

REVUE DE STATISTIQUE APPLIQUÉE

F. BOUTROS-TONI

L'analyse séquentielle exhaustive. Application en médecine

Revue de statistique appliquée, tome 29, n° 4 (1981), p. 31-50

http://www.numdam.org/item?id=RSA_1981__29_4_31_0

© Société française de statistique, 1981, 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/>

L'ANALYSE SEQUENTIELLE EXHAUSTIVE APPLICATION EN MEDECINE

F. BOUTROS-TONI⁽¹⁾

De par son essence l'analyse séquentielle est très différente de l'analyse classique. Lorsqu'on teste l'hypothèse nulle H_0 contre l'hypothèse alternative H_1 nous distinguons deux types d'erreurs :

- L'erreur de première espèce, α , qui est la probabilité de décider H_1 lorsque H_0 est vraie et,
- L'erreur de seconde espèce, β , qui est la probabilité de décider H_0 lorsque H_1 est vraie.

Alors que l'analyse classique fixe à l'avance la taille de l'échantillon d'expérimentation pour donner un verdict au moyen d'un résultat interprété avec un certain seuil d'erreur, la démarche de l'analyse séquentielle est pratiquement inverse, car la taille de l'échantillon n'est pas fixée. Issue de la théorie de la décision l'analyse séquentielle (ou progressive) a été créée par WALD (1947) et BARNARD (1946) et appliquée en médecine par divers auteurs dont ARMITAGE. Lorsque les problèmes biologiques ou médicaux étudiés font intervenir les résultats sous la forme d'une alternative : "bon ou mauvais", "succès ou échec", "meilleur ou moins bon"... , le modèle utilisé est celui de l'urne de BERNOULLI. Mais celle-ci exige qu'après chaque prélèvement il y ait remise (tirage non exhaustif) ou encore que l'échantillon prélevé soit négligeable devant l'effectif total. Or, un malade traité ne devrait plus être remis dans le lot de ceux qui restent et plus particulièrement si la taille de la population dont on tire au hasard l'échantillon est petite. Il n'est pas rare dans l'expérimentation biologique et surtout médicale que la taille de l'échantillon prélevé ne soit pas grande.

La solution correcte dans ce cas est le modèle hypergéométrique. C'est à cet aspect particulier que répond ce travail qui se différencie de celui que M. DUMAS (4a, b, c) a traité dans cette revue, par la solution graphique que nous donnons. Celle-ci tout en étant rigoureuse est d'une grande simplicité. Nous avons standardisé sa représentation.

(1) Centre Hospitalier Universitaire — Faculté de Médecine de COCODY B.P. V-166 Abidjan (Côte d'Ivoire).

I. LE TEST SEQUENTIEL EXHAUSTIF DU RAPPORT DES PROBABILITES

Soit une population de taille N que l'on désire soumettre à un traitement T_0 dont la probabilité de succès est π_0 . Il n'y a qu'une alternative possible avec cette population : bonne réponse ou aucune réponse au traitement T_0 . Dans le premier cas le nombre de succès est $B_0 = N\pi_0$ et dans le second cas le nombre d'échecs est $D_0 = N(1 - \pi_0) = N - B_0$.

Effectuons un tirage au sort sans remise que nous appellerons tirage exhaustif. Soient b le nombre de succès et d celui des échecs au traitement dans l'échantillon prélevé de taille $n = b + d$. Soit P_0 la probabilité que l'échantillon de taille n contienne b succès et $d = n - b$ échecs. Le nombre de cas favorables est égal au produit $\binom{B_0}{b} \cdot \binom{D_0}{d}$, celui des cas possibles étant $\binom{N}{n}$, on a :

$$P_0 = \frac{\binom{B_0}{b} \binom{D_0}{d}}{\binom{N}{n}}$$

Or il y a $\binom{n}{b}$ suites différentes contenant b succès et $d = n - b$ échecs, ayant la même probabilité de se réaliser égale à $p_0 = P_0 / \binom{n}{b}$

$$p_0 = \frac{B_0! D_0!}{(B_0 - b)! (D_0 - d)!} \bigg/ \frac{N!}{(N - n)!} \quad (1)$$

On reconnaît les relations de la distribution hypergéométrique. Prenons, à présent, pour référence le traitement connu T_0 et testons un nouveau traitement T_1 sur une population de taille N . On décide de n'accepter le nouveau traitement que si sa probabilité de succès est égale à une probabilité $\pi_1 > \pi_0$. Cette probabilité π_1 peut être soit fixée à l'avance, soit le plus souvent déterminé par une expérience-pilote.

Avec le traitement T_1 nous voulons tester l'hypothèse nulle H_0 dans laquelle la probabilité de succès est π_0 contre l'hypothèse H_1 dans laquelle la probabilité de succès est π_1 . Dans l'hypothèse H_1 le nombre de succès est $B_1 = N\pi_1$ et celui des échecs est $D_1 = N(1 - \pi_1) = N - B_1$. Soit p_0 la probabilité dans l'hypothèse H_0 et p_1 la probabilité dans l'hypothèse H_1 . Nous admettrons qu'on a une probabilité de rejet inférieure ou égale à α d'avoir $d \leq D_0$ échecs dans l'hypothèse H_0 . De même, on admettra une probabilité d'acceptation inférieure à β d'avoir $d \geq D_1$ ($D_1 < D_0$) échecs dans l'hypothèse H_1 . Avec les notations précédentes p_0 est identique à (1) et on calcule de la même manière :

$$P_1 = \frac{\binom{B_1}{b} \binom{D_1}{d}}{\binom{N}{n}}$$

$$p_1 = \frac{B_1! D_1!}{(B_1 - b)! (D_1 - d)!} \bigg/ \frac{N!}{(N - n)!} \quad (2)$$

Le rapport de vraisemblance est $\frac{p_1}{p_0} = \frac{P_1}{P_0}$ on posera :

$$f = \frac{p_1}{p_0} = \frac{B_1! (B_0 - b)!}{B_0! (B_1 - b)!} \cdot \frac{D_1! (D_0 - d)!}{D_0! (D_1 - d)!} \quad (3)$$

Les détails du test de WALD ne seront pas exposés ici. On les trouvera dans les références bibliographiques (7, 8, 9).

