

REVUE DE STATISTIQUE APPLIQUÉE

C. DUBY

Y. MOUGEY

J. ULMO

Analyse de plans d'expériences à deux facteurs contrôlés quand l'écart type est une fonction connue de la moyenne

Revue de statistique appliquée, tome 23, n° 1 (1975), p. 5-33

http://www.numdam.org/item?id=RSA_1975__23_1_5_0

© Société française de statistique, 1975, 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/>

ANALYSE DE PLANS D'EXPÉRIENCES A DEUX FACTEURS CONTROLÉS QUAND L'ÉCART TYPE EST UNE FONCTION CONNUE DE LA MOYENNE ⁽¹⁾

C. DUBY, Y. MOUGEY, J. ULMO

Les laboratoires industriels sont souvent amenés, pour le développement de leurs produits, à effectuer des expérimentations mettant en oeuvre des méthodes de mesures normalisées. Ces méthodes, bien connues, ont fait l'objet avant publication, d'études au cours desquelles est notamment déterminée la répétabilité du résultat de la mesure. Il se trouve que cette répétabilité — et donc l'écart-type de l'erreur de mesure — est presque toujours donnée en fonction de la valeur de la caractéristique mesurée. On trouvera de nombreux cas de ce type dans les recueils de méthodes d'essais normalisés [0].

Par ailleurs, quand de telles méthodes sont utilisées, les conditions expérimentales sont suffisamment bien contrôlées pour que l'erreur résiduelle soit essentiellement constituée par l'erreur de mesure.

Jusqu'à présent, l'analyse des résultats expérimentaux ainsi obtenus s'effectuait en appliquant les méthodes classiques d'analyse de variance qui — on le sait — supposent constante la variance d'erreur dans le domaine expérimental étudié. La question se pose de savoir dans quelle mesure les conclusions obtenues par une telle pratique sont valables.

Pour répondre à cette question, il faut définir une méthode d'analyse mieux adaptée au problème posé.

Pour cela nous nous placerons dans le cas de 2 facteurs contrôlés en sachant que ce cas peut aisément être étendu.

Si μ_{ij} est la valeur de la caractéristique mesurée pour les niveaux A_i et B_j des facteurs contrôlés A et B, on dispose seulement des mesures y_{ijm} de cette valeur, m variant de 1 à n_{ij} , si on a procédé à $n_{ij} \geq 1$ essais dans les conditions (A_i, B_j)

On peut considérer que les y_{ijm} sont n_{ij} variables aléatoires indépendantes de même loi, de moyenne μ_{ij} et d'écart type $\sigma_{ij} = f(\mu_{ij})$ où f est une fonction connue.

La loi de y_{ijm} sera supposée normale, ce qui paraît légitime compte tenu de la nature des fluctuations résiduelles.

Le modèle statistique adopté est le modèle classique

$$y_{ijm} = \mu_{ij} + \epsilon_{ijm}$$

(1) Article remis en Avril 1974

les erreurs ϵ_{ijm} étant indépendantes entre elles, normales, de moyenne nulle et d'écart type $\sigma_{ij} = f(\mu_{ij})$

Les problèmes que l'on cherche à résoudre sont :

- estimation des paramètres μ_{ij}
- tests d'absence d'effet d'un des facteurs contrôlés et d'absence d'interaction entre les 2 facteurs.

La méthode proposée pour l'estimation des μ_{ij} est la méthode du maximum de vraisemblance. Les propriétés des estimateurs et des tests qu'on a proposés et explicités complètement dans le cas simple où σ_{ij} est fonction affine de la moyenne ne sont vraies qu'asymptotiquement par rapport à l'ensemble des n_{ij} .

C'est pourquoi on a procédé à une étude par simulation de la validité des résultats trouvés et de ceux que donnent les méthodes classiques (qui supposent σ_{ij} constant) dans des conditions expérimentales correspondant aux problèmes qui ont été posés.

PREMIERE PARTIE

RECHERCHE D'UNE METHODE D'ANALYSE DES RESULTATS

I 1 – ESTIMATION DES PARAMETRES μ_{ij}

La méthode à laquelle nous avons pensé est la méthode du maximum de vraisemblance (M.V.) ; c'est du reste, elle qui est utilisée dans le cas classique où :

$$\sigma_{ij} = \sigma = c^{te}$$

I 1.1 – RECHERCHE DES ESTIMATEURS DU MAXIMUM DE VRAISEMBLANCE $\hat{\mu}_{ij}$

Compte-tenu de l'hypothèse d'indépendance des Y_{ijm} , la vraisemblance de l'ensemble des μ_{ij} correspondant à l'ensemble $\vec{y} \in \mathbb{R}^n$ des n résultats y_{ijm} obtenus est

$$\mathcal{L}[y, \{\mu_{ij}/(i, j) \in \mathcal{A}\}] = \prod_{(i, j) \in \mathcal{A}} \left[\frac{1}{(\sqrt{2\pi})^{n_{ij}}} \frac{1}{[f(\mu_{ij})]^{n_{ij}}} \exp - \frac{1}{2} \sum_{m=1}^{n_{ij}} \left(\frac{y_{ijm} - \mu_{ij}}{f(\mu_{ij})} \right)^2 \right]$$

où $\mathcal{A}$ désigne l'ensemble des combinaisons (i, j) effectivement essayées. $\mathcal{A}$ peut n'être qu'une partie de l'ensemble

$$I \times J = \{1, 2, \dots, k\} \times \{1, 2, \dots, l\}$$

des combinaisons possibles. Quand, comme ici, on ne fait aucune hypothèse sur la structure des μ_{ij} , c'est-à-dire sur la nature des effets des traitements (A_i, B_j) , on ne peut estimer que les μ_{ij} tels que $(i, j) \in \mathcal{A}$.

Cette vraisemblance est le produit

$$\prod_{(i,j) \in \alpha} \mathcal{L}(y_{ij1}, y_{ij2}, \dots, y_{ijn_{ij}}; \mu_{ij})$$

des vraisemblances de chacun des μ_{ij} pour l'ensemble des n_{ij} résultats y_{ijm} correspondants.

$\hat{\mu}_{ij}$ est donc l'estimateur du M.V. de μ_{ij} correspondant au sous ensemble d'observations $y_{ij1}, y_{ij2}, \dots, y_{ijn_{ij}}$ effectués avec le seul traitement (A_i, B_j) .

Les $\hat{\mu}_{ij}$ sont chacun fonction des seules observations $y_{ij1}, y_{ij2}, \dots, y_{ijn_{ij}}$ — de même indice (i, j) . Ils sont, par suite, indépendants en probabilité.

I 1.2 — Etude des $\hat{\mu}_{ij}$

I 1.2.1 — Pour des n_{ij} finis

Contrairement à ce qui se passe dans le cas habituel $\sigma_{ij} = c^{te}$, les $\hat{\mu}_{ij}$ ne sont pas des estimateurs exhaustifs des μ_{ij} si σ_{ij} dépend effectivement de μ_{ij} . Ils ne sont par suite pas non plus efficaces comme dans le cas habituel (sans biais et de variance la plus petite possible) [1 chap. 4].

Puisque μ_{ij} n'est fonction que du sous-ensemble $y_{ij1}, y_{ij2}, \dots, y_{ijn_{ij}}$ des observations correspondant au traitement (A_i, B_j) , il suffit de montrer qu'il n'existe pas de résumé exhaustif de ces observations pour μ_{ij} .

On sait en effet (Théorème de Darmais-Koopman [1] et [2]) que si Y est une variable aléatoire réelle dont la densité $g(y, \mu)$ dépend du paramètre réel μ et dont le domaine de variation ne dépend pas de μ , une condition nécessaire et suffisante pour qu'un échantillon de taille $n > 1$ de Y admette un résumé exhaustif pour μ est que g soit de forme exponentielle, c'est-à-dire que l'on puisse écrire

$$\text{Log } [g(y, \mu)] = a(y) \alpha(\mu) + b(y) + \beta(\mu)$$

Ici

$$\text{Log } [g(y_{ij}, \mu_{ij})] = -\text{Log } (\sqrt{2\pi}) - \text{Log } [f(\mu_{ij})] - \frac{1}{2} \left(\frac{y_{ij} - \mu_{ij}}{f(\mu_{ij})} \right)^2$$

ne peut se mettre sous la forme ci-dessus à moins que $f(\hat{\mu}_{ij}) = c^{te}$

I 1.2.2 Existence et propriétés asymptotiques des $\hat{\mu}_{ij}$

Si f est une fonction deux fois continûment dérivable de μ_{ij} sur l'ensemble Θ des valeurs possibles pour μ_{ij} (Θ est l'ensemble des μ tels que $f(\mu) > 0$), ce qu'on supposera, la densité $L_{ij}(y_{ij}, \mu_{ij})$ de Y_{ij} satisfait aux conditions de régularité qui justifient la recherche de μ_{ij} comme solution de l'équation de vraisemblance.

$$(1_{i,j}) \quad \frac{\partial \text{Log } L_{ij}(\mu_{ij})}{\partial \mu_{ij}} = \frac{\partial \text{Log } \mathcal{L}(y_{ij1}, y_{ij2}, \dots, y_{ijn_{ij}}; \mu_{ij})}{\partial \mu_{ij}} = 0 ;$$

on impose en outre la condition :

$$\frac{\partial^2 \text{Log } L_{ij}(\hat{\mu}_{ij})}{\partial \mu_{ij}^2} < 0$$

qui assure un maximum (relatif) de la vraisemblance.

$$(1_{i,j}) \text{ s'écrit : } -n_{ij} \frac{f'(\mu_{ij})}{f(\mu_{ij})} + \sum_{m=1}^{n_{ij}} \frac{(y_{ijm} - \mu_{ij})}{[f(\mu_{ij})]^2} + \sum_{m=1}^{n_{ij}} \frac{(y_{ijm} - \mu_{ij})^2}{[f(\mu_{ij})]^3} f'(\mu_{ij}) = 0$$

On trouvera en [1] et [2] l'énoncé de ces conditions de régularité qui portent sur les variables aléatoires $\frac{\partial}{\partial \mu_{ij}} \text{Log L}(Y_{ij}, \mu_{ij})$ et $\frac{\partial^2}{\partial \mu_{ij}^2} \text{Log L}(Y_{ij}, \mu_{ij})$.

Ces conditions sont réalisées puisque la forme développée de $(1_{i,j})$ montre que ces variables peuvent se mettre sous la forme

$$A(\mu_{ij}) + B(\mu_{ij})(Y_{ij} - \mu_{ij}) + C(\mu_{ij})(Y_{ij} - \mu_{ij})^2$$

avec $A(\mu_{ij})$, $B(\mu_{ij})$, $C(\mu_{ij})$, finies et continues pour tout $\mu_{ij} \in \Theta$.

Dans ces conditions, il existe une suite de racines $\hat{\mu}_{ij}(n_{ij})$ de l'équation $(1_{i,j})$ qui, quand $n_{ij} \rightarrow \infty$ converge presque sûrement vers la vraie valeur μ_{ij}^0 .

Si $\hat{\mu}_{ij}(n_{ij})$ est une suite de racines de l'équation $(1_{i,j})$ qui converge presque sûrement vers μ_{ij} , à partir d'une certaine valeur de n_{ij} , $\hat{\mu}_{ij}(n_{ij})$ réalise un maximum relatif de la vraisemblance correspondant à l'échantillon.

Enfin, toute suite $\hat{\mu}_{ij}(n_{ij})$ de racines de l'équation $(1_{i,j})$ convergeant p.s. vers μ_{ij}^0 , est telle que $\sqrt{n_{ij}}(\hat{\mu}_{ij} - \mu_{ij}^0)$ converge en loi vers une variable normale centrée d'écart-type $\frac{1}{\sqrt{I(\mu_{ij}^0)}}$ ou $I(\mu_{ij})$ est la quantité d'information de Fisher fournie par Y_{ij} sur μ_{ij} .

