

PIERRE THIONET

Sur la distribution exacte du X^2 de Pearson d'une table de contingence

Revue de statistique appliquée, tome 21, n° 4 (1973), p. 5-23

http://www.numdam.org/item?id=RSA_1973__21_4_5_0

© Société française de statistique, 1973, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

SUR LA DISTRIBUTION EXACTE DU X^2 DE PEARSON D'UNE TABLE DE CONTINGENCE

Pierre THIONET

Université de Paris-Dauphine

INTRODUCTION

L'étude directe de la distribution du critère X^2 de Karl PEARSON a fait l'objet ces dernières années d'une littérature abondante [cf. 2 à 13].

On sait que le dit critère est employé notamment comme test de la qualité de l'ajustement et comme test de contingence, nonobstant d'autres usages (test de symétrie, par exemple).

Déjà MANN et WALD [1, 2] avaient recommandé, pour tester la qualité de l'ajustement, un découpage du champ de la variable en k intervalles d'égales probabilités $1/k$, autant que faire se peut.

Une règle d'usage courant, qu'on trouve dans tous les manuels et cours, proscrie l'usage du X^2 s'il n'existe pas au moins *cinq* unités échantillon dans chaque intervalle ; ceci tendrait à ne pas constituer plus de $n/5$ intervalles pour un échantillon de taille n . A vrai dire, cette règle (dont personne n'a trouvé l'origine) tend à permettre la substitution à la distribution (inconnue) du critère X^2 la loi théorique du χ^2 (ki-carré) avec un nombre de degrés de liberté convenablement choisi : $k - I - h$, avec $h = 0$ pour une distribution librement choisie, $h = 1$ si (disons) la moyenne théorique est égale à la moyenne échantillon, etc. . . Ceci pour le test de qualité de l'ajustement. Des règles analogues existent (on y reviendra) pour le test de contingence.

La *règle des cinq unités* a pour justification les nombreuses réussites passées des tests du Ki-carré. Il est douteux que 5 soit un "grand nombre" au point de justifier l'approximation d'une variable binomiale ou multinomiale par une variable de Gauss-Laplace ("normale"), fondement théorique du Ki-carré. En somme, la technique courante réussit beaucoup mieux que la théorie ne le ferait prévoir, — peut-être faute d'une théorie plus fine.

Il a fallu toute l'autorité de COCHRAN [3], 1952, [4], 1954, pour qu'on remette un peu partout en question les *cinq unités* : Cochran préconise notamment la présence d'une unité échantillon unique dans les intervalles situés aux deux bouts de la distribution, intervalles recouvrant une probabilité d'environ 0,05 si le jugement est pris au niveau de signification 5 %, ou une probabilité d'environ 0,01 s'il doit être pris au niveau 1 %.

VESSEREAU [5] 1957-58, WISE [6] 1963-63, KEMPTHORNE [7] 1966 se prononcent avec plus ou moins de conviction pour une solution extrême : constituer $k = n$ intervalles, dotés chacun d'une probabilité proche de $1/k$. En fait ce n'est pas simple : d'après [13], Van Der WAERDEN [8] serait partisan de k compris entre $n/2$ et $n/3$; on ne lit rien de tel dans son manuel (édition Dunod, 1967 page 242) traduction française ; l'intérêt de choisir k grand est souligné, mais de façon fort nuancée. Tout dépend d'abord de la loi qu'on prétend avoir ajustée au commencement ; des résultats établis quand il s'agit d'une loi de GAUSS-LAPLACE ne sauraient être valables pour une loi quelconque. En outre, si l'on aborde le problème en supposant n déjà grand, on devra retrouver les résultats de MANN et WALD *présentés par WILLIAMS [2] 1950 : avec une loi "normale"

Test au niveau 5 %				Test au niveau 1 %			
$n = 200 \dots 500 \dots 1\ 000 \dots 2\ 000$				$n = 200 \dots 500 \dots 1\ 000 \dots 2\ 000$			
$k \text{ opt.} = 31$	45	59	78	$k \text{ opt.} = 27$	39	51	68

C'est donc en dessous de $n = 200$ qu'on se pose des problèmes, alors qu'on est mal armé pour les résoudre approximativement ; ainsi VESSEREAU étudie les moments du X^2 faute de pouvoir en saisir la distribution.

Dans les articles les plus récents, on a recours au calculateur électronique, de façon à établir :

soit de la distribution exacte d'un X^2 , les points 5 %, 1 %

soit les points 5 %, 1 % d'une distribution approchée obtenue par simulation (Monte-Carlo).

Cette fois, avec 10 000 épreuves indépendantes pour chaque jeu de paramètres, on peut admettre que les approximations sont bonnes [9, 10 11].

Une grande partie de ces résultats concernent le cas où $k = n$, la probabilité attachée à chaque intervalle ou case étant égale à $1/k$: autrement dit : *l'espérance mathématique du nombre d'observations par case est égale à 1.*

Objet de la présente étude :

Cherchant dans ces recherches un "créneau", on a remarqué que le cas du test de contingence d'un tableau $L \times C$ n'avait été traité que par Monte Carlo, en s'en tenant aux points 5 % et 1 %, par ROSCOE et BYARS [11] 1971 ; il s'agit d'un travail préparé à Kansas State University par deux psychosociologues (*professors of education*) ; on sait que cette corporation fait en pays anglosaxon un grand usage du Ki-carré.

En revanche, le test de qualité de l'ajustement (autrement dit le tableau à une seule ligne) avait reçu un traitement *exact* par ZAHN et ROBERTS [12] 1971. Voir aussi [13]. Nous avons donc attaqué "les mains nues" le problème oublié (apparemment) de la distribution *exacte* du X^2 d'une table $L \times C$ (à L lignes et C colonnes) dans le cas particulier (mais essentiel) d'une *espérance unitaire par case*.

Composition de ce travail

On a d'abord réglé aisément les cas de $L = 2, C = 2, 3, 4, \dots$. Un résultat valable quel que soit C est apparu.

Puis on a fait $L = 3, C = 3, 4, 5, 6$ (de plus en plus péniblement), enfin $L = C = 4$. Les résultats sont rassemblés dans les tableaux ci-après.

La méthode (mise au point progressivement) sera exposée sur le cas 4×4 ; après quoi on reviendra sur les cas plus simples.

1^{ère} PARTIE – LES TABLEAUX 4×4

Le cas 4×4 : Détails de la méthode :

On se propose de construire (et dénombrer) tous les tableaux possibles ; au total

$$16! / 4! 4! 4! 4! = \underline{63\,063\,000}$$

c'est là le nombre de dispositions différentes qu'on peut obtenir avec 16 objets distincts partagés en 4 lots de 4 (nombre de quadripartitions).

Supposons (une fois pour toutes) les 4 lots placés au pied des 4 colonnes du tableau 4×4 ; considérons l'une de ces colonnes : on peut :

soit placer un objet dans chaque case (de la colonne) : de 24 façons ;
 soit deux objets dans l'une, un dans deux cases, zéro dans la dernière : 12 choix ;
 soit deux objets dans l'une, deux dans une autre, rien ailleurs : 6 choix ;
 soit trois objets dans l'une, un dans une autre case, rien ailleurs : 4 choix ;
 enfin les 4 objets dans une seule case, rien ailleurs : 1 seul cas.

