

REVUE DE STATISTIQUE APPLIQUÉE

A. VESSEREAU

Note sur les plans factoriels 2^p

Revue de statistique appliquée, tome 9, n° 3 (1961), p. 7-9

http://www.numdam.org/item?id=RSA_1961__9_3_7_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/>

NOTE SUR LES PLANS FACTORIELS 2^p

A. VESSEREAU

Ingénieur en Chef des Manufactures de l'État

Une expérience dans laquelle figurent toutes les combinaisons des variantes (ou niveaux) de p facteurs chacun à 2 variantes est dite "expérience factorielle 2^p". Elle conduit à 2^p résultats (qui peuvent être des moyennes si l'expérience comporte des répétitions). Une fraction de l'expérience, dans laquelle q facteurs ne sont pas pris en considération, constitue un plan factoriel 2^{p-q}.

L'analyse statistique d'une expérience factorielle 2^p consiste à calculer et à tester 2^p - 1 "comparaisons indépendantes", chaque comparaison s'exprimant par une fonction linéaire, convenablement choisie, des résultats. Les comparaisons portent le nom d'"effets principaux" et d'"interactions".

Soient par exemple 3 facteurs A, B, C. On symbolise les 2³ = 8 résultats de la façon suivante :

Résultat	Niveau de A	Niveau de B	Niveau de C
(1)	bas	bas	bas
a	haut	bas	bas
b	bas	haut	bas
ab	haut	haut	bas
c	bas	bas	haut
ac	haut	bas	haut
bc	bas	haut	haut
abc	haut	haut	haut

L'"effet principal de A" est la différence entre les résultats "niveau haut" et "niveau bas" de A :

$$A = (a + ab + ac + abc) - ((1) + b + c + bc)$$

$$A = (a - 1) (b + 1) (c + 1)$$

De même, les effets principaux de B et C s'écrivent symboliquement :

$$B = (a + 1) (b - 1) (c + 1)$$

$$C = (a + 1) (b + 1) (c - 1)$$

L'"interaction entre A et B" est la différence entre l'effet de A au niveau haut de B et l'effet de A au niveau bas de B :

$$AB = (abc + ab - bc - b) - (ac + a - c - (1))$$

$$AB = (a - 1) (b - 1) (c + 1)$$

De même, les interactions AC et BC s'écrivent symboliquement :

$$AC = (a - 1) (b + 1) (c - 1)$$

$$BC = (a + 1) (b - 1) (c - 1)$$

Enfin l'interaction du 2ème ordre ABC est la différence entre l'interaction AB au niveau haut de C et l'interaction AB au niveau bas de C. Elle s'écrit symboliquement :

$$ABC = (a - 1) (b - 1) (c - 1)$$

D'une façon générale, avec p facteurs, effets principaux et interactions s'écrivent symboliquement :

$$(a \pm 1) (b \pm 1) (c \pm 1) \dots (1)$$

Pour un effet principal, on met le signe - dans la parenthèse contenant la lettre correspondant au facteur considéré, le signe + dans toutes les autres.

Pour une interaction du 1er ordre (2 facteurs), on met le signe - dans les parenthèses contenant les lettres correspondant aux facteurs considérés, le signe + dans toutes les autres.

Et ainsi de suite jusqu'à l'interaction d'ordre $2^p - 1$ (tous facteurs réunis) qu'on obtient en mettant le signe - dans toutes les parenthèses.