Les estimateurs $\hat{\mu}_{ij}(n_{ij})$ du M.V. des μ_{ij} sont donc asymptotiquement sans biais, asymptotiquement efficaces et asymptotiquement normaux.

I 1.3 – Calcul de $I(\mu_{ij})$

La variance asymptotique de $\hat{\mu}_{ij}(n_{ij})$ est, on l'a vu, égale à $\frac{1}{n_{ij} I(\mu_{ij})}$, borne inférieure de Fréchet, Darmois, Cramer, Rao de la variance de tout estimateur sans biais de μ_{ij} , basé sur un échantillon de taille n_{ij} de Y_{ij} .

$$I(\mu_{ij}) = V \left[\frac{\partial}{\partial \mu_{ij}} \text{Log L}(Y_{ij}, \mu_{ij}) \right] = V \left[\frac{Y_{ij} - \mu_{ij}}{[f(\mu_{ij})]^2} + \frac{(Y_{ij} - \mu_{ij})^2}{[f(\mu_{ij})]^3} f'(\mu_{ij}) \right]$$

V signifiant "variance de".

Comme

$$\text{cov} [(Y_{ij} - \mu_{ij}), (Y_{ij} - \mu_{ij})^2] = E [Y_{ij} - \mu_{ij}]^3 = 0$$

$$I(\mu_{ij}) = \frac{1}{[f(\mu_{ij})]^4} V [Y_{ij} - \mu_{ij}] + \frac{[f'(\mu_{ij})]^2}{[f(\mu_{ij})]^6} V [(Y_{ij} - \mu_{ij})^2]$$

mais

$$V[Y_{ij} - \mu_{ij}] = \sigma_{ij}^2 = [f(\mu_{ij})]^2$$

et comme $\frac{(Y_{ij} - \mu_{ij})^2}{[f(\mu_{ij})]^2}$, carré d'une variable normale centrée réduite suit une loi de χ^2 à 1 degré de liberté, dont on sait que la variance est égale à 2.

$$\frac{V[Y_{ij} - \mu_{ij}]^2}{[f(\mu_{ij})]^4} = 2$$

On a donc

$$I(\mu_{ij}) = \frac{1 + 2 [f'(\mu_{ij})]^2}{[f(\mu_{ij})]^2}$$

I 1.4 – Application au cas particulier $\sigma_{ij} = a(\mu_{ij} - m_o)$ avec a et m_o connus.

La condition $\sigma_{ij} > 0$ implique $\mu_{ij} > m_o$ si $a > 0$ et $\mu_{ij} < m_o$ si $a < 0$

Dans le cas le plus fréquent où on a $a > 0$, l'ensemble des valeurs possibles pour les μ_{ij} est donc l'intervalle ouvert $] m_o, \infty [$

I 1.41 – Il est facile de se ramener au cas où $\sigma_{ij} = a\mu_{ij}$.

En effet si on pose $Z_{ijm} = Y_{ijm} - m_o$ et $\theta_{ij} = \mu_{ij} - m_o$, on a

$$Z_{ijm} \in N(\theta_{ij}, a\theta_{ij}),$$

les Z_{ijm} étant indépendantes.

Puisque

$$\mathcal{L}_{Z, \mu}(z_{ijm}, \mu_{ij}) = \mathcal{L}_{Y, \mu}(y_{ijm}, \mu_{ij})$$

$\mathcal{L}_{Z, \mu}$ et $\mathcal{L}_{Y, \mu}$ désignant respectivement les densités des variables Z_{ij} et Y_{ij} exprimées en fonction de μ_{ij} ,

$$\begin{aligned} \mathcal{L}_{Y, \mu}(y_{ij1}, y_{ij2}, \dots, y_{ijn_{ij}}; \mu_{ij}) &= \mathcal{L}_{Z, \mu}(z_{ij1}, z_{ij2}, \dots, z_{ijn_{ij}}; \mu_{ij}) \\ &= \mathcal{L}_{Z, \theta}(z_{ij1}, z_{ij2}, \dots, z_{ijn_{ij}}; \theta_{ij}) \end{aligned}$$

avec

$$\theta_{ij} = \mu_{ij} - m_o$$

L'estimateur du M.V. de μ_{ij} déduit de l'échantillon des n_{ij} réalisations y_{ijm} de Y_{ij} est donc le même que celui déduit de l'échantillon correspondant des n_{ij} réalisations z_{ijm} de Z_{ij} et on a $\hat{\theta}_{ij} = \hat{\mu}_{ij} - m_o$.

Il suffit donc de considérer les observations z_{ijm} de Z_{ij} d'écart-type $\sigma_{ij} = a\theta_{ij}$

Dans la suite de l'exposé on posera donc $\sigma_{ij} = a\mu_{ij}$ et on désignera par Y_{ij} et μ_{ij} ce qui, dans le paragraphe précédent, était appelé Z_{ij} et θ_{ij}

I 1.4.2 – Détermination de $\hat{\mu}_{ij}$ dans le cas où $\sigma_{ij} = a\mu_{ij}$

$$L_{ij}(\mu_{ij}) = \frac{1}{(\sqrt{2\pi})^{n_{ij}}} \frac{1}{(a\mu_{ij})^{n_{ij}}} \exp \left[-\frac{1}{2} \sum_{m=1}^{n_{ij}} \frac{(y_{ijm} - \mu_{ij})^2}{a^2 \mu_{ij}^2} \right]$$

et l'équation (1_{i,j}) s'écrit

$$n_{ij} a^2 \hat{\mu}_{ij}^2 + \hat{\mu}_{ij} \left(\sum_{m=1}^{n_{ij}} y_{ijm} \right) - \sum_{m=1}^{n_{ij}} y_{ijm}^2 = 0 \quad (2)$$

si $a > 0$ on doit se limiter aux solutions $\hat{\mu}_{ij} > 0$ et si $a < 0$ aux solutions $\hat{\mu}_{ij} < 0$

On voit facilement que l'équation (2) admet 2 racines de signes contraires. On choisira celle qui a le même signe que $\mathcal{A}$

Comme

$$\begin{aligned} L_{ij}(\mu_{ij}) &\rightarrow 0 \quad \text{si} \quad \mu_{ij} \rightarrow 0 \\ L_{ij}(\mu_{ij}) &\rightarrow 0 \quad \text{si} \quad |\mu_{ij}| \rightarrow \infty \end{aligned}$$

et

$$L_{ij}(\mu_{ij}) > 0 \quad \text{pour} \quad a\mu_{ij} \in]0, \infty[$$

la solution choisie réalise un maximum de $L_{ij}(\mu_{ij})$ sur l'ensemble Θ des μ_{ij} tels que $a\mu_{ij} > 0$.

Donc il existe un estimateur du M.V. de μ_{ij} ; c'est la solution de l'équation (2) qui a même signe que a

Il s'exprime en fonction des deux statistiques $\sum_{m=1}^{n_{ij}} Y_{ijm}$ et $\sum_{m=1}^{n_{ij}} Y_{ijm}^2$, tandis que dans le cas classique $\sigma_{ij} = c^{\text{te}}$, l'estimateur du M.V. de μ_{ij} est

$$Y_{ij.} = \frac{1}{n_{ij}} \sum_{m=1}^{n_{ij}} Y_{ijm},$$

et n'est donc fonction que de la seule statistique $\sum Y_{ijm}$.

$$I(\mu_{ij}) = \frac{1 + 2a^2}{a^2 \mu_{ij}} \quad \text{en sorte que la variance asymptotique de } \hat{\mu}_{ij} \text{ est}$$

$$\frac{1}{n_{ij} I(\mu_{ij})} = \frac{a^2}{1 + 2a^2} \frac{\mu_{ij}^2}{n_{ij}}$$

I 2 – RECHERCHE DE TESTS DE L'HYPOTHESE D'ABSENCE D'EFFET D'UN FACTEUR CONTROLE OU D'ABSENCE D'INTERACTION ENTRE LES FACTEURS

I 2.1 – Formulation générale du problème et des méthodes de test envisagées

Soit p le cardinal de $\mathcal{A}$, c'est-à-dire le nombre de combinaisons (A_i, B_j) effectivement essayées et $\underline{\mu} \in \mathbb{R}^p$ le vecteur dont les composantes sont les μ_{ij} , $(i, j) \in \mathcal{A}$ rangés dans l'ordre $\mu_{11}, \mu_{12}, \dots, \mu_{1l}, \mu_{21}, \mu_{22}, \dots, \mu_{2l}, \dots, \mu_{k1}, \dots, \mu_{kl}$

Les hypothèses qu'on se propose de tester peuvent s'exprimer par un ensemble de q relations du type

$$\varphi_1(\underline{\mu}) = \varphi_2(\underline{\mu}) = \dots \varphi_q(\underline{\mu}) = 0$$

où $\varphi_1, \varphi_2, \dots, \varphi_q$ sont des fonctions à valeur réelles de $\underline{\mu}$. C'est ainsi que l'hypothèse H_B d'absence d'effet du facteur B qui s'écrit, si toutes les combinaisons (A_i, B_j) ont été essayées,

$$\mu_{i1} = \mu_{i2} = \mu_{i1} \quad \forall i$$

peut, par exemple se mettre sous la forme

$$\mu_{ij} - \mu_{i1} = 0 \quad \forall j \neq 1 \quad \text{et} \quad \forall i, \text{ soit}$$

$$\varphi_{ij}(\underline{\mu}) = 0 \quad \forall j \neq 1 \quad \text{et} \quad \forall i \quad \text{en posant} \quad \varphi_{ij}(\underline{\mu}) = \mu_{ij} - \mu_{i1}.$$

Soit $\varphi(\underline{\mu})$ le vecteur de R^q de composantes $\varphi_1, \varphi_2, \dots, \varphi_q$.

Quant, à partir d'un échantillon de taille n d'une variable aléatoire X , réelle ou vectorielle dont la densité de probabilité dépend d'un paramètre vectoriel $\underline{\mu}$ on se propose de tester l'hypothèse $H : \varphi(\underline{\mu}) = 0$ contre l'alternative la plus générale soit $K : \varphi(\underline{\mu})$ quelconque, on dispose de deux tests asymptotiques, basés sur l'estimation de $\underline{\mu}$ par la méthode du maximum de vraisemblance (cf. [1] chap. 5 par 7).

Il faut donc supposer que la densité $g(x, \underline{\mu})$ de X vérifie les conditions de régularité à la base des propriétés asymptotiques des estimateurs de M.V. (cf. [1] chap. 4 ou [2]).

Ces tests supposent que les composantes $\varphi_i(\underline{\mu})$ de $\varphi(\underline{\mu})$ satisfont aux conditions suivantes :

- elles possèdent pour tout $\underline{\mu}$ possible des dérivées partielles premières et secondes uniformément continues et bornées
- elles sont fonctionnellement indépendantes, c'est-à-dire que la matrice jacobienne

$$\left\| \frac{D(\varphi_1, \varphi_2, \dots, \varphi_q)}{D(\mu_{11}, \mu_{12}, \dots, \mu_{k1})} \right\|$$

est de rang maximal $q \leq p$ pour toute valeur possible de $\underline{\mu}$

Ces deux tests sont les suivants :

I 2.1.1 — Test du rapport des vraisemblances maximales.

Soit $x^{(n)}$ l'ensemble des n observations de X constituant l'échantillon et

$$L_n(x^{(n)}, \underline{\mu}) \text{ la densité de } x^{(n)}$$

Soit

$$\lambda_x(x^{(n)}) = \frac{\sup_{\underline{\mu} \in H} L_n}{\sup_{\underline{\mu} \in K} L_n}$$

où " $\underline{\mu} \in H$ " signifie que $\underline{\mu}$ satisfait à H
et " $\underline{\mu} \in K$ " signifie que $\underline{\mu}$ satisfait à K et est donc quelconque.