Nous désignerons par f_i ce nombre 24, 12, 6, 4 ou 1 défini sur chaque colonne $i = 1, 2, 3, 4$. Soit f le produit des f_i des 4 colonnes ; f (compris entre $24^4 = 331\,776$ et $1^4 = 1$) est défini, pour chaque tableau T défini comme suit :

Un tableau T est formé de 16 nombres entiers 0, 1, 2, 3 ou 4, disposés sur l'échiquier vide 4×4 de telle sorte que toutes les sommes de ligne et de colonne soient égales à 4.

Il reste à construire les tableaux T.

*Déf: Symboles**: A tout tableau T correspond un symbole $s = \underline{4^a 3^b 2^c 1^d 0^e}$, où :

a, b, c, d, e sont eux-mêmes des entiers 0, 1, 2, 3, 4 ;
 respectivement : le nombre de fois que 4 figure dans T est a,
 " " " " " 3 " " T " b,
 etc . . .

par suite : $\begin{cases} a + b + c + d + e = 16 \text{ (cases)} \\ \text{(i) } 4a + 3b + 2c + d = 16 \text{ (total général de T)} \end{cases}$

Définition du X^2 :

$$\text{car } (4 - 1)^2 = 9, (3 - 1)^2 = 4, (2 - 1)^2 = 1, (0 - 1)^2 = 1$$

$$\frac{9a + 4b + c + e = X^2}{}$$

* Notation employé en théorie des nombres par exemple.

Remarque : les conditions (i) sont nécessaires, non suffisantes pour qu'à un symbole s corresponde bien au moins un tableau T ; car elles n'expriment pas l'équilibre de chaque colonne (ligne), simplement celui du total général.

Déf. : Opération 0 : Considérons l'ensemble d'opérations suivant sur tableaux :

- permutation quelconque des lignes ;
- permutation quelconque des colonnes ;
- transposition lignes colonnes.

Il est clair que toute opération 0 conserve le X^2

Déf. : Type de tableau T : Partant d'un tableau T donné, les tableaux *de même type* sont la classe d'équivalence constituée par tout tableau T' qu'on peut déduire de T par une opération 0.

Proposition : Les tableaux d'un même type ont *même* X^2 et même *symbole*.

Proposition : $\left\{ \begin{array}{l} \text{La même valeur de } X^2 \text{ peut correspondre à plusieurs symboles,} \\ \text{le même symbole peut correspondre à plusieurs types de tableaux.} \end{array} \right.$

Conséquence : Pour un type donné T de tableau, le nombre n de tableaux distincts concernés, que fournissent les opérations 0, est calculable (avec un peu de méthode).

En revanche, pour obtenir toute la collection de types T, on en est réduit à les *construire sans omission* (et à éliminer les redondances).

1^{er} *Exemple* : $s = \underline{2^2 \ 1^{12} \ 0^2}$

$$T = \begin{array}{cccc} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array}$$

C'est le premier tableau ($X^2 = 4$) au delà de 1^{16} ($X^2 = 0$)

Ce tableau T existe ; son symbole est s. Quel est son nombre n ?

Réponse $n = 72$; $f = 12 \times 12 \times 24 \times 24$.

Justifions la valeur de n :

Il y a 6 partitions différentes (2,2) des lignes, 6 des colonnes, donc 36 bi-partitions du tableau. Par exemple celle ci-dessus ; mais le contenu de l'angle Nord-Ouest peut *aussi* être $\begin{array}{cc} 0 & 2 \\ 2 & 0 \end{array}$ (par mutation des 2 premières lignes)

$$\begin{array}{cc} 0 & 2 \\ 2 & 0 \end{array}$$

D'où $n = 36 \times 2$. A noter qu'il s'avère impossible de construire un tableau de symbole s qui ne soit pas du type T précédent ;

2^{ème} *Exemple* :

Passant à $s = \underline{2^3 \ 1^{10} \ 0^3}$, c'est-à-dire $T = \begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 2 \\ 1 & 0 & 2 & 1 \end{array}$

avec $X^2 = 6$

on a : $f = 24 \times 12^3$

$n = 4 \times 4 \times 12$

Preuve :

4 choix pour la ligne des 1, 4 pour la colonne des 1, 12 *tableaux* 3×3 du modèle obtenu en barrant 1^{ère} ligne et 1^{ère} colonne (cf. Tableaux 3×3)
Nota : Là encore, peut-il y avoir *plusieurs types* ?

Pour les tableaux 3×3 , distinguer le type du symbole ne s'imposait guère, car le dénombrement direct est facile : 6 façons de placer les 2 ; 2 places en première ligne pour le 1 (tout le reste est imposé).

On pourrait penser que la factorisation : $2^3 1^{10} 0^3 = 1^8 (2^3 1^2 0^3)$ fournirait un second type ; il n'en est rien : le symbole de tableau 2×4 : $2^3 1^2 0^3$ ne représente aucun tableau valable connu.

Nous n'avons donc trouvé qu'un type de tableau (soit : $1^7 (2^3 1^3 0^3)$)

3^{ème} *Exemple* :

Vient $s = 2^4 1^8 0^4$ (avec $X^2 = 8$)

type 1 : $2 \ 1 \ 1 \ 0 \quad f = 12^4$

$0 \ 2 \ 1 \ 1 \quad n = 216 \quad$ (résultat pour tout tableau

$1 \ 0 \ 2 \ 1 \quad = 24 \times 9 \quad$ (a b b c)

$1 \ 1 \ 0 \ 2$

type 2 : $2 \ 1 \ 0 \ 1$

$2 \ 0 \ 1 \ 1 : n = 576$

$0 \ 2 \ 1 \ 1 \quad = 4 \times 2 \times n'$

$0 \ 1 \ 2 \ 1$

type 3 : $2 \ 0 \ 1 \ 1$

$0 \ 2 \ 1 \ 1 : n = 72$

$2 \ 0 \ 1 \ 1 \quad = 6 \times 2 \times n''$

$0 \ 2 \ 1 \ 1$

n' se réfère au tableau 4×3

(supprimer la 4^{ème} colonne)

n'' se réfère au tableau 4×2

(supprimer les 3^{ème} et 4^{ème} colonnes)

Le type 1 mérite une explication : il y a 24 façons de placer les 2 (les a) mais c'est un dénombrement détaillé qui fait apparaître les 9 cas existant chaque fois pour placer les autres chiffres (ou lettres). (cf. Annexe II)

La liste complète des *types* est donnée en Annexe I.

La Table ci-après donne les résultats finals : distribution de X^2 ; celle du ki-carré à 9 degrés de liberté a été placée en regard, autant que faire se peut pour une confrontation. Elles diffèrent notablement en dessous de $X^2 = 18$ mais, entre 5 % et 1 %, les probabilités cumulées sont très voisines : c'est dire que *les tests courants ne souffrent guère du passage de l'une à l'autre*.

On a calculé $E(X^4)$, trouvé égal à 146,3 au lieu de 99 pour le ki-carré ; on suppose qu'une erreur de calcul se sera glissée.

A toutes fins utiles, une seconde table résumant le calcul de la distribution pour chaque type a aussi été reproduite. (Table II).