Avec les notations précédentes :

$$\frac{\beta}{1 - \alpha} < \frac{p_1}{p_0} < \frac{1 - \beta}{\alpha} \quad (4)$$

— Aucune décision n'est prise lorsque les inégalités restent inchangées et les essais thérapeutiques continuent.

— l'hypothèse H_0 est acceptée lorsque $\frac{\beta}{1 - \alpha} \geq \frac{p_1}{p_0}$

— l'hypothèse nulle H_0 est rejetée si on a $\frac{p_1}{p_0} \geq \frac{1 - \beta}{\alpha}$

On supposera par la suite $B_1 > B_0$ ce qui entraîne $D_1 < D_0$. Nous donnerons une solution mathématique et une solution graphique beaucoup plus intéressante pour le biologiste.

II. SOLUTION MATHEMATIQUE

On pose : $f = f(b, d) = \frac{p_1}{p_0}$, $f_A = \frac{\beta}{1 - \alpha}$ et $f_R = \frac{1 - \beta}{\alpha}$

On considère par la suite les formes logarithmiques de ces fonctions ; la solution mathématique est directement issue des règles de décision précédentes.

a) Les essais progressifs continuent et aucune décision n'est prise si l'on a :

$$\log f_A < \log f < \log f_R$$

b) On accepte l'hypothèse H_0 (et on rejette H_1) dès que l'on a :

$$\log f_A \geq \log f$$

c) On rejette l'hypothèse H_0 (et on accepte H_1) dès que l'on a :

$$\log f \geq \log f_R$$

III. SOLUTION GRAPHIQUE

De (3) on tire :

$$\log \frac{D_1! (D_0 - d)!}{D_0! (D_1 - d)!} = \log f - \log \frac{B_1! (B_0 - b)!}{B_0! (B_1 - b)!} \quad (5)$$

Posons :

$$X = \log \frac{B_1! (B_0 - b)!}{B_0! (B_1 - b)!} \quad (6)$$

$$Y = \log \frac{D_1! (D_0 - d)!}{D_0! (D_1 - d)!} \quad (7)$$

La relation (5) devient :

$$Y = -X + \log f$$

La représentation de cette fonction en coordonnées logarithmiques est une droite parallèle à la deuxième bissectrice.

A chaque essai thérapeutique, il correspond un pas logarithmique pour une valeur particulière de (6) ou (7) suivant qu'il s'agit d'un succès ou d'un échec. Nous allons décrire les pas et les frontières qui vont nous permettre de décider en faveur de H_0 ou de H_1 .

1. Pas logarithmiques

Un point $M(X, Y)$ est tel que $X = \log [\varphi(b)]$ et $Y = \log [\varphi(d)]$ où $\varphi(b)$ et $\varphi(d)$ sont respectivement des fonctions de b et d . De ce fait, les pas consécutifs sont inégaux. A chaque valeur de b ou d variant d'une unité, il correspond respectivement des pas parallèles à X et Y augmentant ou décroissant avec $\varphi(b)$ ou $\varphi(d)$.

a) Les Pas Horizontaux

De (6) on tire $X = X_0 = 0$ (pour $b = 0$), $X = X_1 = X_0 + \log \frac{B_1}{B_0}$ (pour $b = 1$) . . . et d'une façon générale pour $b = k$:

$$X_k = X_{k-1} + \log \frac{B_1 - (k-1)}{B_0 - (k-1)} \quad (8)$$

qui montre que $X_k > X_{k-1}$

b) Les Pas Verticaux

De (7) on tire $Y = Y_0 = 0$ (pour $d = 0$), $Y = Y_1 = Y_0 - \log \frac{D_0}{D_1}$ (pour $d = 1$). . . et d'une façon plus générale pour $d = j$:

$$Y_j = Y_{j-1} - \log \frac{D_0 - (j-1)}{D_1 - (j-1)} \quad (9)$$

Cette relation montre qu'on a toujours :

$$Y_j < Y_{j-1}$$

2. FRONTIERES OBLIQUES (Fig. 1)

Elles sont déterminées par les droites $Y = Y_A$ et $Y = Y_R$ obtenues en remplaçant f respectivement par $f_A = \frac{\beta}{1-\alpha}$ et $f_R = \frac{1-\beta}{\alpha}$

a) La frontière d'acceptation de l'hypothèse H_0 est :

$$Y_A = -X - \log \frac{1-\alpha}{\beta} \quad (10)$$

b) La frontière de rejet de l'hypothèse H_0 est :

$$Y_R = -X + \log \frac{1-\beta}{\alpha} \quad (11)$$

3. FRONTIERES SUPERIEURES ET INFERIEURES (Fig. 1)

a) Les frontières supérieures :

– horizontale pour $d = 0$, $X = X_0 = 0$

– verticale pour $b = 0$, $Y = Y_0 = 0$

b) Les frontières inférieures d'arrêt :

– horizontale pour $d = D_1$, $Y = Y_F$ telle que :

$$Y_F = -\log \frac{D_0!}{D_1! (D_0 - D_1)!} = -\log \binom{D_0}{D_1} \quad (12)$$

– verticales pour $b = B_0$, $X = X_F$ telle que

$$X_F = \log \frac{B_1!}{B_0! (B_1 - B_0)!} = -\log \binom{B_1}{B_0} \quad (13)$$

Si la droite $Y = Y_F$ est atteinte pour $d = D_1$ on est certain qu'avant l'épreuve la population étudiée contenait au moins D_1 individus ne répondant pas au traitement en cours. Si la droite $X = X_F$ est atteinte pour $b = B_0$ on a la certitude que la population contenait au moins B_0 individus répondant au traitement. Dans ces conditions on acceptera l'hypothèse H_0 lorsqu'on aura franchi Y_F et on la rejettera lorsqu'on aura franchi X_F . En effet, le pas vertical suivant sera obtenu pour $d = D_1 + 1$ et le pas horizontal suivant pour : $b = B_0 + 1$.

$$H_O \quad \begin{cases} Y_A = -X - \log \frac{1-\alpha}{\beta} \\ Y_F = -\log(D_{D_1}^O) \end{cases}$$

$$H_1 \quad \begin{cases} Y_R = -X + \log \frac{1-\beta}{\alpha} \\ X_F = \log \left(\frac{B_1}{B_0} \right) \end{cases}$$

$$f_A = \frac{\beta}{1-\alpha} \quad f_R = \frac{1-\beta}{\alpha}$$

5. MARCHÉ A SUIVRE DANS LA METHODE GRAPHIQUE

Dans la figure (1) on a déterminé un hexagone, délimité par :

- l'axe des ordonnées (pour $b = 0$)
- l'axe des abscisses (pour $d = 0$)
- la frontière d'acceptation oblique Y_A
- la frontière de rejet oblique Y_R
- la frontière verticale d'acceptation X_F
- la frontière horizontale inférieure de rejet Y_F .

