

REVUE DE STATISTIQUE APPLIQUÉE

M. BASTIEN

Contribution à un problème d'excentricité

Revue de statistique appliquée, tome 6, n° 4 (1958), p. 41-58

http://www.numdam.org/item?id=RSA_1958__6_4_41_0

© Société française de statistique, 1958, 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/>

CONTRIBUTION A UN PROBLÈME D'EXCENTRICITÉ

par

M. BASTIEN
Statisticien à l'I. R. S. I. D.

RESUME

L'étude numérique dont nous présentons ici l'essentiel du point de vue pratique sous forme de graphiques a pour but d'aider à résoudre le problème suivant :

Des pièces fabriquées industriellement en série présentent chacune une excentricité provoquée par la superposition d'un phénomène systématique et d'un phénomène aléatoire dont l'effet dépend d'un paramètre.

En moyenne, cette excentricité n'est jamais nulle mais elle est d'autant plus grande que le phénomène systématique est plus important.

A l'aide d'un échantillon d'excentricités observées sur un lot de pièces, on désire :

1°) Tester l'absence du phénomène systématique contre sa présence.

2°) S'il y a lieu, estimer alors ce qui subsisterait après élimination du phénomène systématique.

Nous indiquons d'abord un mode d'emploi suivi d'exemples et présentons quelques compléments et quelques justifications mathématiques en annexe.

INTRODUCTION -

Dans de nombreuses situations industrielles, la fabrication d'une pièce de série comporte un stade que l'on peut très schématiquement représenter de la façon suivante : tracer un cercle concentrique intérieur à un cercle préexistant. Le processus de fabrication, à ce stade, ne permet pas d'obtenir une concentricité parfaite à l'échelle de la précision des mesures.

On peut cependant, assez souvent, supposer que l'excentricité constatée est due à deux genres de causes :

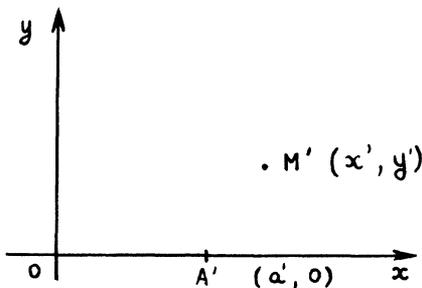
- les unes (vibrations, etc.) ayant une symétrie de révolution et des effets aléatoires dont la variabilité est fonction d'un paramètre;
- les autres (pointage systématiquement défectueux, etc.) ayant un effet constant dans une direction privilégiée de l'appareil-

lage, direction d'ailleurs inconnue, ou du moins non trans-
criptible sur la pièce.

On aboutit finalement au schéma suivant :

A chaque pièce d'une série fabriquée dans des conditions identiques est attachée la valeur de la distance entre les centres des deux cercles d'une section plane adéquate. Ces distances, ou excentricités, constituent donc une population théoriquement infinie. On ne peut mesurer que l'excentricité mais on sait que :

- le centre M' du cercle intérieur est aléatoire autour d'un point A' situé à la distance certaine a' du centre O du cercle extérieur;
- après le choix défini par les notations ci-après d'axes perpendiculaires, les coordonnées x' et y' de M' sont distribuées normalement et indépendamment avec la même variance et ont, pour espérance mathématique, respectivement a' et 0 .



C'est ce schéma qu'il faut trouver légitime pour utiliser les graphiques ci-joints.

On appellera $\frac{a'}{\delta}$, notée a : excentricité systématique réduite et a' , notée $a\delta$: excentricité systématique réelle.

On définit de même, à partir de l'excentricité réelle $OM' = e$, l'excentricité réduite $\frac{e}{\delta}$ que l'on notera r_a , ou, plus simplement, r lorsqu'aucune confusion ne sera possible.

L'excentricité mesurée sur une pièce quelconque de série est donc un individu d'une population caractérisée par les deux seuls paramètres a et δ .

Nous excluons évidemment le cas où ces paramètres seraient susceptibles de variations, c'est-à-dire, le cas d'une fabrication instable. (La stabilité de la fabrication peut, dans une certaine mesure, être vérifiée par la comparaison d'échantillons d'excentricités prélevés pendant différentes périodes assez courtes).

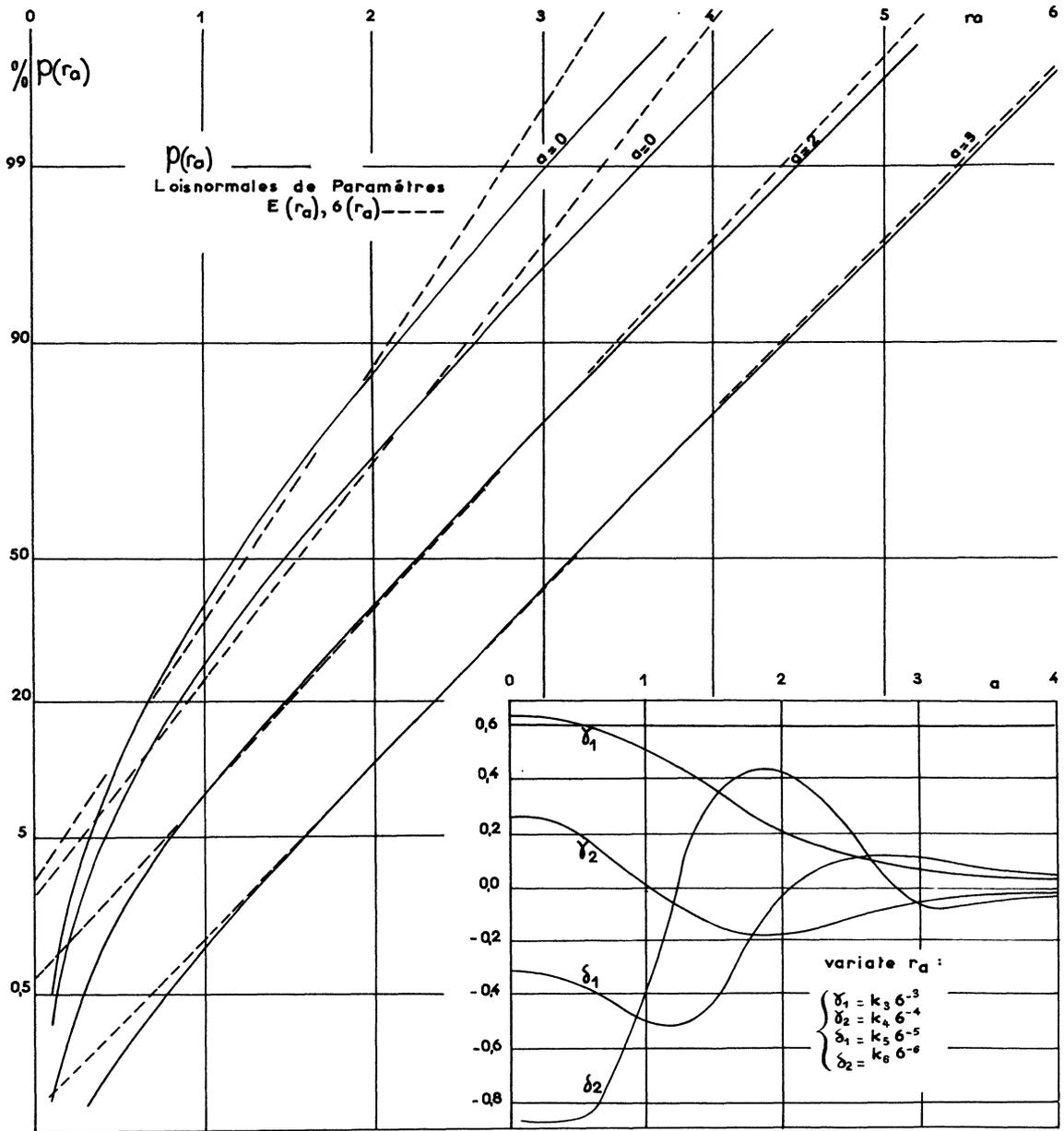
Nous renvoyons pour plus de détails sur la façon dont se pose le problème concret à l'article précédent de M. Rambach.