Table I

Tableaux 4 x 4	nf	X ²		X ²	
		Cumulés	%	%	χ ²
X ² = 0	331 776	63 063 000	100	100	0
4	5 971 968	62 731 224	99.5	99	2.088
6	7 962 624	56 759 256	90.0	90	4.168
8	17 915 904	48 796 632	77.4	80	5.380
10	9 953 280	30 880 728	49.0	50	8.343
12	10 948 608	20 927 448	33.2	30	10.656
14	3 981 312	9 978 840	15.8	20	12.242
16	3 766 176	5 997 528	9.5	10	14.684
18	1 308 672	2 231 352	3.5	5	16.919
20	608 256	922 680	1.5	2	19,679
24	241 920	314 424	0.50	1	21.666
26	55 296	72 504	0.115	0.1	27.877
30	12 288	17 208	0.027		
32	2 592	4 920	0.008		
36	2 304	2 328	0.004		
48	24	24	0.000		

Table II

Tableau 4 x 4 (détail)					
Symbole	X ²	nf	f	n	Référence
1 ¹⁶	0	331.776	24 ⁴	1	a
2 ² 1 ¹² 0 ²	4	5 971.968	24 ² x 12 ²	6 ² x 2	b
2 ³ 1 ¹⁰ 0 ³	6	7 962.624	24 x 12 ³	4 ² x 12	c
2 ⁴ 1 ⁸ 0 ⁴	8	17 915.904	12 ⁴	24 x 9 x 4 (3 types)	d' c'' d'''
2 ⁵ 1 ⁶ 0 ⁵	10	5 971.968	12 ³ x 6	6 ² x 4 x 2 x 2 (2 types)	e' e''
2 ⁶ 1 ⁴ 0 ⁶	12	2 985.984	12 ² x 6 ²	6 ² x 2 ³ x 2 (2 types)	f' f''
2 ⁸ 0 ⁸	16	116.690	6 ⁴	6 x 15	g
3 ¹ 2 ² 1 ⁹ 0 ⁴	10	3 981.312	24 x 12 ² x 4	4 ² x 18	h
3 ¹ 2 ³ 1 ⁷ 0 ⁵	12	7 962.624	12 ³ x 4	4 x 72 x 2 + 4 ² x 3 ² x 4 (types)	i' i''
3 ¹ 2 ⁴ 1 ⁵ 0 ⁶	14	3 981.312	12 ² x 6 x 4	4 ² x 3 ² x 4 x 2 (2 types)	j' j''
3 ¹ 2 ⁵ 1 ³ 0 ⁷	16	995.328	12 x 6 ² x 4	4 ² x 3 ² x 2 ²	k
3 ² 2 ² 1 ⁶ 0 ⁶	16	2 654.208	12 ² x 4 ²	6 ² x 2 ⁴ + 4 x 72 x 2 (2 types)	l' l''
3 ² 2 ³ 1 ⁴ 0 ⁷	18	995.328	12 x 6 x 4 ²	6 ² x 2 x 4 + 12 ² x 4 (2 types)	m' m''
3 ² 2 ⁴ 1 ² 0 ⁸	20	41.472	6 ² x 4 ²	6 ² x 2	n
3 ³ 1 ⁷ 0 ⁶	18	147.456	24 x 4 ³	4 ² x 6	o
3 ³ 2 ¹ 1 ⁵ 0 ⁷	20	442.368	12 x 4 ³	4 ² x 3 ² x 4	p
3 ⁴ 1 ⁴ 0 ⁸	24	55.296	4 ⁴	24 x 9	q
4 ¹ 2 ³ 1 ⁶ 0 ⁶	18	165.888	12 ³ x 1	4 ² x 6	r
4 ¹ 2 ⁴ 1 ⁴ 0 ⁷	20	124.416	12 ² x 6 x 1	4 ² x 3 ²	s
4 ¹ 2 ⁶ 0 ⁹	24	20.736	6 ³ x 1	4 ² x 6	t
4 ¹ 3 ¹ 2 ³ 1 ³ 0 ⁸	24	165.888	12 x 6 x 4 x 1	4 ² x 3 ² x 2 ²	u
4 ¹ 3 ² 2 ¹ 1 ⁴ 0 ⁸	26	55.296	12 x 4 ² x 1	4 ² x 12	v
4 ¹ 3 ³ 1 ³ 0 ⁹	30	12.288	4 ³ x 1	4 ² x 12	w
4 ² 2 ⁴ 0 ¹⁰	32	2.592	6 ² x 1 ²	72	x
4 ² 3 ² 1 ² 0 ¹⁰	36	2.304	4 ² x 1 ²	144	y
4 ⁴ 0 ¹²	48	24	1 ⁴	24	z

Moyens de contrôle :

Le plus simple est constitué par le nombre total de tableaux (63 063 000) qui doit être retrouvé finalement comme total général des $n \times f$ de chaque prototype ; tant que ce total n'est pas dépassé, on doit avoir oublié quelques types ; si on le dépasse, on a du faire quelque double compte en calculant les n . Ceci nous fût très utile, pour 4×4 mais aussi 3×5 et 3×6 .

9! /3! 3! 3!	= 1 680	total général pour le cas 3×3
12! /4! 4! 4!	= 34 650	" " " " " 3×4
15! /5! 5! 5!	= 756 756	" " " " " 3×5
18! /6! 6! 6!	= 17 153 136	" " " " " 3×6

Procédure de calcul

Une fiche est établie pour chaque type de tableau, portant le symbole, la valeur de X^2 , celle de f , celle de n , de $f \times n$.

Les fiches de même symbole sont épinglées.

Les fiches sont établies dans l'ordre (lexicographique) des abcde.

Un tri des fiches sur la valeur de X^2 est opéré ; on totalise les nf de même X^2 .

Enfin les totaux cumulés des Σnf sont effectués dans l'ordre des X^2 décroissants.

Nombre de tableaux :

Chaque tableau T est un tableau de Fréchet (dont les marges sont données).

Le nombre de tableaux distincts peut en principe être calculé par une formule (en fait compliquée) donnée par Guisepe LETI [14] 1970. Il semble que ce soit la somme des n étendue à tous les T . On n'en a pas l'usage ici, sinon comme moyen de contrôle des n (ou de la complétude de l'ensemble des T).

Valeurs de f :

Il est remarquable que, pour un symbole donné, les divers types possèdent le même f ; or qui dit types différents, dit non-identité du contenu des colonnes (lignes), non-invariance dans les opérations 0. D'ailleurs il est déjà remarquable que fn calculé sur les colonnes coïncide avec celui calculé sur les lignes.