Notons que les frontières obliques sont toujours parallèles à la deuxième bissectrice du plan XOY. Les points A_0 et R_0 sont toujours les mêmes, pour α et β donnés. Donc seules les frontières X_F et Y_F sont à déterminer pour chaque plan d'expérience par lecture directe dans les tables pour B_1 et D_1 inférieures à 100 ou en utilisant les logarithmes des factorielles si ces quantités dépassent 100. On peut trouver ces différentes tables dans HALD (Statistical Tables and Formulas ; WILEY).

Les relations générales données pour X_k et Y_j (j et k variant par unités) permettent de tracer graphiquement les pas sans utiliser les factorielles. Nous avons tabulé la fonction $\log \frac{u-k}{v-k}$ ($k = 0, 1, 2, \dots$) [réf. 3c] rendant tout calcul inutile. Les pas correspondant aux différentes valeurs de b sont portés horizontalement et ceux correspondant aux valeurs de d , perpendiculairement.

Les tables de $\log_{10} \frac{1-u}{v}$ ($0,001 \leq u, v \leq 0,10$) sont données plus loin afin de calculer les ordonnées à l'origine.

IV. FORME ASYMPTOTIQUE : LE TEST NON EXHAUSTIF

On peut déduire le rapport de vraisemblance du modèle non-exhaustif de BERNOULLI à partir de la relation (3). En prenant $B_0 = N\pi_0$, $D_0 = N(1 - \pi_0)$, $B_1 = N\pi_1$, $D_1 = N(1 - \pi_1)$ on montre dans (3a ; p. 73-74) que :

$$f = \prod_{i=0}^{i=b-1} \frac{\left(\pi_1 - \frac{i}{N}\right)}{\left(\pi_0 - \frac{i}{N}\right)} \prod_{j=0}^{j=d-1} \frac{\left(1 - \pi_1 - \frac{j}{N}\right)}{\left(1 - \pi_1 - \frac{j}{N}\right)}$$

lorsque $N \rightarrow \infty$ on aboutit au rapport de vraisemblance du test séquentiel non exhaustif :

$$f = \frac{\pi_1^b (1 - \pi_1)^d}{\pi_0^b (1 - \pi_0)^d} \quad (14)$$

Le test séquentiel non exhaustif est tel que

$$\frac{\beta}{1 - \alpha} < f < \frac{1 - \beta}{\alpha} \quad (14 \text{ bis})$$

avec les mêmes règles de décision que précédemment.

V – STATISTIQUE EXHAUSTIVE DES PREFERENCES

A. CAS GENERAL

Un traitement B a été utilisé couramment après avoir été préféré à un traitement C. L'expérimentation qui a conduit à cette décision avait montré une préférence pour B dans une proportion égale à π_0 . On décide de substituer au traitement B un nouveau traitement A à condition que la préférence de A par rapport à B soit dans une proportion $\pi_1 > \pi_0$.

1. Equation Générale

Soit une population d'effectif $2N$ individus (ou N paires d'individus). A la moitié de cet effectif on administre un traitement A, à l'autre moitié le traitement B. Supposons qu'il y ait Δ préférences pour le traitement A.

Dans la population :

- le nombre en faveur de A est $N + \Delta$
- le nombre en faveur de B est $N - \Delta$

Dans les essais progressifs dont le nombre est $2n$ (ou n paires), soit δ le nombre de préférences en faveur du traitement A.

- le nombre en faveur de A est $n + \delta$
- le nombre en faveur de B est $n - \delta$

Il suffit de remplacer dans (3) : $B = N + \Delta$, $D = N - \Delta$, $b = n + \delta$, $d = n - \delta$

Nous voulons tester l'hypothèse H_0 selon laquelle $\Delta = \Delta_0 = N\pi_0$ préférences, contre l'hypothèse H_1 dans laquelle on a $\Delta = \Delta_1 = N\pi_1$ préférences.

On aura :

$$\begin{aligned} B_0 &= N + \Delta_0 & D_0 &= N - \Delta_0 \\ B_1 &= N + \Delta_1 & D_1 &= N - \Delta_1 \end{aligned}$$

Par ailleurs les frontières seront calculées en tenant compte de :

$$\begin{aligned} 0 &\leq n + \delta \leq N + \Delta_0 \\ 0 &\leq n - \delta \leq N - \Delta_1 \end{aligned}$$

De la forme générale étudiée précédemment, on obtient :

$$\begin{aligned} f &= \left\{ \frac{(N + \Delta_1)! [(N + \Delta_0) - (n + \delta)]!}{(N + \Delta_0)! [(N + \Delta_1) - (n + \delta)]!} \right\} \left\{ \frac{(N - \Delta_1)! [(N - \Delta_0) - (n - \delta)]!}{(N - \Delta_0)! [(N - \Delta_1) - (n - \delta)]!} \right\} \\ \log &\left\{ \frac{(N - \Delta_1)! [(N - \Delta_0) - (n - \delta)]!}{(N - \Delta_0)! [(N - \Delta_1) - (n - \delta)]!} \right\} \\ &= \log f - \log \left\{ \frac{(N + \Delta_1)! [(N + \Delta_0) - (n + \delta)]!}{(N + \Delta_0)! [(N + \Delta_1) - (n + \delta)]!} \right\} \end{aligned}$$