GENERALITES SUR LA DISTRIBUTION DE L'EXCENTRICITE -

Le paramètre δ n'intervient que comme un paramètre d'échelle et pour étudier l'allure de la distribution de e en fonction de a , il suffit d'étudier celle de l'excentricité réduite r .

On montre (Cf. Annexe et graphiques 1 et 2) que cette distribution se déforme progressivement :

de la distribution de χ à 2 degrés de liberté
 $(E(\chi) = 1,253; \sigma(\chi) = 0,655)$ lorsque $a = 0$



Graphique 1.

— Fonctions de répartition $P(r_a)$.

---- Lois normales de paramètres $E(r_a)$ et $\sigma(r_a)$.

Carton : premiers coefficients $K_q \sigma^{-q}$ de la variate r_a .

en une distribution très sensiblement normale
 $E(r) \sim a; \sigma(r) \sim 1$ lorsque a dépasse 3.

Ceci peut être d'un certain secours pour les cas où l'on soupçonne le schéma exposé ci-dessus de ne pas être valable. On devra notamment s'abstenir d'effectuer les calculs indiqués ci-dessous lorsque le graphique cumulatif des excentricités montrera une nette dissymétrie négative ou une nette bimodalité. On sera en droit, par contre, de supposer l'existence d'un autre facteur perturbateur que l'excentricité systématique et d'agir en conséquence sur la fabrication pour tenter de l'éliminer.

TEST DE L'HYPOTHESE D'UNE EXCENTRICITE SYSTEMATIQUE NULLE -

Lorsqu'une machine fabrique des pièces dont l'excentricité est jugée trop élevée, on peut se demander :

- si cela provient du phénomène aléatoire seul, dû à la machine en action, qui n'est en général pas susceptible d'amélioration,
- si à ce phénomène aléatoire ne s'ajoute pas l'effet d'une excentricité systématique notable qui pourrait alors être réduite par un montage plus soigné.

On est donc amené à tester sur un échantillon d'excentricités si l'excentricité systématique est nulle.

Nous proposons en annexe d'utiliser à cet effet le coefficient de variation observé sur un échantillon d'excentricités e_1, e_2, \dots, e_n de taille n , c'est-à-dire :

$$V = \frac{s}{\bar{e}} \quad \text{où} \quad \bar{e} = \frac{1}{n} \sum_1^n e_i \quad \text{et} \quad s^2 = \frac{1}{n-1} \sum_1^n (e_i - \bar{e})^2$$

V est une variable aléatoire puisqu'il est calculé sur échantillon et plus l'excentricité systématique réduite a est grande, plus la distribution de V est décalée vers zéro.

On admettra donc que $a = 0$ tant que V ne sera pas trop petit.

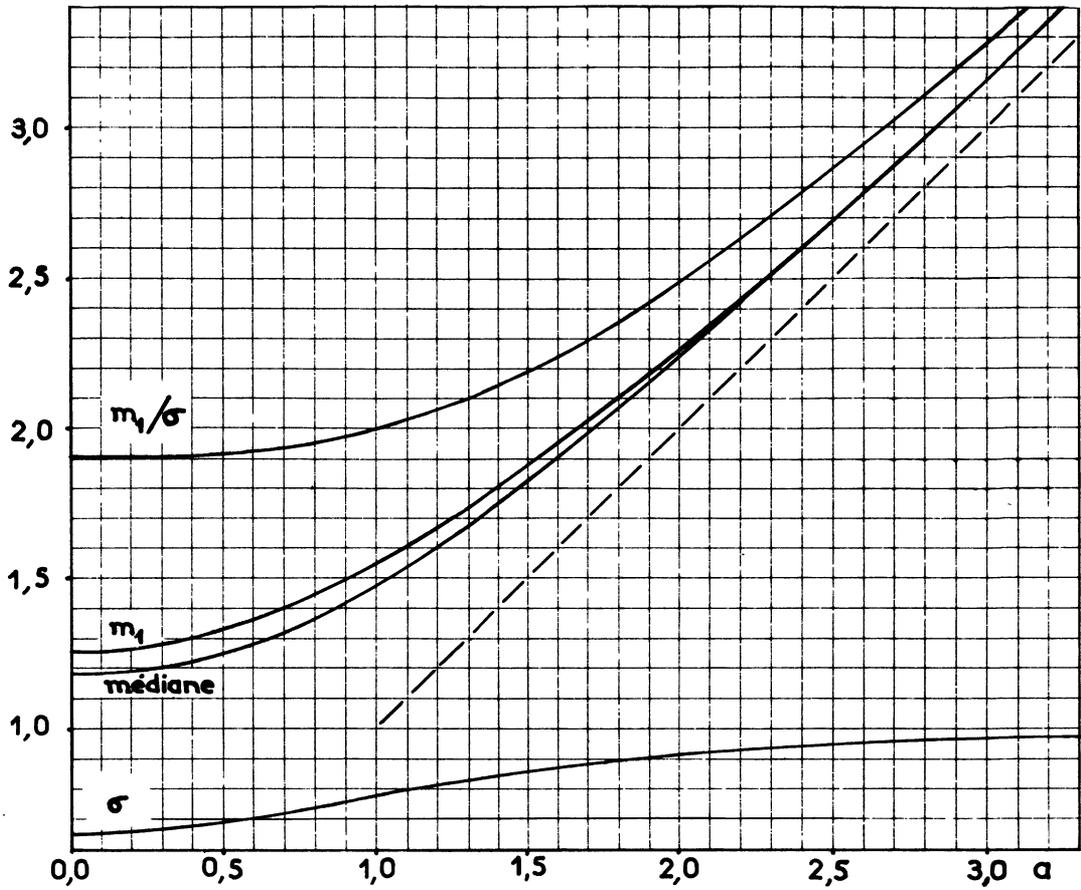
On rejettera, au contraire, l'hypothèse $a = 0$ si V est inférieur à l'une ou l'autre valeur du tableau ci-après, lu à la ligne correspondant à l'effectif n : avec la première valeur, on acceptera le risque 5% de rejeter à tort l'hypothèse $a = 0$ et avec la deuxième valeur, le risque 10%.

