84 individus par niveau de fumure (6 lots et 14 individus par pots).

La plus petite différence significative, au seuil 5%, et en centimètres, vaut :

$$t_{0.975} \sqrt{\frac{2 * CMR}{n}} = 2.042 \sqrt{2 \times 5.619/84} = 0.75$$

Ce qui donne le classement des moyennes :

\bar{x}_5	\bar{x}_7	\bar{x}_4	\bar{x}_1	\bar{x}_2	\bar{x}_6	\bar{x}_3
7.49	8.23	8.44	8.67	9.64	9.69	10.34

Même si toutes les comparaisons deux à deux n'ont que peu d'intérêt dans cet exemple, nous pouvons, à titre d'exemple, appliquer la méthode de Newman Keuls.

Les plus petites amplitudes significatives au seuil 5% et pour 30 degrés de liberté sont :

pour deux populations : $2.89 \sqrt{5.169/84} = 0.75$

pour trois populations : $3.49 \sqrt{5.169/84} = 0.90$

pour quatre populations : $3.85 \sqrt{5.169/84} = 1.00$

pour cinq populations : $4.10 \sqrt{5.169/84} = 1.06$

pour six populations : $4.30 \sqrt{5.169/84} = 1.11$

pour sept populations : $4.46 \sqrt{5.169/84} = 1.15$

On retrouve logiquement pour deux populations la valeur de la plus petite différence significative.

Les moyennes ordonnées des sept niveaux de fumure ont été données précédemment.

La plus grande amplitude ($\bar{x}_3 - \bar{x}_5 = 2.85$) dépasse largement la valeur maximum admissible pour sept populations. Ceci est cohérent avec les résultats de l'analyse de variance. En partant de la moyenne la plus faible, on constate que :

$$\bar{x}_5 + 1.15 = 7.49 + 1.15 = 8.64 < \bar{x}_1$$

Puisque l'amplitude relative à sept populations (1.15) conduit à une valeur inférieure à \bar{x}_1 , il en est de même pour les amplitudes relatives à six, cinq et quatre populations. On peut donc immédiatement considérer le cas de trois moyennes, pour lequel on a :

$$\bar{x}_5 + 0.90 = 7.49 + 0.90 = 8.39 < \bar{x}_4$$

on ne peut donc pas regrouper \bar{x}_5 , \bar{x}_7 et \bar{x}_4 .

Pour deux moyennes, on a :

$$\bar{x}_5 + 0.75 = 7.49 + 0.75 = 8.44 > \bar{x}_7$$

On ne peut donc pas rejeter l'hypothèse $m_5 = m_7$.

De même, en recommençant la même opération à partir de la deuxième moyenne, on a :

$$\bar{x}_7 + 0.90 = 8.23 + 0.90 = 9.13 > \bar{x}_1$$

On ne peut donc pas rejeter l'hypothèse $m_7 = m_4 = m_1$.

De même pour la troisième moyenne :

$$\bar{x}_4 + 1.06 = 8.44 + 1.06 = 9.50 < \bar{x}_2$$

et on n'effectue pas la comparaison de \bar{x}_4 et \bar{x}_1 puisque l'hypothèse $m_4 = m_1$ a déjà été acceptée.

En continuant de la même façon, on obtient successivement :

$$\bar{x}_1 + 1.00 = 8.67 + 1.00 = 9.67 < \bar{x}_6$$

$$\bar{x}_1 + 0.75 = 8.67 + 0.75 = 9.42 < \bar{x}_2$$

et

$$\bar{x}_2 + 0.90 = 9.64 + 0.90 = 10.54 > 10.34$$

ce qui conduit à l'acceptation de l'hypothèse $m_2 = m_6 = m_3$.

Au total, le résultat obtenu est, dans le cas présent, identique à celui qui découle de l'utilisation de la méthode de la PPDS mais le risque de première espèce est nettement inférieur. On remarque en outre que, si les méthodes de comparaisons multiples permettent de ramener le risque d'erreur à une limite raisonnable, elle ne permettent pas d'éviter les difficultés pratiques résultant du chevauchement des groupes de moyennes. Les méthodes de Duncan, Dunnett et Gupta s'appliquent de la même manière.

5 La méthode des contrastes

Dans la plupart des essais concernés, les chercheurs choisissent les niveaux des facteurs en souhaitant effectuer des comparaisons particulières. Par exemple, dans un essai variétal, la variété A correspond à la variété vulgarisée alors que les variétés B et C sont nouvelles. Le chercheur peut vouloir comparer les nouvelles variétés avec l'ancienne ainsi que les nouvelles variétés entre elles. Ceci est un exemple de questions précises que le chercheur peut se poser avant de faire une expérience.

Se demander si le rendement des nouvelles variétés est meilleur revient à poser la question :

Est-ce que $(2 \times \text{moyenne de A} - \text{moyenne de B} - \text{moyenne de C})$ est significativement différent de zéro ?

Ce genre de comparaison s'appelle des contrastes car la somme des coefficients des moyennes :

$$2 \quad -1 \quad -1$$

vaut zéro.

Si un facteur a k niveaux et que le but est de tester un contraste entre ces niveaux avec les coefficients C_1, C_2, \dots, C_k , il faut avoir $\sum_i C_i = 0$ et le contraste vaut :

$$\text{SCE contraste} = \frac{(\sum_i C_i T_i)^2}{r \sum_i C_i^2}$$

où T_i représente la somme des valeurs du niveau i du facteur étudié et r le nombre total de parcelles servant à calculer les totaux (c'est-à-dire le nombre d'observations de chacun des k niveaux du facteur).

Cette SCE a toujours un seul degré de liberté et l'hypothèse nulle est testée en comparant le rapport :

$$F = \text{SCE contraste} / \text{CMR}$$

les degrés de liberté valent 1 pour la SCE contraste et sont ceux de l'erreur pour le CMR.

Exemple 7.1 Soit une expérience en blocs complets randomisés dont le but est de comparer la variété vulgarisée A avec les nouvelles variétés B et C.

Variété	A	B	C
Total	16.0	18.8	18.0
Coefficients	2	-1	-1

$$Q_1 = \text{SCE contraste} = \frac{(2*16.0 - 1*18.8 - 1*18.0)^2}{4*6} = 0.960$$

$$F = \frac{0.960}{0.089} = 10.7 \text{ avec 1 et 6 ddl}$$

Ce test est significatif au seuil 5%.

L'autre contraste évident consiste à comparer les variétés B et C. Les coefficients seront donc

$$0 \quad 1 \quad -1$$

et :

$$Q_2 = \text{SCE contraste} = \frac{(0*16.0 + 1*18.8 - 1*18.0)^2}{4*2} = 0.080$$

ce qui donne un F qui ne peut être significatif car plus petit que 1.

Il est intéressant de noter que la somme des SCE des deux contrastes (Q_1 et Q_2) est égale à la SCE des variétés (voir table 7.1). En fait, pour tester des hypothèses particulières, la SCE des variétés a été divisée en deux SCE.

De manière générale la SCE d'un facteur à k niveaux peut être divisée en k contrastes si les SCE sont orthogonales. Si les contrastes sont orthogonaux, la somme du produit des coefficients de deux contrastes pris au hasard sera toujours égale à zéro. Dans l'exemple précédent il n'y a que deux contrastes, la vérification est donc rapide :

$$(2 * 0) + (-1 * 1) + (-1 * -1) = 0$$

Quand les contrastes sont orthogonaux, les tests effectués sont statistiquement indépendants et le seuil du test (5% ou 1%) est préservé pour tous les tests. Si les comparaisons ne sont pas indépendantes, la somme des SCE des contrastes ne vaut pas la SCE du facteur, les tests ne seront pas indépendants et le seuil du test sera inconnu. Pour un facteur à k niveaux, il n'est pas possible d'avoir plus de k tests indépendants et il est utile de réfléchir un peu avant de les poser.

Les contrastes sont souvent inclus dans le tableau d'analyse de variance (voir la table 7.1)

Tableau 7.1 – table d'analyse de variance

Source	DDL	SCE	CM	F
Inter-blocs	3	0.587	0.196	
Inter-Variétés	2	1.040	0.520	5.84*
A versus B & C	1	0.960	0.960	10.70**
B versus C	1	0.080	0.080	0.89
Résiduelle	6	0.533	0.089	
Total	11	2.160		

Lorsque les niveaux d'un facteur sont quantitatifs, comme par exemple des doses d'engrais ou de phytocides, il peut être intéressant de savoir si la réponse aux différents niveaux du facteur suit une ligne droite, ou une courbe particulière (quadratique ou cubique par exemple).

Les coefficients de ces tests s'appellent les polynômes orthogonaux et leur valeur se trouve facilement dans toute bonne table statistique dans le cas où les intervalles entre les niveaux restent constants. Le tableau ci-dessous donne les plus courants. Ils s'utilisent de la même manière que les autres coefficients.

Tableau 7.2 – polynômes orthogonaux linéaires et quadratiques

Nb DE NIVEAUX	POLYNÔME	COEFFICIENTS : C_i					$\sum_i C_i^2$
3	LINEAIRE	-1	0	+1			2
3	QUADRATIQUE	+1	-2	+1			6
4	LINEAIRE	-3	-1	+1	+3		20
4	QUADRATIQUE	+1	-1	-1	+1		4
5	LINEAIRE	-2	-1	0	+1	+2	10
5	QUADRATIQUE	+2	-1	-2	-1	+2	14

Chapitre 8

Hypothèses de l'analyse de variance

Les techniques d'analyse de variance sont tellement utilisées que l'on oublie parfois dans quels cadres elles s'appliquent. En effet, avant de se précipiter sur la valeur du test F de Fisher, il est indispensable de vérifier que les hypothèses de base sont bien respectées. A savoir :

1. Les effets traitements et les effets blocs sont additifs ;
2. les erreurs expérimentales sont indépendantes ;
3. les erreurs expérimentales ont toutes la même variance ;
4. les erreurs expérimentales suivent une loi normale.

Le non-respect d'une ou plusieurs de ces hypothèses affecte de manière sensible le niveau du test F de Fisher. C'est-à-dire que la probabilité associée à la valeur calculée perd une grande partie de sa signification.

Nous allons voir dans ce paragraphe ce que signifient ces hypothèses. Puis dans les deux paragraphes suivants nous proposerons des remèdes lorsque ces hypothèses ne sont pas vérifiées.

1 Test de normalité

Tous les tests développés en analyse de variance supposent que les observations suivent une loi normale. Cette hypothèse est donc fondamentale et de nombreux tests ont été proposés dans la littérature. Le but de ce paragraphe n'est pas de faire un tour exhaustif de la question mais plutôt de présenter les principaux tests d'ajustement. Nous nous limiterons au cas de la loi normale mais il est facile de tester l'ajustement à une autre loi. Par exemple, l'ajustement à une loi de Weibull dans le cas de la distribution diamétrique d'arbres.

Dans le cas de l'analyse de variance, il est préférable d'effectuer les tests de normalité sur les résidus car leur distribution correspond à la distribution centrée (c'est-à-dire de moyenne nulle) des observations.

D'une façon générale, on considère dans les tests d'ajustement, d'une part une population infinie dont les individus sont classés en p catégories, en fonction d'un critère qualitatif

ou quantitatif, et d'autre part, un échantillon aléatoire simple d'effectif n dont les individus sont classés de la même manière. Le but est de vérifier si la population possède une distribution de probabilité P_1, P_2, \dots, P_p telle que :

$$\sum_{i=1}^p P_i = 1$$

1.1 Test d'ajustement du χ^2

Supposons tout d'abord que la distribution de référence est complètement définie, c'est-à-dire que P_1, P_2, \dots, P_p est connue. Soit n_1, \dots, n_p les effectifs observés dans chacune des p classes ($\sum_i n_i = n$) La distance du χ^2 entre la distribution théorique et la distribution empirique est définie par :

$$\chi_{obs}^2 = \sum_{i=1}^p \frac{(n_i - nP_i)^2}{nP_i} \quad (8.1)$$

Les dénominateurs nP_i (effectifs attendus) donnent une importance relative suffisante aux fréquences n_i et nP_i les plus petites.

Lorsque l'hypothèse nulle d'ajustement à la loi théorique est vraie, la quantité χ_{obs}^2 suit approximativement une distribution χ^2 à $p - 1$ degrés de liberté.

Cette propriété résulte du fait que chacune des p fréquences peut être considérée comme une variable binomiale et donc asymptotiquement comme une loi normale.

La valeur χ_{obs}^2 est nulle lorsque les fréquences observées sont toutes égales aux fréquences attendues, c'est-à-dire lorsqu'il y a concordance complète entre la distribution observée et la distribution théorique. On rejette donc l'hypothèse nulle lorsque la valeur observée est trop élevée, c'est-à-dire lorsque :

$$\chi_{obs}^2 > \chi_{1-\alpha}^2$$

Ce test est toujours unilatéral.

Il est à noter que la distribution est approximative. Cette approximation est considérée comme satisfaisante lorsque les effectifs attendus sont supérieurs ou égaux à 5. Quand cette condition n'est pas remplie, on peut regrouper des classes voisines, de manière à augmenter les fréquences attendues.

Lorsque la distribution théorique n'est pas complètement définie, le ou les paramètres qui caractérisent la distribution théorique doivent être estimés.

Dans le cas de la loi normale, en général, la moyenne et la variance doivent être estimées. Dans le cas de la loi binomiale ou de la loi de Poisson, seule la moyenne est en général estimée.

Si les paramètres estimés correspondent aux paramètres qui minimisent la distance entre la distribution empirique et la distribution de référence, la distance, calculée à l'aide de l'équation (8.1), suit une loi du χ^2 avec $p - k - 1$ degrés de liberté, où k est le nombre de paramètres estimés.

Il est important de noter que les estimateurs de la moyenne et de la variance obtenue en minimisant la formule 8.1 ne sont en général pas les estimateurs classiques du maximum

de vraisemblance. Si les formules classiques sont utilisées, le niveau du test peut être sensiblement altéré.

1.2 Test de Kolmogorov-Smirnov

Le test de Kolmogorov-Smirnov est basé sur la comparaison de la fonction cumulative de fréquence $N(x)$ pour l'échantillon (c'est-à-dire le nombre d'observations inférieures à x) avec la fonction de répartition $F(x)$ pour la population (c'est-à-dire la probabilité qu'une observation soit inférieure à x). De façon plus précise on détermine l'écart maximum en valeur absolue entre ces deux fonctions :

$$|N(x) - F(x)|$$

et on compare cet écart à des valeurs critiques tabulées.

Pour des échantillons d'effectif supérieur à 35 et pour des niveaux de signification respectivement égaux à 0.05 et 0.01, on peut également utiliser les valeurs approchées suivantes :

$$\frac{1.36}{\sqrt{n}} \quad \text{et} \quad \frac{1.63}{\sqrt{n}}$$

On rejettera l'hypothèse nulle, respectivement aux niveaux 5% et 1% lorsque l'écart maximum observé est supérieur ou égal à l'une ou l'autre de ces valeurs.

Les conditions d'application du test de Kolmogorov-Smirnov sont nettement différentes de celles qui caractérisent le test du χ^2 de Pearson. Il est d'une utilisation plus large que le test du χ^2 , car il reste applicable pour des échantillons d'effectif réduit.

Il faut enfin noter que si les paramètres de $F(x)$ sont estimés à partir des données les valeurs critiques de la statistique ne sont pas connues.

1.3 Test de Lin et Mudholkar

Ce test repose sur une idée fort simple : un échantillon (x_1, \dots, x_n) est extrait d'une variable normale X si et seulement si la moyenne et la variance empirique sont indépendantes. Ce résultat est une caractérisation classique de la loi normale.

Il suffit donc de tester la nullité du coefficient de corrélation entre la moyenne et la variance. Ceci implique l'obtention d'un échantillon des variables aléatoires \bar{x} et s^2 . Pour ce faire, on utilise la méthode dite du jackknife.