Posons :

$$X = \log \left\{ \frac{(N + \Delta_1)!}{(N + \Delta_0)!} \frac{[(N + \Delta_0) - (n + \delta)]!}{[(N + \Delta_1) - (n + \delta)]!} \right\}$$

$$Y = \log \left\{ \frac{(N - \Delta_1)!}{(N - \Delta_0)!} \frac{[(N - \Delta_0) - (n - \delta)]!}{[(N - \Delta_1) - (n - \delta)]!} \right\}$$

$$\boxed{Y = \log f - X}$$

2. Frontières

a) Supérieures :

– verticale : $n + \delta = 0 \rightarrow X = X_0 = 0$

– horizontale : $n - \delta = 0 \rightarrow Y = Y_0 = 0$

b) Obliques :

– droite d'acceptation : $Y = Y_A ; f = f_A = \frac{\beta}{1 - \alpha}$

$$Y_A = -\log \frac{1 - \alpha}{\beta} - X$$

– droite de rejet : $Y = Y_R ; f = f_R = \frac{1 - \beta}{\alpha}$

$$Y_R = +\log \frac{1 - \beta}{\alpha} - X$$

c) Inférieures d'arrêt :

– verticale : $n + \delta = N + \Delta_0 \quad X = X_F$

$$X_F = \log \left(\frac{N + \Delta_1}{N + \Delta_0} \right)$$

– horizontale : $n - \delta = N - \Delta_1 \quad Y = Y_F$

$$Y_F = -\log \left(\frac{N - \Delta_0}{N - \Delta_1} \right)$$

3. Les Pas Horizontaux et verticaux : avec $j, k = 0, 1, 2, 3, \dots$

– Pas horizontaux : on pose $r = n + \delta$. On montre que :

$$X_k = X_{k-1} + \log \frac{(N + \Delta_1) - (r - 1)}{(N + \Delta_0) - (r - 1)} ; (\delta = 1, 2, 3, \dots) \quad (15)$$

— Pas verticaux : on pose $s = n - \delta$. On montre que :

$$Y_j = Y_{j-1} - \log \frac{(N - \Delta_0) - (s - 1)}{(N - \Delta_1) - (s - 1)} ; (\delta = -1, -2, -3 \dots) \quad (16)$$

4. Les Règles de décision

Les frontières sont identiques à celles de la figure 1. On change : $B_0 = N + \Delta_0$
 $D_1 = N - \Delta_1$

a) Aucune décision n'est prise lorsque le cheminement ne coupe aucune frontière et l'expérimentation continue.

b) L'hypothèse nulle H_0 est acceptée si l'une des deux frontières Y_A ou BM est atteinte ou traversée.

c) L'hypothèse H_0 est rejetée si l'une des deux frontières Y_R ou CM est atteinte ou traversée.

Graphiquement on portera les pas correspondant aux préférences Δ_0 le long de l'axe des abscisses ; ceux correspondant aux préférences Δ_1 seront portés en ordonnées.

B. CAS PARTICULIER

Considérons le cas où $\Delta_0 = 0$ et $\Delta_1 = \Delta$, ce qui revient à prendre la drogue B comme référence dans le cas général étudié. Le test consiste à décider en faveur de l'hypothèse nulle H_0 suivant laquelle il n'y a aucune préférence entre drogue A et drogue B ($\Delta_0 = 0$) contre l'hypothèse alternative H_1 selon laquelle la drogue A est préférable à la drogue B, le nombre de préférences s'élevant à Δ .

Les formules (15) et (16) se simplifient et on obtient :

$$X_k = X_{k-1} + \log \frac{(N + \Delta) - (r - 1)}{N - (r - 1)} \quad (17)$$

$$Y_j = Y_{j-1} - \log \frac{N - (s - 1)}{(N - \Delta) - (s - 1)} \quad (18)$$

VI. COMPARAISON DES PLANS EXHAUSTIF ET NON EXHAUSTIF

Dans un travail à paraître (3b) nous montrons que le test séquentiel exhaustif est plus économique que le test séquentiel non exhaustif. Nous ne donnerons ici que les résultats essentiels.

1. Plan séquentiel exhaustif (E) approché : figure 2

Dans (3b) on montre que les frontières du plan séquentiel exhaustif peuvent être approchées dans un système de coordonnées cartésiennes par (19) et (20), en posant $B_1 - B_0 = B$

– la frontière $d_{A(E)}$ d'acceptation de l'hypothèse nulle :

$$d_{A(E)} = \left[1 - \left(\frac{\beta}{1-\alpha} \right)^{1/B} \right] \frac{D_0 + D_1}{2} + \left(\frac{\beta}{1-\alpha} \right)^{1/B} \frac{D_0 + D_1}{B_0 + B_1} b \quad (19)$$

– la frontière $d_{R(E)}$ de rejet de l'hypothèse nulle :

$$d_{R(E)} = \left[1 - \left(\frac{1-\beta}{\alpha} \right)^{1/B} \right] \frac{D_0 + D_1}{2} + \left(\frac{1-\beta}{\alpha} \right)^{1/B} \frac{D_0 + D_1}{B_0 + B_1} b \quad (20)$$

Ces deux droites se rencontrent au point P de coordonnées :

$$\left(\frac{B_0 + B_1 + 1}{2} ; \frac{D_0 + D_1 + 1}{2} \right).$$

– les frontières supérieures d'arrêt sont : horizontale d'acceptation de H_0 pour $d = D_0$: (NQ) ; verticale de rejet de H_0 pour $b = B_0$: (MQ)

