

REVUE DE STATISTIQUE APPLIQUÉE

MAURICE DUMAS

L'épreuve séquentielle exhaustive

Revue de statistique appliquée, tome 18, n° 2 (1970), p. 31-43

http://www.numdam.org/item?id=RSA_1970__18_2_31_0

© Société française de statistique, 1970, 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'ÉPREUVE SÉQUENTIELLE EXHAUSTIVE

Maurice DUMAS

La Revue de Statistique Appliquée a publié récemment sous le même titre que ci-dessus, une étude relative à l'adaptation au cas exhaustif de la théorie de l'épreuve séquentielle, non-exhaustive, de Wald.

Des travaux récents ont conduit l'auteur de cette étude à aboutir dans la même voie aux quelques résultats nouveaux exposés ci-après.

SOMMAIRE

- 1 - Introduction
- 2 - Notations et conventions
- 3 - Premiers résultats
- 4 - Allure générale des courbes $C(x,y) = \lambda$
- 5 - Lignes d'acceptation et de rejet
- 6 - Comparaison des épreuves séquentielles exhaustive et non-exhaustive
- 7 - Cas particuliers : risques nuls
- 8 - Règles de l'épreuve séquentielle exhaustive
- 9 - Exemple numérique
- 10 - Plans approchés

1 - INTRODUCTION

Nous avons récemment publié⁽¹⁾ l'état de nos recherches concernant l'adaptation de l'épreuve séquentielle de Wald au cas où l'exhaustivité des prélèvements est prise en considération.

Nous avons alors établi différents points, notés brièvement au n° 3, à l'aide des notations rappelées au n° 2.

Depuis lors, nous avons cherché à préciser le tracé des courbes $C(x,y) = C^{\text{te}}$ servant à définir les lignes d'acceptation et de rejet constituant le plan, et cela nous a amené à revenir sur un point de la théorie exposée ;

(1) L'épreuve séquentielle exhaustive.

Revue de Statistique Appliquée 1969, Vol XVII N° 1

Ce mémoire est simplement désigné dans la suite par le signe bibliographique [1].

corrélativement sont apparus deux avantages : théorie simplifiée et accord plus étroit qu'auparavant réalisé entre les épreuves séquentielles exhaustive et non-exhaustive.

Entre temps, du 3 au 11 Septembre 1969, l'Institut International de Statistique a tenu à Londres sa 37ème Session biennale. Il avait donné pour thème à une de ses sections "Règles d'arrêt optimum et plan séquentiel".

Considérant combien nos travaux se rattachaient à ce thème, nous avons profité de ce que nous avons aimablement été invité à assister à la Session, pour établir, en mai 1969, une communication qui, sous le titre "Epreuve séquentielle et exhaustivité", résumait les résultats publiés en [1]. Mais lors de la présentation orale de notre communication, nous avons pu faire état des résultats nouveaux exposés ci-après.

2 - NOTATIONS ET CONVENTIONS

a) Nous conservons les notations de notre [1], à savoir :

U est l'effectif du lot

p est le pourcentage des défectueux et u est le nombre des défectueux contenus dans le lot

$$Up = u$$

u_1 et u_2 sont les nombres de défectueux correspondant respectivement à α (risque du producteur) et à β (risque du client).

S désigne la différence positive $u_2 - u_1$.

La quantité $C(x,y)$, dite "cote" du point x,y , a pour expression :

$$C(x,y) = \frac{u_2! (U - u_2)! (u_1 - y)! (U - u_1 - x)!}{u_1! (U - u_1)! (u_2 - y)! (U - u_2 - x)!}$$

b) L'exposé est fait en admettant que l'épreuve donne lieu à un cheminement à partir de l'origine sur un plan avec en abscisse x , nombre des bons (et non pas : nombre total des essais) et en ordonnée y , nombre des défectueux.

