

# REVUE DE STATISTIQUE APPLIQUÉE

DANIEL CUTHBERT

**Critique et analyse graphique des résultats de plans  
d'expériences factoriels  $2^p$  (à deux niveaux pour  
chacun des  $P$  facteurs contrôlés)**

*Revue de statistique appliquée*, tome 9, n° 3 (1961), p. 11-26

[http://www.numdam.org/item?id=RSA\\_1961\\_\\_9\\_3\\_11\\_0](http://www.numdam.org/item?id=RSA_1961__9_3_11_0)

© Société française de statistique, 1961, 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/>

# CRITIQUE ET ANALYSE GRAPHIQUE

## DES RÉSULTATS DE PLANS D'EXPÉRIENCES FACTORIELS $2^p$

### (A DEUX NIVEAUX POUR CHACUN DES P FACTEURS CONTROLÉS)

*Il est toujours nécessaire de procéder à l'examen critique des résultats de plans d'expérience  $2^p$  et de leurs facteurs  $2^{p-q}$ .*

*Il est bon de savoir que si les données sont rassemblées dans des conditions qui satisfont toutes les hypothèses à la base de la représentation factorielle habituelle, alors l'analyse n'offre pas de difficultés ; les meilleurs tests sont connus, le niveau de signification réel est le niveau indiqué et la sensibilité des tests peut être déterminée. Mais s'il y a une ou plusieurs valeurs aberrantes dans la série des observations, si la variance de l'erreur n'est pas la même pour toutes les observations ou si le plan de "randomisation" a été modifié par inadvertance (par exemple par l'introduction d'un split-plot), on doit alors pouvoir trouver des moyens de détecter ces défauts. Pour analyser et interpréter les plans d'expérience  $2^p$ , il peut être utile d'étudier la distribution des fréquences cumulées des comparaisons habituelles calculées à partir des données brutes.*

*On donne une méthode pour déceler les effets principaux et les interactions significatifs quand ceux-ci sont en petit nombre devant les "comparaisons ayant le caractère de termes d'erreur" ("error contrasts").*

*Des exemples sont donnés concernant des expériences comportant 16, 32, 64, 128 observations.*

#### I - CONSTRUCTION D'UNE GRILLE SEMI-NORMALE ET EXEMPLE D'APPLICATION.

Les 31 comparaisons calculées à partir des résultats d'un plan d'expérience  $2^{5-0}$  relatif à une production de pénicilline et décrit par Davies (cf. Référence 2, exemple 9.2)(1) sont données classées par valeurs absolues décroissantes dans le tableau I. Les plus grandes valeurs indiquent les effets qui, très vraisemblablement, ne sont pas nuls. Bien entendu nous avons besoin d'une méthode qui nous aide à juger objectivement si les effets les plus grands sont réels ou non.

C'est une question assez importante pour les grands plans factoriels puisque, dans ceux-ci, les comparaisons les plus grandes, même dans le

-----

(1) Voir les références en fin de texte.

cas d'effets nuls, ont une espérance mathématique égale à plusieurs fois la valeur moyenne de l'ensemble des comparaisons. Ainsi, dans une expérience portant sur 32 observations, la valeur du rapport est d'environ 2,4. Si l'on utilise le niveau de signification courant de 0,05 dans de telles expériences, cela conduira, dans plus de la moitié des cas, à estimer, à tort, certains effets comme étant réels. Si ce fait est reconnu et accepté, cela n'est pas grave. On trouvera au chapitre suivant une méthode simple pour maintenir la fréquence des effets faussement significatifs de chaque expérience, au niveau désiré.

S'il n'y avait pas d'effets, alors la distribution des valeurs absolues des comparaisons serait celle de l'étendue ou "range" d'échantillons d'effectif égal à 2<sup>(1)</sup>. La distribution de l'étendue de couples tirés d'une distribution normale est connue et tabulée. Même si la distribution de l'erreur n'est pas normale, on peut s'attendre à ce que la distribution des comparaisons (qui sont des sommes de variables) soit très proche de la normale. Toutefois, on trouvera ci-dessous quelques exceptions intéressantes.

S'il y a un seul effet important, on peut s'attendre à ce que la distribution observée des comparaisons classées suivant leurs valeurs absolues ne soit qu'à peine modifiée, puisqu'un seul terme erroné (celui lié à l'effet important) est à soustraire de la série. On peut s'attendre alors à ce que les (n - 1) comparaisons restantes se comportent comme un échantillon issu d'une population où les effets seraient nuls.

Si, enfin, une proportion considérable des n effets diffère sensiblement de zéro, alors (en l'absence d'information sur l'ordre de grandeur de  $\sigma_R$ , écart-type de l'erreur), l'examen des comparaisons ordonnées ne permet pas d'apprendre grand-chose. En pareil cas, il est manifestement indispensable de disposer d'une estimation indépendante de l'erreur. La question se pose de savoir ce qu'on doit considérer comme une proportion considérable de grandes comparaisons. Nous y répondrons en partie par la suite.

La figure 1 donne un modèle de grille valable dans la plupart des cas. On porte en abscisses les valeurs absolues des comparaisons.

Les grilles spéciales pour n = 15, 31, 63, 127 sont simples à utiliser puisqu'on n'a pas besoin de calculer les fréquences cumulées. Les numéros des rangs sont indiqués directement sur la grille, comme on peut le voir sur la figure 2 pour n = 31. Il suffit d'inscrire directement sur les lignes horizontales de la grille par valeurs décroissantes, les valeurs absolues des comparaisons. Ces grilles sont suffisamment précises pour les utilisations que nous proposons ici.

Un exemple de l'utilisation d'une "grille à échelle semi-normale" est donnée sur la figure 2 avec les résultats expérimentaux du tableau I. La comparaison la plus grande (effet E) est portée sur la ligne 31, et ainsi de suite jusqu'à celle qui suit la plus petite (de grandeur 0,025 et appelée AE) qui est inscrite sur la ligne 2. Il est peu intéressant de porter chacune des comparaisons de la moitié inférieure de l'ensemble, à cause des corrélations très fortes entre les grandeurs de deux comparaisons voisines.

Même en l'absence de connaissances précises sur les variations d'échantillonnage des statistiques d'ordre, on considèrera A, C et E comme très importante. Ces trois effets devront être jugés comme réels, tant à cause

-----  
(1) Voir "Note sur les plans factoriels 2<sup>n</sup>" par A. Vessereau.