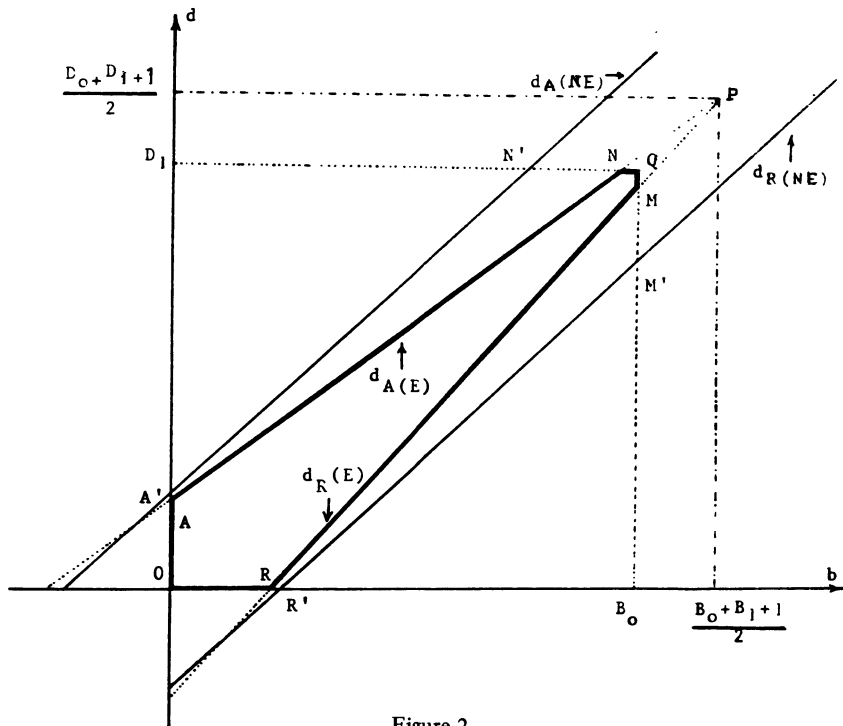


Figure 2

2. Plan séquentiel non exhaustif (NE) : figure 2

Les relations (14) et (14 bis) donnent :

– la droite d'acceptation de H_0 :

$$d_{A(NE)} = \frac{\log \frac{\beta}{1-\alpha}}{\log \frac{1-\pi_1}{1-\pi_0}} + b \frac{\log \frac{\pi_0}{\pi_1}}{\log \frac{1-\pi_1}{1-\pi_0}} \quad (21)$$

— la droite de rejet de H_0 :

$$d_{R(NE)} = \frac{\log \frac{1-\beta}{\alpha}}{\log \frac{1-\pi_1}{1-\pi_0}} + b \frac{\log \frac{\pi_0}{\pi_1}}{\log \frac{1-\pi_1}{1-\pi_0}} \quad (22)$$

Ces deux droites sont croissantes et parallèles.

3. Comparaison des deux plans : figure 2

On montre dans (3b) que :

- a) La pente de $d_{A(NE)}$ est toujours supérieure à celle de $d_{A(E)}$.
- b) L'ordonnée à l'origine A' de $d_{A(NE)}$ est supérieure à celle de $d_{A(E)}$.
- c) L'abscisse à l'origine R' de $d_{R(NE)}$ est supérieure à celle de $d_{R(E)}$.

Le plan séquentiel exhaustif approché est donc entièrement contenu dans le plan séquentiel non exhaustif. Il lui est donc inférieur, d'où son économie.

VII. APPLICATIONS NUMERIQUES (ESSAIS THERAPEUTIQUES)

1^{ère} Application

a) Modèle Hypergéométrique (exhaustif)

Nous voulons tester l'efficacité d'une drogue DR_1 en prenant pour référence une drogue DR_0 qui agit dans une proportion égale à $\pi_0 = 0.55$. On décide de n'accepter la nouvelle drogue DR_1 que si son efficacité dépasse celle de DR_0 dans une proportion $\pi_1 > \pi_0$ dans une population de taille $N = 100$.

Sachant qu'une expérience-pilote a montré que $\pi_1 = 0.75$ on peut écrire avec les notations précédentes :

- dans l'hypothèse $H_0 : B_0 = N\pi_0 = 55 \quad D_0 = 45$
- dans l'hypothèse $H_1 : B_1 = N\pi_1 = 75 \quad D_1 = 25$

Nous accepterons les risques de première et deuxième espèce $\alpha = \beta = 0.05$, ce qui entraîne :

$$\log f_A = -1.28 \quad \text{et} \quad \log f_R = 1.28.$$

Les frontières obliques sont :

$$Y_A = -X - 1.28$$

$$Y_R = -X + 1.28.$$

Lorsque l'efficacité de DR_1 dépasse celle de DR_0 on écrira 1, lorsqu'elle lui est inférieure on écrira - 1. On a obtenu la séquence suivante qui se lit de gauche à droite :

1	- 1	1	- 1	1	1	- 1	- 1	1	1	1	1	1	1	1
1	- 1	1	1	1	1	1	1	- 1	1	- 1	1	1	1	1

↑
 Décision du test exhaustif

↑
 Décision du test non exhaustif.

L'application des formules (8) et (9) permettent le calcul des pas logarithmiques verticaux et horizontaux. Il est pratique de disposer les calculs comme l'indique le tableau 1. Lorsqu'on rencontre dans la séquence précédente 1, b augmente d'une unité, lorsque - 1 est rencontré d augmente d'une unité.

α) La solution graphique (figure 3) montre qu'au 23^e pas la frontière de rejet Y_R est traversée et on conclut au rejet de l'hypothèse nulle H_0 .

La solution mathématique donne d'après le tableau 1 :

$$\log f = X_{18} + Y_5 = 1,460\,221$$

on a donc :

$$\log f > \log f_R = 1,28$$

et l'hypothèse nulle est rejetée d'après les règles de décision données précédemment dans la solution mathématique (II).

On peut vérifier comme dans la solution graphique qu'il était impossible de conclure au pas précédent puisque :

$$\log f = X_{17} + Y_5 = 1,276\,577$$

entraîne

$$\log f_A < \log f < \log f_R.$$

Nos conclusions sont absolument superposables à celles de la solution graphique.