Ce tableau est issu des calculs évoqués en annexe. Pour $n > 25$ on peut le compléter par interpolation linéaire en $n^{-1/2}$ et pour $n > 100$, utiliser les deux expressions :

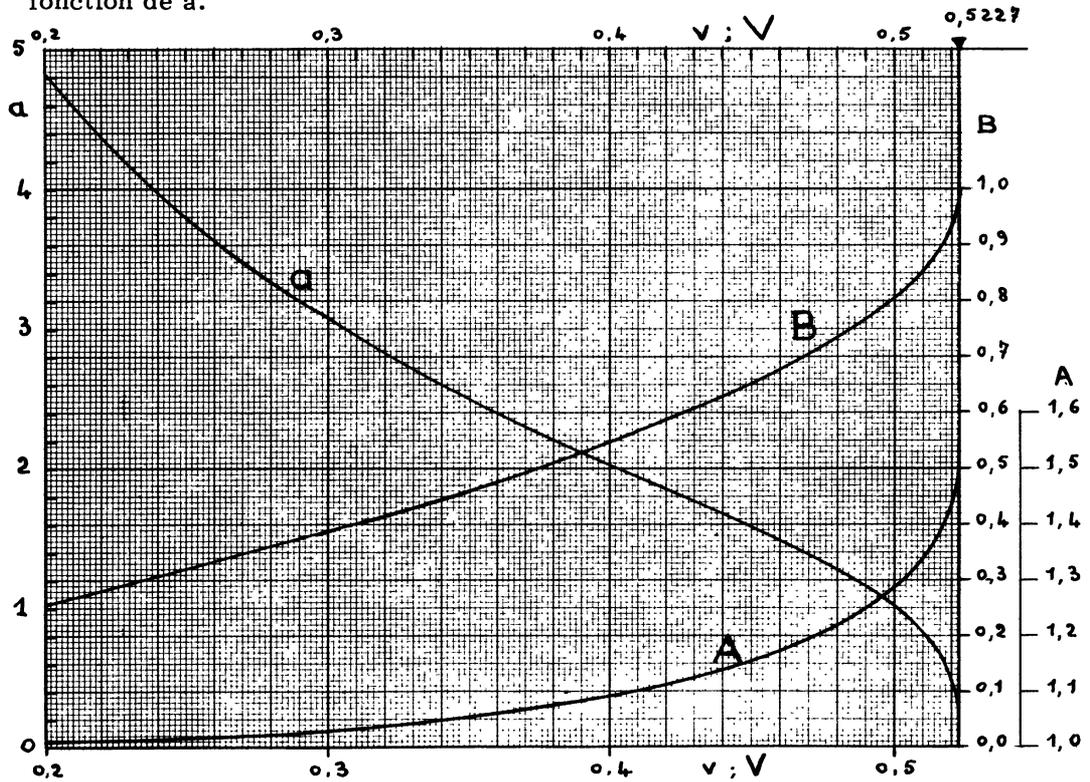
$$\begin{array}{ll} 0,52272 - 0,6108 n^{-1/2} & (\text{risque } \alpha = 5\%) \\ 0,52272 - 0,4760 n^{-1/2} & (\text{risque } \alpha = 10\%) \end{array}$$

L'efficacité (ou puissance) du test croît avec l'effectif n et avec la valeur a inconnue de la population d'excentricités échantillonnée.

Cette notion est développée en annexe, p.56.



Graphique 2 - Médiane, moyenne m_1 , écart-type σ , m_1/σ , de la variatè r_a , en fonction de a .



Graphique 3 - Courbes destinées à l'estimation de a , de δ et de l'excentricité moyenne résiduelle.

Tableau des valeurs de V
au-dessous desquelles on rejette l'hypothèse $a = 0$

| n | Risque $\alpha = 5\%$ | Risque $\alpha = 10\%$ | n | Risque $\alpha = 5\%$ | Risque $\alpha = 10\%$ |
|----|--------------------------|---------------------------|-----|--------------------------|---------------------------|
| 9 | 0,316 | 0,356 | 19 | 0,383 | 0,411 |
| 10 | 0,327 | 0,366 | 20 | 0,386 | 0,414 |
| 11 | 0,337 | 0,373 | 21 | 0,390 | 0,416 |
| 12 | 0,345 | 0,380 | 22 | 0,393 | 0,419 |
| 13 | 0,352 | 0,386 | 23 | 0,396 | 0,421 |
| 14 | 0,359 | 0,391 | 24 | 0,399 | 0,423 |
| 15 | 0,365 | 0,396 | 25 | 0,401 | 0,425 |
| 16 | 0,370 | 0,400 | 50 | 0,437 | 0,454 |
| 17 | 0,375 | 0,404 | 100 | 0,462 | 0,475 |
| 18 | 0,379 | 0,407 | | | |

EVALUATION DES EXCENTRICITES QUE L'ON PEUT ATTENDRE APRES SUPPRESSION DE L'EXCENTRICITE SYSTEMATIQUE -

Lorsque le test ci-dessus est significatif, il conduit à admettre que l'excentricité systématique n'est pas nulle.

Si l'on arrivait à éliminer le phénomène systématique, les excentricités seraient alors diminuées dans leur ensemble.

Avant de tenter cette élimination (par un montage plus soigné des pièces par exemple), on peut désirer chiffrer l'ordre de grandeur de l'amélioration que l'on pourrait ainsi apporter à la fabrication.

L'échantillon e_1, e_2, \dots, e_n qui a servi au test, le permet dans une certaine mesure.

La distribution des excentricités $e_{(a=0)}$ qui seraient dues au phénomène aléatoire seul ne dépend, en effet, que du paramètre δ ($\frac{e_{(a=0)}}{\delta} = r_{(a=0)}$ suit la loi de χ à deux degrés de liberté).

Une estimation δ^* de δ permet donc d'avoir une idée de la distribution de l'excentricité résiduelle $e_{(a=0)}$ à partir de la loi, connue, de $r_{(a=0)}$.

On ne peut estimer δ sans estimer a ; nous proposons en annexe la méthode des moments tant que V est inférieur à 0,52272.

Cette méthode conduit au graphique 3 qui permet, en prenant $V = \frac{s}{e}$ pour abscisse :

- de lire a^* , estimation de l'excentricité systématique réduite;
- d'obtenir $\delta^* = A s$, estimation du paramètre de variabilité du phénomène aléatoire seul;
- d'obtenir $1,253\delta^* = B \bar{e}$, estimation de l'excentricité moyenne résiduelle (qui subsisterait après suppression de l'excentricité systématique).

B est un coefficient qui varie de 0 à 1 lorsque a varie de l'infini à zéro et il chiffre l'amélioration due à la suppression de l'excentricité systématique a.

En appelant quantile P résiduel la valeur que les excentricités résiduelles ne dépasseront qu'en proportion 1 - P, on peut évaluer les quantiles résiduels à l'aide de $B\bar{e}$ et, par exemple :

le quantile 0,995 résiduel par 2,60 ($B\bar{e}$)
le quantile 0,95 résiduel par 1,95 ($B\bar{e}$)

La précision de ces estimations et de ces évaluations dépend évidemment de l'effectif n de l'échantillon prélevé.

Mais elle sera cependant meilleure pour les fortes valeurs de a. Au contraire, pour les faibles valeurs de a, c'est l'effet combiné de a et de δ qui est perceptible. Cet effet ne pourrait être décomposé qu'à l'aide d'échantillons de taille prohibitive.

Exemple :

1) On a relevé dans la zone d'attaque du foret d'une perceuse les excentricités de 50 billettes. Elles ne répondent pas aux tolérances exigées et fournissent :

$$\bar{e} = 1,037 \text{ mm}$$

$$\bar{w} = 0,695 \text{ mm}$$

(chaque étendue w calculée sur groupe de 5).

$$\frac{\bar{e}}{\bar{w}} = 1,49$$

est très grand par rapport à la valeur théorique 0,83 de M. Dick (cf. l'article de M. Rambach).