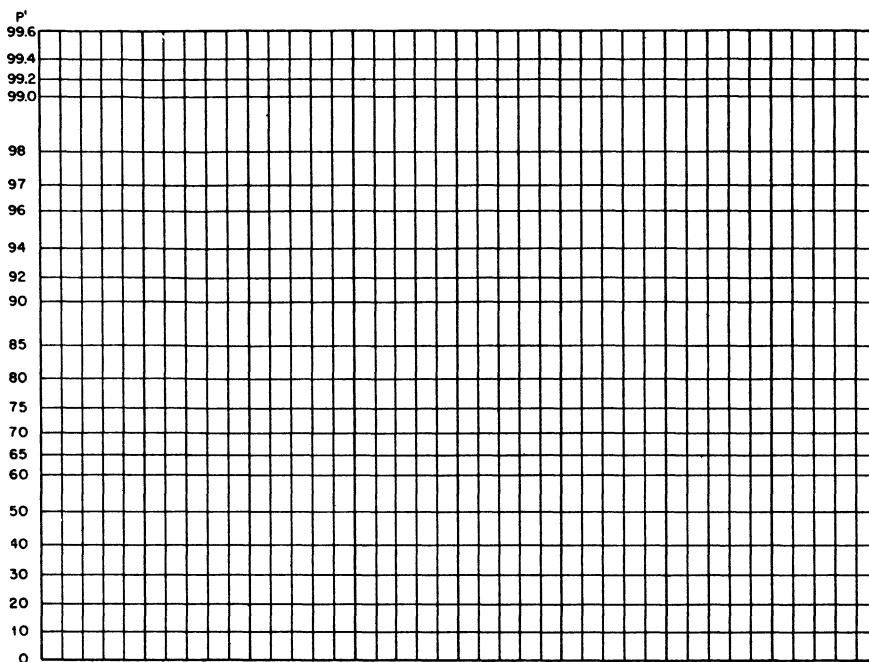


Figure 1 - Grille à échelle semi-normale.

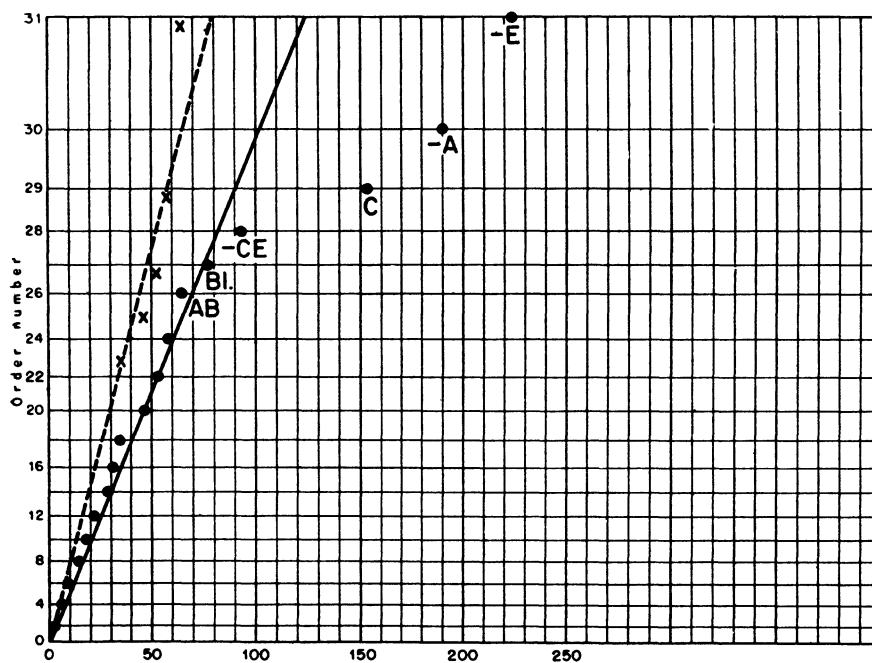


Figure 2 - Grille semi-normale pour un plan factoriel  $2^5$ . Points et ligne en trait plein : 31 comparaisons. Croix et ligne en pointillé : 24 plus petites comparaisons.

Tableau I

Comparaison d'un plan d'expérience factoriel  $2^5$   
 données dans l'ordre décroissant  
 Valeurs multipliées par 100 et arrondies

N°	Effet	Valeur	N°	Effet	Valeur
31	E	224	15	DE	30
30	A	190	14	BE	29
29	C	153	13	BDE	28
28	CE	93	12	ABE	22
27	ABCDE (Blocs)	77	11	ADE	21
26	AB	61	10	BCD	18
25	ABCD	58	9	BCDE	16
24	ACE	58	8	ABDE	11
23	AD	54	7	CDE	12
22	AC	53	6	D	9
21	BC	53	5	BD	7
20	ACDE	47	4	B	6
19	BCE	39	3	CD	4
18	ABD	34	2	AE	2
17	ACD	33	1	ABC	0
16	ABCE	31			

de leurs qualités d'effets principaux, qu'en raison de leurs grandeurs vis-à-vis des autres comparaisons de la série.

Si l'expérience avait été nulle, c'est-à-dire si aucun facteur n'avait eu d'effet, l'écart-type  $\sigma_c$  des sommes intervenant dans les comparaisons, aurait pu être estimé approximativement par la comparaison pour laquelle la probabilité est très proche de 0,683 ou 68,3 %. Pour  $n = 31$ , c'est la 22e comparaison (valeur 0,53) puisque  $(22 - 1/2) 31 = 0,693$ . Mais il ne s'agit manifestement pas ici d'une expérience nulle et on utilisera donc moins de 31 comparaisons pour estimer l'erreur. La valeur 0,53 ne devra pas être prise pour estimer  $\sigma_c$ .

On sait (p. 418 de la référence 2) que CE peut être considéré comme importante d'après des informations antérieures. On sait aussi que les variations d'une semaine à l'autre sont courantes. C'est pourquoi cette expérience a été faite en deux blocs, en confondant l'interaction ABCDE avec l'effet des semaines. Puisque précisément ces deux comparaisons viennent à la suite dans le classement par valeurs absolues décroissantes, il semble opportun de les supprimer de l'ensemble qui doit être utilisé pour estimer l'erreur.

Les deux comparaisons suivantes mesurent AB et ABCD. Ces dernières et toutes les autres plus petites seront (toujours à titre d'essai) considérées comme des comparaisons ayant le caractère de termes d'erreur. On peut, soit reporter les 26 comparaisons sur une grille comme celle de la figure 1, soit relever chacun des 26 points de la figure 2 jusqu'à un point correspondant à la nouvelle fréquence cumulée  $p' = (i - 1/2)/26$ .

Il est remarquable que, même après la suppression de cinq comparaisons, la plus grande restante se trouve encore voisine de l'ancienne ligne supérieure (à 0,981 au lieu de 0,984 obtenue précédemment). Le nouveau tracé est indiqué sur la figure 2 par des croix. L'écart-type des sommes

intervenant dans les comparaisons est maintenant estimé égal à 0,33. L'extrémité supérieure du graphique suggère, sans toutefois le prouver, qu'il a été tronqué. Aucune modification sérieuse n'intervient si l'on décide d'inclure, dans le système des comparaisons d'erreurs, la comparaison due aux blocs. On peut donc utiliser avec confiance la série des 27 comparaisons pour estimer l'erreur.