Si l'hypothèse nulle H est réalisée la statistique

$$Q_n = -2 \text{ Log } \lambda_n(X^{(n)})$$

converge en loi vers un χ^2 à q degrés de liberté quand n tend vers l'infini et les tests de régions de rejet définies par $-2 \log \lambda_n(x^{(n)}) \geq c$ où c est une constante dépendant du seuil de signification α adopté, sont convergents (leur puissance tend vers 1 quand $n \rightarrow \infty$ pour toute valeur de $\underline{\mu}$ ne satisfaisant pas à H).

I 2.1.2 -- Test de Wald

Soit $\hat{\underline{\mu}}_{(n)}$ l'estimateur du M.V. de $\underline{\mu}$ sous l'hypothèse K, correspondant à un échantillon de taille n .

On sait que la matrice $I_n^{-1}(\underline{\mu})$ inverse de la matrice d'information $I_n(\underline{\mu})$ de $\underline{\mu}$ correspondant à un échantillon de taille n ($I_n(\underline{\mu}) = n I_1(\underline{\mu}) = n I(\underline{\mu})$) peut être considérée comme la matrice de covariance asymptotique de $\hat{\underline{\mu}}_{(n)}$.

Pour éviter les notations à double indice, on suppose dans la suite de cet énoncé que $\underline{\mu}$ est de dimension p et a pour composantes μ_i , $i = 1, 2, \dots, p$.

Soit

$$b_{ij}^{(n)}(\underline{\mu}) = \sum_{r=1}^p \sum_{s=1}^p \frac{\partial \varphi_i(\underline{\mu})}{\partial \mu_r} \cdot \frac{\partial \varphi_j(\underline{\mu})}{\partial \mu_s} I_{rs}^{-1}(\underline{\mu})_{(n)}$$

où $I_{rs}^{-1}(\underline{\mu})_{(n)}$ désigne l'élément (r, s) de la matrice $I_n^{-1}(\underline{\mu})$

Les $b_{ij}^{(n)}$ peuvent être interprétés comme les covariances asymptotiques des statistiques $\varphi_i(\hat{\underline{\mu}})$ et $\varphi_j(\hat{\underline{\mu}})$. Cela résulte de l'application de la formule des accroissements finis aux $\varphi_i(\hat{\underline{\mu}}_1, \hat{\underline{\mu}}_2, \dots, \hat{\underline{\mu}}_p)$ au voisinage de $\varphi_i(\mu_1, \mu_2, \dots, \mu_p)$ (Cf. [5], chap. 6, par. a.2).

Soit $B^{(n)}(\underline{\mu})$ la matrice $q \times q$ d'éléments $b_{ij}^{(n)}(\underline{\mu})$

Puisque $\hat{\underline{\mu}}$ a une distribution asymptotiquement normale, il en est de même pour $\underline{\varphi}(\hat{\underline{\mu}})$ avec une moyenne $\underline{\varphi}(\underline{\mu})$ et une matrice de covariance $B^{(n)}(\underline{\mu})$. (Cf. [5], Chap. 6, par a.2).

On en déduit que

$$(\underline{\varphi}(\hat{\underline{\mu}}) - \underline{\varphi}(\underline{\mu}))' [B^{(n)}(\underline{\mu})]^{-1} (\underline{\varphi}(\hat{\underline{\mu}}) - \underline{\varphi}(\underline{\mu}))$$

(où ' est un symbole de transposition) a une distribution asymptotique de χ^2 à q degrés de liberté.

Par suite si $H : \underline{\varphi}(\underline{\mu}) = 0$ est réalisée, $(\underline{\varphi}(\hat{\underline{\mu}}))' [B^{(n)}(\underline{\mu})]^{-1} \underline{\varphi}(\hat{\underline{\mu}})$ a une distribution asymptotique de χ^2 , tandis que si H n'est pas réalisée, elle tend presque sûrement vers l'infini avec n .

$$B^{(n)}(\underline{\mu}) = \frac{1}{n} B^{(1)}(\underline{\mu}) \rightarrow [B^{(n)}(\underline{\mu})]^{-1} = n [B^{(1)}(\underline{\mu})]^{-1}$$

$\underline{\mu}$ étant inconnu ce résultat n'est pas intéressant pour la pratique mais Wald a montré que la statistique

$$W_n(\underline{\mu}) = (\underline{\varphi}(\hat{\underline{\mu}}))' [B^{(n)}(\hat{\underline{\mu}})]^{-1} \underline{\varphi}(\hat{\underline{\mu}})$$

converge aussi en loi vers un χ^2 à q degrés de liberté si H est réalisée et que les tests de régions de rejet définies par $W_n(\hat{\underline{\mu}}) \geq c$, où c est une constante dépendant du seuil de signification α choisi, sont convergents. Si H n'est pas réalisée $W_n(\hat{\underline{\mu}}) \xrightarrow{p.s.} \infty$.

I 2.2. — Application aux plans d'expériences

I 2.2.1 — Cas particulier $n_{ij} = n$ pour tout $(i, j) \in \mathcal{A}$

L'extension est alors immédiate. Il suffit de prendre pour variable X des tests définis précédemment le vecteur $\underline{Y}$ de $\mathbf{R}^p$ de composantes Y_{ij} ; $(i, j) \in \mathcal{A}$

Puisque les Y_{ij} sont indépendants, $\underline{Y}$ suit une loi normale de moyenne $\underline{\mu}$ de matrice de covariance Σ , diagonale, dont l'élément de la ligne (i, j) est $\sigma_{ij}^2 = [f(\mu_{ij})]^2$

I 2.2.2 — Cas général

Quand les n_{ij} ne sont pas tous égaux on ne peut plus définir un n échantillon de la variable vectorielle $\underline{Y}$. Il a donc fallu étendre les deux méthodes de test précédentes à ce cas. Nous nous sommes limités au cas où les composantes Y_{ij} de $\underline{Y}$ sont indépendantes en probabilité.

Pour pouvoir parler de convergence et de propriétés asymptotiques, il fallait définir un mode de croissance de l'ensemble des p valeurs n_{ij} vers l'infini.

On a supposé que la structure du plan restait la même, c'est-à-dire qu'on avait $n_{ij} = v p_{ij}$, les p_{ij} étant des constantes positives ou nulles définissant cette structure et v un entier qu'on fait tendre vers l'infini.

On aura donc $\forall (i, j) \in \mathcal{A}$ c'est-à-dire tel que $p_{ij} > 0$, $n_{ij} \rightarrow \infty$ si et seulement si $v \rightarrow \infty$. Dans ces conditions l'énoncé du test du rapport des vraisemblances maximales n'est pas modifié à condition de remplacer n par v et $x^{(n)}$ par $y^{(v)}$, ensemble des observations

$$y_{ijm}, m(i, j) = 1, 2, \dots, n_{ij} \text{ avec } n_{ij} \propto p_{ij}, \forall (i, j) \in \mathcal{A}$$

I 2.2.3 — Extension du théorème de Wald

Pour voir sous quelle forme il s'étend, il faut faire intervenir de façon plus explicite la taille n de l'échantillon dans l'expression de W_n .

Puisque $I_n(\underline{\mu}) = n I(\underline{\mu})$, $I(\underline{\mu})$ étant l'information sur $\underline{\mu}$ fournie par une seule observation de X .

$$I^{-1}(\underline{\mu}) = \frac{1}{n} I^{-1}(\underline{\mu}) \quad , \quad B^{(n)}(\underline{\mu}) = \frac{1}{n} B(\underline{\mu}) \quad \text{avec} \quad B(\underline{\mu}) = B^{(1)}(\underline{\mu})$$

et

$$W_n(\hat{\underline{\mu}}) = n [\varphi(\hat{\underline{\mu}})]' [(B(\hat{\underline{\mu}}))^{-1}] \varphi(\hat{\underline{\mu}})$$

Dans le cas présent, les composantes Y_{ij} de $\underline{Y}$ étant indépendantes on a $I(\underline{\mu}) =$ matrice diagonale d'élément $I(\mu_{ij})$

Soit $\hat{\underline{\mu}}_v$ le vecteur de composantes $\hat{\mu}_{ij}(n_{ij})$ avec $n_{ij} = v p_{ij}$ on sait que

$$(\hat{\mu}_{ij(v)} - \mu_{ij}) \sqrt{n_{ij}} \xrightarrow{\text{Loi}} N\left(0, \frac{1}{\sqrt{I(\mu_{ij})}}\right)$$

Les $\hat{\mu}_{ij}$ étaient indépendants on en conclut que

$$(\hat{\underline{\mu}}_v - \underline{\mu}) \sqrt{v} \xrightarrow{\text{Loi}} N(0, K(\underline{\mu}))$$

la matrice de covariance $K(\mu)$ étant une matrice diagonale d'éléments $\frac{1}{p_{ij} I(\mu_{ij})}$

Posons

$$B_1(\underline{\mu}) = \left(\frac{\partial \varphi}{\partial \mu}(\underline{\mu}) \right) K(\underline{\mu}) \left(\frac{\partial \varphi}{\partial \mu}(\underline{\mu}) \right)'$$

où $\frac{\partial \varphi}{\partial \mu}(\underline{\mu})$ est la matrice $(q \times p)$ de composantes $\frac{\partial \varphi_r}{\partial \mu_{ij}}$, $r \in \{1, 2, \dots, q\}$

$B_1(\underline{\mu})$ a pour élément (r, s) :

$$b_{1rs} = \sum_{ij} \frac{\partial \varphi_r(\mu)}{\partial \mu_{ij}} \frac{\partial \varphi_s(\mu)}{\partial \mu_{ij}} \frac{1}{p_{ij} I(\mu_{ij})}$$

On peut montrer (cf. [3]) que $v[\varphi(\hat{\mu}_v) - \varphi(\underline{\mu})]' [B_1(\hat{\mu}_v)]^{-1} [\varphi(\hat{\mu}_v) - \varphi(\underline{\mu})]$ converge en loi vers un χ^2 à q degrés de liberté quand $v \rightarrow \infty$.

Si $H : \varphi(\underline{\mu}) = 0$ est réalisée, $W_v(\hat{\mu}_v) = v [\varphi(\hat{\mu}_v)]' [B_1(\hat{\mu}_v)]^{-1} \varphi(\hat{\mu}_v)$ converge donc en loi vers un χ^2 à q degrés de liberté. C'est donc $W_v(\hat{\mu}_v)$ qui remplace la statistique $W_n(\hat{\mu})$ du test de Wald.

Quand H n'est pas réalisée, $W_v(\hat{\mu}) \xrightarrow{\text{p.s.}} \infty$ quand $v \rightarrow \infty$ et les tests de région critique $W_v(\hat{\mu}) \geq c$ sont convergents.

Nous avons choisi d'utiliser le test de Wald, de préférence au test du rapport des vraisemblances maximales, parce que le test de Wald est exclusivement basé sur le calcul de $\hat{\mu}$ estimation du M.V. de μ sous l'hypothèse la plus générale tandis que le test du rapport des vraisemblances maximales nécessite également le calcul de $\hat{\mu}_H$ estimateur de M.V. de μ quand on suppose H réalisée, et de la valeur de la vraisemblance correspondante pour l'échantillon. Ce calcul est notablement plus compliqué que celui de μ parce que les liaisons introduites introduites entre les μ_{ij} par les relations $\varphi(\mu) = 0$ ne permettent plus d'estimer séparément chaque μ_{ij} à partir des seules observations y_{ijm} correspondant à la combinaison (i, j) considérée.