Appelons E_i l'échantillon obtenu par suppression de l'observation x_i et notons \bar{x}_i sa moyenne et s_i^2 sa variance empirique. Notons \bar{x} la moyenne des \bar{x}_i et s^2 la moyenne des s_i^2 . Pour des raisons théoriques on calcule le coefficient de corrélation entre \bar{x} et s^2 :

$$r = \frac{\sum_i (\bar{x}_i - \bar{x})(s_i^2 - s^2)}{\sqrt{\sum_i (\bar{x}_i - \bar{x})^2} \sqrt{\sum_i (s_i^2 - s^2)^2}}$$

Sous l'hypothèse de normalité, le coefficient de corrélation suit asymptotiquement une loi normale $\mathcal{N}(0, \frac{3}{n})$ (et donc $\sqrt{\frac{n}{3}}r \sim \mathcal{N}(0, 1)$).

Pour des petites valeurs de la taille n de l'échantillon, on utilise le fait, établi par Fisher, que la statistique $Z = \log \frac{1+r}{1-r}$ est pratiquement gaussienne d'espérance nulle et de variance $\sigma^2 = \frac{3}{n} - \frac{7.324}{n^2} + \frac{53.005}{n^3}$.

Ce test de normalité se révèle particulièrement performant si la distribution réelle est asymétrique.

Exemple 8.1 Soit un échantillon de taille $n = 10$:

54, 47, 44, 74, 52, 83, 41, 77, 48, 49

On calcule $r = 0.786$. La statistique Z vaut 2.1217 et $\sigma^2 = 0.280$. On a donc $\frac{Z}{\sqrt{\sigma^2}} = 4.01 > 1.96$ de la loi normale au seuil 5%. On est donc dans la région critique et la loi de X n'est pas gaussienne.

Il faut remarquer que si l'on ne fait pas la transformation Z de Fisher, $\sqrt{\frac{n}{3}}r$ prend la valeur 1.435 et on est alors conduit, à tort, à accepter H_0 .

On peut enfin, et quelle que soit la taille n de l'échantillon, utiliser le test fondé sur la statistique Z qui converge en loi vers $\mathcal{N}(0, \frac{3}{n})$.

1.4 Test de Shapiro-Wilks

L'idée consiste à comparer deux statistiques T_1 et T_2 qui, sous l'hypothèse de normalité estiment toutes deux la variance, et sous l'hypothèse alternative estiment des quantités différentes :

$$T_1 = (n-1)S^2 \quad \text{et} \quad T_2 = \left(\sum_{i=1}^{[n/2]} a_n(i)(x_{(n-i+1)} - x_{(i)}) \right)^2$$

où $a_n(i)$ correspond aux scores normaux, $x_{(i)}$ correspond aux observations ordonnées de façon croissante et $[n/2]$ correspond à la partie entière de $n/2$.

La statistique de test est :

$$SW = \frac{T_2}{T_1}$$

Les valeurs de la statistique de tests sont lues dans les tables.

Exemple 8.2 Soit un échantillon de sept observations :

$x_1 = 6, x_2 = 1, x_3 = -4, x_4 = 8, x_5 = -2, x_6 = 5, x_7 = 0$

L'échantillon ordonné est :

$x_{(1)} = -4, x_{(2)} = -2, x_{(3)} = 0, x_{(4)} = 1, x_{(5)} = 5, x_{(6)} = 6, x_{(7)} = 8$

On a : $(n-1)S^2 = 118$

A partir de la table des scores normaux, on calcule :

$$\sqrt{T_2} = 0.6233(8+4) + 0.3031(6+2) + 0.1401(5-0)$$

D'où $SW = 0.97530$

Pour $n = 7$, avec un risque de première espèce de 5%, on lit dans la table la valeur critique du test : 0.979 > 0.97530. Il n'y a donc aucune raison de rejeter l'hypothèse de normalité.

2 Additivité

Les effets de deux facteurs sont dits additifs si l'effet de l'un de ces facteurs reste constant pour tous les niveaux de l'autre facteur.

Par exemple, si nous considérons un facteur traitement et un facteur bloc, cela signifie que l'effet du traitement est le même dans tous les blocs, et, que l'effet bloc est identique pour tous les traitements.

Tableau 8.1 – Exemple théorique 1

Traitement	Bloc		Effet bloc (I-II)
	I	II	
A	180	120	60
B	160	100	60
Effet traitement (A-B)	20	20	

Ceci correspond au modèle classique :

$$\text{observation} = \text{effet moyen} + \text{effet traitement} + \text{effet bloc} + \text{erreur}$$

Le modèle multiplicatif peut être noté :

$$\text{observation} = \text{effet moyen} + \text{effet traitement} \times \text{effet bloc} + \text{erreur}$$

Deux facteurs ont des effets multiplicatifs si leurs effets, exprimés en termes de pourcentage, sont additifs.

Tableau 8.2 – Exemple théorique 2

Traitement	Bloc		Effet bloc	
	I	II	(I-II)	$100 \times \frac{I-II}{II}$
A	180	120	60	50
B	150	100	50	50
Traitement (A-B)	30	20		
$100 \times \frac{A-B}{B}$	20	20		

sonnette d'alarme : Interaction traitement \times bloc

3 Non-indépendance des erreurs

L'hypothèse d'indépendance des erreurs signifie qu'aucune erreur expérimentale n'est dépendante d'une autre. Ceci est généralement respecté si la randomisation a été bien effectuée. Il est donc important de vérifier que deux traitements ne sont pas toujours côte à côte dans toutes les répétitions, et de manière générale que l'on n'observe pas de schéma général d'une répétition à une autre. Une répartition systématique des traitements est souvent à l'origine de la plupart des erreurs dépendantes. Un exemple est présenté dans le tableau 8.3.

Tableau 8.3 – Répartition systématique de cinq traitements (A,B,C,D,E) parmi quatre blocs

A	B	C	D	E
B	C	D	E	A
C	D	E	A	B
D	E	A	B	C

sonnette d'alarme : cartographie des résidus.

Les résidus sont découpés en classes. Une couleur ou un symbole est associé à chacune des classes et les résidus sont reportés sur une carte en fonction de leur localisation spatiale.

4 Variance hétérogène

L'hétérogénéité des variances est la violation des hypothèses de l'analyse de variance la plus fréquente. Nous allons voir des méthodes qui permettent parfois d'y remédier. Néanmoins, comme dans toute médecine un bon diagnostic est indispensable pour prescrire un bon remède.

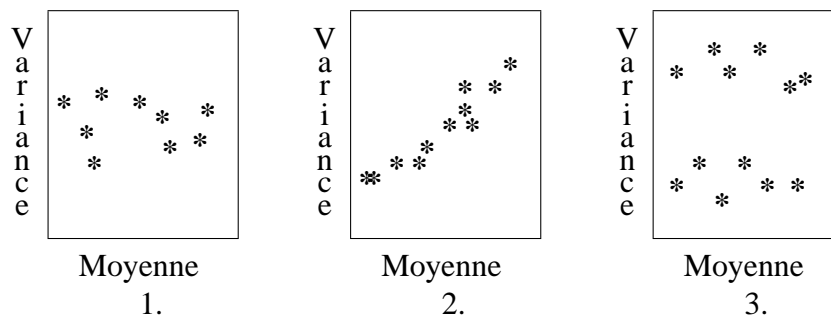
La démarche suivante est donc conseillée :

1. Pour chaque traitement calculer la moyenne, la variance et l'écart type.
2. Faire deux graphes, le premier avec un axe pour la moyenne et un pour la variance, le deuxième avec un axe pour la moyenne et l'autre pour l'écart-type. Il y a autant de points que de traitements.
3. Observer les graphes obtenus pour essayer de voir si les variances sont hétérogènes.

La figure 8.1 illustre trois cas de figure classique :

1. Variance homogène : cas souhaité ;
2. variance ou écart-type proportionnel à la moyenne : voir paragraphe 9 ;
3. variance hétérogène : voir paragraphe 4

Figure 8.1 – Exemple de graphes moyenne versus variance



Chapitre 9

Transformation des données

Le fait que la variance ou l'écart-type soit proportionnel à la moyenne est souvent symptomatique du fait que la distribution de l'erreur expérimentale ne suit pas une loi normale. Par exemple, les données de proportion conduisent souvent à des distributions de type binomial ou Poisson.

Il est nécessaire de vérifier que la transformation a homogénéisé les variances. Si cela est le cas, l'analyse de variance s'effectue sur les données transformées.

Si des comparaisons de moyennes de traitement sont effectuées avec les tests classiques (p.p.d.s., Duncan,...), il est conseillé de travailler uniquement sur les données transformées puis d'effectuer les transformations inverses pour revenir à l'échelle d'origine.

Nous allons maintenant présenter trois transformations souvent utilisées.

1 Transformation logarithmique

La transformation logarithmique est la plus appropriée quand l'écart-type est proportionnel à la moyenne ou que les effets sont multiplicatifs. Cette situation est assez fréquente lorsque des processus de croissance ou de multiplication sont en jeu.

Pour effectuer une transformation logarithmique sur un ensemble de valeurs, il suffit de prendre le logarithme de chaque valeur.

Si le jeu de données comporte des petites valeurs (en pratique inférieures à 1) et surtout des valeurs nulles, il est préférable d'utiliser $\log(x + 1)$ à la place de $\log(x)$.

2 Transformation racine carrée

La transformation racine carrée est la plus appropriée quand la variance est proportionnelle à la moyenne. Cela peut se produire quand, par exemple, l'on compte des événements très rares (ou au contraire très fréquents) en termes de pourcentage. Si les pourcentages observés sont répartis entre 0% et 100%, il peut être préférable d'utiliser la transformation arc sinus.

Tableau 9.1 – Résultat de la transformation logarithmique sur les données de l'exemple théorique 8.2

Traitement	Bloc		Effet bloc (I-II)
	I	II	
A	2.255	2.079	0.176
B	2.176	2.000	0.176
Effet traitement (A-B)	0.079	0.079	

Si les valeurs des données sont petites (en pratique inférieure à 10), et surtout lorsque l'on obtient des valeurs nulles, il est préférable d'utiliser $\sqrt{x + \frac{1}{2}}$ à la place de \sqrt{x} , où x représente la donnée d'origine.

3 Transformations Arc sinus

La transformation Arc sinus est souvent utilisée lorsque les données représentent des proportions ou des pourcentages. Il est cependant important de préciser que toutes les données de pourcentage (ou de proportions) ne nécessitent pas une transformation et que la transformation Arcsinus n'est pas unique. Avant de l'utiliser il est utile de vérifier que la variance (σ^2) et la moyenne (μ) sont reliées par une relation de la forme : $\sigma^2 = \mu(1-\mu)$. La valeur de 0% doit être remplacée par $\frac{1}{4n}$ et la valeur 100% par $100 - \frac{1}{4n}$, où n représente le dénominateur utilisé pour calculer le pourcentage, c'est-à-dire, la taille de la population de référence.

Remarque : très souvent, on ne dispose d'aucune raison théorique permettant de justifier a priori l'utilisation de telle ou telle transformation. Il existe des méthodes empiriques pour rechercher la transformation optimum. Nous renvoyons le lecteur intéressé à la bibliographie.

Il faut cependant noter que les transformations de variables ont l'inconvénient de compliquer sensiblement l'interprétation des résultats.

4 Partitionnement de l'erreur

Il arrive que les variances soient hétérogènes sans qu'aucune relation n'apparaisse entre la variance et la moyenne. Par exemple, une application non uniforme d'un fongicide peut entraîner des différences sensibles de variance d'un traitement à l'autre.

Le plus souvent l'hétérogénéité des variances provient de quelques traitements dont les erreurs associées sont différentes des autres. La procédure suivante est donc conseillée :

1. Identifier les traitements dont les valeurs sont très différentes d'une répétition à l'autre. Pour chacun de ces traitements rechercher les observations dont les valeurs sont très différentes des autres ;
2. Vérifier qu'il n'y a pas eu une erreur de frappe au moment de la saisie des données ou d'absurdité évidente. Dans ce cas il faut soit corriger soit éliminer la donnée suspecte ;
3. Examiner la disposition sur le terrain des valeurs suspectes afin de voir si il n'y a pas de relation entre la disposition géographique et la différence de valeurs observée (facteur bloc mal défini par exemple)
4. Répartir les traitements en groupes (pas trop petit) de variance homogène et effectuer une analyse de variance sur chacun des groupes. La construction des groupes dépend du problème. On peut par exemple avoir des groupes d'espèces ou bien un groupe pour les lignées pures et un groupe pour les hybrides, etc. Pour comparer deux traitements d'un même groupe il n'y a aucun problème. Pour comparer deux traitements de groupes différents il suffit de revenir aux formules et de remplacer la variance résiduelle (σ_r^2) par la moyenne des variances résiduelles des deux blocs concernés ($\frac{1}{2}(\sigma_i^2 + \sigma_j^2)$). Si l'erreur est estimée avec le même nombre de d.d.l. dans les deux groupes, la valeur tabulée du test (notons la t) se lit directement dans la table, sinon on remplace t par :

$$t = \frac{\sigma_i^2 t_i + \sigma_j^2 t_j}{\sigma_i^2 + \sigma_j^2}$$

où t_i et t_j sont les valeurs tabulées pour le nombre de d.d.l. des groupes i et j respectivement.

Exemple 9.1 *Un essai en blocs complets randomisés avec trois blocs a été installé pour tester les rendements de différentes variétés de caféier. Il est apparu que les variances des 11 espèces hybrides étaient très différentes des 17 espèces parents. Les analyses de variances séparées des hybrides et des parents ne font pas apparaître de violations des hypothèses de base et donnent les variances résiduelles suivantes :*

Parents : $\sigma_1^2 = 0.2124$ avec $(17-1)*(3-1)=32$ ddl $(t_{32}(5\%)= 1.696)$

Hybrides : $\sigma_2^2 = 0.4331$ avec 20 ddl $(t_{20}(5\%)= 1.725)$

La plus petite différence significative au seuil 5% vaut :

entre deux parents : $1.696 \times \sqrt{\frac{2}{3} \times 0.2124} = 0.638$

entre deux hybrides : $1.725 \times \sqrt{\frac{2}{3} \times 0.4331} = 0.927$

entre un hybride et un parent :

$$\frac{0.4331 \times 1.725 + 0.2124 \times 1.696}{0.4331 + 0.2124} \times \sqrt{\frac{2}{3} \times \frac{0.4331 + 0.2124}{2}} = 1.716 \times 0.467 = 0.802$$

Ces quelques méthodes sont souvent utiles pour se rapprocher des conditions de l'analyse de variance mais ne sauraient en aucun cas constituer des remèdes miracles toujours

applicables. Si les hypothèses de base ne sont manifestement pas respectées et que les quelques règles ci-dessus ne s'appliquent pas il peut se révéler plus judicieux d'essayer des méthodes non paramétriques au lieu de s'obstiner malgré tout à vouloir faire une analyse de variance classique.

Chapitre 10

Analyse de variance non paramétrique

Nous n'aborderons dans ce chapitre que le cas à un ou deux facteurs, avec au moins trois traitements.

Même si les conditions d'application des tests de l'analyse de variance non paramétrique sont moins restrictives que celle de l'analyse de variance usuelle, il faut cependant noter que les échantillons doivent être aléatoires et indépendants entre eux. De plus, les variables aléatoires considérées sont supposées continues.

L'idée sous-jacente des deux tests de ce paragraphe est d'effectuer une analyse de variance basée sur les rangs, c'est à dire sur les numéros d'ordre des valeurs observées rangées par ordre croissant. Ceci simplifie en général beaucoup les calculs mais provoque évidemment une perte d'information. De ce fait, les analyses de variance non paramétriques sont généralement moins puissantes que les analyses de variance paramétriques correspondantes, dans les conditions où ces dernières sont applicables.

1 Le test des rangs de KRUSKAL-WALLIS : cas d'un facteur

Ce test permet de décider si les différences observées entre t échantillons indépendants sont dues au hasard, ou au fait que les échantillons proviennent de différentes populations. Ce qu'on peut résumer en disant que les traitements sont différents. Ce test s'applique par exemple pour des expériences complètement randomisées.

On commence par remplacer les N observations par leur rang : la plus petite valeur est remplacée par 1, la suivante par 2, etc. La plus grande valeur est remplacée par N . (N = nombre total d'observations dans les t échantillons).