2^{ème} PARTIE – TABLEAUX $2 \times C$ ou $L \times 2$

Tableaux 2×2 :

Il existe 2 types de tableaux : $\begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix}$ & $\begin{matrix} 2 & 0 \\ 0 & 2 \end{matrix}$ ou $\begin{matrix} 0 & 2 \\ 2 & 0 \end{matrix}$

Soit x y z t quatre objets distincts ; on peut les ranger ainsi en tableau :
 6 partitions : xy, yt ; xz, yt ; xt, yz ; zt, xy ; yt, xz ; yz, xt.

pour chacune,

4 dispositions $\begin{pmatrix} x & z \\ y & t \end{pmatrix}$;
 du 1^{er} type $\begin{pmatrix} x & t \\ y & z \end{pmatrix}$;
 $\begin{pmatrix} y & z \\ x & t \end{pmatrix}$;
 $\begin{pmatrix} y & t \\ x & z \end{pmatrix}$;

contre : 2 dispositions
 du 2^e type $\begin{pmatrix} xy & - \\ - & zt \end{pmatrix}$;
 $\begin{pmatrix} - & zt \\ xy & - \end{pmatrix}$;

Distribution de probabilités :

Nous conviendrons que :

1⁴ autrement dit le type $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ a la probabilité 4/6
 2² 0² " " " " $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ a la probabilité 2/6

Tableaux 2 x 3 :

6 objets possèdent 6!/2! 2! 2! = 90 tripartitions (2 2 2). Ne sont possibles que les tableaux (de marge verticale (3 3) suivants :

$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$; $\begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}$ et analogues ; il n'y a que 6!/3! 3! = 20 bipartitions 33

Pour une tripartition donnée horizontale (xy zt uv), il existe 2³ dispositions du 1^{er} de ces 2 types, il en existe 12 du second : 3 choix de la colonne des

uns, 2 permutations chaque fois, choix entre $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ et $\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$ pour le reste. Total :
8 + 12 + = 20

1⁶ (autrement dit le type $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$) a la probabilité 8/20
 2² 1² 0² (" " " " $\begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}$ " " " 12/20

Tableaux 2 x 4 :

8!/4! 4! = 70 = C₈⁴ cas possibles ; 3 types de tableaux 1⁸ : 2⁴ = 16 cas (16/70) ;

2² 1⁴ 0² et 2⁴ 0⁴

K pair : $K = 2k$

$$C_{4k}^{2k} = C_{2k}^{2k} C_{2k}^k + C_{2k}^{2k-2} C_{2k-2}^{k-1} \times 2^2 + \dots + C_{2k}^{2(k-j)} C_{2(k-j)}^{k-j} \times 2^{2j} + \dots + 2^{2k}$$

$$= \sum_{j=0}^k \frac{2^{2j} (2k)!}{(2j)! (k-j)! (k-j)!}$$

K impair : $K = 2k + 1$

$$C_{4k+2}^{2k+1} = C_{2k+1}^{2k} C_{2k}^k \times 2 + C_{2k+1}^{2k-2} C_{2k-2}^{k-1} \times 2^3 + \dots + C_{2k+1}^{2(k-j)} C_{2(k-j)}^{k-j} \times 2^{2j+1}$$

$$\dots + 2^{2k+1} = \sum_{j=0}^k \frac{2^{2j+1} (2k+1)!}{(2j+1)! (k-j)! (k-j)!}$$

Il existe vraisemblablement des démonstrations directes de ces formules, mais un auditeur (dont le nom nous échappe) a bien voulu nous signaler que le dénombrement des trajets équiprobables, dans le problème du scrutin de Désiré ANDRE conduisait aux mêmes formules. Ayant nous-même étudié antérieurement ce problème (Revue de l'Inst. Intern. Statis. 36 : 3, 1968, p. 296-308), nous avons pensé à sa fonction génératrice, mais avons été arrêté par la condition de convergence $|t| < 1/4$, quand on aurait voulu s'en servir (vraisemblablement pour $t = 1/4$)

De même, le dénombrement des tableaux $3 \times C$ (plus généralement $L \times C$), suivant que $C = 3k$, ou $3k + 1$, ou $3k + 2$ donnerait lieu à des développements de

$$\frac{(3C)!}{(C!)^3} \text{ et analogues,}$$

qu'il n'est pas aussi facile d'écrire explicitement.

3^{ème} PARTIE – LES TABLEAUX $3 \times C$

3×3	Symbole	X^2	Type	f	n	nf	Cum	%
	$3^3 0^6$	18	$\begin{matrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{matrix} = A$	1	6	6	6	0,4
	$3^1 2^2 1^2 0^4$	10	$\begin{matrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{matrix} = B$	3×3	$3^2 \times 2$	162	168	10,0
	$2^3 1^3 0^3$	6	$\begin{matrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{matrix} = C$	$3 \times 3 \times 3$	6×2	324	492	29,3
	$2^2 1^5 0^2$	4	$\begin{matrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{matrix} = (HI)$	$6 \times 3 \times 3$	$3^2 \times 2$	972	11 464	87,2
	1^9	0	$\begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix} = (I^3)$	$6 \times 6 \times 6$	1	216	1 680	100
						1 680		

Ici, il n'y a qu'un type par symbole. Comparons au ki-carré :

$$E(X^2) = 4,50 \text{ contre } E(x^2) = 4 \text{ (4 degrés de liberté)}$$

$$V(X^2) = 6,75 \text{ contre } V(x^2) = 8$$

$$E(X^4) = 27 \text{ contre } E(x^4) = 24$$

<u>3 x 4</u>	Symbole	X ²	Types	f	n	nf	Cum
	3 ³ 1 ³ 0 ⁶	18	AI = $\begin{matrix} 3 & 0 & 0 & 1 \\ 0 & 3 & 0 & 1 \\ 0 & 0 & 3 & 1 \end{matrix}$	6	24	144	144
	3 ² 2 ² 1 ² 0 ⁶	16	D = $\begin{matrix} 3 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 1 & 3 \end{matrix}$	3 ²	72	648	792
	3 ¹ 2 ³ 1 ³ 0 ⁵	12	D* = $\begin{matrix} 3 & 1 & 0 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 2 & 2 \end{matrix}$	3 ³	72	1 944	2 736
	3 ¹ 2 ² 1 ⁵ 0 ⁴	10	(BI) = $\begin{matrix} 3 & 0 & 0 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \end{matrix}$	3 ² 6	72	3 888	6 624
	2 ⁴ 1 ⁴ 0 ⁴	8	$\left. \begin{matrix} 2 & 0 & 2 & 0 \\ (H^2) = 0 & 2 & 0 & 2 \\ 1 & 1 & 1 & 1 \\ (H^2)' = 0 & 2 & 0 & 2 \\ 1 & 1 & 2 & 0 \end{matrix} \right\}$	3 ⁴ 3 ⁴	18 72	7 290	13 914
	2 ³ 1 ⁶ 0 ³	6	(CI) = $\begin{matrix} 2 & 1 & 0 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 2 & 1 \end{matrix}$	3 ³ 6	48	7 776	21 690
	2 ² 1 ⁸ 0 ²	4	(HI ²) = $\begin{matrix} 2 & 0 & 1 & 1 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{matrix}$	3 ² 6 ²	36	11 664	33 354
Enfin	1 ¹²	0	(I ⁴) = $\begin{matrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{matrix}$	6 ⁴	1	1 296	34 650