On notera pour l'application pratique du test que v et les p_{ij} ne sont pas connus. Seul les produits $n_{ij} = v p_{ij}$ sont connus. Ceci est sans importance puisque la matrice $v B^{-1}(\hat{\mu})$ qui figure dans l'expression de $W(\hat{\mu})$ est l'inverse de la matrice $\frac{\partial \varphi}{\partial \mu}(\hat{\mu})' T(\hat{\mu}) \left[\frac{\partial \varphi}{\partial \mu}(\hat{\mu}) \right]$, où $T(\hat{\mu})$ est la matrice diagonale d'éléments

$$\frac{1}{n_{ij} I(\hat{\mu}_{ij})}.$$

On a ainsi

$$W(\hat{\mu}) = [\varphi(\hat{\mu})]' \left[\frac{\partial \varphi}{\partial \mu}(\hat{\mu})' T(\hat{\mu}) \left(\frac{\partial \varphi}{\partial \mu}(\hat{\mu}) \right)' \right]^{-1} \varphi(\hat{\mu})$$

1.2.3 – Mise en oeuvre pratique du test de Wald quand les fonctions $\varphi_i(\mu)$ sont des formes linéaires

Les $\varphi_i(\mu)$ sont en effet linéaires quand H exprime l'absence d'effet d'un facteur contrôlé comme on l'a vu en 1.2.1, ou l'absence d'interaction entre les facteurs comme on le verra plus loin. Il est possible dans ce cas d'utiliser un ar-

tifique pour calculer $W(\underline{\hat{\mu}})$ de façon relativement simple en utilisant les résultats de la théorie classique des modèles linéaires appliquée à $\underline{\hat{\mu}}$

En effet, si $\underline{\varphi}(\underline{\mu}) = A\underline{\mu}$ avec A matrice ($q \times p$) de rang q

$$W(\underline{\hat{\mu}}) = (\underline{\varphi}(\underline{\hat{\mu}}))' [A T(\underline{\hat{\mu}}) A']^{-1} \underline{\varphi}(\underline{\hat{\mu}})$$

On a vu qu'on avait asymptotiquement

$$\hat{\mu}_{ij} \in N(\mu_{ij}, \sqrt{t_{ij}(\mu_{ij})}) \quad \text{avec} \quad t_{ij}(\mu_{ij}) = \frac{1}{n_{ij} I(\mu_{ij})}$$

et que les $\hat{\mu}_{ij}$ étaient indépendants.

Nous allons faire comme si on pouvait écrire

$$\forall i, j : \hat{\mu}_{ij} = \mu_{ij} + F_{ij} \quad \text{avec} \quad F_{ij} \in N(0, \sqrt{\tau_{ij}})$$

les τ_{ij} étant connus et les F_{ij} indépendantes

Soit

$$l_{ij} = \frac{\hat{\mu}_{ij}}{\sqrt{\tau_{ij}}}$$

Les variables l_{ij} satisfont au modèle linéaire

$$(3) : l_{ij} = \lambda_{ij} + E_{ij} \quad \text{avec} \quad \lambda_{ij} = \frac{\mu_{ij}}{\sqrt{\tau_{ij}}} \quad \text{et} \quad E_{ij} \in N(0, 1) \text{ et indépendantes}$$

L'hypothèse $H : \underline{\varphi}(\underline{\mu}) = 0$ où les φ_i sont des formes linéaires indépendantes devient $\underline{\psi}(\underline{\lambda}) = 0$ où les ψ_i sont des formes linéaires, indépendantes, de $\underline{\lambda}$.

L'estimateur du M.V. de $\underline{\lambda}$ dans (3) est $\underline{l}$ de composantes l_{ij} et la théorie classique du modèle linéaire permet de tester l'hypothèse linéaire $H : \underline{\psi}(\underline{\lambda}) = 0$ en écrivant que la différence $S_H(\tau)$ des sommes de carrés résiduelle sous H et sous K suit un χ^2 à q degrés de liberté (σ est constant et égal à 1). K est l'hypothèse la plus générale : μ_{ij} quelconques mais tels que $f(\mu_{ij}) > 0$. En fait la somme de carrés résiduelle sous K est nulle puisque λ_{ij} est estimé par l_{ij} et $S_H(\tau)$ est la somme des carrés résiduelles sous H . On sait que

$$S_H(\tau) = (\underline{\psi}(\underline{l}))' [\text{cov}(\underline{\psi}(\underline{l}))]^{-1} \underline{\psi}(\underline{l})$$

(on teste l'hypothèse d'une moyenne nulle pour $\underline{\psi}(1) \in \mathbb{R}^q$ dont la distribution est normale).

Puisque

$$\underline{\varphi}(\underline{\mu}) = \underline{\psi}(\underline{\lambda}) \quad \forall \underline{\mu}, \quad \underline{\psi}(\underline{l}) = \underline{\varphi}(\underline{\hat{\mu}}) \quad \text{et}$$

$$\text{cov}(\underline{\psi}(\underline{l})) = \text{cov}(\underline{\varphi}(\underline{\hat{\mu}})) = \text{cov}(A\underline{\hat{\mu}}) = A \text{cov} \underline{\hat{\mu}} A' = A D_\tau A'$$

en désignant par D_τ la matrice diagonale d'éléments τ_{ij} .

On a donc

$$S_H(\tau) = (\underline{\varphi}(\underline{\hat{\mu}}))' [A D_\tau A']^{-1} \underline{\varphi}(\underline{\hat{\mu}})$$

On voit qu'en prenant

$$\tau_{ij} = t_{ij}(\hat{\mu}_{ij}) \quad \text{soit} \quad D_\tau = T(\underline{\hat{\mu}})$$

on obtient

$$S_H(T(\underline{\hat{\mu}})) = W(\underline{\hat{\mu}})$$

Le processus opératoire de test est donc le suivant

1) On calcule les estimations du M.V., $\hat{\mu}_{ij}$ des μ_{ij} sous l'hypothèse la plus générale K : μ_{ij} quelconques mais tels que $f(\mu_{ij}) > 0$.

2) On procède au test de l'hypothèse linéaire H dans le modèle linéaire

$$\frac{\hat{\mu}_{ij}}{\sqrt{\tau_{ij}}} = \frac{\mu_{ij}}{\sqrt{\tau_{ij}}} + E_{ij}, \quad (3)$$

$E_{ij} \in N(0, 1)$ et indépendantes, en prenant pour valeur des τ_{ij} supposés connus

$t_{ij}(\hat{\mu}_{ij}) = \frac{1}{n_{ij} I(\hat{\mu}_{ij})}$ qu'on désignera simplement par t_{ij} dans la suite de l'exposé.

$W(\hat{\mu})$ est égale à la statistique de test $S_H(T)$, somme des carrés résiduelle sous H, et sa distribution asymptotique sous H est celle de $S_H(T)$ sous H, c'est-à-dire un χ^2 à q degrés de liberté si H est d'ordre q

1.2.3.1 – Application au test de l'hypothèse H_B "pas d'effet du facteur B".

On suppose pour simplifier les notations que le plan est factoriel

$$(n_{ij} \neq 0, \forall (i, j) \in I \times J)$$

L'hypothèse H_B est d'ordre $q = k(l-1)$ et on peut l'exprimer sous la forme

$$\lambda_{i1} \sqrt{t_{i1}} = \lambda_{i2} \sqrt{t_{i2}} = \dots = \lambda_{il} \sqrt{t_{il}} = S_i \quad \forall i \quad (\text{cf. 1 2.1})$$

On a alors pour tout i le modèle linéaire

$$l_{ij} = \frac{S_i}{\sqrt{t_{ij}}} + E_{ij},$$

$E_{ij} \in N(0, 1)$ et indépendantes. L'estimation des moindres carrés $\tilde{S}_i$ de S_i correspondante est solution de l'équation normale

$$\sum_j \left(l_{ij} - \frac{\tilde{S}_i}{\sqrt{t_{ij}}} \right) \frac{1}{\sqrt{t_{ij}}} = 0$$

On a donc

$$\tilde{S}_i = \frac{\sum_j \frac{l_{ij}}{\sqrt{t_{ij}}}}{\sum_j \frac{1}{\sqrt{t_{ij}}}} \quad \text{et} \quad S_H(T) = \sum_{i,j} l_{ij}^2 - \sum_{i,j} \frac{\tilde{S}_i^2}{t_{ij}} = \sum_{i,j} l_{ij}^2 - \sum_i \left[\frac{\left(\sum_j \frac{l_{ij}}{\sqrt{t_{ij}}} \right)^2}{\sum_j \frac{1}{t_{ij}}} \right]$$

En revenant aux $\hat{\mu}_{ij}$, on obtient

$$W(\underline{\hat{\mu}}) = S_H(T) = \sum_{i,j} \frac{\hat{\mu}_{ij}^2}{t_{ij}} - \sum_i \left[\frac{\left(\sum_j \frac{\hat{\mu}_{ij}}{t_{ij}} \right)^2}{\sum_j \frac{1}{t_{ij}}} \right]$$

qui, sous H_B , suit asymptotiquement une loi de γ^2 à $K(l-1)$ degrés de liberté

Remarque : Si l'hypothèse H_B est acceptée on peut prendre $\tilde{S}_i$ pour estimateur de la valeur commune μ_i des μ_{ij} car on peut montrer qu'il converge p.s. vers μ_i sous H_B (cf. [3]). Il semble toutefois plus logique d'estimer μ_i par le M.V. en supposant H_B réalisée, c'est-à-dire que l'ensemble des $n_i = \sum_{j,m} n_{ij}$ observations y_{ijm} , ($j = 1, 2, \dots, l$) à même moyenne μ_i . L'équation de vraisemblance se déduit immédiatement de l'équation $(1_{i,j})$ donnant $\hat{\mu}_{ij}$ en y remplaçant μ_{ij} par μ_i , n_{ij} par n_i et $\sum_m$ par $\sum_{j,m}$.

1.2.3.2 — Test de l'hypothèse d'absence d'interaction entre les deux facteurs

Il n'est pas nécessaire d'avoir, comme dans le cas classique, $n_{ij} > 1$ pour au moins un couple (i, j) pour pouvoir procéder à ce test. Ceci est dû à ce qu'on n'a plus besoin d'estimer séparément l'écart type et la moyenne.

On suppose, là encore, que le plan réalisé est factoriel.

L'hypothèse H_{A+B} à tester peut se formuler

$$\mu_{ij} = \mu_o + \alpha_i + \beta_j \quad \forall (i, j) \quad (4)$$

Elle exprime que le vecteur μ de composantes μ_{ij} appartient au sous espace vectoriel V_{A+B} de R^{kl} défini par (4)

On sait (cf. [4]) que les équations (4) ne définissent pas μ_o , α_i et β_j de façon unique en fonction des μ_{ij} .

Pour que ceux-ci soient définis de façon unique on ajoutera les conditions

$$\sum_i a_i \alpha_i = 0 \quad \sum_j b_j \beta_j = 0 \quad (5)$$

les a_i et b_j étant quelconque mais tels que $\sum_i a_i \neq 0$ et $\sum_j b_j \neq 0$

Les conditions (5) généralisent les conditions de centrage habituelles

$$\sum_i \alpha_i = 0 \quad \sum_j \beta_j = 0$$

et un choix convenable des a_i et des b_j permettra de simplifier un peu la mise en oeuvre du test.

La résolution du système d'équations (4) + (5) en μ_o , α_i et β_j conduit aux expressions

$$\left\{ \begin{array}{l} \mu_o^* = \frac{\sum_{i,j} a_i b_j \mu_{ij}}{\left(\sum_i a_i\right)\left(\sum_j b_j\right)} = \mu_{..}^* \\ \alpha_i^* = \mu_{i.}^* - \mu_{..}^* \quad \text{avec} \quad \mu_{i.}^* = \frac{\sum_j b_j \mu_{ij}}{\sum_j b_j} \\ \beta_j^* = \mu_{.j}^* - \mu_{..}^* \quad \text{avec} \quad \mu_{.j}^* = \frac{\sum_i a_i \mu_{ij}}{\sum_i a_i} \end{array} \right. \quad (6)$$

voir note (1)

(1) L'astérisque indique qu'il s'agit de la solution particulière correspondant au système de poids a_i , b_j retenu.

En posant

$$\varphi_{ij}^*(\mu) = \mu_{ij} - \mu_{i.}^* - \mu_{.j}^* + \mu_{..}^*$$

l'hypothèse H_{A+B} prend la forme $\varphi_{ij}^*(\mu) = 0 \quad \forall (i, j)$

Les φ_{ij}^* ne sont pas linéairement indépendantes. En effet, le noyau de l'application $\underline{\mu} \rightarrow \underline{\varphi}^*(\underline{\mu})$, ensemble des $\underline{\mu}$ satisfaisant les relations (4) et (5) est de dimension $k + l - 1$ puisqu'il dépend de $k + l + 1$ paramètres μ_o, α_i, β_j liés par les deux relations (5). (On peut du reste voir aisément que :

$$\forall i \sum_j b_j \varphi_{ij}^* = 0 \text{ et } \forall j \sum_i \varphi_{ij}^* = 0, k + l - 1$$

de ces relations entraînant la dernière).