L'attribution sans ambiguïté d'un rang à chaque observation nécessite qu'il n'y ait pas d'observations égales. Ce sera théoriquement le cas si les observations sont extraites de distributions continues. Mais dans la pratique l'instrument de mesure peut être incapable de séparer deux observations trop proches.

Les ex-aequo qui se présentent simultanément dans plusieurs échantillons sont affectés d'un rang égal à la moyenne des rangs revenant normalement à ces différentes valeurs. Par exemple, si les sixième et septième valeurs par ordre croissant sont toutes deux égales, les rangs sont conventionnellement : $(6+7)/2 = 6.5$

Puis la somme des rangs de chaque échantillon (ou traitement) est calculée.

Le test de Kruskal-Wallis détermine si les sommes des rangs sont (ou ne sont pas) significativement différentes, c'est-à-dire si les échantillons sont tellement différents qu'il y a peu de chance qu'ils proviennent de la même distribution (et donc que la probabilité que les traitements soient identiques est faible).

A partir de ces sommes, on peut ensuite obtenir la valeur :

$$H = \frac{12}{N(N+1)} \sum_{j=1}^t \frac{R_j^2}{n_j} - 3(N+1)$$

où

- t = nombre d'échantillons ;
- n_j = nombre d'observations dans le $j^{\text{ème}}$ échantillon ;
- $N = \sum_j n_j$, nombre total d'observations ;
- R_j = somme des rangs du $j^{\text{ème}}$ échantillon ;
- $\sum_{j=1}^t$ indique la somme sur les t échantillons.

La présence d'ex-aequo nécessite l'utilisation d'un facteur correctif, la quantité H doit être divisée par :

$$1 - \frac{\sum(k(k^2 - 1))}{N(N^2 - 1)}$$

le signe de sommation étant relatif aux différentes séries d'ex-aequo et k désigne le nombre de termes de ces différentes séries. Par exemple, si parmi 37 observations il y a sept séries de deux ex-aequo et une série de trois ex-aequo, le facteur correctif vaut :

$$1 - \frac{7 \times 2(2^2 - 1) + 1 \times 3(3^2 - 1)}{37(37^2 - 1)} = 0.9987$$

Remarque : le facteur correctif est assez proche de 1. Dans cet exemple, il y a 28 valeurs distinctes.

Pour de petits effectifs, les valeurs tabulées du test de Kruskal-Wallis sont données à la fin du texte.

Lorsque les effectifs sont plus importants, H est distribué comme une variable chi-deux avec $t - 1$ ddl. La règle de décision est donc la suivante : si H est égal ou plus grand que la valeur d'une variable chi-deux avec $t - 1$ ddl au seuil fixé (5% par exemple), on rejettera l'hypothèse d'égalité des traitements, au seuil fixé.

Le test est global, c'est à dire qu'une différence significative signifie qu'au moins un traitement est différent. Pour détecter une différence entre deux traitements, un test de comparaison multiple doit être utilisé.

Par analogie avec la ppds, il est possible de conclure que deux traitements (i et j) seront significativement différents au seuil α si la quantité :

$$Q_{ij} = \frac{\left| \frac{R_i}{n_i} - \frac{R_j}{n_j} \right|}{\sqrt{\frac{N(N+1)}{12} \times \left(\frac{1}{n_j} + \frac{1}{n_i} \right)}}$$

est supérieure à z , valeur de la loi normale au seuil $\alpha_{t(t-1)}$. Dans l'exemple de la table 10.1, au seuil 5% : $\alpha_{t(t-1)} = 0.05/6 = 0.0083$ et $z = 2.394$.

La valeur $\alpha_{t(t-1)}$ provient du fait que $1 + 2 + \dots + t = t(t-1)/2$ et que le test est unilatéral.

Tableau 10.1 – Comparaison des hauteurs des arbres de trois parcelles de forêts : hauteurs observées et rangs

Hauteurs			Rangs		
Parc. 1	Parc. 2	Parc. 3	Parc. 1	Parc. 2	Parc. 3
23.4	22.5	18.9	8	5.5	1
24.4	22.9	21.1	12.5	7	2
24.6	23.7	21.2	16.5	10	3
24.9	24.0	22.1	18	11	4
25.0	24.4	22.5	19	12.5	5.5
26.2	24.5	23.6	23	14.5	9
26.3	25.3	24.5	25	20	14.5
26.8	26.0	24.6	29.5	21	16.5
26.8	26.2	26.2	29.5	23	23
26.9	26.4	26.7	31.5	26	27.5
27.0	26.7		33	27.5	
27.6	26.9		35	31.5	
27.7	27.4		36	34	
	28.5			37	
Totaux			316.5	280.5	106
Nombres d'observations			13	14	10

Exemple 10.1 Avec les données du tableau 10.1, $N = 13 + 14 + 10 = 37$

$$H = \frac{12}{37(37+1)} \times \left(\frac{316.5^2}{13} + \frac{280.5^2}{14} + \frac{106^2}{10} \right) - 3(37+1) = 9.32$$

Il y a sept séries de deux ex-aequo et une série de trois ex-aequo, le facteur correcteur a déjà été calculé. Le H corrigé vaut donc : $9.32/0.9987 = 9.33$
et par comparaison avec la table du chi-deux :

$$\chi_2^2(0.95) = 5.99 \qquad \chi_2^2(0.99) = 9.21$$

l'hypothèse d'identité des trois populations-parents est rejetée même au niveau 1%
Un calcul rapide montre que seuls les traitements 1 et 3 sont significativement différents au seuil 1% ($Q_{13} = 3.019$).

2 Le test de FRIEDMAN : cas de deux facteurs

Ce test permet de décider si les différences observées entre t échantillons répartis en p blocs sont dues au hasard, ou au fait que les traitements sont différents. Ce test s'applique donc surtout pour les expériences en blocs complètement randomisés.

On commence par remplacer les observations par leur rang. L'opération s'effectue indépendamment pour chaque bloc. Le traitement des ex-aequo à l'intérieur d'un bloc reste le même que pour le test de Kruskal-Wallis.

Puis on calcule :

$$T = \frac{12}{tb(t+1)} \sum_{j=1}^t R_j^2 - 3b(t+1)$$

où

- t = nombre d'échantillons ;
- b = nombre de blocs ;
- R_j = somme des rangs du $j^{\text{ème}}$ échantillon ;
- $\sum_{j=1}^t$ indique la somme sur les t échantillons.

Pour un nombre suffisant de blocs ou de traitements, T suit une loi du chi-deux à $t - 1$ ddl. La règle de décision est donc la suivante : si T est plus grand que la valeur chi-deux avec $t - 1$ ddl au seuil fixé (5% par exemple), on rejettera l'hypothèse d'égalité des traitements, au seuil fixé.

Lorsque les données sont peu nombreuses il faut se reporter à la table donnée à la fin du texte.

Le test est global, c'est à dire qu'une différence significative signifie qu'au moins un traitement est différent. Pour détecter une différence entre deux traitements, un test de comparaison multiple doit être utilisé.

Par analogie avec la p.p.d.s., il est possible de conclure que deux traitements (i et j) seront significativement différents au seuil α si la quantité :

$$\frac{|R_i - R_j|}{\sqrt{\frac{bt(t+1)}{6}}}$$

est supérieure à z , valeur de la loi normale au seuil $\alpha/_{t(t-1)}$.

Dans l'exemple du tableau 10.2, au seuil 5% : $\alpha/_{t(t-1)} = 0.05/20 = 0.0025$ et $z = 2.807$.

On en déduit :

$$T = \frac{12}{5 \times 4 \times (5 + 1)} \times (12^2 + 15^2 + 17^2 + 12^2 + 4^2) - 3 \times 4 \times (5 + 1) = 9.8$$

Tableau 10.2 – Comparaison de 5 traitements fongicides : rendements observés et rangs

Traitements	Bloc 1	Bloc 2	Bloc 3	Bloc 4	Totaux
1	527	604	606	533	
2	633	600	650	567	
3	642	708	662	504	
4	623	550	562	667	
5	377	408	500	333	
1	2	4	3	3	12
2	4	3	4	4	15
3	5	5	5	2	17
4	3	2	2	5	12
5	1	1	1	1	4

ce qui conduit à rejeter l'hypothèse d'égalité des traitements au seuil 5%. En effet, pour $b = 4$ et $n = 5$, la valeur critique lue dans la table au seuil 5% est de 7.800.

Un calcul rapide montre que seul les traitements 2 et 3 sont significativement différents du traitement 5 au seuil 5%.

Chapitre 11

Corrélation de rangs

1 Introduction

L'un des objectifs classique de la recherche forestière est d'étudier les liaisons entre différentes variables quantitatives mesurées sur des échantillons représentatifs d'une population donnée. Le but peut être de prédire une variable par une autre, de comprendre le mécanisme d'un phénomène, ou simplement de vérifier la pertinence de différents systèmes de mesure. L'outil le plus souvent utilisé pour mesurer la liaison entre deux variables quantitatives est le coefficient de corrélation linéaire, aussi appelé coefficient de corrélation de (Bravais)-Pearson.

Soit deux variables quantitatives X_1 et X_2 mesurées sur n individus, X_1 prenant les valeurs $(x_{11}, x_{12}, \dots, x_{1i}, \dots, x_{1n})$ et X_2 prenant les valeurs $(x_{21}, x_{22}, \dots, x_{2i}, \dots, x_{2n})$. Le coefficient de corrélation linéaire s'écrit :

$$r(X_1, X_2) = \frac{\sum_{i=1}^n (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)}{\sqrt{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2} \sqrt{\sum_{i=1}^n (x_{2i} - \bar{x}_2)^2}} = \frac{\text{Cov}(X_1, X_2)}{S_1 S_2} \quad (11.1)$$

Il est facile de voir que $-1 \leq r(X_1, X_2) \leq 1$.

Ce coefficient mesure la liaison affine entre les deux variables (X_1, X_2) mais les inférences statistiques sur les coefficients de corrélation de Pearson ne sont valables que sous l'hypothèse de normalité des variables X_1 et X_2 . En pratique pour tester l'hypothèse $r(X_1, X_2) = 0$, il suffit que l'une des deux variables suive une distribution normale. Si cette hypothèse n'est pas vérifiée, il paraît pertinent de tester l'indépendance entre (X_1, X_2) avec un coefficient de rangs, à condition que ces variables soient continues.

2 Coefficient de corrélation de rangs de Spearman

C'est la première statistique sur les rangs qui a été proposée (1904).

Dans un premier temps, les variables étudiées sont transformées en rangs :

$$(X_1, X_2) \longrightarrow (Y_1, Y_2)$$

avec $Y_1 = (1, 2, \dots, n)$ et $Y_2 = (1, 2, \dots, n)$ $y_{1i} = j$ si l'individu i est classé le $j^{\text{ème}}$ pour la variable X_1 et $y_{2i} = j$ si l'individu i est classé le $j^{\text{ème}}$ pour la variable X_2 ($j = 1, \dots, n$).

Par analogie avec l'équation (11.1), le coefficient de corrélation de rangs de Spearman s'écrit :

$$r_s(X_1, X_2) = \frac{\sum_{i=1}^n (y_{1i} - \bar{y}_1)(y_{2i} - \bar{y}_2)}{\sqrt{\sum_{i=1}^n (y_{1i} - \bar{y}_1)^2} \sqrt{\sum_{i=1}^n (y_{2i} - \bar{y}_2)^2}} \quad (11.2)$$

Il est facile de voir que $-1 \leq r_s(X_1, X_2) \leq 1$.

Ce coefficient de corrélation peut également s'écrire en fonction des différences de rangs : $d_i = y_{1i} - y_{2i}$. Après quelques calculs, on obtient l'expression classique :

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}$$

L'intérêt de cette écriture réside dans l'expression de r_s en fonction de d_i^2 , qui peut être assimilé à une distance.

Les valeurs critiques de r_s sont tabulées jusqu'à $n = 30$, mais quand $n \geq 10$, l'hypothèse $r_s = 0$ peut être testée par le test de Student avec :

$$t_{n-2} = r_s \sqrt{\frac{n-2}{1-r_s^2}}$$

S'il existe des ex-aequo, on a $\sum_{i=1}^n (y_{1i} - \bar{y}_1) < \frac{n^2-n}{12}$ et il faut faire intervenir un terme de correction dans l'équation (11.2) et l'on a :

$$\sum_{i=1}^n (y_{ki} - \bar{y}_k) = \frac{n^2 - n}{12} - \sum_{i=1}^n \frac{t_{ki}^3 - t_{ki}}{12} \quad k = 1, 2$$

t_{ki} est le nombre d'ex-aequo au rang i de la variable X_k .

Avec le logiciel SAS, l'option *spearman* dans la procédure *corr* permet directement de calculer le coefficient de corrélation de Spearman.

Enfin, il est à noter que l'efficacité de r_s est de 91% par rapport au r de Pearson.

Exemple 11.1

Ind.	y_1	y_2	d_i	d_i^2
1	2	3	-1	1
2	6	4	2	4
3	5	2	3	9
4	1	1	0	0
5	10	8	2	4
6	9	11	-2	4
7	8	10	-2	4
8	3	6	-3	9
9	4	7	-3	9
10	12	12	0	0
11	7	5	2	4
12	11	9	2	4

$$\sum d_i^2 = 52$$

$$r_s = 1 - \frac{6 \times 52}{12^3 - 12} = 0.82$$

- Valeur seuil de r_s à 5% : 0.506
- valeur seuil de r_s à 1% : 0.712

$$t = 0.82 \sqrt{\frac{12-2}{1-0.82^2}} = 4.53 ; \text{ significatif à 1\% (pour 10 ddl)}$$

3 Coefficient de corrélation de rangs de Kendall

Ce coefficient est une autre mesure de l'association entre deux variables transformées en rangs.

Soient deux variables quantitatives transformées en rangs : $(X_1, X_2) \longrightarrow (Y_1, Y_2)$.

On classe les individus de Y_1 dans l'ordre croissant des rangs : $Y_1 = (y_{(1)}, y_{(2)}, \dots, y_{(n)})$. Pour chaque couple (y_{2i}, y_{2j}) de la variable Y_2 tel que $i < j$, on associe la valeur +1 si $y_{2i} < y_{2j}$ et la valeur -1 si $y_{2i} > y_{2j}$. La somme des C_n^2 valeurs est notée S . Le coefficient de corrélation de rangs de Kendall est défini par :

$$\tau = \frac{S}{C_n^2} = \frac{S}{\frac{1}{2}n(n-1)} \quad (11.3)$$

avec des ex aequo, l'expression de τ devient :

$$\tau = \frac{S}{\sqrt{\frac{1}{2}n(n-1) - T_1} \sqrt{\frac{1}{2}n(n-1) - T_2}}$$

avec $T_k = \frac{1}{2} \sum_i t_i(t_i - 1)$ où t_i est le nombre d'ex aequo au rang i pour la variable X_k ($k = 1, 2$).

Les valeurs critiques de τ sont tabulées jusqu'à $n = 10$. Pour des effectifs plus élevés, on a :

$$z = \frac{\tau}{\sqrt{\frac{2(2n+5)}{9n(n-1)}}} \sim \mathcal{N}(0, 1)$$

L'un des avantages du coefficient de corrélation de Kendall est qu'il permet de calculer des coefficients de corrélation partielle basée sur les rangs. On peut, par exemple, calculer le coefficient de corrélation entre X_1 et X_2 en supprimant l'effet d'une variable X_3 . Dans ce cas, le coefficient de corrélation partielle de Kendall s'écrit :

$$\tau_{X_1 X_2 \cdot X_3} = \frac{\tau_{X_1 X_2} - \tau_{X_1 X_3} \tau_{X_2 X_3}}{\sqrt{1 - \tau_{X_1 X_3}^2} \sqrt{1 - \tau_{X_2 X_3}^2}} \sim \sqrt{\frac{\chi_n^2}{n}}$$

Cette quantité est parfois appelée coefficient ϕ .

Pour deux variables (X_1, X_2) données, les valeurs des coefficients de corrélation de Kendall et Spearman sont différentes mais les seuils de significativité des tests sont bien évidemment identiques.