Ici il y a deux types pour les tableaux baptisés (H²). Comparons au ki-carré :

$$E(X^2) = 6,314 \text{ contre } E(\chi^2) = 6 \text{ (6 degrés de liberté)}$$

$$V(X^2) = 8,8 \text{ contre } V(\chi^2) = 12$$

$$E(X^4) = 44,9 \text{ contre } E(\chi^4) = 48$$

La valeur de n (pour alléger) n'est plus indiquée à partir d'ici

<u>3 x 5</u>	Symbole	X ²	Types	f	nf	Cum
	3 ³ 2 ² 1 ² 0 ⁸	22	(AH) = $\begin{matrix} 3 & 0 & 0 & 1 & 1 \\ 0 & 3 & 0 & 2 & 0 \\ 0 & 0 & 3 & 0 & 2 \end{matrix}$	3 ²	3 240	3 240
	3 ³ 1 ⁶ 0 ⁶	18	(AI ²) = $\begin{matrix} 3 & 0 & 0 & 1 & 1 \\ 0 & 3 & 0 & 1 & 1 \\ 0 & 0 & 3 & 1 & 1 \end{matrix}$	6 ²	2 160	5 400
	3 ² 2 ³ 1 ³ 0 ⁷	18	E = $\begin{matrix} 3 & 0 & 2 & 0 & 0 \\ 0 & 3 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 & 2 \end{matrix}$	3 ³	9 720	15 120
	3 ² 2 ² 1 ⁵ 0 ⁶	16	(DI) = $\begin{matrix} 3 & 1 & 0 & 0 & 1 \\ 0 & 2 & 2 & 0 & 1 \\ 0 & 0 & 1 & 3 & 1 \end{matrix}$	3 ² 6	19 440	34 560
	3 ¹ 2 ⁴ 1 ⁴ 0 ⁶	14	(BH) = $\begin{matrix} 3 & 0 & 0 & 2 & 0 \\ 0 & 2 & 1 & 0 & 2 \\ 0 & 1 & 2 & 1 & 1 \end{matrix}$	3 ⁴		
	(2 types)				58 320	92 880
		14	(BH)' = $\begin{matrix} 3 & 0 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 & 2 \\ 0 & 1 & 2 & 2 & 0 \end{matrix}$	3 ⁴		
	3 ¹ 2 ³ 1 ⁶ 0 ⁵	12	(D*I) = $\begin{matrix} 3 & 1 & 0 & 0 & 1 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 & 1 \end{matrix}$	3 ³ 6	58 320	151 200

Symbole	X ²	Types	f	nf	Cum
3 ¹ 2 ² 1 ⁸ 0 ⁴	10	(BI ²) = 3 0 0 1 1 0 3 0 1 1 0 0 3 1 1	6 ²	58 320	209 520
2 ⁵ 1 ⁵ 0 ⁵	10	(CH) = 2 1 0 1 1 0 2 1 2 0 1 0 2 0 2	3 ⁵	87 480	297 000
2 ⁴ 1 ⁷ 0 ⁴	8	(H ² I) = 2 0 2 0 1 0 2 0 2 1 1 1 1 1 1	3 ⁴ 6	218 700	515 700
(2 types) (comme H ²)	8	(H ² I) ['] = 2 0 1 1 1 1 1 2 0 1 0 2 0 2 1	3 ⁴ 6		
2 ³ 1 ⁹ 0 ³	6	(CI ²) = 2 1 0 1 1 0 2 1 1 1 1 0 2 1 1	3 ³ 6 ²	116 640	632 340
2 ² 1 ¹¹ 0 ²	4	(HI ³) = 2 0 1 1 1 0 2 1 1 1 1 1 1 1 1	3 ² 6 ³	116 640	748 980
1 ¹⁵	0	(I ⁵) = 1 1 1 1 1 1 1 1 1 1	6 ⁵	7 776	756 756

Outre le (H²I), on a deux types (BH) selon que la ligne (11) de H est ou non au regard du (3 0 0) de B.
Comparons au ki-carré :

$$E(X^2) = 8,57 \text{ contre } E(\chi^2) = 8 \text{ (degrés de liberté)}$$

$$V(X^2) = 12,3 \text{ contre } V(\chi^2) = 16$$

$$E(X^4) = 85 \text{ contre } E(\chi^4) = 80$$

3 x 6	Symboles	X ²	Types	nf	Cum
	3 ⁶ 0 ¹²	36	(A ²) évident	90	90
	3 ⁴ 2 ² 1 ² 0 ¹⁰	28	(AB) évident	9 720	9 810
	3 ³ 2 ³ 1 ³ 0 ⁹	24	(AC) = 3 0 0 2 1 0 3 0 2 1 0 0 3 1 0 2	38 880	
	(deux types)				
	idem	F =	3 3 0 0 0 0 0 0 3 1 1 1 0 0 0 2 2 2	9 720	58 410
	3 ³ 2 ² 1 ⁵ 0 ⁸	22	(AHI) évident	116 640	175 050
	3 ² 2 ⁴ 1 ⁴ 0 ⁸	20	(B ²) 3 0 0 3 0 0 & 3 0 0 0 2 1 0 2 1 0 2 1 & 0 2 1 3 0 0 0 1 2 0 1 2 & 0 1 2 0 1 2	21 870	
	(2 types) = (3 2 ² 1 ² 0 ⁴) ²				
	(2 types) = (2 ² 1 ² 0 ²). (3 ² 2 ² 1 ² 0 ⁶) (DH)				
			3 1 0 0 1 1 & 3 1 0 0 2 0 0 2 2 0 0 2 & 0 2 2 0 1 1 0 0 1 3 2 0 & 0 0 1 3 0 2	174 960	546 840
	3 ³ 1 ⁹ 0 ⁶	18	(AI ³) évident	25 920	
	(2 types)	idem	(EI) évident	349 920	922 680
	3 ² 2 ² 1 ⁸ 0 ⁶	16	(DI ²) évident	349 920	
	3 ¹ 2 ⁵ 1 ⁵ 0 ⁷	16	(BC) 3 0 0 2 1 0 0 2 1 0 2 1 0 1 2 1 0 2	524 880	
	= BC = D*H		(D*H) (trois types)	437 400	2 234 880

Symbole	X ²	Types	nf	Cum
		3 0 0 1 1 1		
		0 2 1 1 0 2		
		0 0 2 2 2 0		
		3 0 0 1 0 2		
		0 2 1 1 1 1		
		0 0 2 2 2 0		
		3 0 0 1 0 2		
		0 2 1 : 2 0		
		0 0 2 2 1 1		
3 ¹ 2 ⁴ 1 ⁷ 0 ⁶		(BHI) voir BH	2 099 520	4 334 400
	14	(2 types)		
2 ⁶ 1 ⁶ 0 ⁶		type		
	12	1 ^{ère} ligne de <i>uns</i>	1 487 160	
(= C ² = H ³)		type		
		1 ligne avec 4 <i>uns</i>		
		type		
		chaque ligne a 2 <i>uns</i>		
3 ¹ 2 ³ 1 ⁹ 0 ⁵	12	(D* I ²)	1 049 760	6 871 320
3 ¹ 2 ² 1 ¹¹ 0 ⁴	10	(B I ³)	699 840	
2 ⁵ 1 ⁸ 0 ⁵	10	(CHI)	3 141 280	10 720 440
2 ⁴ 1 ¹⁰ 0 ⁴	8	(H ² I ²)	3 936 600	14 657 040
2 ³ 1 ¹² 0 ³	6	(CI ³)	1 399 680	16 056 720
2 ² 1 ¹⁴ 0 ²	4	(HI ⁴)	1 049 760	17 106 480
1 ¹⁸	0	(I ⁶)	46 656	17 153 136

Comparaison avec le ki-carré ; On pourrait calculer E(X²) et V(X²) et les comparer respectivement à 10 et à 20.

Comparaison des deux queues de distribution (du côté des grands X²)

On donne ci-après les graphiques relatifs aux cas 3 x 5 et 3 x 6. L'accord est en fait *très satisfaisant*. D'ailleurs, pour toute valeur de X², l'accord n'est pas mauvais pour 3 x 6 (mais pas pour 3 x 5).