Dans le plan correspondant (fig. 1), on définit une partie utile limitée par les axes et par les deux droites $y = u_1 + 1$ et $x = U - u_2 + 1$. Soit OBDC le rectangle correspondant.

Sur ce plan on peut porter le point M, de coordonnées

$$U - \frac{u_1 + u_2 - 1}{2} \quad \text{et} \quad \frac{u_1 + u_2 + 1}{2}$$

La droite à 45°, d'équation

$$y = x - U + u_1 + u_2$$

passé à la fois par le point M, à l'extérieur du rectangle, et par le point D, sommet du rectangle ; elle rencontre l'axe des x en E.

3 - PREMIERS RESULTATS

Nous avons montré dans [1] dans quelles conditions l'on était amené à considérer la quantité $C(x,y)$, puis nous avons reconnu que le problème était de déterminer si possible des quantités λ_A et λ_R , fonctions seulement de α et de β , destinées à intervenir dans les équations

$$C(x,y) = \lambda_A \quad \text{et} \quad C(x,y) = \lambda_R$$

définissant respectivement les lignes d'acceptation et de rejet.

Nous avons reconnu également que toute courbe telle que celles-ci avait, au voisinage des axes $x = 0$ et $y = 0$, l'allure générale d'une courbe passant par le point M de la figure 1.

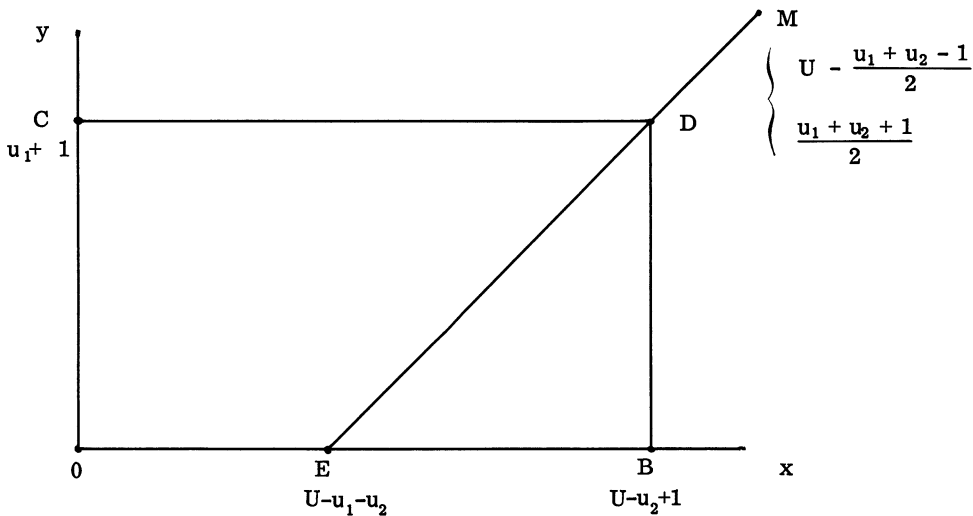


Figure 1 - Partie utile du plan x, y .

D'après la position de M par rapport au rectangle OBDC, nous avons pensé que trois cas de figure étaient à examiner, à savoir :

- le cas faisant l'objet du n° 6 de [1], c'est-à-dire celui où les lignes A et R coupent le rectangle respectivement entre B et D et entre C et D ;

- les deux cas faisant l'objet du n° 7 de [1], c'est-à-dire ceux où ces deux lignes coupent, l'une et l'autre, le rectangle entre B et D ou entre C et D.

Or, la poursuite de l'étude des courbes $C(x,y) = \lambda$ a montré (voir le n° 4 ci-après) que toutes les courbes passaient par le point D, seul cas qui n'avait pas été envisagé dans [1].