Exemple 11.2 Classement de quatre produits par deux juges

Produits	a	b	c	d
Juge 1	3	4	2	1
Juge 2	3	1	4	2

On classe les produits selon l'ordre défini par le juge 1

Produits	d	c	a	b
Juge 1	1	2	3	4
Juge 2	2	4	3	1

Les paires bien classées et mal classées par le juge 2, par rapport au classement du juge 1 sont comptabilisées :

- $2 < 4 \implies +1$
- $2 < 3 \implies +1$
- $2 > 1 \implies -1$
- $4 > 3 \implies -1$
- $4 > 1 \implies -1$
- $3 > 1 \implies -1$

d'où $S = -2$ et $\tau = \frac{S}{\frac{1}{2}n(n-1)} = \frac{-2}{6} = -0.333$

4 Coefficient de concordance de Kendall

Le coefficient de concordance de Kendall permet de mesurer l'association entre $k > 2$ variables et s'écrit :

$$W = \frac{\sum_{i=1}^n (R_i - \bar{R})^2}{\frac{1}{12}k^2(n^3 - n)} \quad (11.4)$$

où k est le nombre de variables, n est le nombre d'individus et R_i est la somme des rangs des k variables décrivant l'individu i .

Avec des ex aequo, cette expression devient :

$$W = \frac{\sum_{i=1}^n (R_i - \bar{R})^2}{\frac{1}{12}k^2(n^3 - n) - k \sum_{j=1}^n T_j}$$

avec $T_j = \frac{\sum_i (t_{ij}^3 - t_{ij})}{12}$ où t_{ij} est le nombre d'ex aequo au rang i de la variable j .

Pour de petits effectifs ($n < 7$), il existe des tables pour tester l'hypothèse $W = 0$, c'est-à-dire s'il y a indépendance entre les différentes variables. Pour des effectifs plus élevés, on a :

$$k(n-1)W \sim \chi_{n-1}^2$$

Chapitre 12

Tests sur les résidus d'une régression

De nombreux ouvrages très pertinents sur la régression ayant été écrits, nous renvoyons le lecteur intéressé à la bibliographie (voir par exemple Tomassone, 1989 ou Draper et Smith, 1981).

Comme pour toutes les techniques statistiques, la validité des tests dépend de certaines hypothèses qu'il est important de vérifier. Les hypothèses de base de la régression sont les mêmes que les hypothèses de l'analyse de variance (voir page 47) et les techniques présentées dans le chapitre 8 s'appliquent. Le problème spécifique à la régression le plus fréquemment observé concerne l'indépendance des résidus lorsqu'une relation d'ordre existe entre les données. Ceci se produit, par exemple, lorsque les données sont mesurées le long de lignes ou lorsque des inventaires sont répétés tous les ans.

Dans la suite de la section, nous considérerons que la relation d'ordre est le temps.

Il est à noter que si ces tests permettent, par des calculs simples le rejet d'une hypothèse, ils ne permettent qu'imparfaitement d'imaginer les remèdes qui peuvent être apportés pour corriger une anomalie.

1 Test de Durbin-Watson

L'idée de base de ce test est de calculer un coefficient d'autocorrélation linéaire de rang 1, c'est-à-dire un coefficient de corrélation entre la valeur du résidu au temps t et la valeur du résidu au temps $t - 1$. La statistique de test s'écrit :

$$d = \frac{\sum_{i=2}^m (e_i - e_{i-1})^2}{\sum_{i=1}^m e_i^2} \quad (12.1)$$

où e_i représentent les résidus de la régression et m le nombre d'observations.

Il est possible de montrer que d est compris entre 0 et 4. En effet, l'équation 12.1 peut se réécrire :

$$d = 2 - 2 \left(\frac{\sum_{i=2}^m e_i e_{i-1}}{\sum_{i=1}^m e_i^2} \right)$$

Cette statistique de test n'a de sens que si l'index des valeurs de la régression représente le temps et si le pas de temps entre e_i et e_{i+1} est à peu près constant.

La table des valeurs est un peu plus compliquée à utiliser que d'habitude. On notera tout d'abord la présence de deux valeurs d_{inf} et d_{sup} pour un seuil de significativité, un nombre m d'observations et un nombre k de variables explicatives. La procédure de test s'utilise comme suit :

Soit H_0 : il n'y a pas de corrélation entre les résidus

- test de H_1 : il existe une corrélation positive entre les résidus
 - Si $d < d_{\text{inf}}$ alors on rejette l'hypothèse H_0 au niveau α
 - Si $d > d_{\text{sup}}$ alors on ne rejette pas l'hypothèse H_0 au niveau α
 - Si $d_{\text{inf}} \leq d \leq d_{\text{sup}}$, on ne peut pas conclure à partir du test.
- test de H_1 : il existe une corrélation négative entre les résidus
 - On n'utilise la même procédure que précédemment en remplaçant d par $(4 - d)$
- test de H_1 : il existe une corrélation entre les résidus
 - Si $d < d_{\text{inf}}$ ou $(4 - d) < d_{\text{inf}}$ on rejette H_0 au niveau 2α
 - Si $d > d_{\text{sup}}$ ou $(4 - d) > d_{\text{sup}}$ on ne rejette pas H_0 au niveau 2α
 - dans les autres cas on ne peut pas conclure.

Ne pas pouvoir conclure est certes un peu frustrant mais il est difficile d'obtenir une conclusion dans ce cas et nous ne développerons donc pas ce sujet dans ce paragraphe.

Exemple 12.1 *A partir des résidus de la régression de 50 hauteurs d'arbres à l'aide d'une variable explicative (par exemple le diamètre), on trouve une valeur de $d = 0.625$. Pour tester l'hypothèse H_0 : il n'y a pas de corrélation entre les résidus contre l'hypothèse H_1 : il existe une corrélation entre les résidus, on lit dans la table : $d_{\text{inf}} = 1.50$ et $d_{\text{sup}} = 1.59$ ($k = 1$ et $m = 50$)
Comme $0.625 < 1.50$, l'hypothèse H_0 est rejetée au seuil $2\alpha = 10\%$.*

2 Test des suites

Ce test est utilisé pour vérifier l'indépendance des observations. Supposons qu'en lançant 20 fois une pièce de monnaie on observe 10 fois de suite le côté pile puis 10 fois de suite le côté face. Ces résultats sont étranges car la probabilité de réalisation de 2 suites est certainement faible si la pièce n'est pas truquée.

De même si l'on avait observé alternativement des piles et des faces, on pourrait avoir des doutes sur la qualité de la pièce car le résultat serait trop systématique.

Supposons que les résidus soient ordonnés dans un ordre chronologique et que la séquence des signes soit la suivante :

+ + - + - - - - + + - + + +

On dira qu'il y a 7 suites :

(++)(-)(+)(- - -)(++)(-)(+ + +)

La statistique de test dépend donc du nombre de piles (n_1), du nombre de faces (n_2) et du nombre de suites (r).

Dans le cas de grands échantillons, on peut tester, avec une table de loi normale réduite, la quantité :

$$Z = \frac{r - m_r}{s_r}$$

avec

$$m_r = \frac{2n_1n_2}{n_1 + n_2} + 1$$
$$s_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

L'application de ce test à l'examen des résidus est aisée. Notons toutefois qu'il ne tient compte que de leur signe et non pas de leur grandeur.

Pour un nombre de + et de - donné, l'hypothèse de répartition aléatoire des résidus est rejetée si le nombre de suites ne se situe pas entre les bornes données dans la table.

Chapitre 13

Tables statistiques

Tableau A – Valeur de la loi normale. Probabilité d'avoir une valeur supérieure.

| z | Prob. | z | Prob. | z | Prob. |
|------|-------|------|-------|------|---------|
| 0.00 | 0.00 | 1.50 | 86.64 | 3.00 | 99.730 |
| 0.05 | 3.99 | 1.55 | 87.89 | 3.05 | 99.771 |
| 0.10 | 7.97 | 1.60 | 89.04 | 3.10 | 99.806 |
| 0.15 | 11.92 | 1.65 | 90.11 | 3.15 | 99.837 |
| 0.20 | 15.85 | 1.70 | 91.09 | 3.20 | 99.863 |
| 0.25 | 19.74 | 1.75 | 91.99 | 3.25 | 99.885 |
| 0.30 | 23.58 | 1.80 | 92.81 | 3.30 | 99.903 |
| 0.35 | 27.37 | 1.85 | 93.57 | 3.35 | 99.919 |
| 0.40 | 31.08 | 1.90 | 94.26 | 3.40 | 99.933 |
| 0.45 | 34.73 | 1.95 | 94.88 | 3.45 | 99.944 |
| 0.50 | 38.29 | 2.00 | 95.45 | 3.50 | 99.953 |
| 0.55 | 41.77 | 2.05 | 95.96 | 3.55 | 99.961 |
| 0.60 | 45.15 | 2.10 | 96.43 | 3.60 | 99.968 |
| 0.65 | 48.43 | 2.15 | 96.84 | 3.65 | 99.974 |
| 0.70 | 51.61 | 2.20 | 97.22 | 3.70 | 99.978 |
| 0.75 | 54.67 | 2.25 | 97.56 | 3.75 | 99.982 |
| 0.80 | 57.63 | 2.30 | 97.86 | 3.80 | 99.986 |
| 0.85 | 60.47 | 2.35 | 98.12 | 3.85 | 99.988 |
| 0.90 | 63.19 | 2.40 | 98.36 | 3.90 | 99.990 |
| 0.95 | 65.79 | 2.45 | 98.57 | 3.95 | 99.992 |
| 1.00 | 68.27 | 2.50 | 98.76 | 4.00 | 99.9937 |
| 1.05 | 70.63 | 2.55 | 98.92 | 4.05 | 99.9949 |
| 1.10 | 72.87 | 2.60 | 99.07 | 4.10 | 99.9959 |
| 1.15 | 74.99 | 2.65 | 99.20 | 4.15 | 99.9967 |
| 1.20 | 76.99 | 2.70 | 99.31 | 4.20 | 99.9973 |
| 1.25 | 78.87 | 2.75 | 99.40 | 4.25 | 99.9979 |
| 1.30 | 80.64 | 2.80 | 99.49 | 4.30 | 99.9983 |
| 1.35 | 82.30 | 2.85 | 99.56 | 4.35 | 99.9986 |
| 1.40 | 83.85 | 2.90 | 99.63 | 4.40 | 99.9989 |
| 1.45 | 85.29 | 2.95 | 99.68 | 4.45 | 99.9991 |

Tableau B – Valeurs critiques de la distribution du chi-deux

| DDL | 99% | 95% | 90% | 70% | 50% | 30% | 10% | 5% | 1% |
|-----|--------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 0.0002 | 0.003 | 0.02 | 0.15 | 0.46 | 1.07 | 2.71 | 3.84 | 6.64 |
| 2 | 0.02 | 0.10 | 0.21 | 0.71 | 1.39 | 2.41 | 4.60 | 5.99 | 9.21 |
| 3 | 0.12 | 0.35 | 0.58 | 1.42 | 2.37 | 3.67 | 6.25 | 7.82 | 11.34 |
| 4 | 0.30 | 0.71 | 1.06 | 2.20 | 3.36 | 4.88 | 7.78 | 9.49 | 13.28 |
| 5 | 0.55 | 1.14 | 1.61 | 3.00 | 4.35 | 6.06 | 9.24 | 11.07 | 15.09 |
| 6 | 0.87 | 1.64 | 2.20 | 3.83 | 5.35 | 7.23 | 10.65 | 12.59 | 16.81 |
| 7 | 1.24 | 2.17 | 2.83 | 4.67 | 6.35 | 8.38 | 12.02 | 14.07 | 18.48 |
| 8 | 1.65 | 2.73 | 3.49 | 5.53 | 7.34 | 9.52 | 13.36 | 15.51 | 20.09 |
| 9 | 2.09 | 3.33 | 4.17 | 6.39 | 8.34 | 10.66 | 14.68 | 16.92 | 21.67 |
| 10 | 2.56 | 3.94 | 4.86 | 7.27 | 9.34 | 11.78 | 15.99 | 18.31 | 23.21 |
| 11 | 3.05 | 4.58 | 5.58 | 8.15 | 10.34 | 12.90 | 17.28 | 19.68 | 24.73 |
| 12 | 3.57 | 5.23 | 6.30 | 9.03 | 11.34 | 14.01 | 18.55 | 21.03 | 26.22 |
| 13 | 4.11 | 5.89 | 7.04 | 9.93 | 12.34 | 15.12 | 19.81 | 22.36 | 27.69 |
| 14 | 4.66 | 6.57 | 7.79 | 10.82 | 13.34 | 16.22 | 21.06 | 23.69 | 29.14 |
| 15 | 5.23 | 7.26 | 8.55 | 11.72 | 14.34 | 17.32 | 22.31 | 25.00 | 30.58 |
| 16 | 5.81 | 7.96 | 9.31 | 12.62 | 15.34 | 18.42 | 23.54 | 26.30 | 32.00 |
| 17 | 6.41 | 8.67 | 10.09 | 13.53 | 16.34 | 19.51 | 24.77 | 27.59 | 33.41 |
| 18 | 7.00 | 9.39 | 10.87 | 15.44 | 17.34 | 20.60 | 25.99 | 28.87 | 34.81 |
| 19 | 7.63 | 10.12 | 11.65 | 15.35 | 18.34 | 21.69 | 27.20 | 30.14 | 36.19 |
| 20 | 8.26 | 10.85 | 12.44 | 16.27 | 19.34 | 22.78 | 28.41 | 31.41 | 37.57 |
| 22 | 9.54 | 12.34 | 14.04 | 18.10 | 21.34 | 24.94 | 30.81 | 33.92 | 40.29 |
| 24 | 10.86 | 13.85 | 15.66 | 19.94 | 23.34 | 27.10 | 33.20 | 36.42 | 42.98 |
| 26 | 12.20 | 15.38 | 17.29 | 21.79 | 25.34 | 29.25 | 35.56 | 38.89 | 45.64 |
| 28 | 13.57 | 16.93 | 18.94 | 23.65 | 27.34 | 31.39 | 37.92 | 41.34 | 48.28 |
| 30 | 14.95 | 18.49 | 20.60 | 25.51 | 29.34 | 33.53 | 40.26 | 43.77 | 50.89 |
| 32 | 16.36 | 20.07 | 22.27 | 27.37 | 31.34 | 35.67 | 42.59 | 46.19 | 53.49 |
| 34 | 17.79 | 21.66 | 23.95 | 29.24 | 33.34 | 37.80 | 44.90 | 48.60 | 56.06 |
| 36 | 19.23 | 23.27 | 25.64 | 31.11 | 35.34 | 39.92 | 47.21 | 51.00 | 58.62 |
| 38 | 20.69 | 24.88 | 27.34 | 32.99 | 37.34 | 42.04 | 49.51 | 53.38 | 61.16 |
| 40 | 22.16 | 26.51 | 29.05 | 34.87 | 39.34 | 44.17 | 51.81 | 55.76 | 63.69 |

Tableau C – Valeur de la loi t de Student. Probabilité d’avoir une valeur supérieure quel que soit le signe

| Degrés de Liberté | 50% | 70% | 90% | 95% | 99% |
|-------------------|------|------|------|-------|-------|
| 1 | 1.00 | 1.96 | 6.31 | 12.71 | 63.66 |
| 2 | 0.82 | 1.39 | 2.92 | 4.30 | 9.92 |
| 3 | 0.76 | 1.25 | 2.35 | 3.18 | 5.84 |
| 4 | 0.74 | 1.19 | 2.13 | 2.78 | 4.60 |
| 5 | 0.73 | 1.16 | 2.02 | 2.57 | 4.02 |
| 6 | 0.72 | 1.13 | 1.94 | 2.45 | 3.71 |
| 7 | 0.71 | 1.12 | 1.90 | 2.37 | 3.50 |
| 8 | 0.71 | 1.11 | 1.86 | 2.31 | 3.36 |
| 9 | 0.70 | 1.10 | 1.83 | 2.26 | 3.25 |
| 10 | 0.70 | 1.09 | 1.81 | 2.23 | 3.17 |
| 11 | 0.70 | 1.09 | 1.80 | 2.20 | 3.11 |
| 12 | 0.70 | 1.08 | 1.78 | 2.18 | 3.06 |
| 13 | 0.69 | 1.08 | 1.77 | 2.16 | 3.01 |
| 14 | 0.69 | 1.08 | 1.76 | 2.14 | 2.98 |
| 15 | 0.69 | 1.07 | 1.75 | 2.13 | 2.95 |
| 16 | 0.69 | 1.07 | 1.75 | 2.12 | 2.92 |
| 17 | 0.69 | 1.07 | 1.74 | 2.11 | 2.90 |
| 18 | 0.69 | 1.07 | 1.73 | 2.10 | 2.88 |
| 19 | 0.69 | 1.07 | 1.73 | 2.09 | 2.86 |
| 20 | 0.69 | 1.06 | 1.72 | 2.09 | 2.84 |