X ²	0	4	8	10	12	14	16	18	20	22	24	28	36
%	99,7	93,6	85,4	40	25	13	5,4	3,2	1,0	0,34	0,06	0,00	0
%	95	90	70	50	30	20	10	5		2	1	0,1	
X ² ₍₁₀₎	3,94	4,9	7,3	9,3	11,8	13,4	16	18,3		21,2	23,2	29,6	

OBSERVATIONS FINALES

I – La désignation figurative des tableaux dans les cas 2 x C et 3 x C

En dehors du cas 4 x 4, nous avons conservé des notations particulières très claires (et dangereuses) désignant des types ou (mieux) des *genres* de tableaux sans qu'il soit nécessaire d'en écrire le détail exact ; en somme les "armes parlantes" d'un tableau. Au début, pour les cas 2 x C ou L x 2, les types

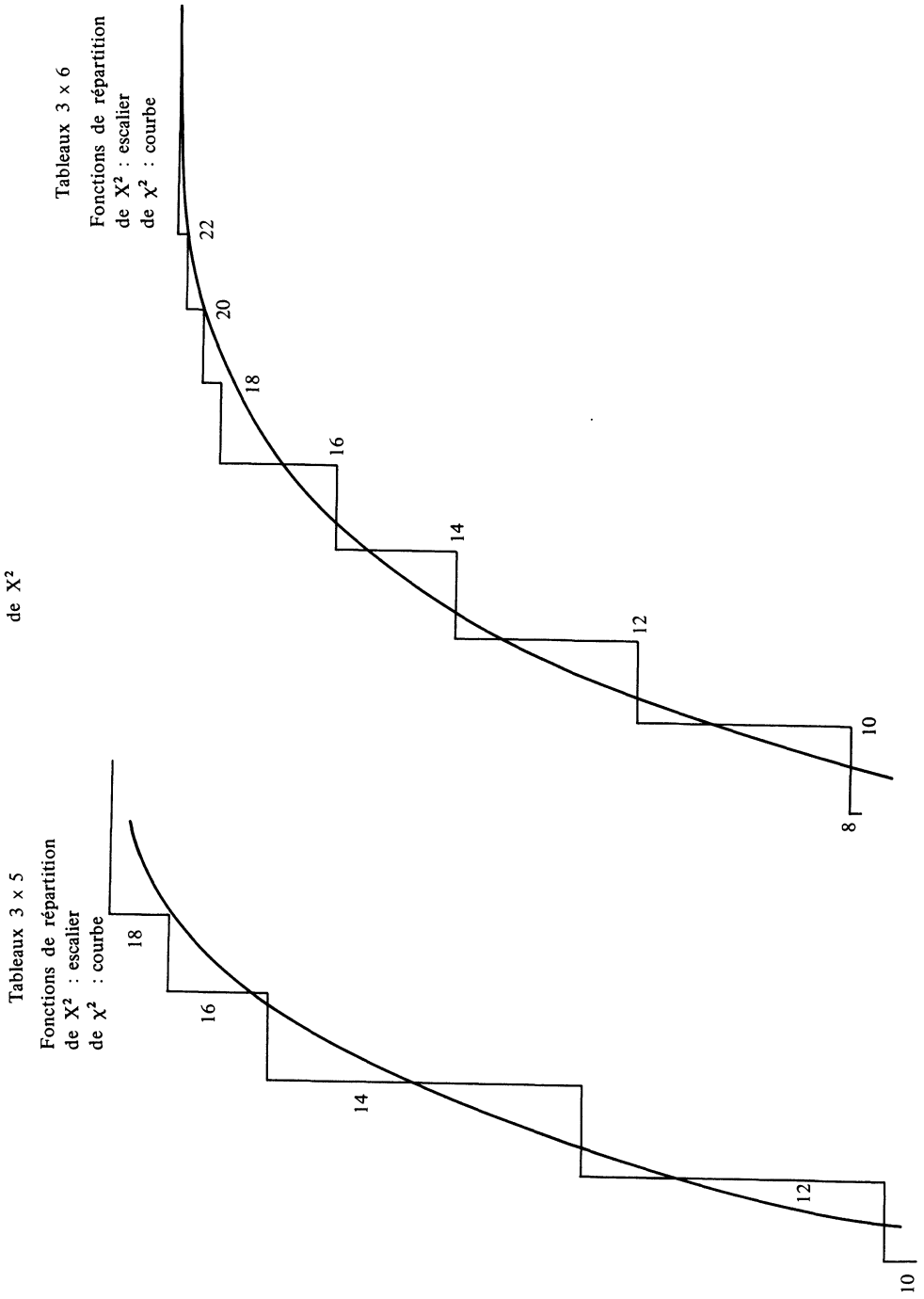
$$\begin{array}{ccc} 1 & 1 & \\ & & 2 & 0 \\ 1 & 1 & & \& & 0 & 2 \end{array}$$

ne posaient aucun problème. La mémoire est soulagée par la convention :

$$\begin{array}{ccc} 0 & 2 & 1 \\ H = 2 & 0 & I = 1 \\ 1 & 1 & 1 \end{array}$$

Ayant défini (à propos de 3 x 3) les types A, B, C, auxquels se joignent HI et I³ (de sens évident), on en a déduit une *construction* de tableaux 3 x 4

— Figure 1 —
 Comparaison pour
 les grandes valeurs
 de χ^2



et un dénombrement, dont les failles sont de deux sortes :

- existence des tableaux D et D* non obtenus ainsi ;
- erreurs dans le dénombrement, en raison des *doubles comptes* ; s'il est bien exact que $n(P \cdot Q) = \lambda n(P) \times n(Q)$ pour beaucoup de tableaux P et Q quand on les juxtapose (au même titre que $f(P \cdot Q) = f(P) \times f(Q)$), ceci est inexact en général. Cette algèbre embryonnaire est assez trompeuse, car par exemple H^2 désigne :

$$\begin{array}{cc} 0 & 2 & 0 & 2 & & 0 & 2 & 1 & 1 \\ \text{soit :} & 2 & 0 & 2 & 0 & , & \text{soit :} & 2 & 0 & 2 & 0 \\ & 1 & 1 & 1 & 1 & & 1 & 1 & 0 & 2 \end{array}$$

Si le nombre des H (désigné par $n(H)$) est manifestement 6, celui des H^2 n'est pas $6^2 = 36$; ceux du type de gauche sont déjà : 18.

3 choix d'une ligne (1 1 1 1) ; 6 partitions de la première des autres lignes entre le signe 0 et le signe 2 (la composition de la 3^{ème} ligne s'en déduit) ; enfin il y a (pour le type de droite) 3 choix de la ligne (2 200) et 6 partitions de cette ligne, 18 choix au total ; or je vois 4 façons distinctes d'écrire les deux autres lignes (à supposer que l'on ait (2 2 0 0) à la première :

$$\begin{array}{cccc} 1 & 0 & 2 & 1 & 1 & 0 & 1 & 2 & 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 2 & 0 & 1 & 2 & 1 & 1 & 0 & 1 & 2 \end{array} \text{ soit } 72 \text{ cas du type de droite}$$

Au total $n(H^2) = 18 + 72 = 90$

Dans le cas où il n'y a pas de doubles comptes, si P est $L \times p$ et Q est $L \times q$ le facteur $\lambda = C_{p+q}^p$ (partition des $p + q$ colonnes entre celles de P et de Q) est correct.