Une estimation  $\hat{\sigma}_R$  de l'erreur sur chaque observation peut être obtenue à partir de la formule évidente :

$$\sigma_R = \sigma_c / \sqrt{n + 1}$$

L'estimation ainsi faite, basée sur 27 degrés de liberté, correspond presque exactement à celle donnée dans la référence 7, estimée avec les 15 degrés de liberté liés aux interactions de 3 et de 4 facteurs.

Un pas en avant vers une estimation plus objective de ce qui, dans les expériences factorielles sans répétition, dépend du hasard et de ce qui est systématique, a été faite par J.W. Tukey (Réf. 5) et aussi par S.C. Pearce (Réf. 4). Il consiste à "faire la liste" de tous les effets, interactions et comparaisons de blocs que l'on suppose vraisemblablement importants, avant que l'expérience ne soit terminée. Il faut se garder d'affecter à une erreur ou à un effet réel les comparaisons correspondantes. Seules les comparaisons qui ne sont pas sur la liste seront étudiées comme susceptibles de constituer l'erreur. Cette suggestion m'apparaît comme un compromis des plus acceptables entre le procédé dangereux qui consiste à appeler erreur tout ce qui ressemble à l'erreur et le procédé prudent qui consiste à ne considérer comme telle que ce qui est déterminé par la stricte "randomisation" de traitements identiques. Il convient toutefois de signaler que cette règle ne modifierait sensiblement aucune des conclusions ou estimations trouvées dans l'expérience dont nous venons de parler.

Pour donner une idée des variations dues à l'échantillonnage qui peuvent se présenter dans des graphiques tels que ceux de la figure 2, on a représenté sur la figure 3 des graphiques semi-normaux pour dix séries de 31 valeurs normales réduites prélevées au hasard (tables de Dixon et Massey (3)) On s'est servi des 31 premiers nombres donnés dans les colonnes 11 à 20. Les valeurs portées en abscisse ne concernent que le premier graphique de chaque série (chaque ligne inclinée devrait partir de zéro). La pente de chaque droite est celle qui correspond à l'écart-type de la population (1,00) et non celle de la meilleure droite ajustée aux points indiqués.

## II - GRAPHIQUES SEMI-NORMAUX STANDARDISES.

Il peut être commode d'utiliser une grille semi-normale avec des limites fixées pour étudier les résultats d'un plan d'expérience  $2^4$ . Si l'on divise les 15 comparaisons calculées et ordonnées par la onzième valeur à partir de la plus petite, on obtient une série de statistiques d'ordre sans dimensions qui, en l'absence d'effets réels devraient se trouver autour d'une droite type comme celle qui est indiquée figure 4a (trait plein). Pour les plans d'expérience  $2^5$ ,  $2^6$  et  $2^7$ , les comparaisons doivent être divisées par la comparaison qui occupe le rang 22, 44 et 88 à partir de la plus petite.

Le tableau II donne pour ces quatre plans d'expérience les écarts limites correspondant aux probabilités 0,05, 0,2 et 0,4 de trouver un effet faussement significatif dans l'expérience.

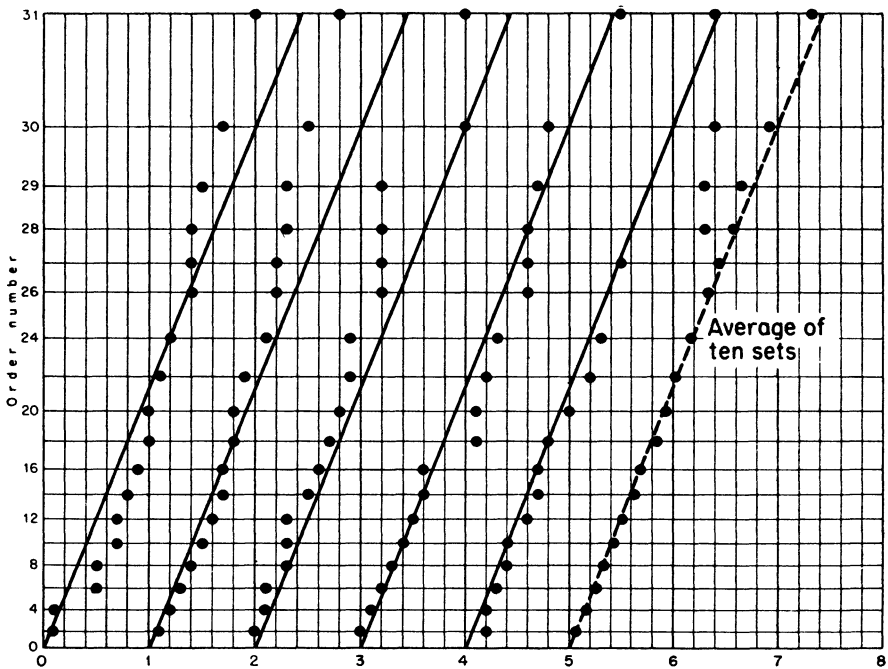
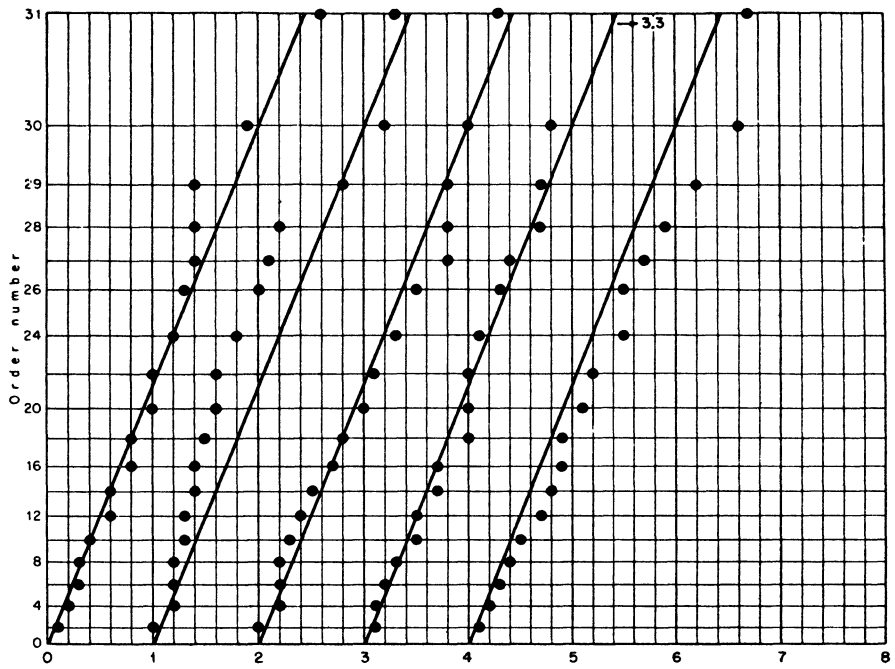


Figure 3 - Graphiques semi-normaux pour 10 séries de 31 variables normales réduites .