Tableau D – Seuil à 5% de la distribution F de Fisher.

| $f_2 \backslash f_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 12 | 24 | >25 |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 161.4 | 199.5 | 215.7 | 224.6 | 230.2 | 234.0 | 238.9 | 243.9 | 249.0 | 254.3 |
| 2 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.37 | 19.41 | 19.45 | 19.50 |
| 3 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.84 | 8.74 | 8.64 | 8.53 |
| 4 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.04 | 5.91 | 5.77 | 5.63 |
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.82 | 4.68 | 4.53 | 4.36 |
| 6 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.15 | 4.00 | 3.84 | 3.67 |
| 7 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.73 | 3.57 | 3.41 | 3.23 |
| 8 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.44 | 3.28 | 3.12 | 2.93 |
| 9 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.23 | 3.07 | 2.90 | 2.71 |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.07 | 2.91 | 2.74 | 2.54 |
| 11 | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 2.95 | 2.79 | 2.61 | 2.40 |
| 12 | 4.75 | 3.88 | 3.49 | 3.26 | 3.11 | 3.00 | 2.85 | 2.69 | 2.50 | 2.30 |
| 13 | 4.67 | 3.80 | 3.41 | 3.18 | 3.02 | 2.92 | 2.77 | 2.60 | 2.42 | 2.21 |
| 14 | 4.60 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.70 | 2.53 | 2.35 | 2.13 |
| 15 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.64 | 2.48 | 2.29 | 2.07 |
| 16 | 4.49 | 3.63 | 3.24 | 3.01 | 2.85 | 2.74 | 2.59 | 2.42 | 2.24 | 2.01 |
| 17 | 4.45 | 3.59 | 3.20 | 2.96 | 2.81 | 2.70 | 2.55 | 2.38 | 2.19 | 1.96 |
| 18 | 4.41 | 3.55 | 3.16 | 2.93 | 2.77 | 2.66 | 2.51 | 2.34 | 2.15 | 1.92 |
| 19 | 4.38 | 3.52 | 3.13 | 2.90 | 2.74 | 2.63 | 2.48 | 2.31 | 2.11 | 1.88 |
| 20 | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.45 | 2.28 | 2.08 | 1.84 |
| 21 | 4.32 | 3.47 | 3.07 | 2.84 | 2.68 | 2.57 | 2.42 | 2.25 | 2.05 | 1.81 |
| 22 | 4.30 | 3.44 | 3.05 | 2.82 | 2.66 | 2.55 | 2.40 | 2.23 | 2.03 | 1.78 |
| 23 | 4.28 | 3.42 | 3.03 | 2.80 | 2.64 | 2.53 | 2.38 | 2.20 | 2.00 | 1.76 |
| 24 | 4.26 | 3.40 | 3.01 | 2.78 | 2.62 | 2.51 | 2.36 | 2.18 | 1.98 | 1.73 |
| 25 | 4.24 | 3.38 | 2.99 | 2.76 | 2.60 | 2.49 | 2.34 | 2.16 | 1.96 | 1.71 |
| 26 | 4.22 | 3.37 | 2.98 | 2.74 | 2.59 | 2.47 | 2.32 | 2.15 | 1.95 | 1.69 |
| 27 | 4.21 | 3.35 | 2.96 | 2.73 | 2.57 | 2.46 | 2.30 | 2.13 | 1.93 | 1.67 |
| 28 | 4.20 | 3.34 | 2.95 | 2.71 | 2.56 | 2.44 | 2.29 | 2.12 | 1.91 | 1.65 |
| 29 | 4.18 | 3.33 | 2.93 | 2.70 | 2.54 | 2.43 | 2.28 | 2.10 | 1.90 | 1.64 |
| 30 | 4.17 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.27 | 2.09 | 1.89 | 1.62 |
| 40 | 4.08 | 3.23 | 2.84 | 2.61 | 2.45 | 2.34 | 2.18 | 2.00 | 1.79 | 1.51 |
| 60 | 4.00 | 3.15 | 2.76 | 2.52 | 2.37 | 2.25 | 2.10 | 1.92 | 1.70 | 1.39 |
| 120 | 3.92 | 3.07 | 2.68 | 2.45 | 2.29 | 2.17 | 2.02 | 1.83 | 1.61 | 1.25 |
| >120 | 3.84 | 2.99 | 2.60 | 2.37 | 2.21 | 2.10 | 1.94 | 1.75 | 1.52 | 1.00 |

f_1 correspond aux degrés de liberté de la plus grande variance.

Tableau E – Seuil à 1% de la distribution F de Fisher.

| $f_2 \backslash f_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 12 | 24 | >25 |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 4052. | 4999. | 5403. | 5625. | 5764. | 5859. | 5982. | 6106. | 6234. | 6366. |
| 2 | 98.50 | 99.00 | 99.17 | 99.25 | 99.30 | 99.33 | 99.37 | 99.42 | 99.46 | 99.50 |
| 3 | 34.12 | 30.82 | 29.46 | 28.71 | 28.24 | 27.91 | 27.49 | 27.05 | 26.60 | 26.12 |
| 4 | 21.20 | 18.00 | 16.69 | 15.98 | 15.52 | 15.21 | 14.80 | 14.37 | 13.93 | 13.46 |
| 5 | 16.26 | 13.27 | 12.06 | 11.39 | 10.97 | 10.67 | 10.29 | 9.89 | 9.47 | 9.02 |
| 6 | 13.74 | 10.92 | 9.78 | 9.15 | 8.75 | 8.47 | 8.10 | 7.72 | 7.31 | 6.88 |
| 7 | 12.25 | 9.55 | 8.45 | 7.85 | 7.46 | 7.19 | 6.84 | 6.47 | 6.07 | 5.65 |
| 8 | 11.26 | 8.65 | 7.59 | 7.01 | 6.63 | 6.37 | 6.03 | 5.67 | 5.28 | 4.86 |
| 9 | 10.56 | 8.02 | 6.99 | 6.42 | 6.06 | 5.80 | 5.47 | 5.11 | 4.73 | 4.31 |
| 10 | 10.04 | 7.56 | 6.55 | 5.99 | 5.64 | 5.39 | 5.06 | 4.71 | 4.33 | 3.91 |
| 11 | 9.65 | 7.20 | 6.22 | 5.67 | 5.32 | 5.07 | 4.74 | 4.40 | 4.02 | 3.60 |
| 12 | 9.33 | 6.93 | 5.95 | 5.41 | 5.06 | 4.82 | 4.50 | 4.16 | 3.78 | 3.36 |
| 13 | 9.07 | 6.70 | 5.74 | 5.20 | 4.86 | 4.62 | 4.30 | 3.96 | 3.59 | 3.16 |
| 14 | 8.86 | 6.51 | 5.56 | 5.03 | 4.69 | 4.46 | 4.14 | 3.80 | 3.43 | 3.00 |
| 15 | 8.68 | 6.36 | 5.42 | 4.89 | 4.56 | 4.32 | 4.00 | 3.67 | 3.29 | 2.87 |
| 16 | 8.53 | 6.23 | 5.29 | 4.77 | 4.44 | 4.20 | 3.89 | 3.55 | 3.18 | 2.75 |
| 17 | 8.40 | 6.11 | 5.18 | 4.67 | 4.34 | 4.10 | 3.79 | 3.45 | 3.08 | 2.65 |
| 18 | 8.28 | 6.01 | 5.09 | 4.58 | 4.25 | 4.01 | 3.71 | 3.37 | 3.00 | 2.57 |
| 19 | 8.18 | 5.93 | 5.01 | 4.50 | 4.17 | 3.94 | 3.63 | 3.30 | 2.92 | 2.49 |
| 20 | 8.10 | 5.85 | 4.94 | 4.43 | 4.10 | 3.87 | 3.56 | 3.23 | 2.86 | 2.42 |
| 21 | 8.02 | 5.78 | 4.87 | 4.37 | 4.04 | 3.81 | 3.51 | 3.17 | 2.80 | 2.36 |
| 22 | 7.94 | 5.72 | 4.82 | 4.31 | 3.99 | 3.76 | 3.45 | 3.12 | 2.75 | 2.31 |
| 23 | 7.88 | 5.66 | 4.76 | 4.26 | 3.94 | 3.71 | 3.41 | 3.07 | 2.70 | 2.26 |
| 24 | 7.82 | 5.61 | 4.72 | 4.22 | 3.90 | 3.67 | 3.36 | 3.03 | 2.66 | 2.21 |
| 25 | 7.77 | 5.57 | 4.68 | 4.18 | 3.86 | 3.63 | 3.32 | 2.99 | 2.62 | 2.17 |
| 26 | 7.72 | 5.53 | 4.64 | 4.14 | 3.82 | 3.59 | 3.29 | 2.96 | 2.58 | 2.13 |
| 27 | 7.68 | 5.49 | 4.60 | 4.11 | 3.78 | 3.56 | 3.26 | 2.93 | 2.55 | 2.10 |
| 28 | 7.64 | 5.45 | 4.57 | 4.07 | 3.75 | 3.53 | 3.23 | 2.90 | 2.52 | 2.06 |
| 29 | 7.60 | 5.42 | 4.54 | 4.04 | 3.73 | 3.50 | 3.20 | 2.87 | 2.49 | 2.03 |
| 30 | 7.56 | 5.39 | 4.51 | 4.02 | 3.70 | 3.47 | 3.17 | 2.84 | 2.47 | 2.01 |
| 40 | 7.31 | 5.18 | 4.31 | 3.83 | 3.51 | 3.29 | 2.99 | 2.66 | 2.29 | 1.80 |
| 60 | 7.08 | 4.98 | 4.13 | 3.65 | 3.34 | 3.12 | 2.82 | 2.50 | 2.12 | 1.60 |
| 120 | 6.85 | 4.79 | 3.95 | 3.48 | 3.17 | 2.96 | 2.66 | 2.34 | 1.95 | 1.38 |
| >120 | 6.64 | 4.60 | 3.78 | 3.32 | 3.02 | 2.80 | 2.51 | 2.18 | 1.79 | 1.00 |

f_1 correspond aux degrés de liberté de la plus grande variance.

Tableau G – Valeurs critiques pour le test de Friedman

| $b = 3$ | | | $b = 4$ | | | $b = 5$ | | | $b = 6$ | | |
|---------|-------|-------|---------|-------|-------|---------|-------|-------|---------|-------|-------|
| t | 5% | 1% | t | 5% | 1% | t | 5% | 1% | t | 5% | 1% |
| 2 | - | - | 2 | 6.000 | - | 2 | 7.600 | 8.000 | 2 | 9.357 | 9.929 |
| 3 | 6.000 | - | 3 | 7.400 | 9.000 | 3 | 8.533 | 10.13 | 3 | 9.857 | 11.76 |
| 4 | 6.500 | 8.000 | 4 | 7.800 | 9.600 | 4 | 8.800 | 11.20 | 4 | 10.39 | 12.82 |
| 5 | 6.400 | 8.400 | 5 | 7.800 | 9.960 | 5 | 8.960 | 11.68 | | | |
| 6 | 7.000 | 9.000 | 6 | 7.600 | 10.20 | 6 | 9.067 | 11.87 | | | |
| 7 | 7.143 | 8.857 | 7 | 7.800 | 10.54 | 7 | 9.143 | 12.11 | | | |
| 8 | 6.250 | 9.000 | 8 | 7.650 | 10.50 | 8 | 9.200 | 12.30 | | | |
| 9 | 6.222 | 9.556 | 9 | 7.667 | 10.73 | 9 | 9.244 | 12.44 | | | |
| 10 | 6.200 | 9.600 | 10 | 7.680 | 10.68 | | | | | | |
| 11 | 6.545 | 9.455 | 11 | 7.691 | 10.75 | | | | | | |
| 12 | 6.500 | 9.500 | 12 | 7.700 | 10.80 | | | | | | |
| 13 | 6.615 | 9.385 | 13 | 7.800 | 10.85 | | | | | | |
| 14 | 6.143 | 9.143 | 14 | 7.714 | 10.89 | | | | | | |
| 15 | 6.400 | 8.933 | 15 | 7.720 | 10.92 | | | | | | |
| 16 | 6.500 | 9.375 | 16 | 7.800 | 10.95 | | | | | | |
| 17 | 6.118 | 9.294 | 17 | 7.800 | 11.05 | | | | | | |
| 18 | 6.333 | 9.000 | 18 | 7.733 | 10.93 | | | | | | |
| 19 | 6.421 | 9.579 | 19 | 7.863 | 11.02 | | | | | | |
| 20 | 6.300 | 9.300 | 20 | 7.800 | 11.10 | | | | | | |
| 21 | 6.095 | 9.238 | 21 | 7.800 | 11.06 | | | | | | |
| 22 | 6.091 | 9.091 | 22 | 7.800 | 11.07 | | | | | | |
| 23 | 6.348 | 9.391 | | | | | | | | | |
| 24 | 6.250 | 9.250 | | | | | | | | | |
| 25 | 6.080 | 8.960 | | | | | | | | | |
| 26 | 6.077 | 9.308 | | | | | | | | | |
| 27 | 6.000 | 9.407 | | | | | | | | | |
| 28 | 6.500 | 9.214 | | | | | | | | | |
| 29 | 6.276 | 9.172 | | | | | | | | | |
| 30 | 6.200 | 9.267 | | | | | | | | | |
| 31 | 6.000 | 9.290 | | | | | | | | | |
| 32 | 6.063 | 9.250 | | | | | | | | | |
| 33 | 6.061 | 9.152 | | | | | | | | | |
| 34 | 6.059 | 9.176 | | | | | | | | | |
| 35 | 6.171 | 9.314 | | | | | | | | | |
| 36 | 6.167 | 9.389 | | | | | | | | | |
| 37 | 6.054 | 9.243 | | | | | | | | | |
| 38 | 6.158 | 9.053 | | | | | | | | | |
| 39 | 6.000 | 9.282 | | | | | | | | | |
| 40 | 6.050 | 9.150 | | | | | | | | | |
| 41 | 6.195 | 9.366 | | | | | | | | | |
| 42 | 6.143 | 9.190 | | | | | | | | | |
| 43 | 6.186 | 9.256 | | | | | | | | | |
| 44 | 6.318 | 9.136 | | | | | | | | | |
| 45 | 6.178 | 9.244 | | | | | | | | | |
| 46 | 6.043 | 9.435 | | | | | | | | | |
| 47 | 6.128 | 9.319 | | | | | | | | | |
| 48 | 6.167 | 9.125 | | | | | | | | | |
| 49 | 6.041 | 9.184 | | | | | | | | | |
| 50 | 6.040 | 9.160 | | | | | | | | | |

Tableau H – Valeurs critiques pour le coefficient de corrélation r_s de Spearman (test unilatéral)

| N | Niveau de significativité | |
|-----|---------------------------|-------|
| | 0.05 | 0.01 |
| 4 | 1.00 | |
| 5 | .900 | 1.000 |
| 6 | .829 | .943 |
| 7 | .714 | .893 |
| 8 | .643 | .833 |
| 9 | .600 | .783 |
| 10 | .564 | .746 |
| 12 | .506 | .712 |
| 14 | .456 | .645 |
| 16 | .425 | .601 |
| 18 | .399 | .564 |
| 20 | .377 | .534 |
| 22 | .359 | .508 |
| 24 | .343 | .485 |
| 26 | .329 | .465 |
| 28 | .317 | .448 |
| 30 | .306 | .432 |