L'application $\underline{\mu} \rightarrow \underline{\varphi}^*(\underline{\mu})$ est donc de rang $kl - (k + l - 1) = (k - 1)(l - 1)$ et ce rang est l'ordre de l'hypothèse H_{A+B} .

V_{A+B} qui est le noyau de l'application φ^* est de dimension $k + l - 1$

Nous allons encore calculer $W(\hat{\mu})$ en testant l'hypothèse H_{A+B} dans le modèle (3). Elle se formule alors

$$\lambda_{ij} = \frac{\mu_o}{\sqrt{t_{ij}}} + \frac{\alpha_i}{\sqrt{t_{ij}}} + \frac{\beta_j}{\sqrt{t_{ij}}} \quad \text{avec} \quad \sum_i a_i \alpha_i = \sum_j b_j \beta_j = 0$$

Le système des équations normales donnant les estimateurs des moindres carrés, $\hat{\mu}_o^*, \hat{\alpha}_i^*, \hat{\beta}_j^*$ sous H_{A+B} et les conditions (5) est

$$\left\{ \begin{array}{l} (0) \quad \sum_{ij} \frac{l_{ij}}{\sqrt{t_{ij}}} = \tilde{\mu}_o^* \sum_{ij} \frac{1}{t_{ij}} + \sum_i \left(\tilde{\alpha}_i^* \sum_j \frac{1}{t_{ij}} \right) + \sum_j \left(\tilde{\beta}_j^* \sum_i \frac{1}{t_{ij}} \right) \\ (i) \quad \sum_j \frac{l_{ij}}{\sqrt{t_{ij}}} = \tilde{\mu}_o^* \sum_j \frac{1}{t_{ij}} + \tilde{\alpha}_i^* \sum_j \frac{1}{t_{ij}} + \sum_j \frac{\tilde{\beta}_j^*}{t_{ij}}, \quad i = 1, 2, \dots, k \\ (j)' \quad \sum_i \frac{l_{ij}}{\sqrt{t_{ij}}} = \tilde{\mu}_o^* \sum_i \frac{1}{t_{ij}} + \sum_i \frac{\tilde{\alpha}_i^*}{t_{ij}} + \tilde{\beta}_j^* \sum_i \frac{1}{t_{ij}}, \quad j = 1, 2, \dots, l \\ \sum_i a_i' \tilde{\alpha}_i^* = 0 \\ \sum_j b_j \tilde{\beta}_j^* = 0 \end{array} \right.$$

Pour résoudre ce système, on peut supprimer les lignes (k) et (l') puisque $(1) + (2) + \dots + (k) = (1)' + (2)' + \dots + (l)' = (0)$ ce qui permet de se ramener à un système de Cramer.

Le procédé ci-dessous permet, par un choix convenable des a_i et b_j , de résoudre ce système d'une manière astucieuse. On peut aussi utiliser une méthode générale (inversion d'une matrice principale du système) qui s'applique quels que soient les poids a_i et b_j . On sait (Cf. [4]) que la somme des carrés résiduel ne dépend de $\tilde{\mu}_o^*, \tilde{\alpha}_i^*, \tilde{\beta}_j^*$ que par l'intermédiaire des estimateurs des moindres carrés $\hat{\mu}_{ij} = \hat{\mu}_o^* + \hat{\alpha}_i^* + \hat{\beta}_j^*$ des μ_{ij} sous H_{A+B} , dans le modèle (3) et que ceux-ci ne dépendent pas du choix du système de poids (a_i, b_j) adopté pour définir μ_o , les α_i et les β_j . On trouvera donc pour $W(\hat{\mu})$ la même expression que ci-dessous (7) en fonction des $\hat{\mu}_{ij}$. Ceci résulte aussi de la théorie générale précé-

dente (1.2.3). De façon plus précise, l'expression de $W(\hat{\underline{\mu}})$ en fonction des $\hat{\mu}_{ij}$, $\tilde{\mu}_o^*$, $\tilde{\alpha}_i^*$, $\tilde{\beta}_j^*$ est aisée à trouver et ne dépend pas du choix des poids a_i et b_j adoptés pour définir $\tilde{\mu}_o^*$, $\tilde{\alpha}_i^*$, $\tilde{\beta}_j^*$. On sait en effet (Cf. [4]) que, dans un modèle linéaire général du type $\underline{Y} = X\underline{\beta} + \underline{E}$ avec $E(\underline{E}) = 0$ et une matrice X non nécessairement de plein rang, la somme des carrés résiduelle peut toujours se mettre sous la forme

$$S = \underline{Y}'\underline{Y} - \hat{\underline{Y}}'\underline{Y} = \underline{Y}'\underline{Y} - \underline{b}'X'\underline{Y} = \underline{Y}'\underline{Y} - \underline{b}'\underline{P}$$

où $\underline{b}$ est un estimateur des moindres carrés d'un $\underline{\beta}$ tel que $E(\underline{Y}) = X\underline{\beta}$, tandis que $\hat{\underline{Y}} = X\underline{b}$ est l'unique estimateur des m.c. de $E(\underline{Y}) = X\underline{\beta}$, et $\underline{P} = X'\underline{Y}$ est le vecteur des seconds membres du système des équations normales définissant les estimateurs des m.c. $\underline{b}$ des $\underline{\beta}$.

On a donc dans le cas présent où $\underline{Y} = \underline{1}$ de composantes $l_{ij} = \frac{\hat{\mu}_{ij}}{\sqrt{t_{ij}}}$ et $\underline{\beta}$ a pour composantes μ_o , les α_i et les β_j :

$$W(\hat{\underline{\mu}}) = S = \sum_{ij} \frac{\hat{\mu}_{ij}^2}{t_{ij}} - \tilde{\mu}_o^* \sum_{ij} \frac{\hat{\mu}_{ij}}{t_{ij}} - \sum_i \tilde{\alpha}_i^* \left(\sum_j \frac{\hat{\mu}_{ij}}{t_{ij}} \right) - \sum_j \tilde{\beta}_j^* \left(\sum_i \frac{\hat{\mu}_{ij}}{t_{ij}} \right) \quad (7)$$

pour tout système $(\tilde{\mu}_o^*, \tilde{\alpha}_i^*, \tilde{\beta}_j^*)$ de solutions des équations (0), (i), (j)'

Si on prend $a_i = \sum_j \frac{1}{t_{ij}}$, $b_j = \sum_i \frac{1}{t_{ij}}$, on obtient :

$$\left\{ \begin{array}{l} (0) \quad \tilde{\mu}_o^* = \frac{\sum_{ij} \frac{l_{ij}}{\sqrt{t_{ij}}}}{\sum_{ij} \frac{1}{t_{ij}}} = \frac{\sum_{ij} \frac{\hat{\mu}_{ij}}{t_{ij}}}{\sum_{ij} \frac{1}{t_{ij}}} \\ (i) \quad a_i \tilde{\alpha}_i^* + \sum_j \frac{1}{t_{ij}} (\tilde{\beta}_j^* + \tilde{\mu}_o^*) = \sum_j \frac{l_{ij}}{\sqrt{t_{ij}}} \\ (j)' \quad \sum_i \frac{\tilde{\alpha}_i^*}{t_{ij}} + b_j (\tilde{\beta}_j^* + \tilde{\mu}_o^*) = \sum_i \frac{l_{ij}}{\sqrt{t_{ij}}} \\ \sum_i a_i \tilde{\alpha}_i^* = 0 \\ \sum_j b_j \tilde{\beta}_j^* = 0 \end{array} \right.$$

En éliminant $\tilde{\beta}_j^* + \tilde{\mu}_o^*$ dans les équations (i) pour i de 1 à $k-1$, et en remplaçant l_{ij} par $\frac{\hat{\mu}_{ij}}{\sqrt{t_{ij}}}$, on obtient pour l'ensemble de ces équations et de l'équation $\sum_i a_i \tilde{\alpha}_i^* = 0$, l'expression :

$$A \tilde{\underline{\alpha}}^* = \underline{C} \quad \text{avec} \quad \tilde{\underline{\alpha}}^* = (\tilde{\alpha}_1^*, \tilde{\alpha}_2^*, \dots, \tilde{\alpha}_k^*), \quad (8)$$

A matrice $(k \times k)$ d'éléments a_{ih} définis par

$$\forall i \neq k \quad \left\{ \begin{array}{l} a_{ii} = a_i - \sum_j \frac{1}{t_{ij}^2 b_j} \\ a_{ih} = - \sum_j \frac{1}{t_{ij} t_{hj} b_j} \quad \forall h \neq i \\ a_{kh} = a_h \quad \forall h \end{array} \right.$$

$$\text{et } C' = (c_1, c_2, \dots, c_k) \quad \text{avec} \quad c_i = \sum_j \frac{\hat{\mu}_{ij}}{t_{ij}} - \sum_{i'j} \frac{\hat{\mu}_{i'j}}{t_{ij} t'_{ij} b_j} \quad \forall i \neq k$$

$$\text{et} \quad c_k = 0$$

Les $\tilde{\beta}_j^*$ peuvent être exprimés en fonction des $\tilde{\alpha}_i^*$ et de $\tilde{\mu}_o^*$ par

$$\tilde{\beta}_j^* = \frac{1}{b_j} \left(\sum_i \frac{\hat{\mu}_{ij}}{t_{ij}} - \sum_i \frac{\tilde{\alpha}_i^*}{t_{ij}} \right) - \tilde{\mu}_o^*.$$

et $W(\hat{\mu})$ donné par l'expression (7) ci-dessus a une loi asymptotique de $\chi^2_{(k-1)(l-1)}$ si H_{AB} est réalisée.

Remarque : Estimation des effets des facteurs contrôlés sous l'hypothèse d'additivité

On trouvera dans [3] en annexe, les équations de la vraisemblance pour estimer μ_o, α_i, β_j , à partir des observations y_{ijm} . Il semble que même dans le cas simple $\sigma_{ij} = a\mu_{ij}$, il faille, pour les résoudre, utiliser des méthodes itératives. C'est la raison pour laquelle on peut attacher quelque intérêt à la proposition qui suit.

Les $\hat{\mu}_{ij}$ ne peuvent être utilisés à la place des μ_{ij} dans les expressions (6) pour déduire des estimateurs d'effets, μ_o, α_i, β_j définis par un système de poids indépendants des n_{ij} et des $\hat{\mu}_{ij}$, par exemple par les conditions habituelles (9) $\sum_i \alpha_i = \sum_j \beta_j = 0$ ($a_i = b_j = 1$), puisque $\hat{\mu}$ n'a aucune raison d'appartenir à V_{A+B} c'est-à-dire de satisfaire les équations (4). Par ailleurs les estimateurs $\tilde{\mu}_o^*, \tilde{\alpha}_i^*, \tilde{\beta}_j^*$ correspondant aux poids $a_i = \sum_j \frac{1}{t_{ij}}$ et $b_j = \sum_i \frac{1}{t_{ij}}$ adoptés ne présentent pas grand intérêt.

En effet, les effets $\mu_o^*, \alpha_i^*, \beta_j^*$, qu'ils soient définis qu'étant donnés les poids a_i, b_j et les poids utilisés dépendent des effectifs n_{ij} et des μ_{ij} puisque $t_{ij} = \frac{1}{n_{ij} I(\hat{\mu}_{ij})}$ (Le fait que les poids augmentent indéfiniment avec les n_{ij} n'est pas gênant puisque les résultats sont inchangés si on remplace ces poids par des poids qui leur sont proportionnels. On peut donc remplacer les n_{ij} par les p_{ij} dans les expressions de ces poids).