4 - ALLURE GENERALE DES COURBES $C(x,y) = \lambda$

a) Une étude non détaillée plus loin permet d'arriver à une bonne connaissance des courbes $C(x,y) = \lambda$. Nous en retenons particulièrement les points suivants, valables pour toute courbe de cette famille :

1 - La courbe coupe l'axe des x en un point d'abscisse x_1 , inférieure à $U - u_2 + 1$ et éventuellement négative : x_1 , est donné par

$$\frac{(U - u_2)! (U - u_1 - x_1)!}{(U - u_1)! (U - u_2 - x_1)!} = \lambda \text{ d'où } x_1 \approx \left[U - \frac{u_1 + u_2 - 1}{2} \right] (1 - \lambda^{1/s})$$

En ce point la courbe admet pour tangente la droite qui - pour $x_1 < U - u_1 - u_2$, ce qui est le cas général, et en particulier pour $x_1 \leq 0$ - passe très approximativement par le point M_1 de coordonnées

$$U - \frac{u_1 + u_2 - 1}{1} \quad \text{et} \quad \frac{u_1 + u_2 + 1}{2} \times \frac{1}{1 + \frac{S^2 - 1}{3(u_1 + u_2 + 1)^2}}$$

2 - Quel que soit λ , la courbe passe par le point D. En ce point D, la tangente à la courbe a une pente bien déterminée, à savoir

$$\frac{1}{\lambda} \frac{u_2!}{u_1!} \frac{(U - u_2)!}{(U - u_1)!} ;$$

on se rend compte que très généralement cette pente pourra être considérée comme nulle, de sorte que, pratiquement, la tangente est la droite CD.

3 - L'allure générale de la courbe est celle indiquée par la figure 2, en fonction de la valeur de λ .

b) Les seules indications qui précèdent sont de nature à permettre de tracer approximativement la courbe elle-même.

Il est bon cependant de préciser cette courbe notamment par le calcul des points suivants :

1 - point où la courbe a pour ordonnée u_1 ; car au voisinage de ce point, comme il a été dit en B.4 de [1], se fait sentir l'influence du minimum que présente la fonction $\Gamma(n)$ pour n compris entre 1 et 2 ;

2 - si x est négatif, point où la courbe traverse l'axe des y ; l'ordonnée y_1 de ce point est donnée par :

$$\frac{u_2!}{u_1!} \frac{(u_1 - y_1)!}{(u_2 - y_1)!} = \lambda \quad \text{d'où} \quad y_1 \approx \frac{u_1 + u_2 + 1}{2} \left(1 - \frac{1}{\lambda^{1/s}} \right)$$