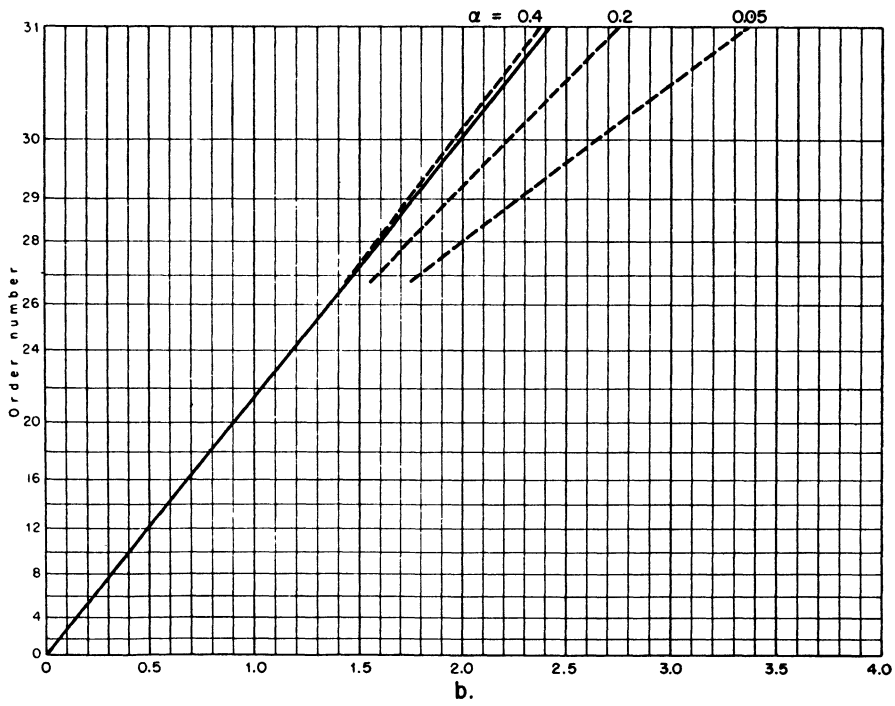
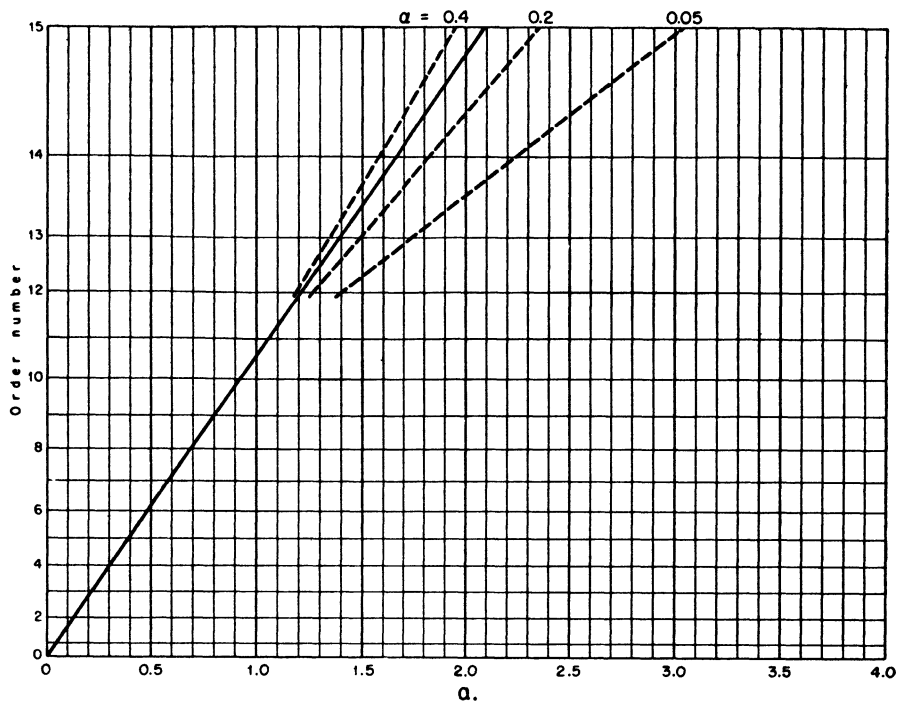


Figure 4a et b - Graphiques semi-normaux standardisés pour 15 et 31 comparaisons.



Tableau II

n	$\tau$	Ecartes limites pour les graphiques semi-normaux standardisés : d		
		$\alpha = 0.05$	$\alpha = 0.20$	$\alpha = 0.40$
15	2.09	0.94	0.26	-0.14
31	2.42	0.94	0.33	-0.05
63	2.69	0.66	0.23	-0.04
127	2.88	0.66	0.25	-0.00

n = nombre des comparaisons orthogonales

$\tau$  = espérance mathématique de  $t_n^{(0)}$  = comparaison standardisée de rang n en l'absence de tout effet

$\alpha$  = fréquence des effets faussement significatifs par expérience

d = valeurs tabulées, écarts limites à droite de la droite type.

Les points correspondants se placent approximativement sur des droites comme le montre la figure 4a.

Dans des expériences ayant pour but de rechercher l'existence de facteurs, il apparaît souhaitable d'utiliser comme valeurs pour  $\alpha$ , 0.4 ou même 0.8, sachant cependant que cela produira un certain nombre de "fausses alertes". Ce procédé se justifie par une meilleure sensibilité pour la mise en évidence des faibles effets. Au cours des études ultérieures, les effets ainsi sélectionnés par erreur auront toutes chances d'être mis à l'épreuve. Par contre les effets réels qui seraient omis ne seront vraisemblablement plus étudiés ; s'ils sont nombreux, la perte d'information sur leur action combinée peut être importante.

La figure 4b représente la grille semi-normale standardisée pour n = 31.

### III - QUELQUES PLANS D'EXPERIENCES $2^{p-q}$ .

La figure 5 représente sur une grille semi-normale standardisée les résultats effectifs de 8 plans d'expériences, où p - q est égal à 5.

Pour gagner de la place, quatre graphiques ont été tracés sur chaque grille. Les valeurs portées en abscisses doivent naturellement être diminuées chaque fois de 1 en passant d'un graphique au suivant. Les comparaisons standardisées sont données dans le tableau III.

La figure 6 représente sur la grille appropriée six exemples de plans d'expériences  $2^4$ . Les comparaisons standardisées correspondantes classées dans l'ordre croissant sont indiquées dans le tableau IV.

La figure 7 montre trois expériences comportant 64 observations.

La figure 8 montre deux expériences comportant 128 observations.

Ces expériences ont été choisies comme étant d'assez "bons" exemples ne comportant pas les complications que nous mentionnons ci-après.

