

# REVUE DE STATISTIQUE APPLIQUÉE

A. VESSEREAU

## **Une approximation utile pour la construction d'un plan de contrôle par variable lorsque l'écart-type est inconnu**

*Revue de statistique appliquée*, tome 33, n° 1 (1985), p. 5-11

[http://www.numdam.org/item?id=RSA\\_1985\\_\\_33\\_1\\_5\\_0](http://www.numdam.org/item?id=RSA_1985__33_1_5_0)

© Société française de statistique, 1985, 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/>

# UNE APPROXIMATION UTILE POUR LA CONSTRUCTION D'UN PLAN DE CONTROLE PAR VARIABLE LORSQUE L'ECART-TYPE EST INCONNU

A. VESSEREAU

La construction d'un plan de contrôle par variable, lorsque le caractère contrôlé obéit à une loi normale, est facile lorsque l'écart-type  $\sigma$  est connu ; lorsqu'il est inconnu on doit utiliser la loi de t décentrée, dont il n'existe que des tables incomplètes et de consultation malaisée. La méthode exposée au paragraphe 3 de cette note est d'application facile ; imaginée par Takao ENKAWA, elle a été publiée en japonais dans "Quality - Vol. 10, N° 4 - 1980".

Nous adressons nos remerciements au Professeur ENKAWA qui a bien voulu nous autoriser à l'exposer dans une langue plus accessible aux lecteurs de cette revue.

## 1. LES PARAMETRES DU PLAN DE CONTROLE LORSQUE $\sigma$ EST CONNU

Lorsqu'on fixe une seule limite de tolérance (inférieure  $T_i$  ou supérieure  $T_S$ ) une unité est défectueuse si la mesure  $x$  obtenue sur celle-ci est, suivant le cas, inférieure à  $T_i$  ou supérieure à  $T_S$ . Le contrôle s'effectue en prélevant au hasard un échantillon de  $n$  unités dans la population (ou lot) contrôlée ;  $\bar{x}$  désignant la moyenne des  $n$  mesures  $x_i$  obtenues et  $m$  la moyenne (inconnue) du lot, la quantité  $\sqrt{n} \frac{\bar{x} - m}{\sigma}$  suit la loi normale réduite.

La condition d'acceptation du lot s'écrit :

$$\bar{x} \geq T_i + K \sigma \quad \text{ou} \quad \bar{x} \leq T_S - K \sigma$$

$K$  est la "constante d'acceptation"

Lorsqu'on se donne des conditions :

$$\left. \begin{array}{l} p_A = \text{proportion d'unités défectueuses dans le lot} \\ \alpha = \text{risque de rejeter un lot de la qualité } p_A \end{array} \right\} \text{ Risque du fournisseur}$$

$$\left. \begin{array}{l} p_R > p_A = \text{proportion d'unités défectueuses dans le lot} \\ \beta = \text{risque d'accepter un lot de la qualité } p_R \end{array} \right\} \text{ Risque du client}$$

l'effectif  $n$  de l'échantillon à prélever et la constante d'acceptation  $K$  s'obtiennent par les formules classiques ci-après (on prend pour  $n$  la valeur entière immédiatement supérieure à celle donnée par la formule)

$$(1) \left\{ \begin{array}{l} n = \left[ \frac{u_{1-\alpha} + u_{1-\beta}}{u_{1-p_A} - u_{1-p_R}} \right]^2 \\ K = \frac{u_{1-\alpha} u_{1-p_R} + u_{1-\beta} u_{1-p_A}}{u_{1-\alpha} + u_{1-\beta}} \end{array} \right.$$

( $u_{1-\alpha} \dots u_{1-p_R}$  sont les fractiles d'ordres  $(1 - \alpha) \dots, (1 - p_R)$  de la variable normale réduite).