TABLEAU 1

Rang des observations	X	Y	log f
1 obs (b = 1, d = 0)	$X_1 = X_0 + \log \frac{75}{55}$	$Y_0 = 0$	$X_1 = 0.134699$
2 (b = 1, d = 1)	X_1	$Y_1 = Y_0 - \log \frac{45}{25}$	$X_1 + Y_1 = -0.120574$
3 (b = 2, d = 1)	$X_2 = X_1 + \log \frac{74}{54}$	Y_1	$X_2 + Y_1 = 0.016264$
4 (b = 2, d = 2)	X_2	$Y_2 = Y_1 - \log \frac{44}{24}$	$X_2 + Y_2 = -0.246978$

Rang des observations		X	Y	log f
5	(b = 3, d = 2)	$X_3 = X_2 + \log \frac{73}{53}$	Y_2	$X_3 + Y_2 = -0.107931$
6	(b = 4, d = 2)	$X_4 = X_3 + \log \frac{72}{52}$	Y_2	$X_4 + Y_2 = 0.033398$
7	(b = 4, d = 3)	X_4	$Y_3 = Y_2 - \log \frac{43}{23}$	$X_4 + Y_3 = -0.238343$
8	(b = 4, d = 4)	X_4	$Y_4 = Y_3 - \log \frac{42}{22}$	$X_4 + Y_4 = -0.519170$
9	(b = 5, d = 4)	$X_5 = X_4 + \log \frac{71}{51}$	Y_4	$X_5 + Y_4 = -0.375481$
10	(b = 6, d = 4)	$X_6 = X_5 + \log \frac{70}{50}$	Y_4	$X_6 + Y_4 = -0.229353$
11	(b = 7, d = 4)	$X_7 = X_6 + \log \frac{69}{49}$	Y_4	$X_7 + Y_4 = -0.080700$
12	(b = 8, d = 4)	$X_8 = X_7 + \log \frac{68}{48}$	Y_4	$X_8 + Y_4 = 0.070513$
13	(b = 9, d = 4)	$X_9 = X_8 + \log \frac{67}{47}$	Y_4	$X_9 + Y_4 = 0.224490$
14	(b = 10, d = 4)	$X_{10} = X_9 + \log \frac{66}{46}$	Y_4	$X_{10} + Y_4 = 0.381276$
15	(b = 11, d = 4)	$X_{11} = X_{10} + \log \frac{65}{45}$	Y_4	$X_{11} + Y_4 = 0.540977$
16	(b = 12, d = 4)	$X_{12} = X_{11} + \log \frac{64}{44}$	Y_4	$X_{12} + Y_4 = 0.703704$
17	(b = 12, d = 5)	X_{12}	$Y_5 = Y_4 - \log \frac{41}{21}$	$X_{12} + Y_5 = 0.413139$
18	(b = 13, d = 5)	$X_{13} = X_{12} + \log \frac{63}{43}$	Y_5	$X_{13} + Y_5 = 0.579011$
19	(b = 14, d = 5)	$X_{14} = X_{13} + \log \frac{62}{42}$	Y_5	$X_{14} + Y_5 = 0.748153$
20	(b = 15, d = 5)	$X_{15} = X_{14} + \log \frac{61}{41}$	Y_5	$X_{15} + Y_5 = 0.920699$
21	(b = 16, d = 5)	$X_{16} = X_{15} + \log \frac{60}{40}$	Y_5	$X_{16} + Y_5 = 1.096790$
22	(b = 17, d = 5)	$X_{17} = X_{16} + \log \frac{59}{39}$	Y_5	$X_{17} + Y_5 = 1.276577$
23	(b = 18, d = 5)	$X_{18} = X_{17} + \log \frac{58}{38}$	Y_5	$X_{18} + Y_5 = 1.460221$

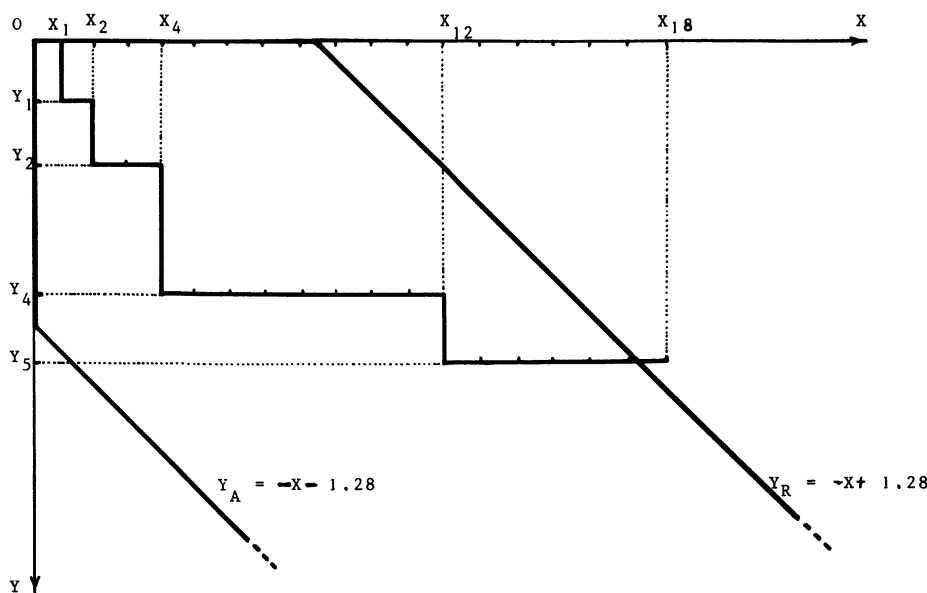


Figure 3

b) Comparaison avec le modèle de BERNOULLI (Fig. 4)

L'application des relations (21) et (22) aux données précédentes entraîne :

$$d_{A(NE)} = 0,527 b + 5,02$$

$$d_{R(NE)} = 0,527 b - 5,02$$

Ces droites sont portées sur la figure 4. Sur celle-ci le cheminement est tel que lorsqu'on rencontre 1 dans la séquence qui précède, le pas est porté parallèlement à l'axe des abscisses, lorsqu'on rencontre -1, le pas est porté parallèlement à l'axe des ordonnées. On voit que la décision de rejeter H_0 a lieu plus tardivement, sept pas (en pointillé sur la figure 4) plus loin que dans le test séquentiel exhaustif. Ceci vérifie l'économie de ce dernier sur le modèle non exhaustif de Bernoulli.

c) Le plan exhaustif approché (Fig. 4)

On peut se demander si l'application du plan exhaustif approché, donne une décision sensiblement différente du plan exhaustif exact du paragraphe (a).