L'écart-type des excentricités est $s = 0,356$ mm et le coefficient de variation $V = 0,343$. On en déduit :

- au seuil de 5%, l'excentricité systématique est significative,
- le graphique 3, lu à la valeur $V = 0,343$, conduit aux estimations :

excentricité systématique réduite : $a^* = 2,6$
paramètre de variabilité du phénomène
aléatoire seul : $\delta^* = 0,37$
excentricité moyenne résiduelle escomptée : $0,47$ mm

On modifia la fabrication par un serrage approprié de la billette à percer mais l'on se contenta de mesurer sur 50 billettes la différence maximum des épaisseurs en trois points seulement d'une section plane dans la zone d'attaque du foret. Chaque mesure est donc une valeur par défaut de l'excentricité et il ne serait pas correct de comparer l'échantillon de ces 50 nouvelles mesures (elles sont toutes inférieures à 0,65 mm et leur moyenne est 0,28 mm) aux prévisions ci-dessus.

Néanmoins, cet échantillonnage fit présumer que la fabrication était suffisamment améliorée. Elle donne actuellement satisfaction.

II) Afin de mieux illustrer le principe du test et des prévisions, nous avons fabriqué deux échantillons de 50 excentricités répondant au schéma théorique :

- le premier échantillon fournit :

$$\begin{cases} \bar{e} = 237,42 \\ s = 92,88 \end{cases} \quad \text{d'où } V = 0,391$$

L'excentricité systématique est significative au risque 5% et les estimations sont :

$$\begin{cases} a^* = 2,1 \text{ pour l'excentricité systématique réduite,} \\ \delta^* = 1,08 \text{ s} = 100 \text{ pour le paramètre de variabilité du phénomène aléatoire seul,} \\ a^*\delta^* = 210 \text{ pour l'excentricité systématique réelle,} \\ 0,53 \bar{e} = 126 \text{ pour l'excentricité moyenne résiduelle,} \\ 2,60 \times 126 = 328 \text{ pour le quantile } 0,995 \text{ résiduel} \end{cases}$$

- le deuxième échantillon fournit :

$$\begin{cases} \bar{e} = 146,48 \\ s = 75,34 \end{cases} \quad V = 0,514 \text{ n'est pas significatif.}$$

Une seule des 50 excentricités dépasse la valeur 328.

L'estimation de δ est $\delta^* = 1,37$ $s = 103$. Il n'y a pas lieu de s'intéresser à la valeur $a^* = 0,74$ puisque le test n'est pas significatif. Une telle excentricité demanderait un échantillon de taille prohibitive pour être décelée.

Notons que si l'on était sûr qu'il n'y a pas excentricité systématique, il faudrait estimer δ par $\sqrt{\frac{\sum e^2}{2n}} = 116^{(1)}$.

(1) Le premier échantillon a été tiré d'une population où $a = 2$; $\delta = 100$ et le deuxième d'une population où $a = 0$; $\delta = 100$.

ANNEXE

La ligne directrice des calculs fut la suivante :

- Etude de la loi de $r_{(a)}$, et de ses premiers cumulants (due à M. Chouchand de la Compagnie Française des Métaux).
- Choix d'une statistique :
qui élimine le paramètre δ et puisse fournir une estimation de a ,
dont on puisse approcher la loi de distribution à l'aide des cumulants de $r_{(a)}$,
qui soit d'une utilisation pratique relativement simple.
- Etude de la loi approchée de cette statistique à l'aide de ses cumulants :
pour $a = 0$ afin de définir une région de rejet de l'hypothèse "absence de
phénomène systématique;
pour $a \neq 0$ mais, plus brièvement, afin d'étudier la puissance du test de
l'hypothèse $a = 0$ en fonction de a .

LOIS ET MOMENTS DE L'EXCENTRICITE -

La loi de $r_{(a)}$ est la loi de $\sqrt{x^2 + y^2}$ où x et y sont distribués normalement et indépendamment avec la variance unité et des espérances mathématiques valant respectivement a et 0 .

Nous rappelons très brièvement les résultats théoriques suivants :

La densité du couple (x, y) est $\frac{1}{2\pi} e^{-\frac{(x-a)^2}{2}} e^{-y^2/2}$ d'où l'on tire la densité de r :

$$\frac{dP(r)}{dr} = e^{-a^2/2} e^{-r^2/2} r \sum_{n=0}^{\infty} \frac{1}{(n!)^2} \left(\frac{ar}{2}\right)^{2n}$$

car

$$\frac{1}{\pi} \int_0^{\pi} e^{ar \cos \theta} d\theta = I_0(ar) = \sum_{n=0}^{\infty} \frac{1}{(n!)^2} \left(\frac{ar}{2}\right)^{2n} \quad (\text{fonction de Bessel})$$

La fonction de répartition de r peut s'écrire :

$$P(r) = e^{-a^2/2} \sum_{n=0}^{\infty} \left(\frac{a^2}{2}\right)^n \frac{1}{n!} \text{Prob}\left\{\chi^2_{\nu=2n+2} \leq r^2\right\}$$

Le moment d'ordre q est :

$$m_q = e^{-a^2/2} \sum_{n=0}^{\infty} \left(\frac{a^2}{2}\right)^n \frac{2^{q/2}}{(n!)^2} \Gamma\left(n + \frac{q}{2} + 1\right)$$

On peut utiliser cette expression de m_q pour montrer que c'est une solution de l'équation différentielle :

$$\alpha \frac{d^2 m_q}{d\alpha^2} + (2\alpha + 1) \frac{d m_q}{d\alpha} - q m_q = 0 \quad \text{où } \alpha = \frac{a^2}{4}$$

et que m_q peut alors s'écrire :

$$\left\{ \begin{array}{l} \text{q pair : } m_q = 2^{q/2} \left(\frac{q}{2}\right)! \left\{ 1 + \frac{q}{(1!)^2} \alpha + \frac{q(q-2)}{(2!)^2} \alpha^2 + \dots + \frac{q(q-2)\dots 6.4.2}{\left[\left(\frac{q}{2}\right)!\right]^2} \alpha^{q/2} \right\} \\ \text{q impair : } m_q = q(q-2)\dots 3.1 \sqrt{\frac{\pi}{2}} \left\{ 1 + \frac{q}{(1!)^2} \alpha + \dots + \frac{q(q-2)\dots 3.1}{\left[\left(\frac{q+1}{2}\right)!\right]^2} \alpha^{q/2} + q(q-2)\dots 3.1 S_q \right\} \\ \text{où } S_q = \sum_{n=\frac{q+1}{2}+1}^{\infty} (-1)^{n-\frac{q+1}{2}} \frac{1.3.5\dots(2n-q-2)}{(n!)^2} \alpha^n \end{array} \right.$$

La série alternée et convergente S_q permet de calculer des valeurs suffisamment approchées de m_q lorsque q est impair. On en déduit les valeurs numériques des cumulants par les formules habituelles.

On trouve que la loi de $r_{(a)}$ est déjà assez voisine de la normale pour $a > 2$ et que cette approximation où :

$$\left\{ \begin{array}{l} E(r_a) \sim a + \frac{1}{2a} \sim \text{Médiane de } r_a \\ \sigma(r_a) \sim 1 - \frac{1}{4a^2} \end{array} \right.$$

s'améliore très rapidement lorsque a croît (Cf. graphiques 1 et 2).

Rappelons que r_a est la variate d'une population où $\delta = 1$ et posons

$$\left\{ \begin{array}{l} E(r_a) = m_1 \\ \sigma(r_a) = \sigma \end{array} \right.$$

ESTIMATIONS DE L'EXCENTRICITE SYSTEMATIQUE REDUITE a ET DU PARAMETRE δ DE VARIABILITE DU PHENOMENE ALEATOIRE SEUL -

On dispose de l'échantillon e_1, e_2, \dots, e_n . Ces excentricités sont tirés indépendamment de la même population caractérisée par les paramètres a et δ qu'il s'agit d'estimer.