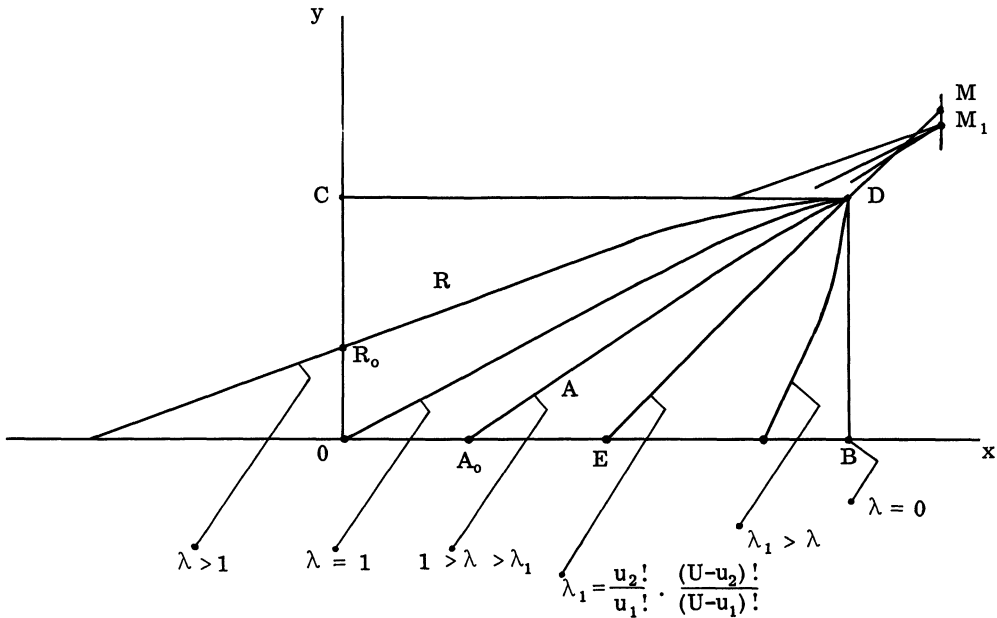


Figure 2 - Courbes $C(x, y) = \lambda$ pour différentes valeurs de λ .

5 - LIGNES D'ACCEPTATION ET DE REJET

a) Les équations des lignes d'acceptation et de rejet sont respectivement de la forme :

$$C(x, y) = \lambda_A < 1 \quad \text{et} \quad C(x, y) = \lambda_R > 1$$

D'après l'étude en cause au n° 4, ces deux courbes se présentent l'une par rapport à l'autre comme les courbes A et R de la figure 2 ; éventuellement, la courbe A est remplacée par la droite DE ou même par une courbe située entre DE et DB (l'une de ces courbes est notée sur la fig. 2), mais en tout état de cause les deux courbes A et R ont en commun le point D.

La recherche des valeurs λ_A et λ_R en fonction de α et de β se fait comme au n° 6 de [1] mais les quantités ε_1 et ε_2 qui y sont considérées sont, d'après ce qui précède, à prendre l'une et l'autre égales à zéro ; de la sorte on arrive aux valeurs suivantes :

$$\lambda_A = \frac{\beta}{1 - \alpha} \quad \text{et} \quad \lambda_R = \frac{1 - \beta}{\alpha}$$

b) Les lignes A et R définies comme il vient d'être dit peuvent être déterminées avec toute la précision que l'on voudra, en s'aidant des indications de l'annexe B de [1]. Les calculs correspondants sont réduits au minimum si l'on tient compte de ceci :

- d'une part, seules sont à considérer les valeurs entières de y , et pour chacune d'elles il suffit que le calcul conduise à déterminer deux valeurs entières consécutives de x donnant à $C(x, y)$ des valeurs qui encadrent, suivant la ligne en cause, soit λ_A soit λ_R .

- d'autre part, pour chacune des déterminations ci-dessus, on dispose d'une valeur de x pouvant servir de bon point de départ à une recherche par approximations successives : cette valeur de x est la valeur lue sur le graphique $C(x,y) = \lambda$ tracée très simplement d'après les seules indications du n° 4.

c) L'exposé des paragraphes a et b ci-dessus donne au problème de l'épreuve séquentielle exhaustive une solution qui, par rapport à celle faisant l'objet de [1], apparaît à la fois comme plus satisfaisante et plus simple :

- plus satisfaisante, puisque les valeurs des rapports de vraisemblances à considérer, à savoir $\beta/(1-\alpha)$ et $(1-\beta)/\alpha$, sont les mêmes que dans le cas de l'épreuve séquentielle non-exhaustive ;

- plus simple, puisque, notamment, les n°s 6 et 7 de [1] qui envisageaient plusieurs cas, deviennent sans objet : un seul cas est à considérer et ce cas est traité exactement comme s'il s'agissait de l'épreuve séquentielle non-exhaustive.

Corrélativement, les n°s 8 et 9 de [1] seraient à modifier plus ou moins profondément : en raison de leur importance, leurs sujets sont repris en 6 et 8 ci-après ; quelques autres points de [1] mériteraient de subir des retouches légères mais nous ne jugeons pas utile d'explicitier ces retouches : l'absence d'indication à leur sujet ne saurait gêner le lecteur.