Dans le cas particulier – très souvent adopté – où  $\alpha = 0,05$  et  $\beta = 0,10$  les formules précédentes s'écrivent, en posant  $p_A = p_{95}$ ,  $p_R = p_{10}$

$$(2) \begin{cases} n = \frac{8,5644}{(u_{1-p_{95}} - u_{1-p_{10}})^2} \\ K = 0,4379 u_{1-p_{95}} + 0,5621 u_{1-p_{10}} \end{cases}$$

Les formules générales (1) déterminent entièrement la courbe d'efficacité du plan –  $P_a(p)$  désignant la probabilité d'acceptation pour la qualité  $p$ , on a :

$$(3) P_a(p) = F[\sqrt{n}(u_{1-p} - K)] \quad (F = \text{fonction de répartition de la variable normale réduite})$$

Le rapprochement des formules (1) et (2) montre que si l'on se donne des conditions quelconques ( $\alpha, p_A$ ), ( $\beta, p_R$ ) on peut toujours les remplacer par des conditions équivalentes ( $\alpha = 0,05$  ;  $p_{95}$ ), ( $\beta = 0,10$  ;  $p_{10}$ ) au moyen des relations :

$$(4) \begin{cases} u_{1-p_{95}} = \frac{1}{u_{1-\alpha} + u_{1-\beta}} [u_{1-\alpha} u_{1-p_R} + u_{1-\beta} u_{1-p_A} + 1,6449 (u_{1-p_A} - u_{1-p_R})] \text{ d'où } p_{95} \\ u_{1-p_{10}} = \frac{1}{u_{1-\alpha} + u_{1-\beta}} [u_{1-\alpha} u_{1-p_R} + u_{1-\beta} u_{1-p_A} - 1,2816 (u_{1-p_A} - u_{1-p_R})] \text{ d'où } p_{10} \end{cases}$$

### Remarques

Lorsqu'on fixe pour limites de tolérance,  $T_i$  et  $T_S$ , deux cas sont à considérer :

Limites dites "séparées" – Les conditions ( $\alpha, p_A$ ), ( $\beta, p_R$ ) peuvent être différentes suivant l'importance que l'on attache au dépassement de l'une ou de l'autre limite. Ce cas n'est pas distinct de celui d'une limite unique : le lot n'est accepté que s'il l'est à la fois pour  $T_i$  et pour  $T_S$ .

Limites dites "combinées" – Une unité est défectueuse lorsque  $x$  est extérieur à l'intervalle ( $T_i, T_S$ ) – Ce cas, plus complexe, ne sera pas traité dans ce qui suit. Toutefois, les limites peuvent être considérées comme "séparées" si l'écart-type  $\sigma$  est suffisamment petit par rapport à  $T_S - T_i$  (en tout état de cause lorsque  $T_S - T_i > 6\sigma$ )

## 2. LES PARAMETRES DU PLAN DE CONTROLE LORSQUE $\sigma$ EST INCONNU – UNE APPROXIMATION CLASSIQUE

Le cas " $\sigma$  inconnu" est le plus fréquent dans le contrôle d'un lot unique. Lorsque le contrôle porte sur une série de lots d'un même fournisseur, si l'on a constaté que les estimations de  $\sigma$  sont stables d'un lot à l'autre, on peut, à

partir de celles-ci, calculer une estimation précise de l'écart-type permettant d'admettre que, lors du contrôle des lots suivants,  $\sigma$  est connu et appliquer les formules (1) à (4).

Dans le cas général,  $\sigma$  inconnu est remplacé par son estimation :

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n - 1}}$$

On définit la condition d'acceptation du lot par :

$$\bar{x} \geq T_i + K' s \quad \text{ou} \quad \bar{x} \leq T_S - K' s$$

La quantité  $\sqrt{n} \frac{\bar{x} - m}{s}$  ne suivant pas la loi normale réduite, mais une loi de t décentrée, les expressions de  $n$  et  $K$  données par les formules (1) et (2) ne sont pas valables. On désignera dans ce qui suit par  $n'$  et  $K'$  l'effectif d'échantillon et la constante d'acceptation lorsque  $\sigma$  est estimé par  $s$ .

La méthode classique exposée ci-après consiste à calculer les valeurs ( $n$ ,  $K$ ) au moyen des formules (2) ( $\sigma$  connu) applicables au cas particulier ( $\alpha = 0,05$  ;  $p_{95}$ ), ( $\beta = 0,10$  ;  $p_{10}$ ). Si l'on a choisi deux autres couples ( $\alpha$ ,  $p_A$ ), ( $\beta$ ,  $p_R$ ), on se ramène à ce cas particulier au moyen des formules (4).