Exemple : (AC) : $n(A) = 6$; $n(C) = 12$; $C_6^3 = 20$; $n(AC) = 20 \times 6 \times 12 = 1440$
avec $f = 27$, on trouve bien : $nf = 38880$

Dans le cas de (AB), on trouve ainsi un $n(AB)$ deux fois trop grand :

$$n(A) = 6 ; n(B) = 18 ; C_6^3 = 20 ; 20 \times 6 \times 18 = 19440 = 9720 \times 2$$

La cause en est l'existence de 2 colonnes identiques comprenant un 3 et deux 0, l'une dans A et l'autre dans B.

Dans le cas de H^2 , le cas de gauche (une ligne de *uns*) était caractérisé par une factorisation horizontale en un tableau 2×4 et un 1×4 .

La double factorisation est le cas le plus décourageant : C^2 et H^3 ne sont pas distincts (encore qu'on puisse douter qu'ils soient confondus) ; de même BC et D^*H .

Néanmoins, il faut noter que, passant de 3×5 à 3×6 , le seul tableau nouveau (le seul *premier* si l'on accepte - avec prudence - l'analogie arithmétique) est F.

Il semble raisonnable de conjecturer qu'on a épuisé ainsi les types de $3 \times K$, et exprimer l'idée que la distribution du X^2 pour $K = 3k, 3k + 1, 3k + 2$ est

à présent (en théorie) calculable sans intervention de nouveaux symboles. Ceci devrait (bien entendu) être discuté.

On n'a pas su résoudre le cas 4×4 au moyen de cette algèbre, qui n'est pas sans analogie avec celle des polymoments et polycumulants. On a su s'en passer. Mais notons aussi qu'on n'a pas essayé de traiter le cas $4 \times K$, $K > 4$; on a cherché avant tout, alors, une technique susceptible d'être programmée.

II – Sur la contingence et les critères de contingence

Le X^2 n'est jamais que le plus banal des critères de contingence. On a utilisé ces dernières années le critère inspiré par la *divergence* J de Küllback-Leibner.

Dans le cas précis où nous nous sommes placé, il s'agit de la somme des

$$(n_{ij} - 1) \text{Log} (n_{ij}/1) = n_{ij} \text{Log} n_{ij} - \text{Log} n_{ij}$$

si le tableau T a pour éléments les n_{ij} . En d'autres termes : tout tableau possédant au moins un élément nul serait affligé d'une valeur infinie du J , situation manifestement absurde. (car tous les tableaux de $X^2 > 0$ ont des éléments nuls)

Plus généralement, le tableau de données est ce qu'il est ; le représenter par son symbole ($:$ par une certaine partition du nombre entier $L \times C$ ou total n) signifie que tous les tableaux possédant le même symbole sont tenus (arbitrairement) pour *équivalents*. Les classes d'équivalence ainsi définies correspondent-elles bien à la *notion concrète* de contingence ? c'est douteux.

Bien plus, le calcul de X^2 ou tout autre critère sert à *ordonner* ces classes, peut-être arbitrairement ; les classes distinctes ayant même X^2 sont bloquées sans ménagement ; rare pour les petits tableaux, le cas devient plus fréquent si L et C augmentent.

Les *types* de tableaux, tels que nous les définissons, correspondent à la notion courante de contingence entre deux caractères dont les classes n'ont aucun ordre privilégié ; on peut dire avec l'école italienne de GINI : *deux caractères non-connexes*, en ce sens qu'aucune connexion n'existe entre les classes ; ce sont les "vrais" caractères qualitatifs, ceux pour lesquels aucune métrique n'est concevable pour les quantifier ; par exemple le caractère : *nationalité* ; déjà le caractère *catégorie socio-professionnelle* est parfois quantifié par un quelconque revenu moyen. Pour étudier une table de contingence entre variables découpées par intervalles (tranches d'âge ou de revenus), il est douteux que 2 tableaux de même type soient vraiment *pensés comme équivalents*.

Or notre analyse en types distincts pour un même symbole (et symboles distincts pour une même valeur du critère X^2) est déjà plus *fine* que la seule analyse par le X^2 qui (comme on dit) met tout dans le même sac.

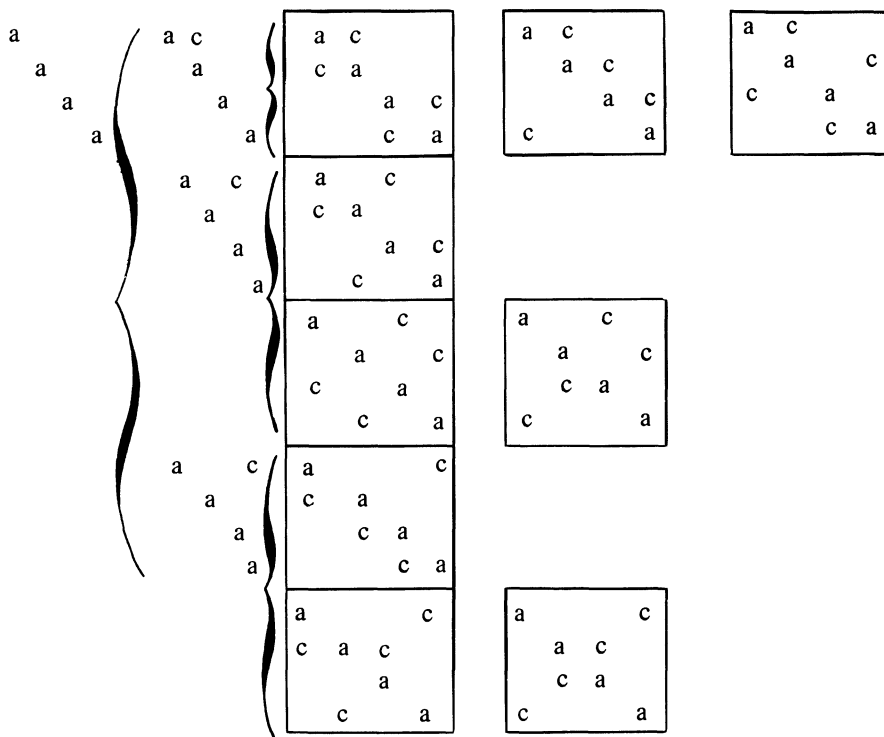
Ces remarques ne doivent pas décourager les chercheurs de critère de contingence mais plutôt les inciter à étudier lequel serait le mieux approprié chaque fois.