6 - COMPARAISON DES EPREUVES SEQUENTIELLES EXHAUSTIVE ET NON-EXHAUSTIVE

La figure 3 représente schématiquement sur un même système d'axes ox , oy , et pour de mêmes valeurs des risques α et β associés respectivement à $u_1 = Up_1$ et $u_2 = Up_2$ d'une part le contour OA_0DR_0 , et les points M et M_1 , de l'épreuve séquentielle exhaustive, et d'autre part les droites de Wald tracées à partir du point A'_0 sur ox et R'_0 sur oy . La figure 5 porte les mêmes points, calculés dans un cas concret.

On remarquera que A_0 et R_0 sont situés entre 0 et respectivement A'_0 et R'_0 : cette disposition est absolument générale.

Les points D et M sont à l'intérieur des droites de Wald ; nous pensons qu'il en est très généralement ainsi puisque :

- les droites de Wald ont, d'après le N.B. du N° 4 a de [1], une pente voisine de $0,5(p_1 + p_2)$;

- la droite OM a une pente égale à $(u_1 + u_2 + 1)/[2U - u_1 - u_2 + 1]$, soit sensiblement $0,5(p_1 + p_2)$;

- la droite OD a une pente égale à $(u_1 + 1)/U - u_2 + 1$, valeur quelque peu supérieure à p_1 .

Dans ces conditions, le contour OA_0DR_0 est "toujours" entièrement à l'intérieur des droites de Wald : il en résulte que l'épreuve séquentielle exhaustive est plus économique que l'épreuve séquentielle non-exhaustive.

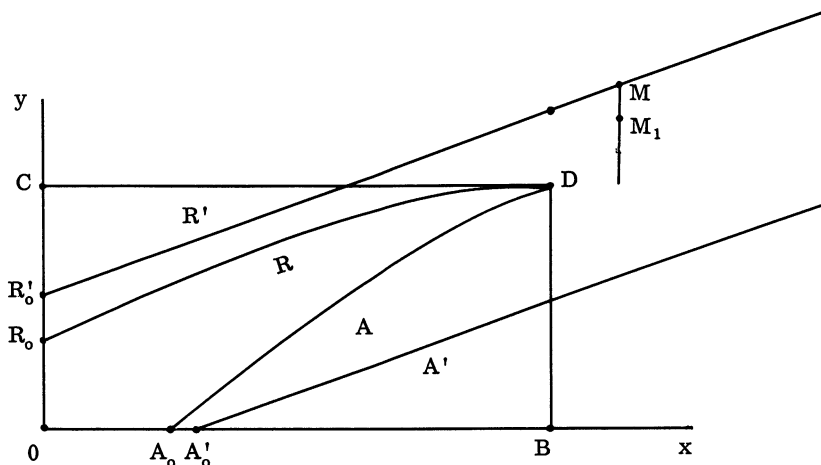


Figure 3 - Comparaison des épreuves.

b) Quelles que soient les valeurs de U , de u_1 et de u_2 et par suite aussi de p_1 et de p_2 , il est possible de calculer les points des lignes A et R de l'épreuve séquentielle exhaustive, et aussi ceux des deux droites de Wald, correspondant au cas de la non-exhaustivité. Mais tandis que les premières sont toujours valables, les droites de Wald ne le sont que sous la condition que le nombre d'essais soit petit par rapport à U : l'intérêt de l'épreuve séquentielle exhaustive est surtout grand lorsque l'on peut prévoir que cette condition ne sera pas remplie.

c) Avec l'épreuve séquentielle non-exhaustive, convenir d'un tronquage est une obligation ; avec l'épreuve séquentielle exhaustive, on n'est jamais dans une telle obligation, et il est bon de ne pas avoir recours à un tronquage puisque dans toute épreuve séquentielle, exhaustive ou non-exhaustive, ce recours soulève quelques difficultés concernant les valeurs des risques encourus réellement.