Par contre les $\tilde{\mu}_{ij} = \tilde{\mu}_o^* + \tilde{\alpha}_i^* + \tilde{\beta}_j^*$ qui sont les solutions du problème)

$$\underline{\mu}^o \text{Min}_{\underline{\mu} \in V_{A+B}} \left[\sum_{ij} \frac{(\hat{\mu}_{ij} - \mu_{ij}^o)^2}{t_{ij}} \right] = \underline{\mu}^o \text{Min}_{\underline{\mu} \in V_{A+B}} \left[\sum_{ij} (\hat{\mu}_{ij} - \mu_{ij}^o)^2 I(\hat{\mu}_{ij}) p_{ij} \right] \quad (10)$$

ne dépendent pas du système (a_i, b_j) de poids adopté pour définir les $\mu_o^*, \alpha_i^*, \beta_j^*$. (10) définit en effet $\underline{\mu}$ comme l'élément de V_{A+B} le plus proche de $\hat{\mu}$ pour la métrique euclidienne définie par la matrice diagonale d'éléments $\sqrt{I(\hat{\mu}_{ij}) p_{ij}}$

On peut alors utiliser les expressions (6) dans lesquelles les μ_{ij} sont remplacés par les $\tilde{\mu}_{ij}$ pour en déduire les estimateurs $\tilde{\mu}_o, \tilde{\alpha}_i, \tilde{\beta}_j$ des effets μ_o, α_i, β_j définis par (9) (On notera que si $\underline{\mu} \in V_{A+B}$, on a $\tilde{\mu}_{ij} = \hat{\mu}_{ij} \forall (i, j)$). On obtient ainsi des estimateurs convergents des μ_o, α_i, β_j si H_{AB} est réalisée puisque, comme on va le voir, les $\tilde{\mu}_{ij}$ sont, dans ce cas, des estimateurs convergents des μ_{ij} .

Démonstration de la convergence des $\tilde{\mu}_{ij}$ vers les μ_{ij} si H_{AB} est réalisée

On a

$$\lim_{v \rightarrow \infty} \sum_{ij} (\hat{\mu}_{ij} - \tilde{\mu}_{ij})^2 I(\hat{\mu}_{ij}) p_{ij} = \lim_{v \rightarrow \infty} \underline{\mu}^o \text{Min}_{\underline{\mu} \in V_{A+B}} \left[\sum_{ij} (\hat{\mu}_{ij} - \mu_{ij}^o)^2 I(\hat{\mu}_{ij}) p_{ij} \right]$$

Mais

$$\lim_{v \rightarrow \infty} (\text{p.s.}) \sum_{ij} (\hat{\mu}_{ij} - \mu_{ij})^2 I(\hat{\mu}_{ij}) p_{ij} = 0 \quad \text{avec} \quad \underline{\mu} \in V_{A+B},$$

si H_{AB} est réalisée, puisque $\mu_{ij} \xrightarrow{\text{p.s.}} \mu_{ij}$ et $\lim (\text{p.s.}) I(\hat{\mu}_{ij}) = I(\mu_{ij})$ où $I(\mu_{ij})$ est fini.

On a donc

$$\lim_{v \rightarrow \infty} (\text{p.s.}) \sum_{ij} (\hat{\mu}_{ij} - \tilde{\mu}_{ij})^2 I(\hat{\mu}_{ij}) p_{ij} = 0$$

ce qui implique

$$(\hat{\mu}_{ij} - \tilde{\mu}_{ij}) \xrightarrow{\text{p.s.}} 0 \quad \forall (i, j)$$

puisque

$$I(\hat{\mu}_{ij}) p_{ij} > 0 \quad \forall (i, j)$$

et par suite

$$\tilde{\mu}_{ij} \xrightarrow{\text{p.s.}} \mu_{ij} \quad \forall (i, j)$$

L'interprétation géométrique du raisonnement précédent est évidente : puisque $\hat{\mu} \xrightarrow{\text{p.s.}} \underline{\mu} \in V_{A+B}$, le point de V_{A+B} le plus proche de $\hat{\mu}$ pour une métrique non dégénérée quelconque converge nécessairement vers $\underline{\mu}$.

I 2.4 – Application du test de Wald au cas où $\alpha_j = a(\mu_{ij} - m_o)$

On peut ici encore se ramener au cas $\sigma_{ij} = a \mu_{ij}$. On suppose donc dans la suite que $m_o = 0$

On a

$$t_{ij} = \frac{1}{n_{ij} I(\hat{\mu}_{ij})} = \frac{a^2}{1 + 2a^2} \frac{\hat{\mu}_{ij}^2}{n_{ij}}$$

et par suite :

1.2.4.1 – Test de l'hypothèse H_B

$$\text{La statistique de test est } W_B(\hat{\mu}) = \frac{1 + 2a^2}{a^2} \left[n - \sum_j \left(\frac{\sum_i \left(\frac{n_{ij}}{\hat{\mu}_{ij}^2} \right)^2}{\sum_i \frac{n_{ij}}{\hat{\mu}_{ij}^2}} \right) \right]$$

1.2.4.2 – Test de l'hypothèse d'additivité des effets de A et B

Les poids sont :

$$a_i = \frac{1 + 2a^2}{a^2} \sum_j \frac{n_{ij}}{\hat{\mu}_{ij}^2}, \quad b_j = \frac{1 + 2a^2}{a^2} \sum_i \frac{n_{ij}}{\hat{\mu}_{ij}^2}$$

et les expressions qui interviennent dans les éléments de A et C (équation (8)) sont :

$$a_{ih} = - \frac{1 + 2a^2}{a^2} \sum_j \frac{n_{ij} n_{hj}}{\hat{\mu}_{ij}^2 \hat{\mu}_{hj}^2 \sum_{i'} \frac{n_{i'j}}{\hat{\mu}_{i'j}^2}} \quad i \neq h,$$

$$a_{ii} = \frac{1 + 2a^2}{a^2} \sum_j \left(\frac{n_{ij}}{\hat{\mu}_{ij}^2} - \frac{n_{ij}^2}{\hat{\mu}_{ij}^4 \sum_{i'} \frac{n_{i'j}}{\hat{\mu}_{i'j}^2}} \right) \quad i \neq k$$

$$\frac{\hat{\mu}_{ij}}{t_{ij}} = \frac{1 + 2a^2}{a^2} \cdot \frac{n_{ij}}{\hat{\mu}_{ij}}, \quad \sum_{i'} \frac{\hat{\mu}_{ij}}{t_{ij} t_{i'j} b_j} = \frac{1 + 2a^2}{a^2} \sum_{i', j} \frac{n_{ij} n_{i'j}}{\hat{\mu}_{ij}^2 \hat{\mu}_{i'j} \sum_{i''} \frac{n_{i''j}}{\hat{\mu}_{i''j}^2}}$$

$$a_{kh} = a_h \quad \forall h,$$

et par suite

$$c_i = \frac{1 + 2a^2}{a^2} \left[\sum_j \frac{n_{ij}}{\hat{\mu}_{ij}^2} - \sum_j \frac{n_{ij} \left(\sum_{i'} \frac{n_{i'j}}{\hat{\mu}_{i'j}^2} \right)}{\hat{\mu}_{ij}^2 \left(\sum_{i'} \frac{n_{i'j}}{\hat{\mu}_{i'j}^2} \right)} \right] \quad i \neq k$$

$$\tilde{\mu}_{\cdot o}^* = \frac{\sum_{ij} \frac{n_{ij}}{\hat{\mu}_{ij}^2}}{\sum_{ij} \frac{n_{ij}}{\hat{\mu}_{ij}^2}}, \quad \tilde{\beta}_{\cdot j}^* = \frac{\sum_i \frac{n_{ij}}{\hat{\mu}_{ij}^2}}{\sum_i \frac{n_{ij}}{\hat{\mu}_{ij}^2}} - \frac{\sum_{ij} \frac{n_{ij}}{\hat{\mu}_{ij}^2}}{\sum_{ij} \frac{n_{ij}}{\hat{\mu}_{ij}^2}}$$

et

$$W_{AB}(\hat{\mu}) = \frac{1 + 2a^2}{a^2} \left[n - \tilde{\mu}_{\cdot o}^* \sum_{ij} \left(\frac{n_{ij}}{\hat{\mu}_{ij}^2} \right) - \sum_i \tilde{\alpha}_{\cdot i}^* \left(\sum_j \frac{n_{ij}}{\hat{\mu}_{ij}^2} \right) - \sum_j \tilde{\beta}_{\cdot j}^* \left(\sum_i \frac{n_{ij}}{\hat{\mu}_{ij}^2} \right) \right]$$

I 3 – CONCLUSION DE LA 1^{ère} PARTIE

I 3.1 – Estimation

On a la possibilité d'estimer les moyennes μ_{ij} correspondant aux combinaisons (A_i, B_j) sous l'hypothèse la plus générale par la méthode du maximum de vraisemblance.

Les estimateurs $\hat{\mu}_{ij}$ sont indépendants en probabilité et $\hat{\mu}_{ij}$ ne dépend que des observations correspondant au traitement (A_i, B_j). Dans le cas où

$$\sigma_{ij} = a(\mu_{ij} - m_o), \hat{\mu}_{ij} - m_o$$

est la solution de même signe que a d'une équation du second degré.

Les $\hat{\mu}_{ij}$ existent toujours et convergent presque sûrement vers les valeurs vraies. Ils sont asymptotiquement efficaces et normaux. Dans l'hypothèse d'absence d'effet d'un facteur, B par exemple, les moyennes μ_i peuvent être estimées de la même façon.

Dans l'hypothèse $H_{A \cdot B}$ d'additivité des effets de A et B, les estimateurs du M.V. des effets α_i et β_j des facteurs sont donnés par des équations complexes faisant intervenir l'ensemble des observations qu'il faut résoudre par des méthodes itératives. On peut déduire d'éléments intervenant dans le test de $H_{A \cdot B}$ des estimateurs convergents des paramètres μ_o, α_i, β_j du modèle. Ils sont fonction des seuls $\hat{\mu}_{ij}$.