Avec les données précédentes les frontières du plan séquentiel exhaustif approché s'écrivent après application des relations (19) et (20).

$$d_{A(E)} = 0,465 b + 4,79$$

$$d_{R(E)} = 0,624 b - 5,55$$

Ces droites sont portées en pointillé sur la figure 4 et vérifient que le plan séquentiel exhaustif est inférieur au plan non exhaustif. Il est remarquable que dans cette application, la décision de rejeter H_0 soit prise au 23^e pas, exactement comme dans le plan exhaustif du paragraphe (a).

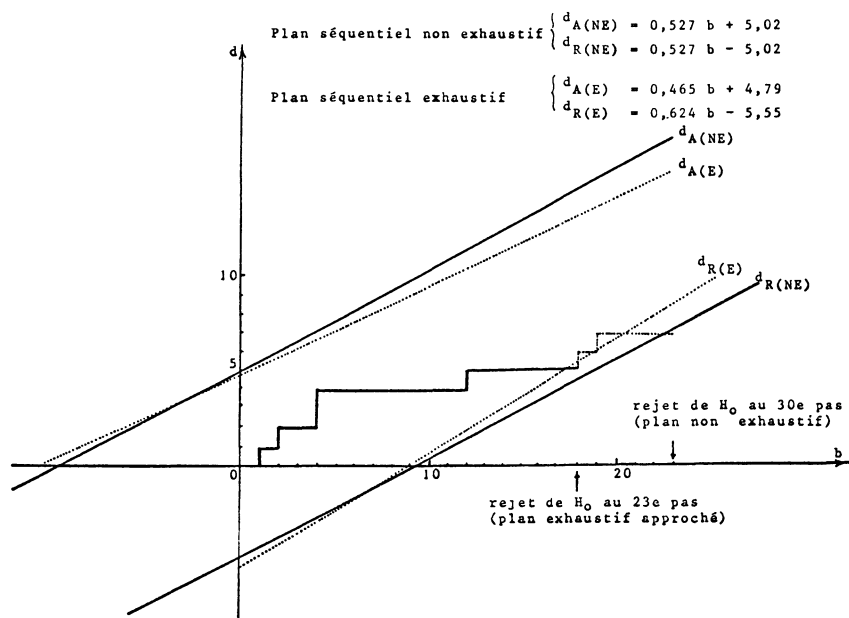


Figure 4

2^e Application

Nous donnerons une application de la statistique exhaustive des préférences.

On veut comparer l'efficacité de deux drogues DR_0 et DR_1 . L'expérimentation a lieu sur $2N$ malades. Une moitié reçoit DR_0 , l'autre reçoit DR_1 . Une expérience-pilote a montré que DR_1 s'est montrée plus efficace que DR_0 dans une proportion s'élevant à $\pi = 0,65$. Nous voulons tester l'hypothèse nulle H_0 selon laquelle DR_1 n'est pas préférable à DR_0 contre l'hypothèse alternative H_1 selon laquelle DR_1 est préférable à DR_0 . Nous accepterons les risques de première et deuxième espèce $\alpha = \beta = 0,05$. On prendra $N = 100$.

L'application des relations du cas particulier du paragraphe VB donne :

– les frontières obliques :

$$Y_A = -X - 1,28$$

$$Y_R = -X + 1,28$$

– les pas logarithmiques sont donnés par (17) et (18) dans lesquelles $r = n + \delta$ et $s = n - \delta$, l'effectif prélevé étant n et δ est tel que lorsque l'efficacité est en faveur de DR_1 on a $\delta = +1$, lorsqu'elle est en faveur de DR_0 on a $\delta = -1$.

On a obtenu la séquence suivante des essais thérapeutiques, où 1 désigne l'efficacité en faveur de DR_1 et -1 celle en faveur de DR_0 :

décision de rejeter H_0
↓

1	1	1	-1	1	1	1	1	1	1	-1	-1	1	-1
1	1	1	-1	1	-1	1	-1	1	1	-1	-1	1	-1

Le nombre de préférence en faveur de DR_1 est $\Delta = 100 \times 0,65 = 65$. Les pas s'écrivent, sachant que $X_0 = Y_0 = 0$.

$$X_k = X_{k-1} + \log \frac{100 + 65 - [(n + \delta) - 1]}{100 - [(n + \delta) - 1]}$$

$$Y_j = Y_{j-1} - \log \frac{100 - [(n - \delta) - 1]}{35 - [(n - \delta) - 1]}$$

On place les pas en faveur de DR_0 parallèlement à Y et ceux en faveur de DR_1 parallèlement à X.

Il est pratique de disposer les calculs selon le tableau 2. Il y a eu 18 comparaisons ayant nécessité $2 \times 18 = 36$ essais. On peut vérifier que l'on a bien $(n + \delta) + (n - \delta) = 2n$. En effet, au pas X_{15} on a $n + \delta = 30$ et au pas Y_3 on a $n - \delta = 6$ ce qui donne bien $2n = 36$.

En fait, la solution graphique montre que la décision de rejeter H_0 est prise au 9^e pas et on conclut que la drogue DR_1 est plus efficace que DR_0 (figure 5).

TABLEAU 2

DROGUE DR_1 (n essais)				DROGUE DR_0 (n essais)			
n	δ	$n + \delta$	X	n	δ	$n - \delta$	Y
1	1	2	$X_1 = 0,219209$	1	- 1	2	$Y_1 = - 0,464156$
2	2	4	$X_2 = 0,222743$				
3	3	6	$X_3 = 0,226400$				
4	4	8	$X_4 = 0,230174$				
5	5	10	$X_5 = 0,234083$				
6	6	12	$X_6 = 0,238131$				
7	7	14	$X_7 = 0,232454$				
8	8	16	$X_8 = 0,246672$				
9	9	18	$X_9 = 0,251184$				
10	10	20	$X_{10} = 0,255868$				
11	11	22	$X_{11} = 0,265798$	2	- 2	4	$Y_2 = - 0,4816218$
12	12	24	$X_{12} = 0,265798$				
13	13	26	$X_{13} = 0,271067$	3	- 3	6	$Y_3 = - 0,500600$
14	14	28	$X_{14} = 0,276556$				
15	15	30	$X_{15} = 0,282281$				

VIII. COMMENTAIRES

Le test séquentiel exhaustif tel qu'il est élaboré n'exige pas plus de calcul que celui du modèle de Bernoulli. Il a été standardisé. Les frontières obliques sont toujours parallèles à la deuxième bissectrice et pour α et β données, ces frontières

sont toujours identiques. Les pas sont aisément programmables avec de simples calculatrices de poche. D'ailleurs la solution mathématique elle-même ne fait nullement intervenir les factorielles en utilisant les formules (8) et (9) comme on l'a montré dans le premier exemple numérique d'application.