En cas d'épreuve séquentielle non-exhaustive, le partage entre acceptation et rejet se fait à l'intersection de la droite de tronquage et de la droite parallèle aux droites de Wald passant par l'origine.

Si en cas d'épreuve séquentielle exhaustive, un tronquage était prévu, le même partage se ferait sur la "ligne neutre", ligne de la même famille que les lignes A et R et passant par l'origine ; l'équation de la ligne neutre est donc

$$C(x,y) = 1$$

Sur la figure 4 a été porté notamment le point Z d'intersection de la ligne neutre et de la droite de tronquage

$$x + y = n$$

Bien entendu, le choix de n , tel que les risques α et β , prévus dans la définition du plan, soient approximativement conservés en cas de tronquage, implique une étude théorique qui n'est pas autrement évoquée dans la présente note.

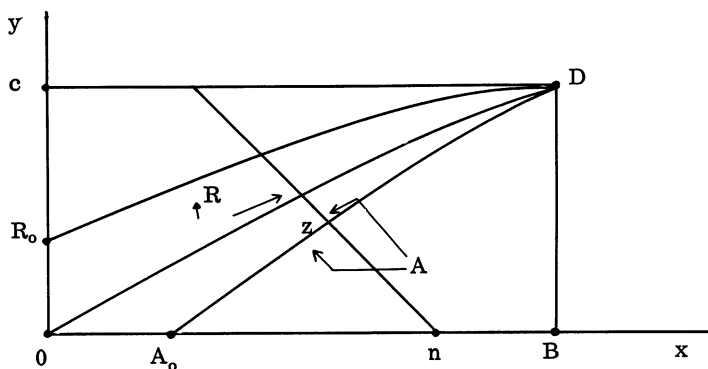


Figure 4 - Epreuve tronquée.

7 - CAS PARTICULIERS : RISQUES NULS

a) Si $\alpha = 0$, la règle de l'épreuve séquentielle exhaustive s'applique sans difficulté : la ligne de rejet R coïncide alors avec la droite CD ($y = u_1 + 1$) de la figure 1.

A ce cas limite est en tout point assimilable celui où α est assez petit pour que la ligne R coupe l'axe des y entre les deux ordonnées u_1 et $u_1 + 1$, et par suite, reste comprise entre ces deux ordonnées.

La relation $C(x,y) = \lambda$ qui est satisfaite pour le point $x = 0$, $y = u_1$, conduit à écrire :

$$\frac{u_2!}{u_1!} \frac{1}{(u_2 - u_1)!} = \frac{1 - \beta}{\alpha}$$

d'où la condition :

$$\alpha < (1 - \beta) \frac{u_1!}{u_2!} (u_2 - u_1)!$$

Ainsi, rigoureusement, moyennant cette seule condition, le risque du producteur est nul puisque en cas d'épreuve d'un lot contenant au plus u_1 défectueux, le cheminement ne peut en aucun cas atteindre la ligne de rejet.

Une condition correspondant à la précédente peut être établie dans le cadre de l'épreuve séquentielle non-exhaustive, mais seulement à partir du moment où le tronquage a été fixé.

b) Si $\beta = 0$, une remarque analogue à celle du a ci-dessus peut être faite : c'est cette fois le risque du client qui est nul.

Il suffit d'ailleurs pour cela que β soit voisin de 0 ; plus précisément, on doit avoir

$$\beta < (1 - \alpha) \frac{(U - u_2)!}{(U - u_1)!} (u_2 - u_1)!$$

c) Exemple : $U = 10$; $u_1 = 1$; $u_2 = 2$;
 $\alpha = 0,05$; $\beta = 0,10$.

On vérifie que ces valeurs de α et de β sont l'une et l'autre assez petites pour que les risques du producteur et du client soient l'un et l'autre nuls.

On constaterait dans ce cas qu'il n'y a pas de points d'arrêt (acceptation ou rejet) à l'intérieur du rectangle.