Tableau I – Probabilités associées au coefficient de corrélation τ de Kendall

| τ | N | | | | τ | N | | |
|--------|------|-------|---------|----------|--------|-------|--------|-----------|
| | 4 | 5 | 8 | 9 | | 6 | 7 | 10 |
| 0 | .625 | .592 | .548 | .540 | 1 | .500 | .500 | .500 |
| 2 | .375 | .408 | .452 | .460 | 3 | .360 | .386 | .431 |
| 4 | .167 | .242 | .360 | .381 | 5 | .235 | .281 | .364 |
| 6 | .042 | .117 | .274 | .306 | 7 | .136 | .191 | .300 |
| 8 | | .042 | .199 | .238 | 9 | .068 | .119 | .242 |
| 10 | | .0083 | .138 | .179 | 11 | .028 | .068 | .190 |
| 12 | | | .089 | .130 | 13 | .0083 | .035 | .146 |
| 14 | | | .054 | .090 | 15 | .0014 | .015 | .108 |
| 16 | | | .031 | .060 | 17 | | .0054 | .078 |
| 18 | | | .016 | .038 | 19 | | .0014 | .054 |
| 20 | | | .0071 | .022 | 21 | | .00020 | .036 |
| 22 | | | .0028 | .012 | 23 | | | .023 |
| 24 | | | .00087 | .0063 | 25 | | | .014 |
| 26 | | | .00019 | .0029 | 27 | | | .0083 |
| 28 | | | .000025 | .0012 | 29 | | | .0046 |
| 30 | | | | .00043 | 31 | | | .0023 |
| 32 | | | | .00012 | 33 | | | .0011 |
| 34 | | | | .000025 | 35 | | | .00047 |
| 36 | | | | .0000028 | 37 | | | .00018 |
| | | | | | 39 | | | .000058 |
| | | | | | 41 | | | .000015 |
| | | | | | 43 | | | .0000028 |
| | | | | | 45 | | | .00000028 |

Tableau J – Valeurs critiques pour le coefficient de concordance de Kendall

| k | N | | | | |
|---------------------|-------|-------|-------|--------|--------|
| | 3 | 4 | 5 | 6 | 7 |
| Valeurs au seuil 5% | | | | | |
| 3 | | | 64.4 | 103.9 | 157.3 |
| 4 | | 49.5 | 88.4 | 143.3 | 217.0 |
| 5 | | 62.6 | 112.3 | 182.4 | 276.2 |
| 6 | | 75.7 | 136.1 | 221.4 | 335.2 |
| 8 | 48.1 | 101.7 | 183.7 | 299.0 | 453.1 |
| 10 | 60.0 | 127.8 | 231.2 | 376.7 | 571.0 |
| 15 | 89.8 | 192.9 | 349.8 | 570.5 | 864.9 |
| 20 | 119.7 | 258.0 | 468.5 | 764.4 | 1158.7 |
| Valeurs au seuil 1% | | | | | |
| 3 | | | 75.6 | 122.8 | 185.6 |
| 4 | | 61.4 | 109.3 | 176.2 | 265.0 |
| 5 | | 80.5 | 142.8 | 229.4 | 343.8 |
| 6 | | 99.5 | 176.1 | 282.4 | 422.6 |
| 8 | 66.8 | 137.4 | 242.7 | 388.3 | 579.9 |
| 10 | 85.1 | 175.3 | 309.1 | 494.0 | 737.0 |
| 15 | 131.0 | 269.8 | 475.2 | 758.2 | 1129.5 |
| 20 | 177.0 | 364.2 | 641.2 | 1022.2 | 1521.9 |

Tableau K – Valeurs critiques du test de Newman et Keuls au seuil 5%

| k^p | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|----------|------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 2 | 6.08 | 8.33 | 9.8 | 10.88 | 11.74 | 12.44 | 13.03 | 13.54 | 13.99 | 14.39 | 14.75 | 15.08 | 15.38 | 15.65 | 15.91 | 16.14 | 16.37 | 16.57 | 16.77 |
| 3 | 4.50 | 5.91 | 6.82 | 7.50 | 8.04 | 8.48 | 8.885 | 9.18 | 9.46 | 9.72 | 9.95 | 10.15 | 10.35 | 10.52 | 10.69 | 10.84 | 10.98 | 11.11 | 11.24 |
| 4 | 3.93 | 5.04 | 5.76 | 6.29 | 6.71 | 7.05 | 7.35 | 7.60 | 7.83 | 8.03 | 8.21 | 8.37 | 8.52 | 8.66 | 8.79 | 8.91 | 9.03 | 9.13 | 9.23 |
| 5 | 3.64 | 4.60 | 5.22 | 5.67 | 6.03 | 6.33 | 6.58 | 6.80 | 6.99 | 7.17 | 7.32 | 7.47 | 7.60 | 7.72 | 7.83 | 7.93 | 8.03 | 8.12 | 8.21 |
| 6 | 3.46 | 4.34 | 4.90 | 5.30 | 5.63 | 5.90 | 6.12 | 6.32 | 6.49 | 6.65 | 6.79 | 6.92 | 7.03 | 7.14 | 7.24 | 7.34 | 7.43 | 7.51 | 7.59 |
| 7 | 3.34 | 4.16 | 4.68 | 5.06 | 5.36 | 5.61 | 5.82 | 6.00 | 6.16 | 6.30 | 6.43 | 6.55 | 6.66 | 6.76 | 6.85 | 6.94 | 7.02 | 7.10 | 7.17 |
| 8 | 3.26 | 4.04 | 4.53 | 4.89 | 5.17 | 5.40 | 5.60 | 5.77 | 5.92 | 6.05 | 6.18 | 6.29 | 6.39 | 6.48 | 6.57 | 6.65 | 6.73 | 6.80 | 6.87 |
| 9 | 3.20 | 3.95 | 4.41 | 4.76 | 5.02 | 5.24 | 5.43 | 5.59 | 5.74 | 5.87 | 5.98 | 6.09 | 6.19 | 6.28 | 6.36 | 6.44 | 6.51 | 6.58 | 6.64 |
| 10 | 3.15 | 3.88 | 4.33 | 4.64 | 4.91 | 5.12 | 5.30 | 5.46 | 5.60 | 5.72 | 5.83 | 5.93 | 6.03 | 6.11 | 6.19 | 6.27 | 6.34 | 6.40 | 6.47 |
| 11 | 3.11 | 3.82 | 4.26 | 4.57 | 4.82 | 5.03 | 5.20 | 5.35 | 5.49 | 5.61 | 5.71 | 5.81 | 5.90 | 5.98 | 6.06 | 6.13 | 6.20 | 6.27 | 6.33 |
| 12 | 3.08 | 3.77 | 4.20 | 4.51 | 4.75 | 4.95 | 5.12 | 5.27 | 5.39 | 5.51 | 5.61 | 5.71 | 5.80 | 5.88 | 5.95 | 6.02 | 6.09 | 6.15 | 6.21 |
| 13 | 3.06 | 3.73 | 4.15 | 4.45 | 4.69 | 4.88 | 5.05 | 5.19 | 5.32 | 5.43 | 5.53 | 5.63 | 5.71 | 5.79 | 5.86 | 5.93 | 5.99 | 6.05 | 6.11 |
| 14 | 3.03 | 3.70 | 4.11 | 4.41 | 4.64 | 4.83 | 4.99 | 5.13 | 5.25 | 5.36 | 5.46 | 5.55 | 5.64 | 5.71 | 5.79 | 5.85 | 5.91 | 5.97 | 6.03 |
| 15 | 3.01 | 3.67 | 4.08 | 4.37 | 4.59 | 4.78 | 4.94 | 5.08 | 5.20 | 5.31 | 5.40 | 5.49 | 5.57 | 5.65 | 5.72 | 5.78 | 5.85 | 5.90 | 5.96 |
| 16 | 3.00 | 3.65 | 4.05 | 4.33 | 4.56 | 4.74 | 4.90 | 5.03 | 5.15 | 5.26 | 5.35 | 5.44 | 5.52 | 5.59 | 5.66 | 5.73 | 5.79 | 5.84 | 5.90 |
| 17 | 2.98 | 3.63 | 4.02 | 4.30 | 4.52 | 4.70 | 4.86 | 4.99 | 5.11 | 5.21 | 5.31 | 5.39 | 5.47 | 5.54 | 5.61 | 5.67 | 5.73 | 5.79 | 5.84 |
| 18 | 2.97 | 3.61 | 4.00 | 4.28 | 4.49 | 4.67 | 4.82 | 4.96 | 5.07 | 5.17 | 5.27 | 5.35 | 5.43 | 5.50 | 5.57 | 5.63 | 5.69 | 5.74 | 5.79 |
| 19 | 2.96 | 3.59 | 3.98 | 4.25 | 4.47 | 4.65 | 4.79 | 4.92 | 5.04 | 5.14 | 5.23 | 5.31 | 5.39 | 5.46 | 5.53 | 5.59 | 5.65 | 5.70 | 5.75 |
| 20 | 2.95 | 3.58 | 3.96 | 4.23 | 4.45 | 4.62 | 4.77 | 4.90 | 5.01 | 5.11 | 5.20 | 5.28 | 5.36 | 5.43 | 5.49 | 5.55 | 5.61 | 5.66 | 5.71 |
| 24 | 2.92 | 3.53 | 3.90 | 4.17 | 4.37 | 4.54 | 4.68 | 4.81 | 4.92 | 5.01 | 5.10 | 5.18 | 5.25 | 5.32 | 5.38 | 5.44 | 5.49 | 5.55 | 5.59 |
| 30 | 2.89 | 3.49 | 3.85 | 4.10 | 4.30 | 4.46 | 4.60 | 4.72 | 4.82 | 4.92 | 5.00 | 5.08 | 5.15 | 5.21 | 5.27 | 5.33 | 5.38 | 5.43 | 5.47 |
| 40 | 2.86 | 3.44 | 3.79 | 4.04 | 4.23 | 4.39 | 4.52 | 4.63 | 4.73 | 4.82 | 4.90 | 4.98 | 5.04 | 5.11 | 5.16 | 5.22 | 5.27 | 5.31 | 5.36 |
| 60 | 2.83 | 3.40 | 3.74 | 3.98 | 4.16 | 4.31 | 4.44 | 4.55 | 4.64 | 4.73 | 4.81 | 4.88 | 4.94 | 5.00 | 5.06 | 5.11 | 5.15 | 5.20 | 5.24 |
| 120 | 2.80 | 3.36 | 3.68 | 3.92 | 4.10 | 4.24 | 4.36 | 4.47 | 4.56 | 4.64 | 4.71 | 4.78 | 4.84 | 4.90 | 4.95 | 5.00 | 5.04 | 5.09 | 5.13 |
| ∞ | 2.77 | 3.31 | 3.63 | 3.86 | 4.03 | 4.17 | 4.29 | 4.39 | 4.47 | 4.55 | 4.62 | 4.68 | 4.74 | 4.80 | 4.85 | 4.89 | 4.93 | 4.97 | 5.01 |

p : nombre de populations

k : nombre de degrés de liberté

Tableau L – Valeurs critiques du test de Duncan au seuil 5%

| k^p | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|----------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 2 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 | 6.09 |
| 3 | 4.50 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 | 4.52 |
| 4 | 3.93 | 4.01 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 | 4.03 |
| 5 | 3.64 | 3.75 | 3.80 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 | 3.81 |
| 6 | 3.46 | 3.59 | 3.65 | 3.68 | 3.69 | 3.70 | 3.70 | 3.70 | 3.70 | 3.70 | 3.70 | 3.70 | 3.70 | 3.70 | 3.70 | 3.70 | 3.70 | 3.70 | 3.70 |
| 7 | 3.34 | 3.48 | 3.55 | 3.59 | 3.61 | 3.62 | 3.63 | 3.63 | 3.63 | 3.63 | 3.63 | 3.63 | 3.63 | 3.63 | 3.63 | 3.63 | 3.63 | 3.63 | 3.63 |
| 8 | 3.26 | 3.40 | 3.48 | 3.52 | 3.55 | 3.57 | 3.58 | 3.58 | 3.58 | 3.58 | 3.58 | 3.58 | 3.58 | 3.58 | 3.58 | 3.58 | 3.58 | 3.58 | 3.58 |
| 9 | 3.20 | 3.34 | 3.42 | 3.47 | 3.50 | 3.52 | 3.54 | 3.54 | 3.55 | 3.55 | 3.55 | 3.55 | 3.55 | 3.55 | 3.55 | 3.55 | 3.55 | 3.55 | 3.55 |
| 10 | 3.15 | 3.30 | 3.38 | 3.43 | 3.47 | 3.49 | 3.51 | 3.52 | 3.52 | 3.53 | 3.53 | 3.53 | 3.53 | 3.53 | 3.53 | 3.53 | 3.53 | 3.53 | 3.53 |
| 11 | 3.11 | 3.26 | 3.34 | 3.40 | 3.44 | 3.46 | 3.48 | 3.49 | 3.50 | 3.51 | 3.51 | 3.51 | 3.51 | 3.51 | 3.51 | 3.51 | 3.51 | 3.51 | 3.51 |
| 12 | 3.08 | 3.23 | 3.31 | 3.37 | 3.41 | 3.44 | 3.46 | 3.47 | 3.48 | 3.49 | 3.50 | 3.50 | 3.50 | 3.50 | 3.50 | 3.50 | 3.50 | 3.50 | 3.50 |
| 13 | 3.06 | 3.20 | 3.29 | 3.49 | 3.39 | 3.42 | 3.44 | 3.46 | 3.47 | 3.48 | 3.48 | 3.49 | 3.49 | 3.49 | 3.49 | 3.49 | 3.49 | 3.49 | 3.49 |
| 14 | 3.03 | 3.18 | 3.27 | 3.33 | 3.37 | 3.40 | 3.43 | 3.44 | 3.46 | 3.47 | 3.48 | 3.48 | 3.48 | 3.48 | 3.49 | 3.49 | 3.49 | 3.49 | 3.49 |
| 15 | 3.01 | 3.16 | 3.25 | 3.31 | 3.34 | 3.39 | 3.41 | 3.43 | 3.45 | 3.46 | 3.47 | 3.47 | 3.48 | 3.48 | 3.48 | 3.48 | 3.48 | 3.48 | 3.48 |
| 16 | 3.00 | 3.14 | 3.24 | 3.30 | 3.34 | 3.38 | 3.40 | 3.42 | 3.44 | 3.45 | 3.46 | 3.47 | 3.47 | 3.47 | 3.48 | 3.48 | 3.48 | 3.48 | 3.48 |
| 17 | 2.99 | 3.13 | 3.22 | 3.29 | 3.33 | 3.37 | 3.39 | 3.41 | 3.43 | 3.44 | 3.45 | 3.46 | 3.47 | 3.47 | 3.47 | 3.48 | 3.48 | 3.48 | 3.48 |
| 18 | 2.97 | 3.12 | 3.21 | 3.28 | 3.32 | 3.36 | 3.38 | 3.41 | 3.42 | 3.44 | 3.45 | 3.45 | 3.46 | 3.47 | 3.47 | 3.47 | 3.47 | 3.47 | 3.47 |
| 19 | 2.96 | 3.11 | 3.20 | 3.27 | 3.31 | 3.35 | 3.38 | 3.40 | 3.42 | 3.43 | 3.44 | 3.45 | 3.46 | 3.46 | 3.47 | 3.47 | 3.47 | 3.47 | 3.47 |
| 20 | 2.96 | 3.01 | 3.19 | 3.26 | 3.30 | 3.34 | 3.37 | 3.39 | 3.41 | 3.42 | 3.44 | 3.45 | 3.45 | 3.46 | 3.46 | 3.47 | 3.47 | 3.47 | 3.47 |
| 24 | 2.92 | 3.01 | 3.16 | 3.23 | 3.28 | 3.32 | 3.35 | 3.38 | 3.39 | 3.41 | 3.42 | 3.43 | 3.44 | 3.45 | 3.46 | 3.46 | 3.47 | 3.47 | 3.47 |
| 30 | 2.89 | 3.04 | 3.13 | 3.20 | 3.25 | 3.29 | 3.32 | 3.35 | 3.37 | 3.39 | 3.41 | 3.42 | 3.43 | 3.44 | 3.45 | 3.46 | 3.46 | 3.47 | 3.47 |
| 40 | 2.86 | 3.01 | 3.10 | 3.17 | 3.22 | 3.27 | 3.30 | 3.33 | 3.35 | 3.37 | 3.39 | 3.41 | 3.42 | 3.43 | 3.44 | 3.45 | 3.46 | 3.46 | 3.46 |
| 60 | 2.83 | 2.98 | 3.07 | 3.14 | 3.20 | 3.24 | 3.28 | 3.31 | 3.33 | 3.36 | 3.37 | 3.39 | 3.41 | 3.42 | 3.42 | 3.44 | 3.45 | 3.46 | 3.46 |
| 120 | 2.80 | 2.95 | 3.05 | 3.12 | 3.17 | 3.22 | 3.25 | 3.29 | 3.31 | 3.34 | 3.36 | 3.38 | 3.40 | 3.41 | 3.42 | 3.44 | 3.45 | 3.46 | 3.70 |
| ∞ | 2.77 | 2.92 | 3.02 | 3.09 | 3.15 | 3.19 | 3.23 | 3.27 | 3.29 | 3.32 | 3.34 | 3.36 | 3.38 | 3.40 | 3.41 | 3.43 | 3.44 | 3.45 | 3.46 |