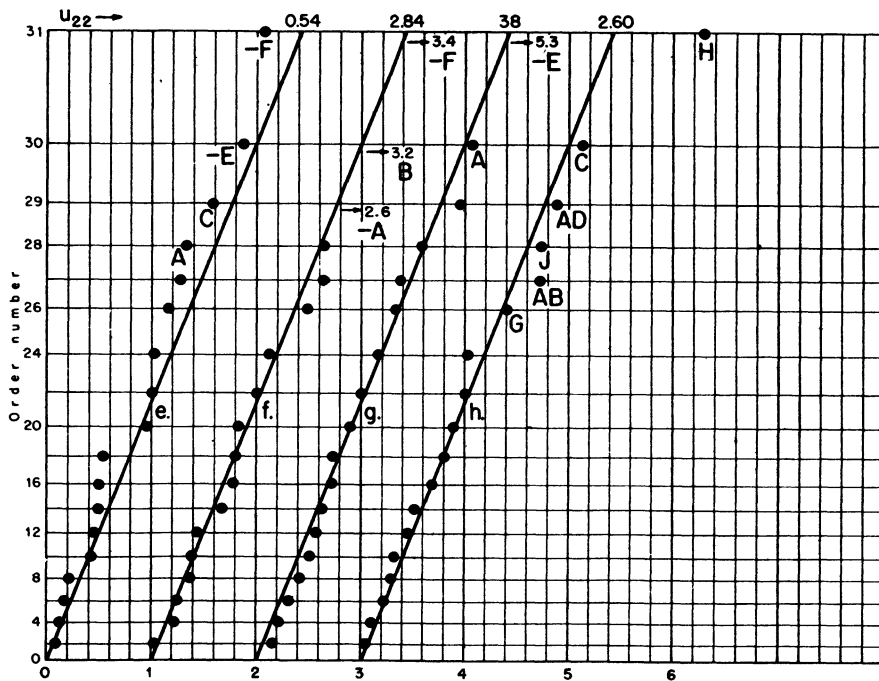
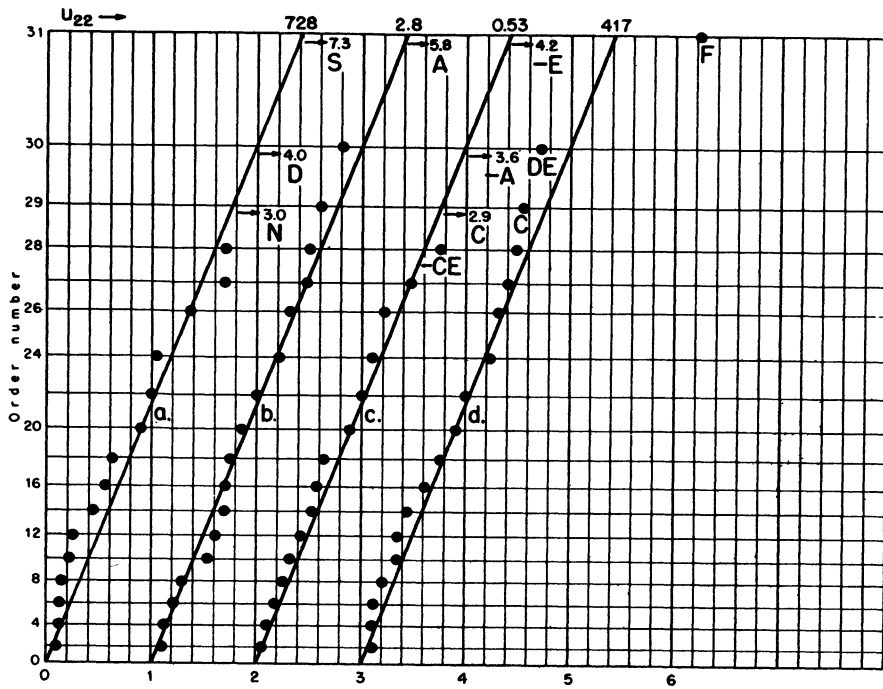


Figure 5 - Grilles semi-normales standardisées pour 8 plans d'expérience  $2^{p-q}$  avec  $p - q = 5$ .

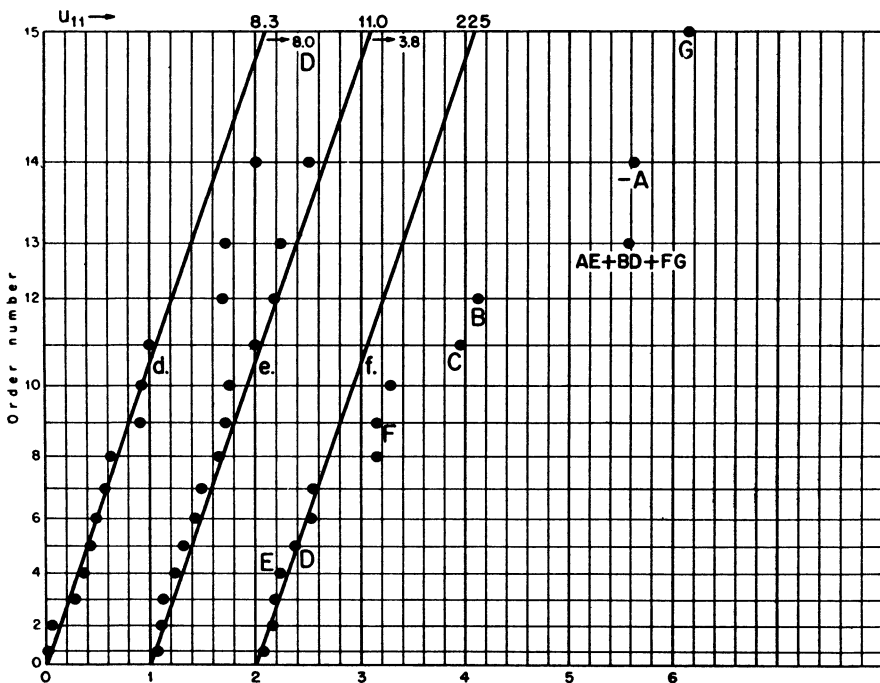
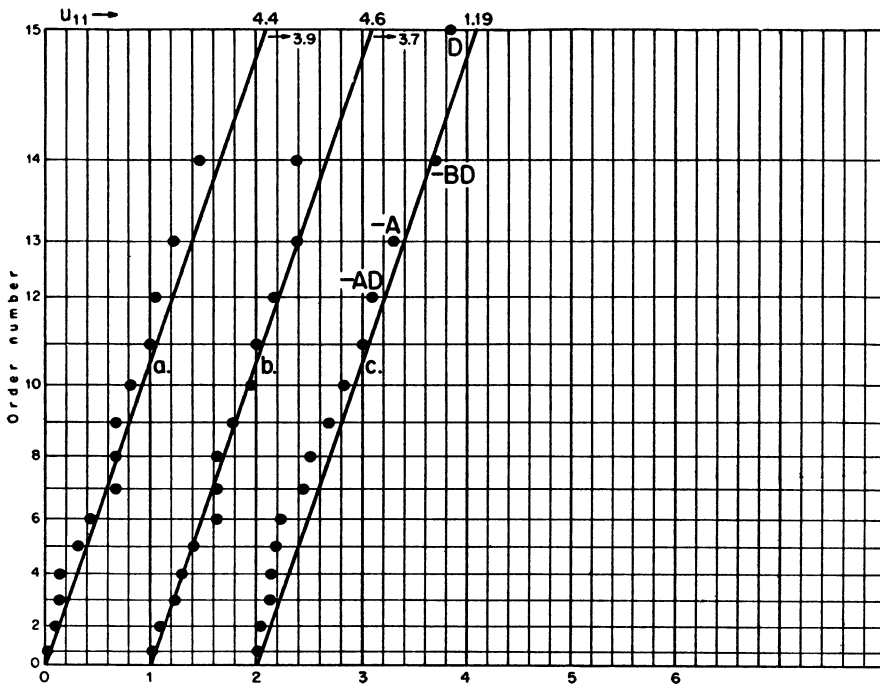


Figure 6 - Grilles semi-normales standardisées pour 6 plans d'expérience  $2^{p-q}$  avec  $p - q = 4$ .