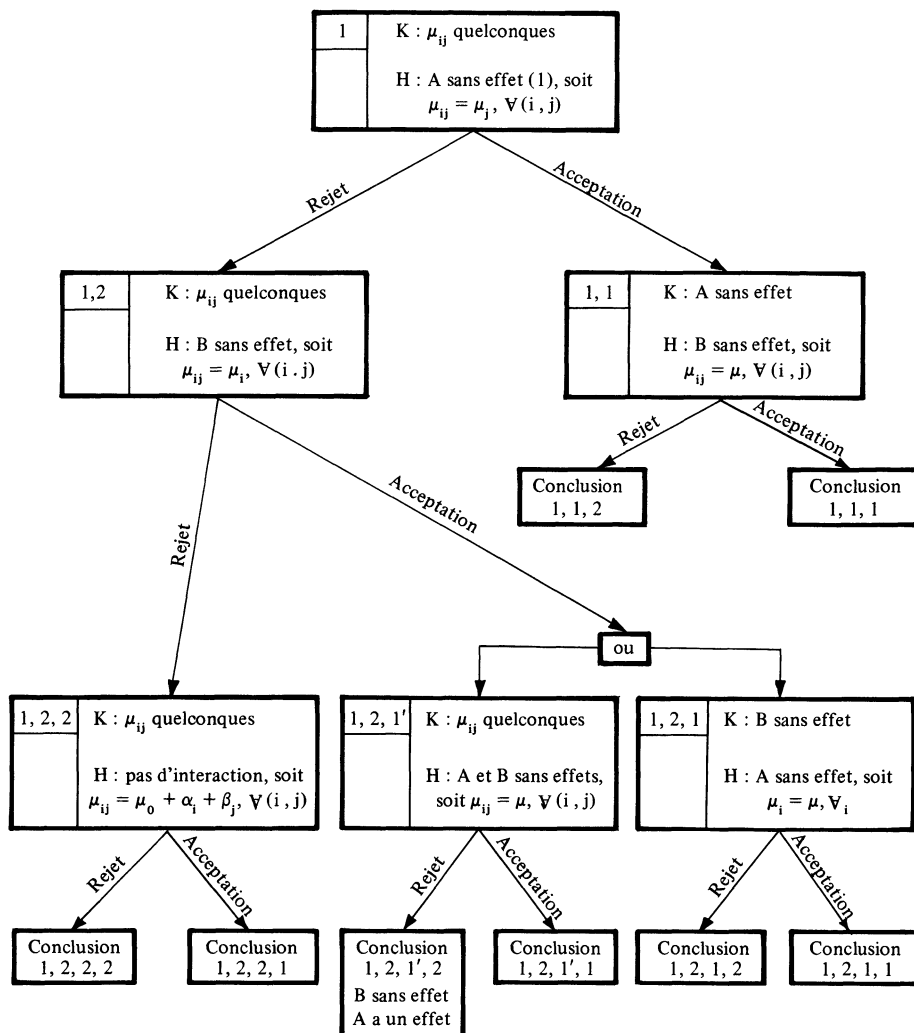
I.3.2 – Tests

On a la possibilité d'utiliser un test asymptotique pour tester les hypothèses d'absence d'effet d'un facteur ou d'absence d'interaction (même dans le cas d'un plan factoriel sans répétitions) et des hypothèses plus générales de la forme $\varphi_i(\mu) = 0$ où les φ_i sont des fonctions dérivables des μ_{ij} , fonctionnellement indépendantes. Deux méthodes ont été envisagées : le test du rapport des vraisemblances maximales et le test de Wald. Il a fallu généraliser ces tests au cas de nombres de répétitions n_{ij} inégaux en fixant, puisqu'il s'agit de propriétés asymptotiques, le processus de croissance vers l'infini des n_{ij} (On a posé $n_{ij} = v p_{ij}$, p_{ij} finis, $v \rightarrow \infty$). Dans les deux cas, la statistique utilisée converge vers un χ^2 à q degrés de liberté si q est l'ordre de l'hypothèse H à tester (nombre de relations fonctionnellement indépendantes entre les μ_{ij} imposées par H).

Le test de Wald a été préféré parce qu'il n'exige pas le calcul des estimateurs du M.V. des paramètres sous H, mais seulement sous l'hypothèse la plus générale. La statistique $W(\hat{\mu})$ utilisée dans ce test peut : dans le cas d'une hypothèse linéaire être calculée assez aisément grâce à un artifice permettant de se ramener à la somme des carrés résiduelle sous H dans un modèle linéaire classique.

Le problème de succession des tests se pose avec un acuité d'autant plus grande que les estimateurs du M.V. des μ_{ij} sous l'hypothèse d'absence d'interaction sont difficilement accessibles. Il paraît donc assez peu opportun de tester d'abord l'hypothèse $H_{A \cdot B}$, si on ne lui porte pas un intérêt essentiel.

Un schéma possible de tests pourra être celui de la figure 1. On y notera une possibilité de contradiction si le chemin parcouru est le chemin (1, 2, 1, 1).



LEGENDE : K : hypothèse générale
 H : hypothèse nulle à tester

(1) A est le facteur qui a, à priori, la plus grande probabilité d'être sans effet

Figure 1 — Schéma possible de tests

C'est pourquoi on pourra préférer le test (1, 2, 1') au test (1, 2, 1). Si on préfère le test (1, 2, 1') et si (1, 2, 1', 1) se produit, il convient d'examiner attentivement les résultats obtenus au cours des différents tests avant de conclure.

Enfin, il ne faut pas oublier qu'estimations et tests reposent sur des méthodes asymptotiques. C'est la raison pour laquelle il a paru nécessaire de procéder à l'étude par simulation dont rend compte la deuxième partie de cette note.

DEUXIEME PARTIE

ETUDE PAR SIMULATION DE LA VALIDITE A DISTANCE FINIE DES RESULTATS ASYMPTOTIQUES DE I, QUAND L'ECART TYPE EST FONCTION LINEAIRE DE LA MOYENNE.

*
* *
*

II – 1. INTRODUCTION – CHOIX DES CONDITIONS DE LA SIMULATION

La méthode de test proposée repose essentiellement sur le résultat suivant :

$\forall (i, j)$ les variables $\hat{z}_{ij} = \frac{\hat{\mu}_{ij} - \mu_{ij}}{\sqrt{t(\hat{\mu}_{ij})}}$ sont asymptotiquement normales réduites. Le premier objectif de la simulation est donc de voir si l'hypothèse (N) d'une loi normale réduite pour les $\hat{z}_{ij}$ est acceptable avec des valeurs habituelles pour les n_{ij} , nombres de répétitions, et pour les paramètres μ_{ij} (on se placera dans le cas $m_0 = 0$ puisqu'on peut toujours s'y ramener).

Il résulte de la présentation de la méthode que si $\forall (i, j)$ (N) est acceptable, $W_H(\hat{\mu})$ suit, si l'hypothèse H est réalisée, une loi de χ^2 . Il est donc inutile d'étudier expérimentalement de façon approfondie la distribution de $W_H(\hat{\mu})$ sous H, si (N) est acceptable. Par contre, il se peut que (N) ne soit pas acceptable mais que $W_H(\hat{\mu})$ suive néanmoins approximativement un χ^2 sous H. Il faut donc envisager la possibilité d'étudier expérimentalement la distribution de $W_H(\hat{\mu})$ sous H et plus particulièrement d'estimer :

$\omega = \Pr(W_H(\hat{\mu}) > \chi_\alpha^2)$ vrai risque de première espèce correspondant au risque nominal α .

Choix des plans

Compte tenu de ces considérations, on a retenu trois types de plans caractérisés par :

A : pas d'effet du facteur A, soit $\mu_{ij} = \mu_{.j} \quad \forall (i, j)$

B : pas d'effet moyen de B, mais présence d'une interaction proportionnelle à l'effet moyen de A, soit : $\mu_{ij} = \mu_{i.} + k_j \mu_{i.}$

C : pas d'interaction, soit $\mu_{ij} = \mu_{i.} + \mu_{.j}$

Le nombre de niveaux retenu pour chaque facteur est le même pour les trois plans, soit 5 pour A et 6 pour B. Le nombre n total des observations est également le même.

L'exemple qui a servi de guide dans le choix des moyennes concerne les teneurs en soufre de produits pétroliers.

L'intervalle de variation retenu pour les μ_{ij} est l'intervalle 0,3, 3,0.

Les trois plans sont définis par les valeurs suivantes des μ_{ij} :

A : $\mu_{ij} = \mu_j$ et $\mu_j = 0,3 ; 0,5 ; 0,8 ; 2,0 ; 2,5 ; 3,0$.

B :

$\mu_i \backslash k_j$	0,3	0,5	0,8	2,0	3,0
0,24	0,372	0,620	0,992	2,48	3,72
0,13	0,339	0,565	0,904	2,26	3,39
0,03	0,309	0,515	0,824	2,06	3,09
0	0,300	0,500	0,800	2,00	3,00
- 0,1	0,270	0,450	0,720	1,80	2,70
- 0,16	0,252	0,420	0,672	1,68	2,52

C : $\mu_{ij} = \mu_i + \mu_j$

$\mu_i \backslash \mu_j$	0,1	0,3	0,5	0,7	1
0,2	0,3	0,5	0,7	0,9	1,2
0,5	0,6	0,8	1,0	1,2	1,5
0,9	1,0	1,2	1,4	1,6	1,9
1,2	1,3	1,5	1,7	1,9	2,2
1,5	1,6	1,8	2,0	2,2	2,5
2	2,1	2,3	2,5	2,7	3,0

Pour chacun de ces trois types de plans, on a retenu trois types de répartition des n_{ij} (nombre de répétitions pour la combinaison (i,j)) et ce, pour un même nombre d'observations n :

a : plan complètement équilibré : $n_{ij} = \text{cte}$

b : plan peu déséquilibré

c : plan très déséquilibré : on a, en randomisant les lignes puis les colonnes d'un tableau des n_{ij} , établi à priori, avec une table de nombres au hasard, attribué aléatoirement chaque n_{ij} à une combinaison (i , j) afin d'être sûr de ne pas avoir privilégié certaines combinaisons.

Le nombre total d'observations prend 2 valeurs :

$$n = 150$$

et

$$n = 90$$

(sauf dans le cas c où on a pris $n = 105$ pour avoir un déséquilibre comparable à celui du cas $n = 150$ sans avoir aucun n_{ij} nul). Ceci donne :

a : $n_{ij} = 5 \quad \forall (i, j)$

b :

$\begin{array}{c c} j \\ \hline i \end{array}$	1	2	3	4	5	6
1	5	5	6	5	5	5
2	5	5	5	6	6	5
3	6	4	6	5	5	5
4	5	4	5	5	5	6
5	3	5	5	4	4	5

$n_{ij} = 3 \quad \forall (i, j)$

$\begin{array}{c c} j \\ \hline i \end{array}$	1	2	3	4	5	6
1	3	3	4	3	3	3
2	3	3	3	4	3	3
3	4	2	4	3	3	3
4	3	2	3	3	4	3
5	1	3	3	2	3	3

c :

$\begin{array}{c c} j \\ \hline i \end{array}$	1	2	3	4	5	6
1	7	5	4	2	8	3
2	6	8	1	6	3	6
3	11	2	5	3	8	1
4	4	9	4	6	4	5
5	10	1	7	5	4	2

$\begin{array}{c c} j \\ \hline i \end{array}$	1	2	3	4	5	6
1	5	3	2	2	6	2
2	4	6	1	4	3	4
3	8	1	3	3	6	1
4	3	7	2	4	2	3
5	8	1	5	3	2	1

On a ainsi pour une valeur de n , 9 plans que l'on appellera Aa, Ab, etc. . .

Choix du paramètre a

Dans l'exemple des dosages de soufre d'une huile, on a :

$$a = 0,08$$

ce qui correspond à une erreur relative à 95 % de $\frac{2\sigma}{\mu} = 2a = 16 \%$

Devant les résultats satisfaisants obtenus avec cette valeur relativement faible de a , il a été décidé d'essayer des valeurs plus grandes, soit : $a = 0,5$ ce qui est très grand, puisque l'erreur relative à 95 % est alors de 100 % (1), et $a = 0,2$ ce qui donne $\frac{2\sigma}{\mu} = 40 \%$.

De plus, on peut se demander s'il est vraiment nécessaire d'avoir recours à méthode spécifique et s'il n'est pas possible, au moins dans le cas où $a \approx 0,08$ d'utiliser la méthode classique d'analyse des plans d'expérience correspondant à l'hypothèse (N') : σ inconnu mais constant. Comme le rejet de (N') n'implique pas que les statistiques classiques $\tilde{S}_H$ ne soient pas sous H distribuées sensiblement suivant des lois F de Fisher-Snedecor, on a estimé en même temps que ϖ , les vrais risques de première espèce d'un test F classique pour le même risque nominal α .

Enfin, même si la méthode spécifique est nettement supérieure à la méthode classique, elle présente l'inconvénient de reposer sur la connaissance de a . C'est pourquoi on a également étudié la sensibilité de la méthode à la valeur adoptée pour a , en remplaçant la vraie valeur $a = 0,08$ par des valeurs $\hat{a} = 0,06$; $0,07$; $0,09$ et $0,10$ dans les formules donnant $\hat{\mu}_{ij}$ et $t(\hat{\mu}_{ij})$.

(1) Avec $a = 0,5$, $\mu - 2\sigma = 0$ quelque soit μ ; on a donc avec l'hypothèse de normalité des erreurs : $P_r(Y < 0) = 0,025$ ce qui paraît un maximum si on sait que Y doit être positif ou nul.