Si la première des formules (2) donne  $n \geq 20$ , on a, avec une très bonne approximation :

$$(5) \quad \begin{cases} K' = K \\ n' = \left(1 + \frac{K^2}{2}\right)n \end{cases}$$

(La constante d'acceptation est inchangée, l'effectif d'échantillon est multiplié par  $\left(1 + \frac{K^2}{2}\right)$ )

$$\text{La condition } n = \frac{8,5644}{(u_{1-p_{95}} - u_{1-p_{10}})^2} \geq 20 \text{ peut s'écrire}$$

$$u_{1-p_{95}} - u_{1-p_{10}} \leq 0,654$$

Ce qui exige que la différence  $p_{10} - p_{95}$  (ou le "rapport de discrimination"  $DS = \frac{p_{10}}{p_{95}}$ ) soit suffisamment petit. Le tableau ci-après donne la valeur maximale de  $p_{10}$  pour différentes valeurs de  $p_{95}$ .

$p_{95\%}$	0,05	0,10	0,20	0,30	0,40	0,50	0,60	0,70	0,80	0,90	1,0
max $p_{10\%}$	0,42	0,74	1,3	1,8	2,3	2,7	3,2	3,6	4,0	4,3	4,7
DS	8,4	7,4	6,5	6,0	5,8	5,4	5,3	5,1	5,0	4,8	4,7

$p_{95\%}$	2,0	3,0	4,0	5,0	6,0	7,0	8,0	9,0	10,0	15
max $p_{10\%}$	8,1	11,0	13,6	16,1	18,4	20,5	22,6	24,6	26,5	35,1
DS	4,1	3,7	3,4	3,2	3,1	2,9	2,8	2,7	2,6	2,3

### Exemple 1

$$p_A = p_{95} = 0,30\% \quad p_R = p_{10} = 1,50\%$$

$$u_{1-p_{95}} = 2,7478 \quad u_{1-p_{10}} = 2,1701$$

Les formules (2) ( $\sigma$  connu) donnent :

$$n = 25,6 \text{ (soit 26)} \quad K = 2,42$$

Ce plan est voisin du plan figurant dans la norme internationale ISO 3951 lettre code M, NQA = 0,25, où  $p_{95} = 0,29\%$ ,  $p_{10} = 1,50\%$  (plan défini par  $n = 25$  ;  $K = 2,43$ )

Les formules (5) sont applicables et donnent, pour  $\sigma$  inconnu :

$$n' = 100 \quad K' = 2,42$$

La norme internationale ISO 3951 donne pour ce plan :

$$n' = 100 \quad K' = 2,43$$

### Exemple 2

$$p_A = p_{95} = 0,20\% \quad p_R = p_{10} = 4\%$$

$$u_{1-p_{95}} = 2,8782 \quad u_{1-p_{10}} = 1,7507$$

Les formules (2) ( $\sigma$  connu) donnent :

$$n = 6,74 \text{ soit } 7 \quad K = 2,24$$

Ce plan correspond à celui de la norme ISO 3951 défini par la lettre-code I et NQA = 0,25 :  $n = 7, K = 2,25$  (dans ce plan  $p_{95} = 0,20\%$ ,  $p_{10} = 3,99\%$ )  $n$  étant très inférieur à 20, les formules (5) –  $\sigma$  inconnu, ne sont pas applicables. (On peut toutefois constater qu'elles donneraient des valeurs voisines des valeurs plus exactes obtenues par la méthode exposée ci-après – Voir Exemple 3).

## 3. LES PARAMETRES DU PLAN DE CONTROLE LORSQUE $\sigma$ EST INCONNU – METHODE DE T. ENKAWA

Cette méthode s'applique quelle que soit la valeur de  $n$ , même très petite.

Comme précédemment on part des valeurs  $n$  (non arrondie) et  $K$  correspondant au plan  $\sigma$  inconnu défini par ( $\alpha = 0,05$  ;  $p_{95}$ ) ( $\beta = 0,10$ ,  $p_{10}$ ) – formules

(2). Cette première approximation est notée  $n^{(1)} = n$ , on calcule ensuite une deuxième approximation par

$$n^{(2)} = \left[ 1 + \frac{3n^{(1)} K^2}{6n^{(1)} - 8} \right] n$$

Une nouvelle itération donne :

$$n^{(3)} = \left[ 1 + \frac{3n^{(2)} K^2}{6n^{(2)} - 8} \right] n$$

Si la valeur arrondie à l'entier immédiatement supérieur de  $n^{(3)}$  est la même que pour  $n^{(2)}$ , on prend celle-ci pour valeur de  $n'$ . Sinon, on procède à une nouvelle itération à partir de  $n^{(3)}$  etc.

En général, la deuxième itération ( $n^{(3)}$ ) est suffisante, la suivante servant de vérification.