On peut résumer l'information contenue dans l'échantillon par les deux premiers moments empiriques dont l'étude théorique est plus facile que celle de statistiques d'ordre.

$$\text{Soient donc } \bar{e} = \frac{1}{n} \sum_1^n e_i \quad \text{et} \quad s^2 = \frac{1}{n-1} \sum_1^n (e_i - \bar{e})^2$$

Pour éliminer le paramètre δ , qui n'est du point de vue théorique qu'un paramètre d'échelle, il est nécessaire d'envisager une fonction de \bar{e}/s .

Les résultats généraux sur les moments d'une moyenne et d'une variance empiriques et ceux de l'étude de la loi de r_a indiquent que pour n assez grand, \bar{e} et s suivent approximativement une loi binormale (avec corrélation puisque la loi de e est dissymétrique) et que $E(\bar{e}) = E(e) = m_1 \delta \rightarrow a \delta$
 $E(s) \sim \sigma(e) = \sigma \delta \rightarrow \delta$ } lorsque a croît

$E \left\{ f \left(\frac{\bar{e}}{s} \right) \right\}$ ne sera donc pas trop différent, pour un effectif suffisant, de $f \left[\frac{E(r_a)}{\sigma(r_a)} \right] = f \left(\frac{m_1}{\sigma} \right)$ et notamment, pour $a > 2$ ou 3 , de $f(a)$.

Quelques essais artificiels et quelques considérations théoriques et pratiques conduisent finalement à donner la préférence au coefficient de variation empirique $V = \frac{s}{\bar{e}}$.

$$v_a = \frac{\sigma(r_a)}{E(r_a)} = \frac{\sigma}{m_1}$$

est une fonction de la seule excentricité systématique réduite a et ce serait le coefficient de variation d'une population d'excentricités où l'excentricité systématique réduite serait a , quel que soit le paramètre δ .

v ne peut être plus grand que $0,52272$ et prend cette valeur pour $a = 0$.

Nous proposons donc :

1°) Lorsque $V \leq 0,52272$, d'utiliser la méthode des moments :

Possédant les deux premiers moments empiriques \bar{e} et s^2 , on recherche les valeurs a^* et δ^* des paramètres telles que :

$$\begin{cases} E(e) = \delta \cdot E(r) = \bar{e} \\ \sigma(e) = \delta \cdot \sigma(r) = s \end{cases}$$

Ces deux égalités peuvent aussi s'écrire $\left\{ \begin{array}{l} v^* = V \\ \delta^* = \frac{s}{\sigma^*} = \frac{\bar{e}}{m_1^*} \end{array} \right.$

où v^*, σ^* et m_1^* sont respectivement le coefficient de variation, l'écart-type et la moyenne théoriques de l'excentricité réduite r lorsque $a = a^*$.

Ces deux égalités peuvent se résoudre facilement par un procédé graphique en construisant les courbes représentatives de a et de $\frac{1}{\sigma}$ en fonction de $v = \frac{\sigma}{m_1}$.

L'excentricité moyenne résiduelle est estimée par :

$$1,253 \delta^* = \left(\frac{1,253}{m_1^*} \right) \bar{e}$$

Le graphique 3 donne a ; $A = \frac{1}{\sigma}$; $B = \frac{1,253}{m_1}$ en fonction du coefficient de variation vrai v .

Il suffira de le lire en prenant $v = v^* = V$ pour obtenir a^* ; $\frac{1}{\sigma^*}$ et $\frac{1,253}{m_1^*}$ et trouver les estimations désirées.

2°) Lorsque $V > 0,52272$, a est estimé par zéro et les deux moments ne fourniront pas les mêmes valeurs $\frac{s}{0,655}$ et $\frac{\bar{e}}{1,253}$. On peut donc convenir d'estimer δ comme si a était effectivement nul par

$$\sqrt{\frac{\sum e^2}{2n}} \sim s \cdot \sqrt{\frac{1+V^2}{2V^2}}$$

ETUDE DE LA LOI DE V, COEFFICIENT DE VARIATION EMPIRIQUE -

Cette étude se révèle alors nécessaire pour tester que l'excentricité systématique réduite a la valeur "a".

δ n'intervient pas et nous pouvons le choisir égal à 1.

La loi exacte de V nous a paru trop difficile à étudier pratiquement et nous nous sommes contentés d'une approximation qui nous semble bonne.

On peut, en effet, approcher :

$$V^q = \left(\frac{S}{\bar{r}}\right)^q = \left(\frac{\sigma}{m_1}\right)^q \left(1 + \frac{s^2 - \sigma^2}{\sigma^2}\right)^{q/2} \left(1 + \frac{\bar{r} - m_1}{m_1}\right)^{-q}$$

par un développement limité. Son espérance mathématique donne une évaluation du moment non centré d'ordre q de V en fonction des moments centrés $\mu_{ij}(s^2, \bar{r})$ du couple (s^2, \bar{r}) . Notons que V a des moments car V est compris entre 0 et \sqrt{n} . (Cf. Cramer, p. 357).

Kendall (volume I, chapitre 11) donne les cumulants $K_{ij}(s^2, \bar{r})$ en fonction de l'effectif n et des cumulants de la variate initiale r.

On peut les écrire sous la forme :

$$K_{ij}(s^2, \bar{r}) = \left\{ K_{2i+j} + A_1(K_k^\alpha K_p^\beta \dots) \right\} n^{-(i+j-1)} + \left\{ A_2(K_k^\alpha K_p^\beta \dots) \right\} n^{-(i+j)} \\ + \left\{ A_3(K_k^\alpha K_p^\beta \dots) \right\} n^{-(i+j+1)} + \dots$$

où les K à un seul indice (second membre) sont les cumulants de r,

$k, p, \dots; \alpha, \beta, \dots$ sont des entiers tels que $k + \alpha + p + \beta + \dots = 2i + j$
 $A_1(K_k^\alpha K_p^\beta \dots)$ etc. sont des expressions contenant les $K_k^\alpha K_p^\beta \dots$, sauf K_{2i+j} .

En passant des $K_{ij}(s^2, \bar{r})$ aux $\mu_{ij}(s^2, \bar{r})$, on calcule une évaluation de $E(V^q)$ sous la forme d'un développement selon les puissances de n^{-1} .

Signalons que $\mu_{ij}(s^2, \bar{r}) = 0(n^{-p})$ où p est égal à $\frac{i+j}{2}$ ou $\frac{i+j+1}{2}$ selon que i + j est pair ou impair.

On obtient finalement des développements limités des cumulants $K_q(V) = 0(n^{q+1})$ et des coefficients $K_q(V) \sigma^{-q}(V) = 0(n^{-q/2+1})$.