 p : nombre de populations k : nombre de degrés de liberté

Tableau M – Valeurs critiques du test de Dunett au seuil 5%

| k^p | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 15 | 20 |
|----------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 5 | 2.57 | 3.03 | 3.29 | 3.48 | 3.62 | 3.73 | 3.82 | 3.90 | 3.97 | 4.03 | 4.09 | 4.14 | 4.26 | 4.42 |
| 6 | 2.45 | 2.86 | 3.10 | 3.26 | 3.39 | 3.49 | 3.57 | 3.64 | 3.71 | 3.76 | 3.81 | 3.86 | 3.97 | 4.11 |
| 7 | 2.36 | 2.75 | 2.97 | 3.12 | 3.24 | 3.33 | 3.41 | 3.47 | 3.53 | 3.58 | 3.63 | 3.67 | 3.78 | 3.91 |
| 8 | 2.31 | 2.67 | 2.88 | 3.02 | 3.13 | 3.22 | 3.29 | 3.35 | 3.41 | 3.46 | 3.50 | 3.54 | 3.64 | 3.76 |
| 9 | 2.26 | 2.61 | 2.81 | 2.95 | 3.05 | 3.14 | 3.20 | 3.26 | 3.32 | 3.36 | 3.40 | 3.44 | 3.53 | 3.65 |
| 10 | 2.23 | 2.57 | 2.76 | 2.89 | 2.99 | 3.07 | 3.14 | 3.19 | 3.24 | 3.29 | 3.33 | 3.36 | 3.45 | 3.57 |
| 11 | 2.20 | 2.53 | 2.72 | 2.84 | 2.94 | 3.02 | 3.08 | 3.14 | 3.19 | 3.23 | 3.27 | 3.30 | 3.39 | 3.50 |
| 12 | 2.18 | 2.50 | 2.68 | 2.81 | 2.90 | 2.98 | 3.04 | 3.09 | 3.14 | 3.18 | 3.22 | 3.25 | 3.34 | 3.45 |
| 13 | 2.16 | 2.48 | 2.65 | 2.78 | 2.87 | 2.94 | 3.00 | 3.06 | 3.10 | 3.14 | 3.18 | 3.21 | 3.29 | 3.40 |
| 14 | 2.14 | 2.46 | 2.63 | 2.75 | 2.84 | 2.94 | 2.97 | 3.02 | 3.07 | 3.11 | 3.14 | 3.18 | 3.26 | 3.36 |
| 15 | 2.13 | 2.44 | 2.61 | 2.73 | 2.82 | 2.89 | 2.95 | 3.00 | 3.04 | 3.08 | 3.12 | 3.15 | 3.23 | 3.33 |
| 16 | 2.12 | 2.42 | 2.59 | 2.71 | 2.80 | 2.87 | 2.91 | 2.97 | 3.02 | 3.06 | 3.09 | 3.12 | 3.20 | 3.30 |
| 17 | 2.11 | 2.41 | 2.58 | 2.69 | 2.78 | 2.85 | 2.90 | 2.95 | 3.00 | 3.03 | 3.07 | 3.10 | 3.18 | 3.27 |
| 18 | 2.10 | 2.40 | 2.45 | 2.68 | 2.76 | 2.83 | 2.89 | 2.94 | 2.98 | 3.01 | 3.05 | 3.08 | 3.16 | 3.25 |
| 19 | 2.09 | 2.39 | 2.55 | 2.66 | 2.75 | 2.81 | 2.87 | 2.92 | 2.96 | 3.00 | 3.03 | 3.06 | 3.14 | 3.23 |
| 20 | 2.09 | 2.38 | 2.54 | 2.65 | 2.73 | 2.80 | 2.86 | 2.90 | 2.95 | 2.98 | 3.02 | 3.05 | 3.12 | 3.22 |
| 24 | 2.06 | 2.35 | 2.51 | 2.61 | 2.70 | 2.76 | 2.81 | 2.86 | 2.90 | 2.94 | 2.97 | 3.00 | 3.07 | 3.16 |
| 30 | 2.04 | 2.32 | 2.47 | 2.58 | 2.66 | 2.72 | 2.77 | 2.82 | 2.86 | 2.89 | 2.92 | 2.95 | 3.02 | 3.11 |
| 40 | 2.02 | 2.29 | 2.44 | 2.54 | 2.62 | 2.68 | 2.73 | 2.77 | 2.81 | 2.85 | 2.87 | 2.90 | 2.97 | 3.06 |
| 60 | 2.00 | 2.27 | 2.41 | 2.51 | 2.58 | 2.64 | 2.69 | 2.73 | 2.77 | 2.80 | 2.83 | 2.86 | 2.92 | 3.00 |
| 120 | 1.98 | 2.24 | 2.38 | 2.47 | 2.55 | 2.60 | 2.65 | 2.69 | 2.73 | 2.76 | 2.79 | 2.81 | 2.87 | 2.95 |
| ∞ | 1.96 | 2.21 | 2.35 | 2.44 | 2.51 | 2.57 | 2.61 | 2.65 | 2.69 | 2.72 | 2.74 | 2.77 | 2.83 | 2.91 |

p : nombre de populations

k : nombre de degrés de liberté

Tableau N – Valeurs des scores normaux $a_n(i)$

| i | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1 | 0.7071 | 0.7071 | 0.6872 | 0.6646 | 0.6431 | 0.6233 | 0.6052 | 0.5888 | 0.5739 | |
| 2 | | 0.0000 | 0.1677 | 0.2413 | 0.2806 | 0.3031 | 0.3164 | 0.6244 | 0.3291 | |
| 3 | | | | 0.0000 | 0.0875 | 0.1401 | 0.1743 | 0.1976 | 0.2141 | |
| 4 | | | | | | 0.0000 | 0.0561 | 0.0947 | 0.1224 | |
| 5 | | | | | | | | 0.0000 | 0.0399 | |
| i | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 1 | 0.5601 | 0.5475 | 0.5359 | 0.5251 | 0.5150 | 0.5056 | 0.4968 | 0.4886 | 0.4808 | 0.4734 |
| 2 | 0.3315 | 0.3325 | 0.3325 | 0.3318 | 0.3306 | 0.3290 | 0.3273 | 0.3253 | 0.3232 | 0.3211 |
| 3 | 0.2260 | 0.2347 | 0.2412 | 0.2460 | 0.2495 | 0.2521 | 0.2540 | 0.2553 | 0.2561 | 0.2565 |
| 4 | 0.1429 | 0.1586 | 0.1707 | 0.1802 | 0.1878 | 0.1939 | 0.1988 | 0.2027 | 0.2059 | 0.2085 |
| 5 | 0.0695 | 0.0922 | 0.1099 | 0.1240 | 0.1353 | 0.1447 | 0.1524 | 0.1587 | 0.1641 | 0.1686 |
| 6 | 0.0000 | 0.0303 | 0.0539 | 0.0727 | 0.0880 | 0.1005 | 0.1109 | 0.1197 | 0.1271 | 0.1334 |
| 7 | | | 0.0000 | 0.0240 | 0.0433 | 0.0593 | 0.0725 | 0.0837 | 0.0932 | 0.1013 |
| 8 | | | | | 0.0000 | 0.0196 | 0.0359 | 0.0496 | 0.0612 | 0.0711 |
| 9 | | | | | | | 0.0000 | 0.0163 | 0.0303 | 0.0422 |
| 10 | | | | | | | | | 0.0000 | 0.0140 |
| i | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| 1 | 0.4643 | 0.4593 | 0.4552 | 0.4493 | 0.4450 | 0.4407 | 0.4366 | 0.4328 | 0.4291 | 0.4254 |
| 2 | 0.3185 | 0.3156 | 0.3126 | 0.3098 | 0.3069 | 0.3043 | 0.3018 | 0.2992 | 0.2968 | 0.2944 |
| 3 | 0.2578 | 0.2571 | 0.2561 | 0.2554 | 0.2543 | 0.2533 | 0.2522 | 0.2510 | 0.2499 | 0.2487 |
| 4 | 0.2119 | 0.2131 | 0.2139 | 0.2145 | 0.2148 | 0.2151 | 0.2152 | 0.2151 | 0.2150 | 0.2148 |
| 5 | 0.1736 | 0.1764 | 0.1787 | 0.1807 | 0.1822 | 0.1836 | 0.1848 | 0.1857 | 0.1864 | 0.1870 |
| 6 | 0.1399 | 0.1443 | 0.1480 | 0.1512 | 0.1539 | 0.1563 | 0.1584 | 0.1601 | 0.1616 | 0.1630 |
| 7 | 0.1092 | 0.1150 | 0.1201 | 0.1245 | 0.1283 | 0.1316 | 0.1346 | 0.1372 | 0.1395 | 0.1415 |
| 8 | 0.0804 | 0.0878 | 0.0941 | 0.0997 | 0.1046 | 0.1089 | 0.1128 | 0.1162 | 0.1192 | 0.1219 |
| 9 | 0.0530 | 0.0618 | 0.0696 | 0.0764 | 0.0823 | 0.0876 | 0.0923 | 0.0965 | 0.1002 | 0.1036 |
| 10 | 0.0263 | 0.0368 | 0.0459 | 0.0539 | 0.0610 | 0.0672 | 0.0728 | 0.0778 | 0.0822 | 0.0962 |
| 11 | 0.0000 | 0.0122 | 0.0228 | 0.0321 | 0.0403 | 0.0476 | 0.0540 | 0.0598 | 0.0650 | 0.0697 |
| 12 | | | 0.0000 | 0.0107 | 0.0200 | 0.0284 | 0.0358 | 0.0424 | 0.0483 | 0.0537 |
| 13 | | | | | 0.0000 | 0.0094 | 0.0178 | 0.0253 | 0.0320 | 0.0381 |
| 14 | | | | | | | 0.0000 | 0.0084 | 0.0159 | 0.0227 |
| 15 | | | | | | | | | 0.0000 | 0.0076 |

$a_n(i) = E(X^{(i)})$ où $X^{(i)}$ est la $i^{\text{ème}}$ observation ordonnée d'un échantillon de taille n extrait de la loi $\mathcal{N}(0, 1)$:

$$X^{(1)} \leq \dots \leq X^{(i-1)} \leq X^{(i)} \leq X^{(i+1)} \leq \dots \leq X^{(n)}$$

Tableau O – Valeur critique du test de Shapiro-Wilks

| n | 0.01 | 0.02 | 0.05 | 0.10 | 0.50 | 0.90 | 0.95 | 0.98 | 0.99 |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 3 | 0.753 | 0.756 | 0.767 | 0.789 | 0.959 | 0.998 | 0.999 | 1.000 | 1.000 |
| 4 | 0.687 | 0.707 | 0.748 | 0.792 | 0.935 | 0.987 | 0.992 | 0.996 | 0.997 |
| 5 | 0.686 | 0.715 | 0.762 | 0.806 | 0.927 | 0.979 | 0.986 | 0.991 | 0.993 |
| 6 | 0.713 | 0.743 | 0.788 | 0.826 | 0.927 | 0.974 | 0.981 | 0.986 | 0.989 |
| 7 | 0.730 | 0.760 | 0.803 | 0.838 | 0.928 | 0.972 | 0.979 | 0.985 | 0.988 |
| 8 | 0.749 | 0.778 | 0.818 | 0.851 | 0.932 | 0.972 | 0.978 | 0.984 | 0.987 |
| 9 | 0.764 | 0.791 | 0.829 | 0.859 | 0.935 | 0.972 | 0.978 | 0.984 | 0.986 |
| 10 | 0.781 | 0.806 | 0.842 | 0.869 | 0.938 | 0.972 | 0.978 | 0.983 | 0.986 |
| 11 | 0.792 | 0.817 | 0.850 | 0.876 | 0.940 | 0.973 | 0.979 | 0.984 | 0.986 |
| 12 | 0.805 | 0.828 | 0.859 | 0.883 | 0.943 | 0.973 | 0.979 | 0.984 | 0.986 |
| 13 | 0.814 | 0.837 | 0.866 | 0.889 | 0.945 | 0.974 | 0.979 | 0.984 | 0.986 |
| 14 | 0.825 | 0.846 | 0.874 | 0.895 | 0.947 | 0.975 | 0.980 | 0.984 | 0.986 |
| 15 | 0.835 | 0.855 | 0.881 | 0.901 | 0.950 | 0.975 | 0.980 | 0.984 | 0.987 |
| 16 | 0.844 | 0.863 | 0.887 | 0.906 | 0.952 | 0.976 | 0.981 | 0.985 | 0.987 |
| 17 | 0.851 | 0.869 | 0.892 | 0.910 | 0.954 | 0.977 | 0.981 | 0.985 | 0.987 |
| 18 | 0.858 | 0.874 | 0.897 | 0.914 | 0.956 | 0.978 | 0.982 | 0.986 | 0.988 |
| 19 | 0.863 | 0.879 | 0.901 | 0.917 | 0.957 | 0.978 | 0.982 | 0.986 | 0.988 |
| 20 | 0.868 | 0.884 | 0.905 | 0.920 | 0.959 | 0.979 | 0.983 | 0.986 | 0.988 |
| 21 | 0.873 | 0.888 | 0.908 | 0.923 | 0.960 | 0.870 | 0.983 | 0.987 | 0.989 |
| 22 | 0.878 | 0.892 | 0.911 | 0.926 | 0.961 | 0.980 | 0.984 | 0.987 | 0.989 |
| 23 | 0.881 | 0.895 | 0.914 | 0.925 | 0.962 | 0.981 | 0.984 | 0.987 | 0.989 |
| 24 | 0.884 | 0.898 | 0.916 | 0.930 | 0.963 | 0.981 | 0.984 | 0.987 | 0.989 |
| 25 | 0.888 | 0.901 | 0.918 | 0.931 | 0.964 | 0.981 | 0.985 | 0.988 | 0.989 |
| 26 | 0.891 | 0.904 | 0.920 | 0.933 | 0.965 | 0.982 | 0.985 | 0.988 | 0.989 |
| 27 | 0.894 | 0.906 | 0.923 | 0.935 | 0.965 | 0.982 | 0.985 | 0.988 | 0.990 |
| 28 | 0.896 | 0.908 | 0.924 | 0.936 | 0.966 | 0.982 | 0.985 | 0.988 | 0.990 |
| 29 | 0.898 | 0.910 | 0.926 | 0.937 | 0.966 | 0.982 | 0.985 | 0.988 | 0.990 |
| 30 | 0.900 | 0.912 | 0.927 | 0.939 | 0.967 | 0.983 | 0.985 | 0.988 | 0.990 |
| 31 | 0.902 | 0.914 | 0.929 | 0.940 | 0.967 | 0.983 | 0.986 | 0.988 | 0.990 |
| 32 | 0.904 | 0.915 | 0.930 | 0.941 | 0.968 | 0.983 | 0.986 | 0.988 | 0.990 |
| 33 | 0.906 | 0.917 | 0.931 | 0.942 | 0.968 | 0.983 | 0.986 | 0.988 | 0.990 |
| 34 | 0.908 | 0.919 | 0.933 | 0.943 | 0.969 | 0.983 | 0.986 | 0.988 | 0.990 |
| 35 | 0.910 | 0.920 | 0.934 | 0.944 | 0.969 | 0.984 | 0.986 | 0.988 | 0.990 |
| 36 | 0.912 | 0.922 | 0.935 | 0.945 | 0.970 | 0.984 | 0.986 | 0.988 | 0.990 |
| 37 | 0.914 | 0.924 | 0.936 | 0.946 | 0.970 | 0.984 | 0.987 | 0.989 | 0.990 |
| 38 | 0.916 | 0.925 | 0.938 | 0.947 | 0.971 | 0.984 | 0.987 | 0.989 | 0.990 |
| 39 | 0.917 | 0.927 | 0.939 | 0.948 | 0.971 | 0.984 | 0.987 | 0.989 | 0.991 |
| 40 | 0.919 | 0.928 | 0.940 | 0.949 | 0.972 | 0.982 | 0.987 | 0.989 | 0.991 |
| 41 | 0.920 | 0.929 | 0.941 | 0.950 | 0.972 | 0.982 | 0.987 | 0.989 | 0.991 |
| 42 | 0.922 | 0.930 | 0.942 | 0.951 | 0.972 | 0.982 | 0.987 | 0.989 | 0.991 |
| 43 | 0.923 | 0.932 | 0.943 | 0.951 | 0.973 | 0.985 | 0.987 | 0.990 | 0.991 |
| 44 | 0.924 | 0.933 | 0.944 | 0.952 | 0.973 | 0.985 | 0.987 | 0.990 | 0.991 |
| 45 | 0.926 | 0.934 | 0.945 | 0.953 | 0.973 | 0.985 | 0.988 | 0.990 | 0.991 |