8 - REGLE DE L'EPREUVE SEQUENTIELLE EXHAUSTIVE

a) Données : U est l'effectif du lot soumis à l'épreuve en vue de son acceptation ou de son rejet.

Par convention entre producteur et client, les risques suivants sont admis :

- risque au plus égal à α de rejeter un lot qui contiendrait au plus u_1 défectueux ;

- risque au plus égal à β d'accepter un lot qui contiendrait plus de u_2 défectueux.

b) Plan. Le plan de l'épreuve est déterminé à partir du moment où sont connues les lignes d'acceptation A et de rejet R .

Il est suffisant pour la connaissance de ces lignes de déterminer pour chacune des valeurs entières de y , de 0 à u_1 , deux valeurs entières $x_\alpha(y)$ et $x_\beta(y)$ telles que :

$$C(x_\beta, y) > \frac{\beta}{1 - \alpha} > C(1 + x_\beta, y) ; C(x_\alpha, y) > \frac{1 - \beta}{\alpha} > C(1 + x_\alpha, y)$$

avec

$$C(x, y) = \frac{u_2!}{u_1!} \frac{(U - u_2)!}{(U - u_1)!} \frac{(u_1 - y)!}{(u_2 - y)!} \frac{(U - u_1 - x)!}{(U - u_2 - x)!}$$

N.B. - La construction graphique du n° 4 ci-dessus permet, pour chacune des lignes A et R , d'obtenir des valeurs approchées de x_α ou de x_β , en prenant λ égal suivant le cas à $\beta/(1-\alpha)$ ou à $(1-\beta)/\alpha$.

c) Préparation. Il convient de traduire le résultat de la recherche des lignes A et R de l'une des manières suivantes :

(i) - on porte sur un graphique le domaine limité par les conditions

$$0 \leq x \leq U - u_2 + 1 \quad ; \quad 0 \leq y \leq u_1 + 1$$

puis à l'intérieur de ce domaine :

- les points de la ligne A définis par $1 + x_\beta(y)$;
- les points de la ligne R définis par $x_\alpha(y)$;
- la ligne R qui est la ligne brisée ayant ces derniers points pour sommets.

(ii) - on établit une table donnant pour $0 \leq y \leq u_1$ les valeurs de $1 + x_\beta(y)$, ligne A, et de $x_\alpha(y)$, ligne R.

d) Exécution et décision. En cours d'épreuve, après chaque essai, considérer le nombre x des bons et le nombre y des défectueux obtenus jusqu'alors, puis :

- si la ligne A est atteinte : accepter le lot ;
- si la ligne R est atteinte ou dépassée, à y donné, du côté des valeurs de x inférieures à $x_\alpha(y)$: rejeter le lot.

e) Tronquage éventuel. Si un tronquage est retenu pour $x + y = n$, il y a lieu de calculer quelques points de la ligne neutre d'équation

$$C(x,y) = 1$$

de manière à déterminer les coordonnées du point Z d'intersection de la ligne neutre et de la droite de tronquage $x + y = n$.

La préparation est complétée en conséquence.

Les lignes A et R comprennent alors l'une et l'autre une partie de la droite de tronquage jusqu'au point Z (fig. 4), ou un tracé équivalent.

9 - EXEMPLE NUMERIQUE

Le graphique de la figure 5 représente avec de mêmes échelles les résultats obtenus à partir des données suivantes, se correspondant.