II – 2. METHODES DE SIMULATION

Il a été décidé de réaliser 100 simulations des résultats des 9 plans obtenus pour l'étude de la distribution des variables ; $\hat{z}_{ij} = \frac{\hat{\mu}_{ij} - \mu_{ij}}{\sqrt{t(\hat{\mu}_{ij})}}$ et $z_{ij} = \frac{\hat{\mu}_{ij} - \mu_{ij}}{\sqrt{t(\mu_{ij})}}$ qui a plus de chances d'être normale réduite que $\hat{z}_{ij}$.

Pour chacune de ces variables, nous aurons donc un échantillon de taille 100.

On a testé l'hypothèse (N) par un test de Kolmogorov qui est basé sur la statistique

$$K_n = \sup_{x \in \mathbb{R}} |F_o(x) - \hat{F}_n(x)| \quad \text{où} \quad F_o(x) \text{ et } \hat{F}_n(x)$$

désignent respectivement les fonctions de répartition théorique et empirique pour un échantillon de taille n d'une variable aléatoire réelle, continue X (cf [2]).

Les observations d'une variable normale réduite X utilisées pour en déduire les réalisations Y_{ijm} des Y_{ij} ont été obtenues à partir de nombres au hasard sur $(0, 1)$ engendrés par la méthode des congruences, par la méthode "directe" (cf [6]) : si U_1 et U_2 sont deux variables uniformes sur $(0, 1)$, indépendantes en probabilité :

$$X_1 = (-2 \log U_1)^{1/2} \cos 2\pi U_2$$

$$X_2 = (-2 \log U_1)^{1/2} \sin 2\pi U_2$$

sont deux variables normales réduites indépendantes. On a testé l'indépendance de la suite de nombres ainsi obtenue par des calculs de coefficients d'intercorrélation d'ordres 1 et 2, et on a également testé que l'espérance et la variance de la suite de ces nombres étaient suffisamment proches respectivement de 0 et de 1.

On ne s'est pas préoccupé des qualités de la suite des nombres au hasard sur $(0, 1)$, car c'est la suite des nombres issus de la distribution normale qui nous intéresse et, de plus, certains défauts de la première suite (en particulier défaut d'indépendance des nombres de la suite) sont atténués quand on passe des variables uniformes aux variables normales.

C'est d'ailleurs la même suite de nombres, de distribution uniforme sur $(0, 1)$, qui a toujours servi afin de comparer uniquement les méthodes d'estimation et de test pour les différents plans, sans être gêné par les aléas de la suite de nombres au hasard.

Pour l'estimation des risques de première espèce, on a fait 800 simulations afin que cette estimation soit bonne.

II – 2.1 – Résultats de la simulation pour l'étude des estimations

Pour chaque combinaison (i, j) on a deux variables z_{ij} et $\hat{z}_{ij}$ pour lesquelles on teste si elles sont normales réduites ou non. Les plans envisagés contenant 30 combinaisons (i, j) , pour chaque plan on a fait 30 tests pour z_{ij} et 30 tests pour $\hat{z}_{ij}$. Ces tests ont été faits chacun à un niveau de 5 %.

Pour chaque plan, ces 30 tests sont indépendants puisque chacun d'entre eux fait intervenir une partie différente de la suite de nombres au hasard engendrée comme on l'a vu précédemment. Mais les séries de 30 tests ne sont pas indépendantes entre elles car, lorsqu'on change de type de plan, on reprend la même suite de nombres au hasard que pour le plan précédent.

Pour un plan, on peut donc dire que le nombre de rejets sur les 30 tests effectués suit une loi binomiale de paramètres $n = 30$, $p = 5\%$ (niveau du test), si on a effectivement la loi normale réduite pour les variables étudiées. La valeur moyenne de cette loi est 1,5. Ceci signifie, que si on a effectivement la loi normale réduite, on obtient en moyenne 1,5 rejets uniquement à cause du risque de première espèce.

Nous avons compté pour chaque plan le nombre de rejets sur les 30 tests effectués sur chacune des variables z_{ij} et $\hat{z}_{ij}$.

Nous avons aussi calculé le nombre moyen de rejets pour l'ensemble des 9 plans et il nous a paru intéressant de comparer ce nombre moyen à 1,5 sans faire réellement d'analyse statistique puisque l'on a constaté que les différentes séries de 30 tests n'étaient pas indépendantes.

On peut récapituler les résultats obtenus dans les tableaux suivants.⁽¹⁾

II – 2.1.1. – Résultats pour $\alpha = 0,08$ et $n = 150$

Tableau I
Nombre de rejets sur 30 tests de la variable z_{ij}

plan \ $\hat{\alpha}$	0,06	0,07	<u>0,08</u>	0,09	0,10
Aa	10	4	4	5	11
Ab	9	3	3	7	13
Ac	9	2	0	3	11
Ba	9	2	1	3	10
Bb	9	1	1	3	5
Bc	8	4	3	3	9
Ca	11	2	0	2	10
Cb	15	3	1	2	7
Cc	10	3	1	5	12
Nombre de rejets moyen	10	2,6	1,5	3,4	9,7

L'accord avec la loi normale réduite est bon quand α est bien estimé. Si α est mal estimé, il devient assez mauvais.

(1) On peut signaler que, pour une variable binomiale $B(n = 30, p = .5\%)$, $P_r(x \geq 3) = 6,7\%$.

Tableau II
Nombre de rejets sur 30 tests de la variable $\hat{z}_{ij}$

plan \ $\hat{a}$	0,06	0,07	<u>0,08</u>	0,09	0,10
Aa	11	3	4	5	12
Ab	8	2	3	8	17
Ac	8	2	0	4	11
Ba	7	2	2	3	13
Bb	8	1	2	3	6
Bc	8	3	3	4	13
Ca	10	3	1	2	11
Cb	13	3	2	3	10
Cc	12	4	1	5	14
Nombre de rejets moyen	9,4	2,5	2	4,1	11,9

L'accord avec la loi normale réduite est un peu moins bon pour $\hat{z}_{ij}$ que pour z_{ij} , ce qui n'est pas étonnant puisque dans $\hat{z}_{ij}$ intervient la variance estimée de $\hat{\mu}_{ij}$ au lieu de la valeur exacte de cette variance.

L'examen des deux tableaux I et II tend à montrer que la structure du plan (plus ou moins grand déséquilibre repéré par les lettres a, b, c) ne joue pas sur l'accord avec la loi normale réduite.

De plus, on a localisé les rejets de la loi normale réduite par rapport aux valeurs données aux μ_{ij} et aux valeurs des n_{ij} . Là encore, il n'y a apparemment aucun lien entre la valeur de μ_{ij} et la présence ou l'absence d'un rejet, ainsi qu'entre la valeur de n_{ij} et la présence ou l'absence d'un rejet.

II – 2.1.2 – Résultats pour $a = 0,08$, $n = 90$ et $\hat{a} = 0,08$

Tableau III
Nombre de rejets sur 30 tests pour z_{ij} et $\hat{z}_{ij}$

plan	Nombre de rejets pour z_{ij}	Nombre de rejets pour $\hat{z}_{ij}$
Aa	2	2
Ab	2	3
Ac	0	0
Ba	2	2
Bb	1	1
Bc	3	3
Ca	1	1
Cb	0	0
Cc	3	4
Moyennes	1,5	1,8

Le nombre total d'observations 150 ou 90 joue peu sur le résultat. Il semble qu'avec $n = 90$, on soit assez proche des conditions asymptotiques.

II – 2.1.3 – Résultats pour $a = 0,5$ et $a = 0,2$, $n = 150$

Tableau IV
Nombre de rejets sur 30 tests pour z_{ij} et $\hat{z}_{ij}$

plan \ a	0,5		0,2	
	z_{ij}	$\hat{z}_{ij}$	z_{ij}	$\hat{z}_{ij}$
Aa	11	20	4	4
Ab	9	19	4	5
Ac	10	16	5	6
Ba	3	13	1	1
Bb	7	11	0	2
Bc	7	18	4	7
Ca	7	13	2	2
Cb	4	14	2	4
Cc	10	15	5	9
Moyennes	7,1	13,4	3	4,4

Lorsque a prend de grandes valeurs, l'accord avec la loi normale réduite est assez mauvais. Il est encore assez bon pour $a = 0,2$.

De tous les tableaux précédents, il ressort qu'apparemment la forme du plan (type A, B ou C) n'influe pas sur la présence ou l'absence de rejets.

II – 2.1.4 – Conclusion de la simulation pour les estimations

Les propriétés asymptotiques des $\hat{z}_{ij}$ semblent assez vite acquises pour $a = 0,08$. Pour $a = 0,2$, elles restent acceptables pour des n_{ij} de l'ordre de 5. Le facteur qui a le plus d'influence sur l'accord avec la loi asymptotique est la valeur de a et de la qualité de son estimation.

II – 3: RESULTATS DE LA SIMULATION POUR LE TEST D'ABSENCE D'INTERACTION.

Dans cette deuxième partie de la simulation, on s'intéresse essentiellement aux statistiques de tests $W(\hat{\mu})$ et $\mathcal{F}_H$. Les résultats de la première partie de la simulation étant assez encourageants pour la rapidité de convergence de la méthode spécifique, nous nous sommes contentés de procéder, pour chaque plan dans le cas $n = 150$, $a = 0,08$, à un test au niveau 5 % de l'hypothèse d'absence d'interaction avec la statistique de test classique $\mathcal{F}_H$ et avec $W(\hat{\mu})$.

Nous avons engendré 800 fois chaque plan afin d'avoir une bonne estimation des vrais risques de première espèce dans les deux cas et de les comparer.

L'estimation du risque de première espèce sera le pourcentage de rejets (sur les 800 tests) quand il n'y a pas d'interaction, c'est-à-dire pour les plans de type A et C. Pour le plan B, on a une estimation de la puissance des tests, mais cela n'a pas grand sens de les comparer car, pour le test classique, avec la statistique $\mathcal{F}_H$ on teste en même temps que l'hypothèse d'absence d'interaction, l'hypothèse d'égalité des variances.

Tableau V

Estimation des risques de première espèce $n = 150$, $\alpha = 0,08$

Plan Méthode	Aa	Ab	Ac	Ca	Cb	Cc
Méthode classique	12,8 %	13 %	30 %	10 %	10 %	22,2 %
Méthode nouvelle	5,7 %	6,7 %	7,6 %	5 %	5 %	7 %

On peut voir sur ces résultats que le test au niveau 5 % d'absence d'interaction proposé par la nouvelle méthode donne un risque de première espèce dans tous les cas, très proche de 5 %. Tandis que le test par la méthode classique donne un risque de première espèce toujours très supérieur à 5 % puisqu'on le trouve égal au minimum à 10 %.

Le test classique est aussi beaucoup plus sensible au déséquilibre du plan, car on voit le risque de première espèce augmenter très fortement quand les nombres de répétitions des combinaisons des traitements deviennent très différents entre eux. (plans Ac et Cc).

III – CONCLUSION

Bien qu'étant basée sur des considérations asymptotiques, la méthode de test que nous venons de proposer, dans le cas où la variance des observations est fonction connue de leur espérance, apparaît donc bien plus fiable que la méthode classique.

Remerciements

Les auteurs tiennent à exprimer leurs remerciements à la Société Shell Française et tout spécialement à Monsieur Daniel Benoist pour leur aide matérielle et morale pendant toute la durée de cette étude.

REFERENCES BIBLIOGRAPHIQUES

- [0] *Méthodes d'essais des produits pétroliers* Ed. Afnor
- [1] J. ULMO et J. BERNIER – *Éléments de décision statistique*. Presses Universitaires de France 1973
- [2] C. FOURGEAUD et A. FUCHS – *Statistique*. Dunod 1967

- [3] Y. MOUGEY — *Plans d'expériences à deux facteurs contrôlés dans lesquels la variance est fonction connue de la moyenne*. Mémoire de DEA de mathématiques. Université de Paris Sud (Orsay). Mai 1972
- [4] H. SCHEFFE — *The analysis of variance*. Wiley 1969
- [5] C. R. RAO — *Linear statistical inference and its applications*. Wiley 1968
- [6] ABRAMOWITZ and STEGUN — *Handbook of mathematical functions*. Dover 1968