Nous avons voulu orienter les applications sur les essais thérapeutiques, en insistant sur l'aspect très pratique de ce qu'on a appelé la Statistique Exhaustive des Préférences. Sous cette forme nous avons appliqué le test dans de multiples domaines de la physiologie et plus particulièrement dans l'étude du temps de réponse à un stimulus auditif dans deux situations différentes (3a). Par comparaison avec le modèle de BERNOULLI la décision a toujours été prise plus tôt avec le test exhaustif. Ce test, de par sa nature possède une frontière d'arrêt au bout de laquelle, une décision doit être prise alors qu'avec le test non exhaustif, elle n'existe pas à l'origine. Et pour tronquer l'épreuve statistique non exhaustive certains auteurs (1c) ont abouti à un compromis d'ordre pratique.

On a de plus, montré l'économie que présente le test séquentiel exhaustif par rapport au modèle de BERNOULLI, problème d'ailleurs déjà évoqué par M. DUMAS (4b). De plus ARMITAGE (1c) a déjà montré que l'analyse non exhaustive était plus économique que l'analyse classique. Le modèle de la distribution hypergéométrique apparaît donc le plus avantageux.

Loin de nous l'idée de jeter le discrédit sur le modèle de Bernoulli qui pourra être utilisé lorsque l'échantillon prélevé n'excède pas le dixième de l'effectif de la population.

TABLE DE $\log_{10} \frac{1-u}{v}$

$\begin{smallmatrix} u \\ \backslash \\ v \end{smallmatrix}$	0.001	0.01	0.02	0.025	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
0.001	2.9996	2.9956	2.9912	2.9890	2.9868	2.9823	2.9777	2.9731	2.9685	2.9638	2.9590	2.9542
0.01	1.9996	1.9956	1.9912	1.9890	1.9868	1.9823	1.9777	1.9731	1.9685	1.9638	1.9590	1.9542
0.02	1.6986	1.6946	1.6902	1.6880	1.6858	1.6813	1.6767	1.6721	1.6675	1.6628	1.6580	1.6532
0.025	1.6017	1.5977	1.5933	1.5911	1.5889	1.5844	1.5798	1.5752	1.5706	1.5659	1.5611	1.5563
0.03	1.5225	1.5185	1.5141	1.5119	1.5097	1.5052	1.5006	1.4960	1.4914	1.4867	1.4819	1.4771
0.04	1.3975	1.3935	1.3891	1.3869	1.3847	1.3802	1.3756	1.3710	1.3664	1.3617	1.3569	1.3521
0.05	1.3006	1.2966	1.2922	1.2900	1.2878	1.2933	1.2787	1.2741	1.2695	1.2648	1.2600	1.2552
0.06	1.2214	1.2174	1.2130	1.2108	1.2086	1.2041	1.1995	1.1949	1.1903	1.1856	1.1808	1.1760
0.07	1.1545	1.1505	1.1461	1.1439	1.1417	1.1372	1.1326	1.1280	1.1234	1.1187	1.1139	1.1091
0.08	1.0965	1.0925	1.0881	1.0859	1.0837	1.0792	1.0746	1.0700	1.0654	1.0607	1.0559	1.0511
0.09	1.0454	1.0414	1.0370	1.0348	1.0326	1.0281	1.0235	1.0189	1.0143	1.0096	1.0048	1.0000
0.10	0.9996	0.9956	0.9912	0.9890	0.9868	0.9823	0.9777	0.9731	0.9685	0.9638	0.9590	0.9542

BIBLIOGRAPHIE

- [1] ARMITAGE P. :
 - a) Restricted sequential Procedures. *Biometrika*, **44**, 9-26.
 - b) Numerical studies in the sequential estimation of a binomial parameter. *Biometrika*, **45**, 1-15 (1958).
 - c) Sequential Medical trials. Blackwell, Oxford (1960)
 - d) Sequential Medical trials, some comments on F. J. Anscombe's *J. Amer. Stat. Ass.*, **58**, 384-387.
 - e) Statistical Methods in Medical Research. Chap. 15 : Sequential Methods. Blackwell (1971).
- [2] BARNARD G.A. — Sequential Tests in Industrial Statistics. *Journ. Roy. Stat. Soc. suppl.*, **8**, 1-26 (1946).
- [3] BOUTROS-TONI F. :
 - a) Le Test Séquentiel Exhaustif et sa forme asymptotique Thèse de Doctorat en Biologie Humaine (Mathématiques et Statistiques). N° d'ordre : 2 (1972) Université de Bordeaux II.
 - b) Economie du test séquentiel exhaustif. *Annales de l'Université d'Abidjan*. A paraître.
 - c) Tables pour le test séquentiel exhaustif. *Ann. Univ. Abidjan*, 1981.
- [4] DUMAS M. :
 - a) L'épreuve séquentielle exhaustive. *Revue de Stat. Appliquée*. **17** (n° 1) ; 5-40 ; (1969).
 - b) L'épreuve séquentielle exhaustive. *Revue de Stat. Appliquée*. **18** (n° 2) ; 31-43 (1970).
 - c) Tables commentées des épreuves séquentielles exhaustives et non exhaustives. 1971. DUNOD.
- [5] HALD A. — Statistical Tables and Formulas. WILEY, 1970.
- [6] NEYMAN J. and PEARSON E.S. — On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika*, **20A**, part I : 175-240 ; part II : 263-294.
- [7] SCHWARTZ D., FLAMANT R., LELLOUCH J. — *L'essai thérapeutique chez l'homme*. Chap. XVI., 1970 FLAMMARION.
- [8] WALD A. — *Sequential Analysis*, WILEY 1966.
- [9] WETHERILL G.B. — *Méthodes séquentielles en Statistiques*, 1969 DUNOD.