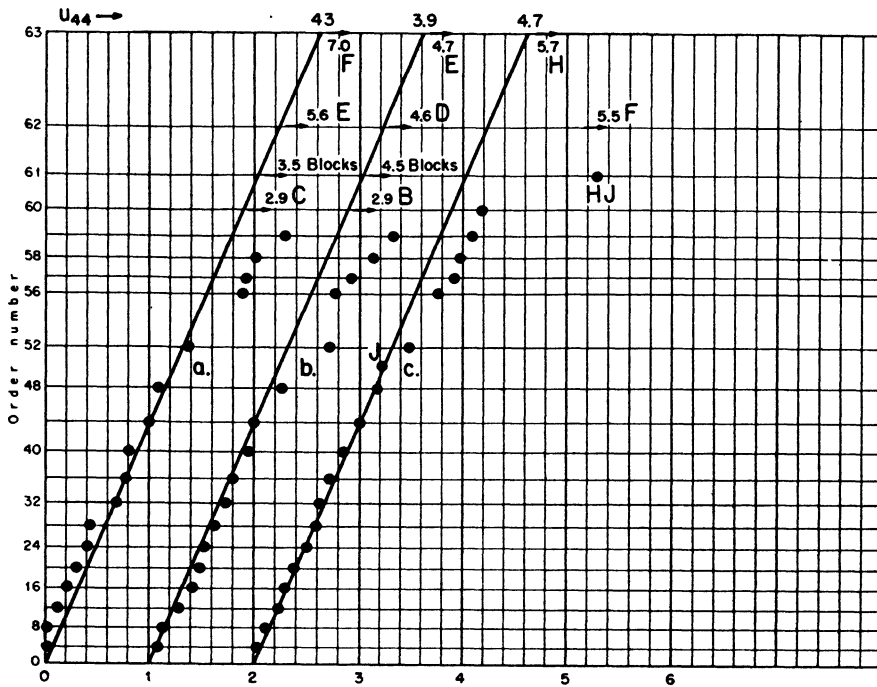


Figure 7 - Grilles semi-normales standardisées pour 3 plans d'expérience  $2^{p-q}$  avec  $p - q = 6$ .

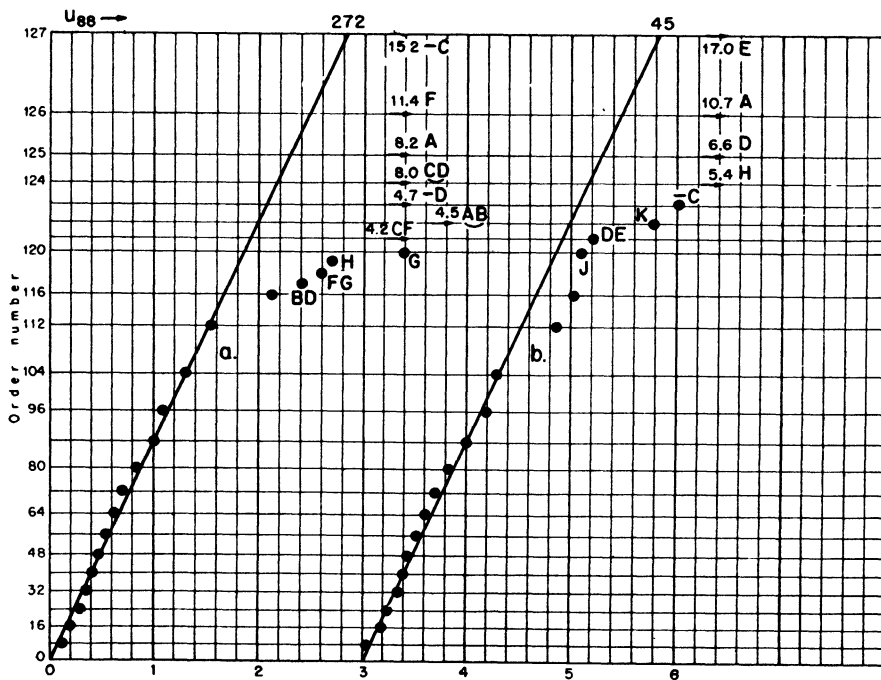


Figure 8 - Grilles semi-normales standardisées pour 2 plans d'expérience  $2^{p-q}$  avec  $p - q = 7$ .

Tableau III  
Comparaisons standardisées de 8 plans d'expériences 2  
données par valeurs croissantes.  
(voir figure 5).

Série Nature	a $2^{5-0}$	b $2^{6-1}$	c $2^{5-0}$	d $2^{6-1}$	e $2^{8-3}$	f $2^{6-1}$	g $2^{11-6}$	h $2^{15-10}$
$u_{22}$ i	728	2.8	3.3	417	.538	2.84	38	2.60
2	.08	.11	.05	.12	.07	.03	.16	.05
4	.13	.14	.11	.12	.12	.22	.21	.10
6	.13	.21	.18	.12	.16	.25	.32	.22
8	.14	.30	.25	.20	.21	.37	.42	.28
10	.22	.54	.33	.36	.42	.39	.53	.32
12	.24	.61	.43	.36	.44	.44	.58	.45
14	.43	.68	.54	.44	.47	.66	.63	.52
16	.56	.71	.58	.61	.49	.79	.74	.68
18	.62	.75	.64	.76	.54	.80	.74	.80
20	.90	.86	.88	.92	.95	.82	.90	.89
22	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
24	1.04	1.21	1.09	1.23	1.02	1.13	1.16	1.03
26	1.37	1.32	1.21	1.32	1.16	1.47	1.32	1.40
27	1.70	1.47	1.46	1.40	1.26	1.63	1.37	1.72
28	1.70	1.50	1.76	1.48	1.33	1.64	1.58	1.73
29	2.96	1.61	2.92	1.56	1.58	2.59	1.95	1.88
30	3.98	1.82	3.58	1.72	1.86	3.24	2.06	2.14
31	7.32	5.75	4.25	3.23	2.07	3.38	5.32	3.30

Tableau IV  
Comparaisons standardisées de 6 plans d'expériences 2<sup>4</sup>  
données par valeurs croissantes.  
(voir figure 6).

approximativement  $u_8$  et non  $u_{11}$  à cause de 3 effets importants.