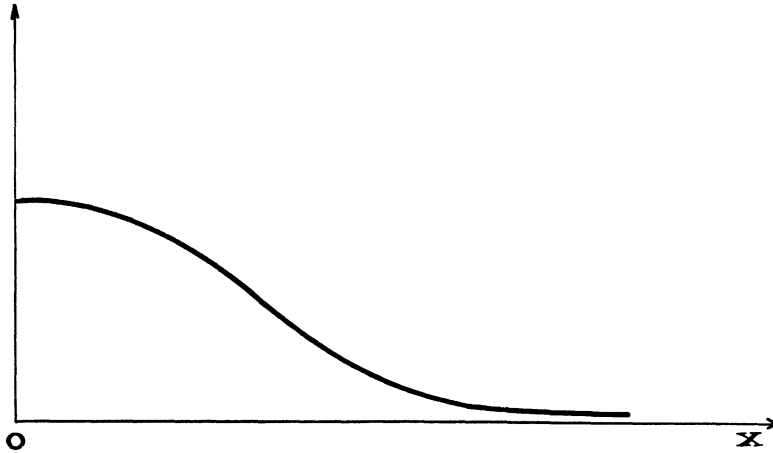
On vérifie facilement que les $2^p - 1$ combinaisons linéaires de résultats ainsi formées sont mutuellement orthogonales.

Les 2^p résultats étant supposés indépendants, normalement distribués, avec même variance σ^2 , chaque effet principal et chaque interaction est également une variable normale, dont la variance est $\sigma_c^2 = 2^p \sigma^2$. Naturellement, les valeurs moyennes de ces variables ne sont pas les mêmes si certains des facteurs ont une action réelle.

Supposons qu'aucun des facteurs n'ait d'action, et prenons pour les effets principaux et les interactions les valeurs absolues des expressions précédemment définies (on observera qu'on a : $AB = -BA$, $ABC = BCA = CAB = -ACB = -CBA = -BAC$, etc.). Alors chaque effet principal et chaque interaction est une variable dont la loi de distribution est représentée par la partie comprise entre 0 et $+\infty$ d'une variable normale de moyenne nulle et variance σ_c^2 ; cette loi de distribution (loi semi-normale) a pour densité de probabilité :

(1) Nous prenons ici les définitions de l'auteur. Généralement, effets principaux et interactions sont définis comme des différences moyennes ramenées à l'unité d'un résultat :

$$\frac{1}{2^{p-1}} (a \pm 1) (b \pm 1) (c \pm 1) \dots$$



$$f(x) = \frac{2}{\sigma_c \sqrt{2}} e^{-\frac{x^2}{2\sigma_c^2}} \quad (0 < x < +\infty)$$

Cette loi n'est autre que la loi de l'étendue w_2 (range) de deux valeurs extraites indépendamment d'une loi normale de moyenne nulle et écart-type $\frac{\sigma_c}{\sqrt{2}}$.

La fonction de distribution de la loi de w_2 a été mise en tables, de sorte qu'il est possible de porter sur un "papier de probabilité" les graduations suivantes :

abscisses : valeurs de la variable (0 à $+\infty$)

ordonnées : valeurs de la fonction de distribution d'une "loi semi-normale" correspondant aux valeurs de la variable portées en abscisses.

Ce papier pourra s'utiliser comme le papier à "échelle gaussienne" qui permet de juger de la normalité d'une distribution expérimentale (méthode de la droite de Henry). En particulier, si les points expérimentaux correspondant aux valeurs trouvées pour les effets principaux et les interactions d'une expérience factorielle 2^p se répartissent au voisinage immédiat d'une ligne droite, on pourra conclure qu'aucun des facteurs étudiés n'a d'action.

Dans une loi semi-normale, la valeur de la fonction de distribution pour $x = \sigma_c$ est :

$$\int_0^{\sigma_c} \frac{2}{\sigma_c \sqrt{2}} e^{-\frac{x^2}{2\sigma_c^2}} = \int_{-\sigma_c}^{\sigma_c} \frac{1}{\sigma_c \sqrt{2}} e^{-\frac{x^2}{2\sigma_c^2}} = 0,683$$

ce qui permet d'estimer graphiquement σ_c au moyen du papier de probabilité défini ci-dessus.

Lorsqu'un plan d'expérience 2^p ne comporte pas de répétition, l'erreur attachée à un résultat doit être estimée à partir des comparaisons (interactions ou effets principaux) considérées comme n'ayant pas d'existence réelle ; les valeurs trouvées pour ces comparaisons sont alors traitées comme des "termes d'erreur" (error contrasts).