ANNEXE I – DETAIL DES TYPES (CAS 4 × 4)

(a)	$\begin{array}{c c} 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array}$	(b)	$\begin{array}{c c} 2 & 0 & 1 & 1 \\ \hline 0 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array}$	(c)	$\begin{array}{c c} 1 & 1 & 1 & 1 \\ \hline 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 2 \\ 1 & 0 & 2 & 1 \end{array}$	(d')	$\begin{array}{c c} 2 & 1 & 1 & 0 \\ \hline 0 & 2 & 1 & 1 \\ 1 & 0 & 2 & 1 \\ 1 & 1 & 0 & 2 \end{array}$	(d'')	$\begin{array}{c c} 2 & 1 & 0 & 1 \\ \hline 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \end{array}$	(d''')	$\begin{array}{c c} 2 & 0 & 1 & 1 \\ \hline 0 & 2 & 1 & 1 \\ 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \end{array}$
(e')	$\begin{array}{c c} 2 & 2 & 0 & 0 \\ \hline 2 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{array}$	(e'')	$\begin{array}{c c} 2 & 2 & 0 & 0 \\ \hline 1 & 1 & 0 & 2 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{array}$	(f')	$\begin{array}{c c} 2 & 0 & 2 & 0 \\ \hline 0 & 2 & 0 & 2 \\ 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \end{array}$	(f'')	$\begin{array}{c c} 2 & 2 & 0 & 0 \\ \hline 0 & 0 & 2 & 2 \\ 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \end{array}$	(g)	$\begin{array}{c c} 2 & 2 & 0 & 0 \\ \hline 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 \end{array}$	(h)	$\begin{array}{c c} 1 & 1 & 1 & 1 \\ \hline 1 & 3 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 1 & 0 & 1 & 2 \end{array}$
(i')	$\begin{array}{c c} 1 & 1 & 1 & 1 \\ \hline 3 & 1 & 0 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 2 & 2 \end{array}$	(i'')	$\begin{array}{c c} 3 & 1 & 0 & 0 \\ \hline 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 2 & 1 & 1 \end{array}$	(j')	$\begin{array}{c c} 3 & 1 & 0 & 0 \\ \hline 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 2 & 2 & 0 \end{array}$	(j'')	$\begin{array}{c c} 3 & 1 & 0 & 0 \\ \hline 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & 2 \end{array}$	(k)	$\begin{array}{c c} 3 & 1 & 0 & 0 \\ \hline 1 & 1 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 2 & 2 \end{array}$		
(l')	$\begin{array}{c c} 3 & 0 & 1 & 0 \\ \hline 0 & 3 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{array}$	(l'')	$\begin{array}{c c} 1 & 1 & 1 & 1 \\ \hline 3 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \\ 0 & 0 & 2 & 2 \end{array}$	(m')	$\begin{array}{c c} 3 & 0 & 1 & 0 \\ \hline 0 & 3 & 1 & 0 \\ 1 & 1 & 0 & 2 \\ 0 & 0 & 2 & 2 \end{array}$	(m'')	$\begin{array}{c c} 3 & 1 & 0 & 0 \\ \hline 0 & 3 & 1 & 0 \\ 1 & 0 & 1 & 2 \\ 0 & 0 & 2 & 2 \end{array}$	(n)	$\begin{array}{c c} 3 & 1 & 0 & 0 \\ \hline 1 & 3 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 \end{array}$	(o)	$\begin{array}{c c} 1 & 1 & 1 & 1 \\ \hline 1 & 3 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 1 & 0 & 0 & 3 \end{array}$
(p)	$\begin{array}{c c} 2 & 0 & 1 & 1 \\ \hline 0 & 1 & 0 & 3 \\ 1 & 3 & 0 & 0 \\ 1 & 0 & 3 & 0 \end{array}$	(q)	$\begin{array}{c c} 3 & 1 & 0 & 0 \\ \hline 1 & 3 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{array}$	(r)	$\begin{array}{c c} 4 & 0 & 0 & 0 \\ \hline 0 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{array}$	(s)	$\begin{array}{c c} 4 & 0 & 0 & 0 \\ \hline 0 & 2 & 2 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 2 \end{array}$	(t)	$\begin{array}{c c} 4 & 0 & 0 & 0 \\ \hline 0 & 2 & 2 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 2 & 0 & 2 \end{array}$	(u)	$\begin{array}{c c} 4 & 0 & 0 & 0 \\ \hline 0 & 3 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & 2 \end{array}$
(v)	$\begin{array}{c c} 4 & 0 & 0 & 0 \\ \hline 0 & 3 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \end{array}$	(w)	$\begin{array}{c c} 4 & 0 & 0 & 0 \\ \hline 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 1 & 0 & 3 \end{array}$	(x)	$\begin{array}{c c} 4 & 0 & 0 & 0 \\ \hline 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 \end{array}$	(y)	$\begin{array}{c c} 4 & 0 & 0 & 0 \\ \hline 0 & 4 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{array}$	(z)	$\begin{array}{c c} 4 & 0 & 0 & 0 \\ \hline 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{array}$		

ANNEXE II – NOMBRE DES TABLEAUX (a b b c)

Supposons les (a) placés, par exemple sur la 1^{ère} diagonale ; plaçons les (c) de toutes les façons possibles : 1^{ère} ligne, 2^{ème} ligne, 3^{ème} et 4^{ème} lignes.



: 9 tableaux distincts du même type.

Ce résultat est employé quand : $a = 3, b = 0, c = 1$
 et quand : $a = 2, b = 1, c = 0$.

REFERENCES

- [1] MANN (H.B.) & WALD (Abraham) – On the Choice of the Number of Intervals in the Application of the Chi-Square test, *A.M.S.* 13, 3 (sept. 1942) 306-317 (& Selected Papers in Stat. and Prob., 1955).
- [2] WILLIAMS (C. Arthur) Jr. – On the Choice of the Number and Width of Classes for the Chi-Square Test of Goodness of Fit, *J.A.S.A.*, 45 (march 1950) 77-86.
- [3] COCHRAN (W.G.) – The Chi-Square Test of Goodness of Fit, *A.M.S.*, 23,3 (sept. 1952) 315-345.

- [4] idem, Some Methods for Strengthening the Common χ^2 Tests, *Biometrics*, 10, (dec. 1954) 417-451.
- [5] VESSEREAU (André) – Sur les conditions d'application du criterium χ^2 de Pearson, *Bull. Inst. Statis.*, 36 (1958) 87-101.
- [6] WISE (M.E.) – Multinomial Probabilities and the X^2 and χ^2 Distributions, *Biometrika*, 50 (June 1963), 145-154.
- [7] KEMPTHORNE (O) – The classical Problem of Inference : Goodness of Fit, in J. NEYMAN ed, 5^e Berkeley Symposium Berkeley, California 1966, 235-49.
- [8] VAN DER WAERDEN (B.L.) – Statistique Mathématique, traduction Guinchat-Degenne, Dunod, Paris, 1967.
- [9] SLAKTER (Malcolm J.) – A Comparison of the Pearson Chi-Square and Kolmogorov Goodness-of-Fit Tests with Respect to Validity, *J.A.S.A.* 60 (sept. 1965) 854-858.
- [10] idem, Accuracy of an Approximation to the Power of the Chi-Square of Fit Test with Small but Equal Expected Frequencies, *J.A.S.A.* 63, (sept. 1968) 912-918.
- [11] ROSCOE (John T.) & BYARS (Jackson A.) – An Investigation of the Restraints with Respect to Sample Size Commonly Imposed on the Use of Chi-Square Statistic, *J.A.S.A.* 66 (décem. 1971) 755-759.
- [12] ZAHN (Douglas A.) & ROBERTS (Gail C.) – Exact X^2 Criterium Tables with Cell Expectations One : An Application to Coleman's Measure of Consensus, *J.A.S.A.*, 66 (march 1971) 145-148.
- [13] GOOD (I.J.), GOVER (T.N.) & MITCHELL (G.J.) – Exact Distributions for X^2 and for Like – lihood – Ratio Statistic for the Equiprobable Multinomial Distribution, *J.A.S.A.* 65, (march 1970) 267-283.
- [14] LETI (Giuseppe) – La distribuzione delle tabelle della classe di Fréchet, *Metron*, 28, (1970) 86-121.