On peut ensuite approcher la fonction de répartition de V par un développement de Gram-Charlier du type Edgeworth (Cf. Cramer, Chapitre 17) :

$$F(x) \sim \left\{ G(x) \right\} - \left\{ \frac{\gamma_1}{3!} g^{(2)} \right\} + \left\{ \frac{\gamma_2}{4!} g^{(3)} + \frac{1}{2} \left(\frac{\gamma_1}{3!} \right)^2 g^{(5)} \right\} - \left\{ \frac{\delta_1}{5!} g^{(4)} \right\} + \left(\frac{\gamma_1}{3!} \right) \left(\frac{\gamma_2}{4!} \right) g^{(6)} + \frac{1}{6} \left(\frac{\gamma_1}{3!} \right)^3 g^{(8)} + \dots$$

où : F(x) est la fonction de répartition de $\frac{V - E(V)}{\sigma(V)}$; γ_1, γ_2 et δ_1 les coefficients de dissymétrie, d'aplatissement et le coefficient $K_5 \sigma^{-5}$ de la loi de V;
 $G(t), g(t), g^{(j)}$ sont la fonction de répartition, la densité, la dérivée d'ordre j de la densité pour la variable normale réduite t.

Les accolades sont successivement en $n^0, n^{-1/2}, n^{-1}, n^{-3/2} \dots$

Tous ces résultats sont évidemment fonction de a .

Nous donnons en exemple sur le graphique 4 les fonctions de répartition de V sur papier normal lorsque $a = 0$ pour $n = 10$ et pour $n = 25$, calculées à l'aide des 4 accolades ci-dessus. Un échantillon de 49 valeurs de V pour lesquelles $a = 0$ et $n = 10$ y figure aussi.

L'approximation normale de la loi de V suffit pour beaucoup d'applications pratiques dès que $n = 10$ avec $E(V) \sim v_a$ et $\sigma(V) \sim \frac{v_a}{\sqrt{2n}}$.

On peut améliorer cette approximation en prenant le premier terme correctif de la série de Gram-Charlier, soit $F(x) \sim G(x) - \frac{Y_1}{3!} g^{(2)}$.

Les graphiques 3 et 5 donnent en fonction de a les paramètres nécessaires à cette approximation plus précise.

TEST DE L'HYPOTHESE D'UNE EXCENTRICITE SYSTEMATIQUE NULLE -

Le phénomène systématique est, par définition, absent si $a = 0$, ou encore si le coefficient de variation vrai est $v_o = 0,522723$.

Le test de l'hypothèse $a = 0$ est transposé en un test sur V coefficient de variation observé.

Ce test est unilatéral à gauche puisqu'on rejetera l'hypothèse si V est trop petit.

- Valeurs de rejet.

V est aléatoire d'un échantillon à l'autre et il s'agit de préciser le "trop petit".

Appelons $V_o(\alpha)$ la valeur au-dessous de laquelle il faut rejeter l'hypothèse $a = 0$, avec un risque de première espèce α , lorsque l'effectif est n .

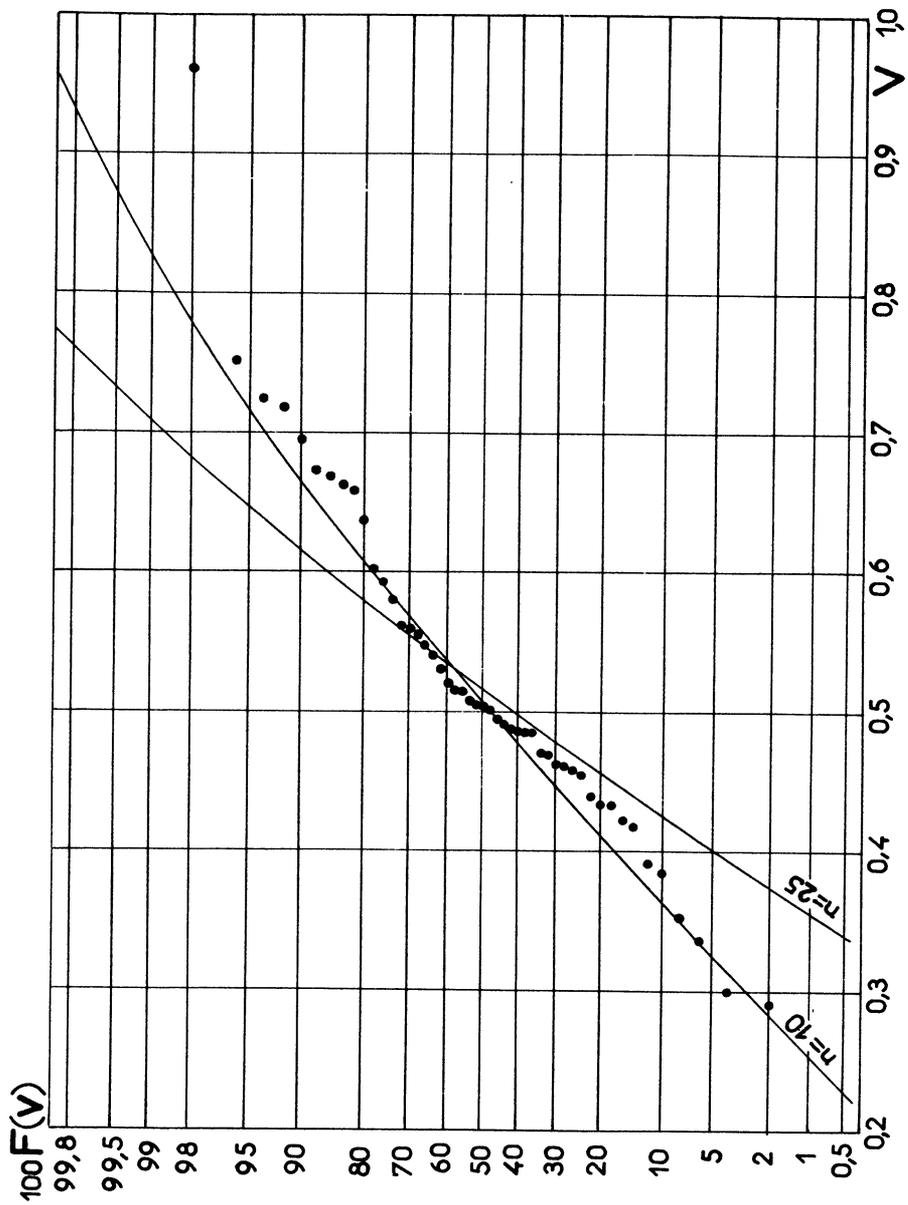
Les moments de V_o ont été calculés jusqu'aux termes en n^{-4} compris, ce qui permet d'utiliser le développement d'Edgeworth ci-dessus jusqu'aux termes en $n^{-3/2}$ de la quatrième accolade et d'obtenir le développement de $V_o(\alpha)$ selon les puissances de $\theta = n^{-1/2}$ jusqu'au terme en $\theta^4 = n^{-2}$ compris en écrivant que :