Série Nature	a $2^{4-0}$	b $2^{4-0}$	c $2^{4-0}$	d $2^{5-1}$	e $2^{5-1}$	f $2^{7-3}$
$u_{11}$ i	4.4	4.6	1.19	8.3	11.	225*
1	.02	.02	.01	.01	.07	.09
2	.09	.09	.04	.05	.11	.17
3	.14	.24	.13	.28	.13	.19
4	.14	.30	.14	.37	.24	.23
5	.32	.41	.18	.43	.33	.38
6	.43	.63	.23	.47	.43	.53
7	.66	.63	.45	.57	.48	.55
8	.66	.63	.51	.63	.65	1.16
9	.66	.78	.68	.92	.72	1.16
10	.82	.96	.83	.93	.76	1.27
11	1.00	1.00	1.00	1.00	1.00	1.96
12	1.05	1.17	1.09	1.70	1.18	2.12
13	1.23	1.39	1.29	1.72	1.24	3.58
14	1.46	1.39	1.69	2.01	1.52	3.62
15	3.87	3.67	1.84	8.03	3.81	4.16

#### IV - UTILISATION DE GRAPHIQUES SEMI-NORMAUX POUR L'ANALYSE CRITIQUE DES DONNEES.

Les graphiques semi-normaux semblent être sensibles à certains types de défauts dans les données expérimentales et sont parfois utiles pour déceler ces défauts.

L'étude des quatre types de défauts suivants nous a permis d'acquérir une certaine expérience :

- a) Une valeur erronée
- b) Deux valeurs erronnées ou plus
- c) Introduction par mégarde d'un split-plot
- d) Distribution d'erreur lognormale.

Seuls les deux premiers cas seront examinés ici.

a) Une valeur erronée - Une seule valeur très aberrante est facilement repérée par l'examen des données brutes, s'il n'y a qu'un ou deux facteurs dont les effets soient notables. Mais si les facteurs (ou interactions) influents sont en plus grand nombre, et si l'erreur qui apporte la valeur aberrante ne dépasse pas environ  $2\sigma_0$ , il peut y avoir plusieurs résultats qui dépassent le résultat aberrant et qui, de ce fait, le cachent. En outre, la plupart des comparaisons utilisées pour estimer l'erreur seront majorées par le résultat aberrant, de telle sorte que les effets réels ne seront peut-être pas décelés.

La valeur erronée apparaîtra d'un seul côté dans chaque comparaison et accroîtra, en moyenne, sa valeur absolue. La partie (a) de la figure 9 montre un graphique semi-normal des données du tableau V.

Données et comparaisons avec une observation erronée (plan  $4^2 \times 2^2/4$ .  
(voir figure 9).

B, C, BC employés pour désigner un facteur quantitatif à 4 niveaux également espacés.

D, F, DF employés pour un facteur qualitatif à 4 niveaux.

\* bcef a été modifiée de - 7,0

Tableau V

Conditions expérimentales	Données brutes non corrigées	Comparaisons non corrigées	Comparaisons corrigées
(1)	17.7	297.5	290.5
aef	15.6	-7.5	-.5
be	21.0	49.9	42.9
abf	22.0	6.5	13.5
cf	13.1	-7.3	-41.3
ace	13.1	-2.7	4.3
bcef*	28.7	25.9	18.9
abc	22.8	-5.5	1.5
def	19.5	-10.5	-3.5
ad	17.4	6.5	-.5
bdf	19.2	-20.1	-13.1
abde	20.0	12.1	5.1
cde	15.1	-10.1	-3.1
acdf	12.3	6.9	-.1
bcd	18.2	-5.3	1.7
abcdef	21.8	12.5	5.5

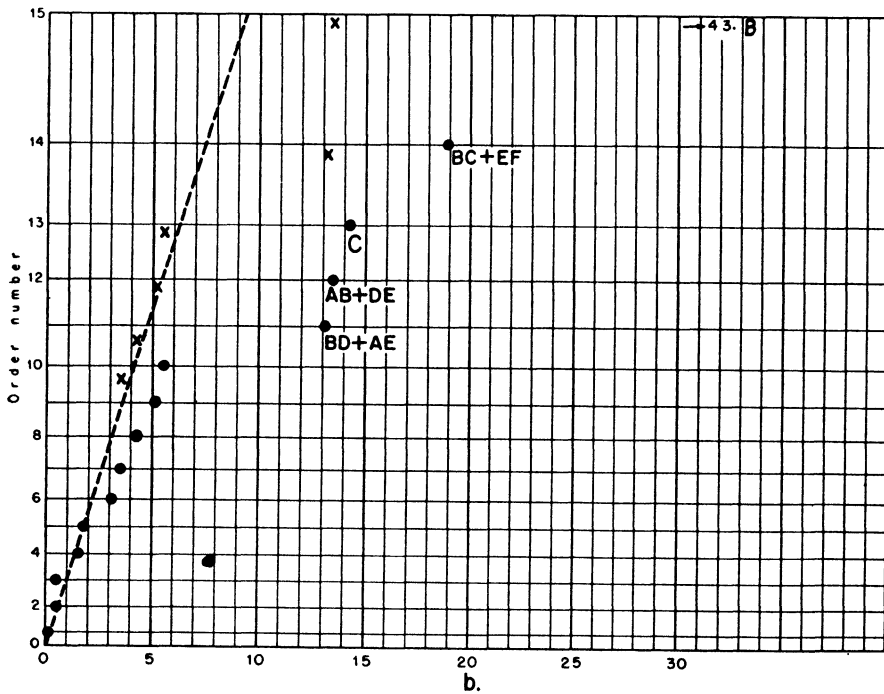
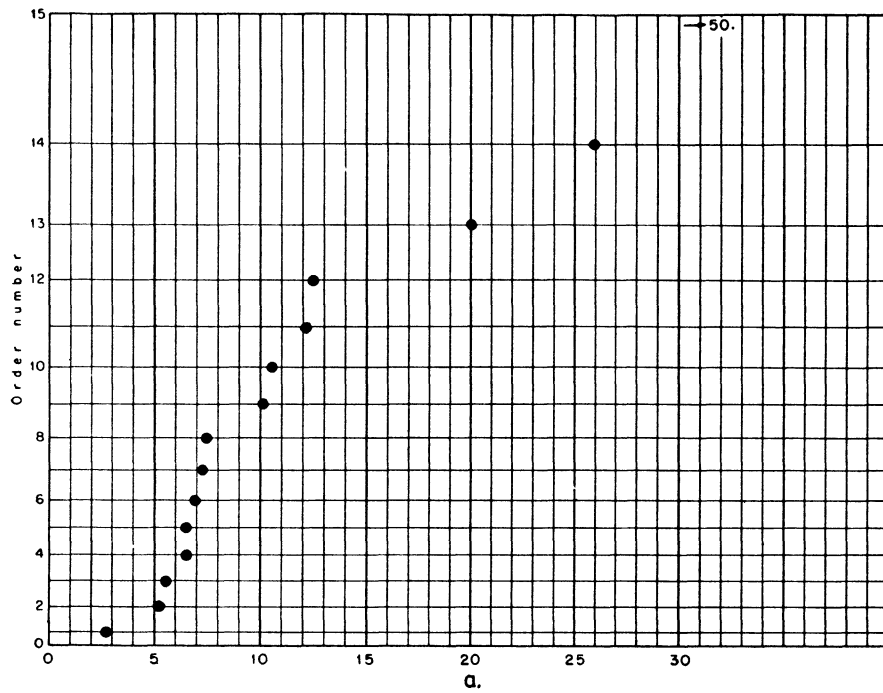


Figure 9 - Effet d'une valeur erronée sur un graphique semi-normal  
 a) comparaisons résultant des observations  
 b) comparaisons après correction d'une observation.