Tableau P – Valeur critique du test de Durbin-Watson au seuil de 5%

| n | $k = 1$ | | $k = 2$ | | $k = 3$ | | $k = 4$ | | $k = 5$ | |
|-----|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| | d_{inf} | d_{sup} | d_{inf} | d_{sup} | d_{inf} | d_{sup} | d_{inf} | d_{sup} | d_{inf} | d_{sup} |
| 15 | 1.08 | 1.36 | 0.95 | 1.54 | 0.82 | 1.75 | 0.69 | 1.97 | 0.56 | 2.21 |
| 16 | 1.10 | 1.37 | 0.98 | 1.54 | 0.86 | 1.73 | 0.74 | 1.93 | 0.62 | 2.15 |
| 17 | 1.13 | 1.38 | 1.02 | 1.54 | 0.90 | 1.71 | 0.78 | 1.90 | 0.67 | 2.10 |
| 18 | 1.16 | 1.39 | 1.05 | 1.53 | 0.93 | 1.69 | 0.82 | 1.87 | 0.71 | 2.06 |
| 19 | 1.18 | 1.40 | 1.08 | 1.53 | 0.97 | 1.68 | 0.86 | 1.85 | 0.75 | 2.02 |
| 20 | 1.20 | 1.41 | 1.10 | 1.54 | 1.00 | 1.68 | 0.90 | 1.83 | 0.79 | 1.99 |
| 21 | 1.22 | 1.42 | 1.13 | 1.54 | 1.03 | 1.67 | 0.93 | 1.81 | 0.83 | 1.96 |
| 22 | 1.24 | 1.43 | 1.15 | 1.54 | 1.05 | 1.66 | 0.96 | 1.80 | 0.86 | 1.94 |
| 23 | 1.26 | 1.44 | 1.17 | 1.54 | 1.08 | 1.66 | 0.99 | 1.79 | 0.90 | 1.92 |
| 24 | 1.27 | 1.45 | 1.19 | 1.55 | 1.10 | 1.66 | 1.01 | 1.78 | 0.93 | 1.90 |
| 25 | 1.29 | 1.45 | 1.21 | 1.55 | 1.12 | 1.66 | 1.04 | 1.77 | 0.95 | 1.89 |
| 26 | 1.30 | 1.46 | 1.22 | 1.55 | 1.14 | 1.65 | 1.06 | 1.76 | 0.98 | 1.88 |
| 27 | 1.32 | 1.47 | 1.24 | 1.56 | 1.16 | 1.65 | 1.08 | 1.76 | 1.01 | 1.86 |
| 28 | 1.33 | 1.48 | 1.26 | 1.56 | 1.18 | 1.65 | 1.10 | 1.75 | 1.03 | 1.85 |
| 29 | 1.34 | 1.48 | 1.27 | 1.56 | 1.20 | 1.65 | 1.12 | 1.74 | 1.05 | 1.84 |
| 30 | 1.35 | 1.49 | 1.28 | 1.57 | 1.21 | 1.65 | 1.14 | 1.74 | 1.07 | 1.83 |
| 31 | 1.36 | 1.50 | 1.30 | 1.57 | 1.23 | 1.65 | 1.16 | 1.74 | 1.09 | 1.83 |
| 32 | 1.37 | 1.50 | 1.31 | 1.57 | 1.24 | 1.65 | 1.18 | 1.73 | 1.11 | 1.82 |
| 33 | 1.38 | 1.51 | 1.32 | 1.58 | 1.26 | 1.65 | 1.19 | 1.73 | 1.13 | 1.81 |
| 34 | 1.39 | 1.51 | 1.33 | 1.58 | 1.27 | 1.65 | 1.21 | 1.73 | 1.15 | 1.81 |
| 35 | 1.40 | 1.52 | 1.34 | 1.58 | 1.28 | 1.65 | 1.22 | 1.73 | 1.16 | 1.80 |
| 36 | 1.41 | 1.52 | 1.35 | 1.59 | 1.29 | 1.65 | 1.24 | 1.73 | 1.18 | 1.80 |
| 37 | 1.42 | 1.53 | 1.36 | 1.59 | 1.31 | 1.66 | 1.25 | 1.72 | 1.19 | 1.80 |
| 38 | 1.43 | 1.54 | 1.37 | 1.59 | 1.32 | 1.66 | 1.26 | 1.72 | 1.21 | 1.79 |
| 39 | 1.43 | 1.54 | 1.38 | 1.60 | 1.33 | 1.66 | 1.27 | 1.72 | 1.22 | 1.79 |
| 40 | 1.44 | 1.54 | 1.39 | 1.60 | 1.34 | 1.66 | 1.29 | 1.72 | 1.23 | 1.79 |
| 45 | 1.48 | 1.57 | 1.43 | 1.62 | 1.38 | 1.67 | 1.34 | 1.72 | 1.29 | 1.78 |
| 50 | 1.50 | 1.59 | 1.46 | 1.63 | 1.42 | 1.67 | 1.38 | 1.72 | 1.34 | 1.77 |
| 55 | 1.53 | 1.60 | 1.49 | 1.64 | 1.45 | 1.68 | 1.41 | 1.72 | 1.38 | 1.77 |
| 60 | 1.55 | 1.62 | 1.51 | 1.65 | 1.48 | 1.69 | 1.44 | 1.73 | 1.41 | 1.77 |
| 65 | 1.57 | 1.63 | 1.54 | 1.66 | 1.50 | 1.70 | 1.47 | 1.73 | 1.44 | 1.77 |
| 70 | 1.58 | 1.64 | 1.55 | 1.67 | 1.52 | 1.70 | 1.49 | 1.74 | 1.46 | 1.77 |
| 75 | 1.60 | 1.65 | 1.57 | 1.68 | 1.54 | 1.71 | 1.51 | 1.74 | 1.49 | 1.77 |
| 80 | 1.61 | 1.66 | 1.59 | 1.69 | 1.56 | 1.72 | 1.53 | 1.74 | 1.51 | 1.77 |
| 85 | 1.62 | 1.67 | 1.60 | 1.70 | 1.57 | 1.72 | 1.55 | 1.75 | 1.52 | 1.77 |
| 90 | 1.63 | 1.68 | 1.61 | 1.70 | 1.59 | 1.73 | 1.57 | 1.75 | 1.54 | 1.78 |
| 95 | 1.64 | 1.69 | 1.62 | 1.71 | 1.60 | 1.73 | 1.58 | 1.75 | 1.56 | 1.78 |
| 100 | 1.65 | 1.69 | 1.63 | 1.72 | 1.61 | 1.74 | 1.59 | 1.76 | 1.57 | 1.78 |

Tableau Q – Valeur critique du test du nombre de paires au seuil de 5%

| | | Valeurs inférieures | | | | | | | | | | | | | | | | | | |
|-------|-------|---------------------|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|
| n_1 | n_2 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 2 | 2 | | | | | | | | | | | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 3 | 3 | | | | | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 |
| 4 | 4 | | | | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 |
| 5 | 5 | | | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 5 |
| 6 | 6 | | 2 | 2 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 6 |
| 7 | 7 | | 2 | 2 | 3 | 3 | 3 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 6 |
| 8 | 8 | | 2 | 3 | 3 | 3 | 4 | 4 | 5 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 7 | 7 | 7 | 7 |
| 9 | 9 | | 2 | 3 | 3 | 4 | 4 | 5 | 5 | 5 | 6 | 6 | 6 | 7 | 7 | 7 | 7 | 8 | 8 | 8 |
| 10 | 10 | | 2 | 3 | 3 | 4 | 5 | 5 | 5 | 6 | 6 | 7 | 7 | 7 | 7 | 8 | 8 | 8 | 8 | 9 |
| 11 | 11 | | 2 | 3 | 4 | 4 | 5 | 5 | 6 | 6 | 7 | 7 | 7 | 8 | 8 | 8 | 9 | 9 | 9 | 9 |
| 12 | 12 | 2 | 2 | 3 | 4 | 4 | 5 | 6 | 6 | 7 | 7 | 7 | 8 | 8 | 8 | 9 | 9 | 9 | 10 | 10 |
| 13 | 13 | 2 | 2 | 3 | 4 | 5 | 5 | 6 | 6 | 7 | 7 | 8 | 8 | 9 | 9 | 9 | 10 | 10 | 10 | 10 |
| 14 | 14 | 2 | 2 | 3 | 4 | 5 | 5 | 6 | 7 | 7 | 8 | 8 | 9 | 9 | 9 | 10 | 10 | 10 | 11 | 11 |
| 15 | 15 | 2 | 3 | 3 | 4 | 5 | 6 | 6 | 7 | 7 | 8 | 8 | 9 | 9 | 10 | 10 | 11 | 11 | 11 | 12 |
| 16 | 16 | 2 | 3 | 4 | 4 | 5 | 6 | 6 | 7 | 8 | 8 | 9 | 9 | 10 | 10 | 11 | 11 | 11 | 12 | 12 |
| 17 | 17 | 2 | 3 | 4 | 4 | 5 | 6 | 7 | 7 | 8 | 9 | 9 | 10 | 10 | 11 | 11 | 11 | 12 | 12 | 13 |
| 18 | 18 | 2 | 3 | 4 | 5 | 5 | 6 | 7 | 8 | 8 | 9 | 9 | 10 | 10 | 11 | 11 | 12 | 12 | 13 | 13 |
| 19 | 19 | 2 | 3 | 4 | 5 | 6 | 6 | 7 | 8 | 8 | 9 | 10 | 10 | 11 | 11 | 12 | 12 | 13 | 13 | 13 |
| 20 | 20 | 2 | 3 | 4 | 5 | 6 | 6 | 7 | 8 | 9 | 9 | 10 | 10 | 11 | 12 | 12 | 13 | 13 | 13 | 14 |

| | | Valeurs supérieures | | | | | | | | | | | | | | | | | | |
|-------|-------|---------------------|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| n_1 | n_2 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 2 | 2 | | | | | | | | | | | | | | | | | | | |
| 3 | 3 | | | | | | | | | | | | | | | | | | | |
| 4 | 4 | | | | 9 | 9 | | | | | | | | | | | | | | |
| 5 | 5 | | | 9 | 10 | 10 | 11 | 11 | | | | | | | | | | | | |
| 6 | 6 | | | 9 | 10 | 11 | 12 | 12 | 13 | 13 | 13 | 13 | | | | | | | | |
| 7 | 7 | | | | 11 | 12 | 13 | 13 | 14 | 14 | 14 | 14 | 15 | 15 | 15 | | | | | |
| 8 | 8 | | | | 11 | 12 | 13 | 14 | 14 | 15 | 15 | 16 | 16 | 16 | 16 | 17 | 17 | 17 | 17 | 17 |
| 9 | 9 | | | | | 13 | 14 | 14 | 15 | 16 | 16 | 16 | 17 | 17 | 18 | 18 | 18 | 18 | 18 | 18 |
| 10 | 10 | | | | | 13 | 14 | 15 | 16 | 16 | 17 | 17 | 18 | 18 | 18 | 19 | 19 | 19 | 20 | 20 |
| 11 | 11 | | | | | 13 | 14 | 15 | 16 | 17 | 17 | 18 | 19 | 19 | 19 | 20 | 20 | 20 | 21 | 21 |
| 12 | 12 | | | | | 13 | 14 | 16 | 16 | 17 | 18 | 19 | 19 | 20 | 20 | 21 | 21 | 21 | 22 | 22 |
| 13 | 13 | | | | | | 15 | 16 | 17 | 18 | 19 | 19 | 20 | 20 | 21 | 21 | 22 | 22 | 23 | 23 |
| 14 | 14 | | | | | | 15 | 16 | 17 | 18 | 19 | 20 | 20 | 21 | 22 | 22 | 23 | 23 | 23 | 24 |
| 15 | 15 | | | | | | 15 | 16 | 18 | 18 | 19 | 20 | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 25 |
| 16 | 16 | | | | | | | 17 | 18 | 19 | 20 | 21 | 21 | 22 | 23 | 23 | 24 | 25 | 25 | 25 |
| 17 | 17 | | | | | | | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 23 | 24 | 25 | 25 | 26 | 26 |
| 18 | 18 | | | | | | | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 25 | 26 | 26 | 27 |
| 19 | 19 | | | | | | | 17 | 18 | 20 | 21 | 22 | 23 | 23 | 24 | 25 | 26 | 26 | 27 | 27 |
| 20 | 20 | | | | | | | 17 | 18 | 20 | 21 | 22 | 23 | 24 | 25 | 25 | 26 | 27 | 27 | 28 |

Références

- [1] Cochran, W.G. et Cox, G.M. (1950) : *Experimental design* John Wiley, New York.
- [2] Dagnélie P. (1970) : *Théorie et méthodes statistiques*. Les Presses Agronomiques de Gembloux.
- [3] Dervin, C. (1990) : *Comment interpréter les résultats d'une analyse factorielle des correspondances ?* ITCF
- [4] Draper, N. and Smith, H. (1981) : *Applied Regression Analysis (2nd Edition)*. John Wiley, New York.
- [5] Gomez, K.A. et Gomez, A.A. (1984) : *Statistical Procedures of Agricultural Research (2nd Edition)*. John Wiley, New York.
- [6] Gouet, J.P. et Philippeau, G. (1986) : *Comment interpréter les résultats d'une analyse de variance* ITCF
- [7] Mead, R. and Curnow, R.N. (1983) : *Statistical Methods in Agriculture and Experimental Biology*. Chapman and Hall, London.
- [8] Philippeau, G. (1986) : *Comment interpréter les résultats d'une analyse en composantes principales* ITCF
- [9] Saporta, G. (1978) : *Théories et méthodes de la statistique* Société des éditions Technip
- [10] Snedecor, G.W. (1956) : *Statistical methods* John Wiley, New York.
- [11] Steel, G.D. et Torrie, J.H. (1981) : *Principles and Procedures of Statistics : A Biometrical Approach (2nd Edition)*. McGraw-Hill, London.
- [12] Tomassone R. (1989) : *Comment interpréter les résultats d'une régression linéaire ?* ITCF
- [13] Tomassone R. (1988) : *Comment interpréter les résultats d'une analyse factorielle discriminante ?* ITCF
- [14] Vessereau A. (1960) : *Méthodes statistiques en biologie et en agronomie*. Nouvelle encyclopédie agricole.