$$F\left\{\frac{V_o(\alpha) - E(V_o)}{\sigma(V_o)}\right\} = \alpha$$

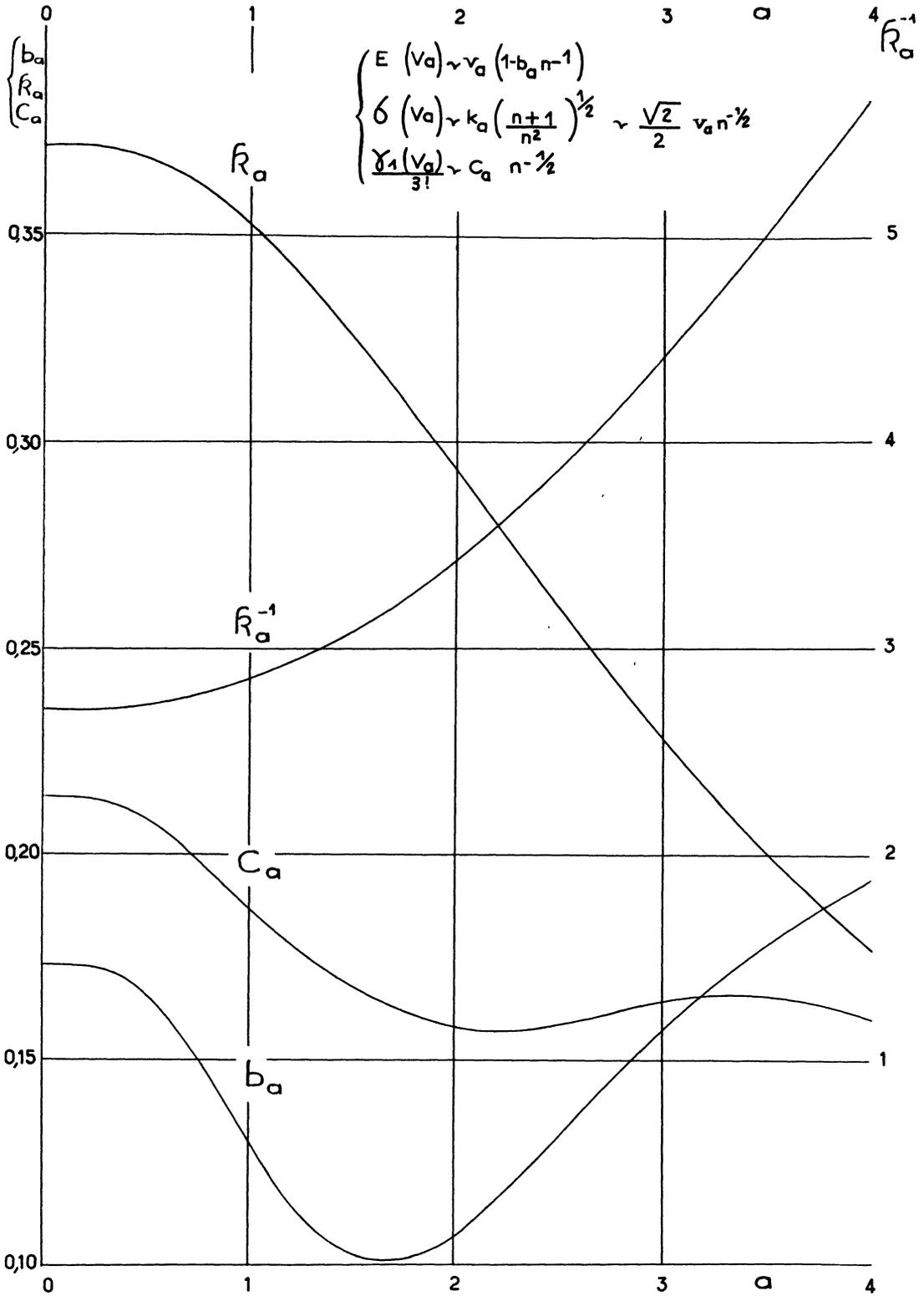
Nous avons choisi 2 valeurs pour le risque α : $G(-1,645) \sim 5\%$ et $G(-1,282) \sim 10\%$. Il vient, avec $\theta = n^{-1/2}$:

$$V_o(0,05)_n \sim 0,522.723 - 0,610.827.\theta + 0,044.896.\theta^2 - 0,080.125.\theta^3 - 0,443.478.\theta^4$$

$$V_o(0,10)_n \sim 0,522.723 - 0,476.036.\theta - 0,039.171.\theta^2 - 0,018.014.\theta^3 - 0,263.184.\theta^4$$



Graphique 4 - Fonctions de répartition approchées de V pour $a = 0$ ($n = 10$ et $n = 25$). Echantillon de 49 valeurs de V ($a = 0$; $n = 10$).



Graphique 5 - Pour la variate $V_a = \frac{s}{e}$. $E(V_a) \sim v_a (1 - b_a \times n^{-1})$. $\sigma(V_a) \sim k_a \times n^{-1/2}$.
 $\frac{Y_1(V_a)}{3!} \sim C_a \times n^{-1/2}$.

- Puissance du test⁽¹⁾.

En notant $V_a(\frac{P}{n})$ le quantile de la distribution de V , sous l'hypothèse a , correspondant à la probabilité P , la puissance du test est $1 - \beta$ définie par :

$$V_a(1 - \beta) = V_a(\alpha)$$

L'étude de la fonction de répartition approchée de V_a faite antérieurement permet alors d'évaluer $1 - \beta$.

Notons que :

1°) Nous avons approché la fonction de répartition de $V_{a=2}$ par le développement d'Edgeworth limité aux 3 premières accolades (il tient compte de la dissymétrie et de l'aplatissement de V_a mais fait intervenir les 8 premiers cumulants de r_a).

2°) Nous avons approché la fonction de répartition de V_a pour d'autres valeurs de a , jusqu'à $a = 4$ par le développement d'Edgeworth limité aux 2 premières accolades (il ne tient compte que de la dissymétrie de V_a mais ne nécessite le calcul que des 6 premiers cumulants de r_a).

L'ensemble des résultats obtenus conduit aux graphiques 6 et 7.

Par exemple, en utilisant le graphique 6 où $\alpha = 5\%$ (on accepte de rejeter 5 fois sur 100 en moyenne l'hypothèse $a = 0$ quand c'est elle qui est vraie), on voit :