1 - Epreuve exhaustive : $U = 50$; $u_1 = 8$; $u_2 = 16$ d'où $u_2 - u_1 = 8$;
 $\alpha = 0,05$; $\beta = 0,10$

La règle du n° 8 conduit tous calculs faits aux valeurs suivantes :

ordonnée y :	0	1	2	3	4	5	6	7	8	9
<u>Ligne d'acceptation</u>										
$1 + x_\beta(y)$:	10	13	15	18	20	22	25	27	30	35
<u>Ligne de rejet</u>										
$x_\alpha(y)$:	(-17)	(-13)	(-8)	(-3)	1	6	11	16	21	35

Sur la figure, ces points sont notés et sont reliés entre eux. Sont notés également :

- le point D de coordonnées : 35 et 9 ;
- le point M de coordonnées : 38,5 et 12,5 ;
- le point M_1 de coordonnées approximatives : 38,5 et 12,1 ;
- le point origine de la ligne d'acceptation A ($x \approx 9,44$; $y = 0$) et la tangente en ce point à la courbe $C(x,y) = \beta/(1-\alpha) = 0,1053$; cette tangente passe par M_1 .

2 - Epreuve non-exhaustive : $p_1 = 8/50 = 0,16$ et $p_2 = 0,32$ (pour U suffisamment grand) ; $\alpha = 0,05$; $\beta = 0,10$.

Les formules classiques conduisent (fig. 5) aux droites A' et R', de pente 0,304, ayant pour origines respectives :

- sur $y = 0$, le point d'abscisse 10,60 soit, pratiquement, 11 ;
- sur $x = 0$, le point d'ordonnée 4,16 soit, pratiquement, 5.

10 - PLANS APPROCHES

En fin du n° B 5 de [1], il a été fait allusion à un plan approché pouvant dans certains cas être préféré au plan rigoureux, en raison de sa simplicité ; en l'occurrence, ce plan approché était représenté graphiquement par deux droites issues de M et aboutissant à des points convenablement choisis sur les axes ox et oy ; les résultats notés au n° 5 ci-dessus conduisent d'ailleurs à remplacer pour le choix de ces derniers points β par $\beta/(1-\alpha)$ et $1/\alpha$ par $(1-\beta)/\alpha$: c'est là une modification très légère pratiquement.

Ces mêmes résultats conduisent à envisager comme pouvant constituer une nouvelle approximation du plan rigoureux (n° 8), le plan dont la représentation graphique serait, pour chacune des lignes A et R, celle obtenue d'après les indications du n° 4 ; seuls donc les points sur l'axe des x , les coordonnées de M_1 et, de préférence, aussi les points correspondant à $y = u_1$ seraient à calculer.

Si l'on part du même point de l'axe des x , le premier plan approché est défini par la droite qui joint ce point au point M, tandis que le nouveau plan approché est défini par une courbe dont la tangente en ce même point passe par le point M_1 : comme M_1 est très voisin de M, ces deux plans commencent par être eux-mêmes très voisins l'un de l'autre. Mais à partir du moment où y est notable, le nouveau plan approché suit manifestement la réalité de plus près que le fait le premier et paraît, par suite, devoir être préféré à celui-ci, à défaut du calcul exact.

Exemple. Avec les données de l'exemple du n° 9, le premier plan approché ci-dessus aurait conduit à retenir les lignes brisées ayant pour sommets les points suivants :

ordonnée y :	0	1	2	3	4	5	6	7	8	9
ligne d'acceptation A :	10	12	15	17	19	22	24	26	29	35
ligne de rejet R :	"	"	"	"	1	5	10	14	18	35

Ces points sont voisins de ceux, plus exacts, indiqués au n° 9, mais ils sont obtenus comme résultats de la construction graphique, particulièrement simple, consistant à tracer à partir de M (coordonnées 38,5 et 12,5) les deux droites aboutissant respectivement :

ligne A : au point de $y = 0$ d'abscisse égale à

$$\left[U - \frac{u_1 + u_2 - 1}{2} \right] \left[1 - \left(\frac{\beta}{1-\alpha} \right)^{1/u_2 - u_1} \right] = 9,44 \quad ;$$

ligne R : au point de $x = 0$ d'ordonnée égale à

$$\left[\frac{u_1 + u_2 + 1}{2} \right] \left[1 - \left(\frac{\alpha}{1-\beta} \right)^{1/u_2 - u_1} \right] = 3,80 .$$