Sur ce graphique il est évident que les plus petites comparaisons s'écartent de l'origine, ce qui laisse supposer une erreur de 7 environ dans toutes les comparaisons. Un examen, plus approfondi des données, fait penser que la mesure "bcef" est erronée. Si l'on remplace cette valeur de 28.7 par 21.7 chaque comparaison est modifiée de  $\pm 7$ . Il est inutile de calculer à nouveau les comparaisons à partir des données. Le tableau M de la référence (2) donne le signe de la correction requise pour chaque combinaison factorielle et pour chaque effet.

Un graphique semi-normal des comparaisons corrigées est donné figure 9, partie b. Si l'on considère B, BC et C comme des effets réels, les 12 comparaisons restantes donnent le graphique de l'erreur indiqué par les croix et la droite en pointillé sur la même figure. On peut penser que l'une des comparaisons AB ou BD, ou les deux, mesurent également des effets réels, mais les données ne sont pas suffisantes pour répondre à cette question.

Les valeurs anormales qui ne sont dues, ni aux erreurs aléatoires habituelles, ni aux effets des facteurs que l'on a fait volontairement varier, représentent un risque sérieux dans les plans d'expérience factoriels  $2^{p-q}$ . Elles sont plus nuisibles au plan d'expérience que les données manquantes, parce qu'elles ne sont pas faciles à identifier. Ces valeurs ne sont d'ailleurs pas aussi rares qu'on pourrait l'espérer. Si ma propre expérience est caractéristique, elles apparaissent avec des fréquences relatives allant de 0.1 à 0.01, selon la complexité des conditions expérimentales et l'expérience de l'expérimentateur.

Beaucoup de chercheurs ont l'impression que les plans d'expérience  $2^{p-q}$  doivent échouer à cause de la présence connue d'interactions d'ordre élevé, c'est-à-dire de surfaces de réponses extrêmement courbées dans la zone considérée. A mon sens, ce sentiment est dû, pour une grande part, à l'apparition de valeurs erronées qui n'ont pas été identifiées comme telles.

b) Deux valeurs erronées ou plus - Deux résultats aberrants ayant à peu près le même degré d'aberrance concourent ensemble à majorer les valeurs absolues de la moitié des comparaisons dans un plan d'expérience  $2^{p-q}$ . L'allure de la figure 6 suggère qu'il s'est produit quelque chose de ce genre. Si les deux résultats n'ont pas le même degré d'aberrance alors toutes les comparaisons seront modifiées : la moitié sera augmentée de la différence entre les écarts à la normale des deux valeurs fausses, et l'autre moitié de la somme des deux écarts. De même, si trois résultats aberrants interviennent, le quart des comparaisons subira l'augmentation maximum.

## V - ECHEC DES GRAPHIQUES SEMI-NORMAUX

Il est inutile de prévenir les statisticiens expérimentés que l'emploi des graphiques semi-normaux suggérés ici abonde en erreurs subjectives qu'il n'est pas proposé pour remplacer de façon générale l'analyse de variance et que son utilisation routinière peut conduire à des catastrophes. Des statisticiens plus optimistes et moins expérimentés pourront retirer des exemples probants que nous venons de donner ici l'impression qu'on leur offre une panacée qui n'échoue pratiquement pas. Ce n'est pas le cas.

L'auteur a construit des graphiques semi-normaux pour tous les plans d'expérience  $2^{p-q}$  donnés dans les ouvrages généraux et les revues aisément accessibles. Plus des neuf dixièmes de ces plans d'expérience donnent des tracés linéaires acceptables pour les plus petites comparaisons. Plusieurs



centaines d'expériences industrielles non publiées ont conduit à une même proportion. On n'a remarqué aucune différence spécifique entre les expériences publiées et celles qui ne l'ont pas été.

La régularité d'un graphique n'est pas une garantie de perfection. On peut s'attendre à des erreurs de deux sortes. Un graphique linéaire (aucun effet n'étant mis en évidence de façon certaine) peut quelque fois nous conduire à surestimer la grandeur de l'erreur et, par suite, à laisser passer un certain nombre d'effets réels. Une ligne extrêmement irrégulière peut nous amener à faire une autre sorte d'erreur qui consiste à juger comme étant réels des effets qui, en fait, sont négligeables. Une erreur variable d'une série à l'autre introduit des corrélations entre les comparaisons. Ceci conduira souvent à un graphique irrégulier avec des points groupés. Aucun résultat quantitatif n'est donné ici, mais il faut signaler que quelques expériences ont montré des tracés très irréguliers.

L'expérimentateur qui connaît très exactement l'erreur attachée à chaque série, qui sait quels sont les facteurs influents et qu'elle est l'importance de leurs influences, peut très bien concevoir et réaliser un plan d'expérience  $2^{p-q}$  dans lequel presque tous les effets principaux et de nombreuses interactions sont grands vis-à-vis de l'erreur. Dans de pareils cas, les graphiques semi-normaux sont de peu d'utilité.

## VI - CONCLUSIONS.

Nous venons d'exposer l'utilisation d'un graphique de la distribution des fréquences cumulées des comparaisons orthogonales habituelles calculées à partir de plans d'expériences  $2^{p-q}$  (pour  $p - q$  allant de 4 à 7). Ce graphique présente quelque intérêt pour révéler certains types généraux de défauts dans des données provenant d'expériences industrielles. Quand une faible proportion seulement de la totalité des comparaisons a un effet, ce graphique peut être employé pour juger de la réalité des plus grands effets trouvés.

## REFERENCES

- [1] - BIRNBAUM A. - "On the Analysis of Factorial Experiments without replication" Technometrics, 1, 4 (Nov. 1959).
- [2] - DAVIES O. L., Editor - "Design and Analysis of Industrial experiments", Second Edition, Oliver and Boyd, London, 1956.
- [3] - DIXON W. J., and MASSEY F. J. - "Introduction to Statistical Analysis" Second Edition, Mc Graw-Hill, 1957.
- [4] - PEARCE S. C. - Technical Communication No. 23, Commonwealth Bureau of Horticulture and Plantation Crops, East Malling, Kent, England.
- [5] - TUKEY J. W. - Personal communication.