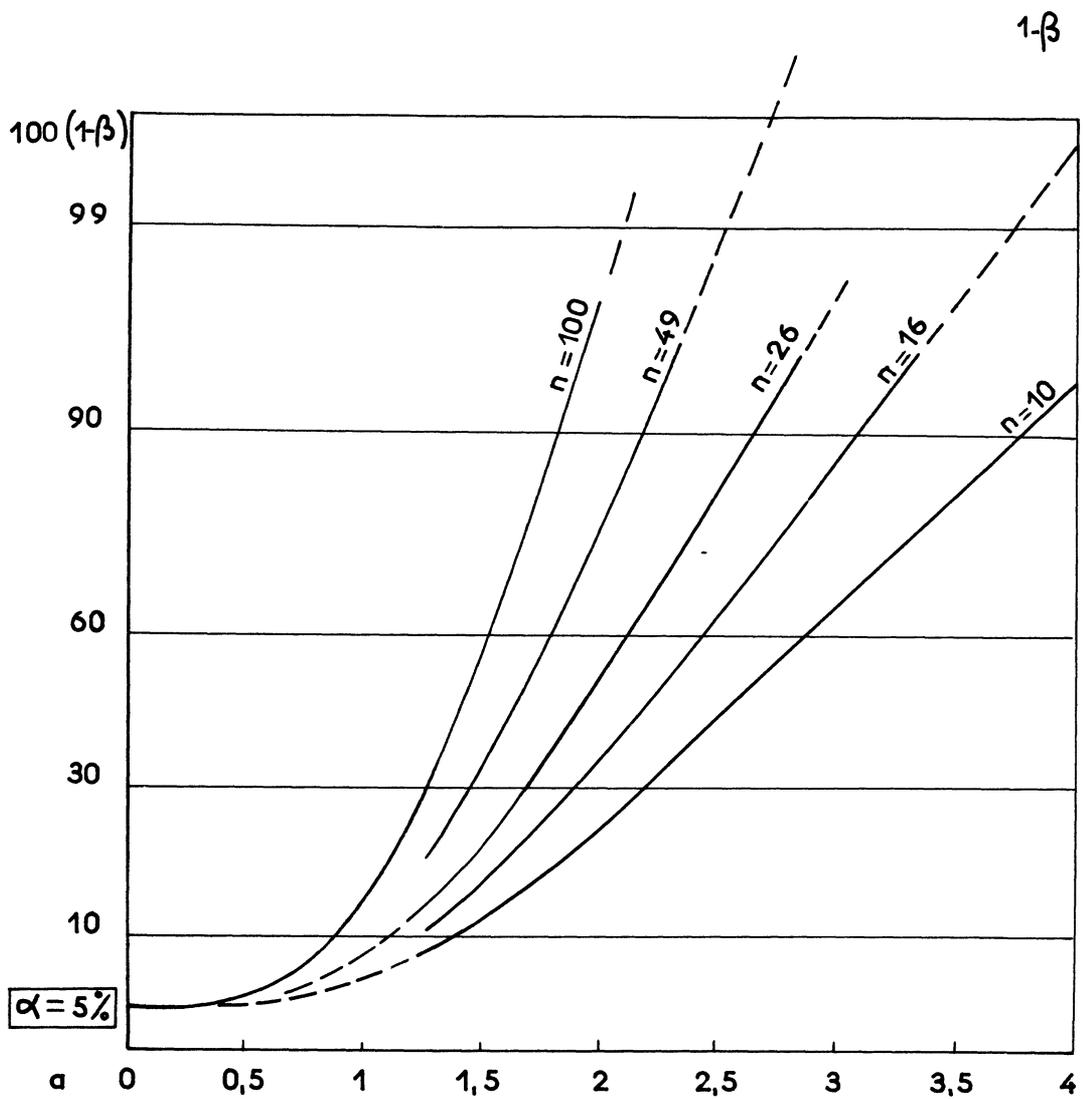
- qu'il faut un effectif de 100 pour déceler 56 fois sur 100 en moyenne une excentricité systématique non nulle quand on échantillonne dans une population où l'excentricité systématique réduite est, en fait, 1,5.
- qu'un effectif de 16 permettra de déceler 1 fois au moins sur 2 en moyenne une excentricité systématique non nulle, mais seulement si l'on échantillonne dans des populations où l'excentricité systématique réduite est, en fait, supérieure ou égale à 2,25.

Nous pensons qu'on peut prendre un risque α plus ou moins grand selon qu'on effectue le test en vue d'améliorer la fabrication ou en vue de vérifier que la fabrication mise au point précédemment ne se détériore pas.

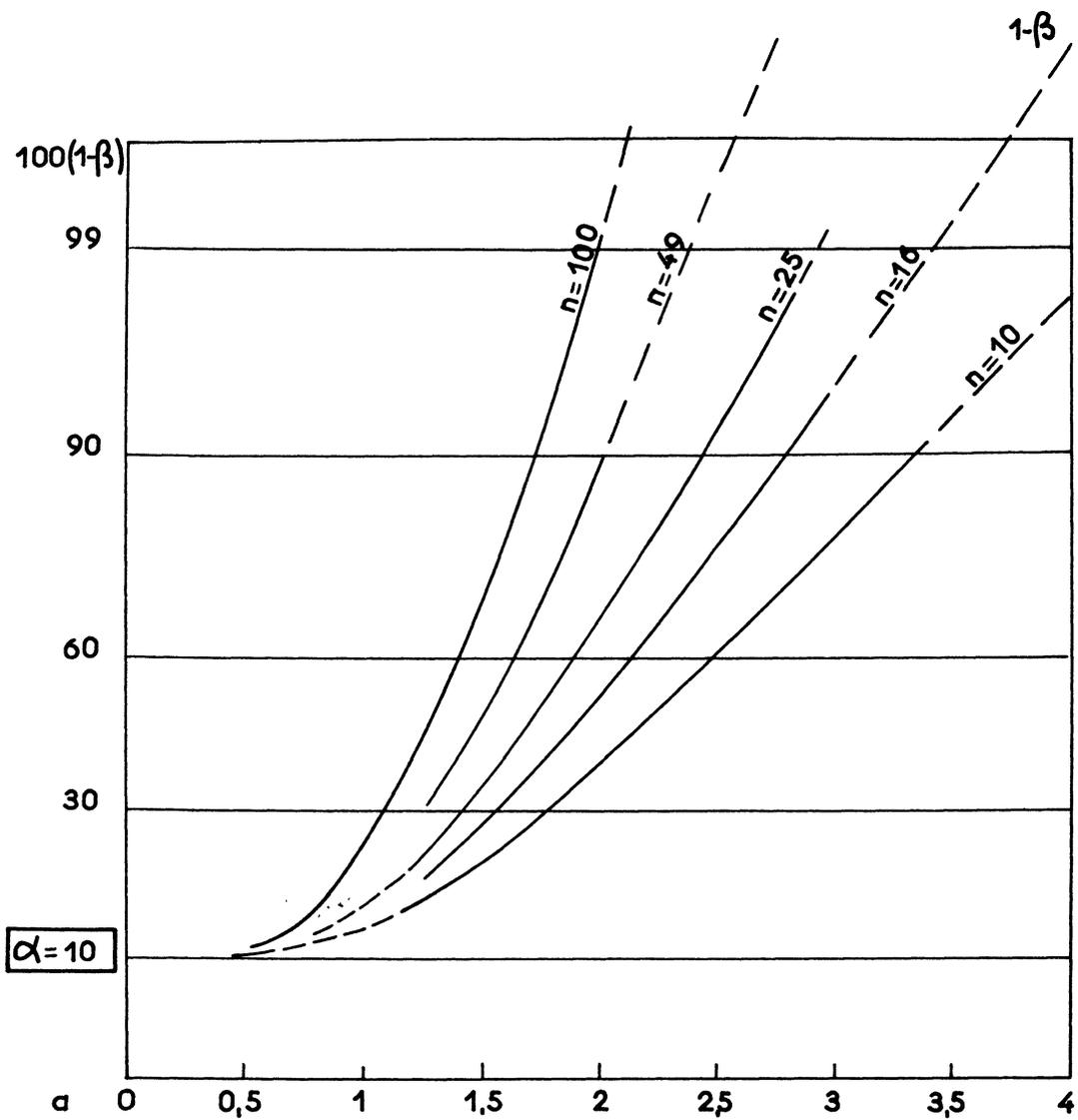
Remarque :

Les graphiques 3 et 5 permettent l'étude, moins précise, de l'hypothèse : "a quelconque non nul", du moins jusqu'à $a = 4$.

(1) Nous rappelons que la puissance $1 - \beta$ du test de l'hypothèse simple H_0 contre l'hypothèse composite $H_1(H_0 + H_1$ constituant l'ensemble des hypothèses admissibles et la région de rejet de H_0 étant préalablement définie) est la probabilité de rejeter H_0 lorsque c'est une des hypothèses simples composant H_1 qui est vraie. β , probabilité d'accepter H_0 dans les mêmes conditions, est le risque de deuxième espèce. Ici H_0 est $a = 0$, H_1 est $a > 0$; toute valeur positive de a définit une hypothèse simple de H_1 et la puissance est fonction de ce paramètre.



Graphique 6 - Puissance du test de l'hypothèse $a = 0$ en fonction de a , pour $\alpha = 5\%$. L'échelle des ordonnées est gaussienne.



Graphique 7 - Puissance du test de l'hypothèse $a = 0$ en fonction de a , pour $\alpha = 10\%$
L'échelle des ordonnées est gaussienne.