A partir de la valeur non arrondie de  $n'$  on calcule ensuite la constante d'acceptation  $K'$  par la formule :

$$K' = \sqrt{\frac{3n' - 3}{3n' - 4}} K$$

Si l'on veut tracer la courbe d'efficacité – valeur de la probabilité d'acceptation  $P_a$  pour toute valeur de  $p$  – on applique la formule suivante :

$$P_a = F \left[ \frac{u_{1-p} - \sqrt{\frac{3n' - 4}{3n' - 3}} K'}{\sqrt{\frac{1}{n'} + \frac{K'^2}{3(n' - 1)}}} \right] = F \left[ \frac{u_{1-p} - K}{\sqrt{\frac{1}{n'} + \frac{K'^2}{3(n' - 1)}}} \right]$$

Le "point d'indifférence" de la courbe correspondant à  $[p_{50}, P_a = 0,50]$  est défini par :

$$p_{50} = 1 - F \left[ \sqrt{\frac{3n' - 4}{3n' - 3}} K' \right] = 1 - F(K)$$

### Exemple 3

(Données de l'exemple 2 au paragraphe 2)

$$p_{95} = 0,20 \% \quad p_{10} = 4 \%$$

On a trouvé  $n^{(1)} = n = 6,74$ ,  $K = 2,24$

Les itérations successives donnent :

$$n^{(2)} = 27,8$$

$$n^{(3)} = 24,5$$

$$n^{(4)} = 24,6$$

On prendra donc  $n' = 25$ , puis

$$K' = \sqrt{\frac{3 \times 24,6 - 3}{3 \times 24,6 - 4}} \times 2,24 = 2,26$$

Les valeurs [ $n' = 25$ ,  $K' = 2,26$ ] sont celles que l'on trouve effectivement dans la norme ISO 3951 (lettre-code I, NQA = 0,25)

Le point d'indifférence  $P_a = 0,50$  correspond à :

$$p = p_{50} = 1 - F(2,24) = 1,25 \%$$

(ISO 3951 donne  $p_{50} = 1,28 \%$ )

#### Exemple 4

La norme ISO 3951 donne, pour la lettre-code D et NQA = 1, les plans "pratiquement équivalents" (c'est à dire ayant des courbes d'efficacité voisines), les paramètres suivants :

$$\begin{cases} \sigma \text{ connu} & n = 2 & K = 1,42 \\ \sigma \text{ estimé} & n' = 5 & K' = 1,53 \end{cases}$$

Lorsqu'on consulte les tables donnant différents points de la courbe d'efficacité considérée commune à ces deux plans on trouve :

$$\begin{cases} \alpha = 0,05 & p_{95} = 0,0046 & (0,46 \%) \\ \beta = 0,10 & p_{10} = 0,284 & (28,4 \%) \end{cases}$$

Il est vraisemblable, bien que cela n'est pas précisé dans la norme, que les valeurs de  $p_{95}$  et  $p_{10}$  ont été calculées pour le cas " $\sigma$  inconnu, estimé par s", c'est-à-dire  $n' = 5$   $K' = 1,53$ .

En effet, lorsque  $\sigma$  est connu, les relations (2) donnent, pour  $n = 2$ ,  $K = 1,42$  les valeurs suivantes :

$$\begin{aligned} p_{95} &= 0,049 & (0,49 \%) & \text{très voisin de } 0,046 \\ p_{10} &= 0,304 & (30,4 \%) & \text{assez nettement différent de } 0,284 \end{aligned}$$

Si l'on part du plan  $n = 2$ ,  $K = 1,42$  ( $\sigma$  connu) la méthode de T. ENKAWA donne les valeurs successives :

$$\begin{aligned} n^{(1)} &= n = 2 \\ n^{(2)} &= 8,049 \\ n^{(3)} &= 4,416 \\ n^{(4)} &= 4,888 \\ n^{(5)} &= 4,772 \\ n^{(6)} &= 4,798 \\ n^{(7)} &= 4,792 \\ n^{(8)} &= 4,793 \end{aligned}$$

On voit qu'avec la plus petite possible de la "valeur de départ"  $n = 2$ , la valeur de  $n'$  est pratiquement stabilisée à la quatrième itération  $n^{(5)}$  et qu'elle l'est définitivement à la sixième ( $n^{(7)}$ )

Prenant  $n' = 4,79$  (donc pratiquement  $n' = 5$ ) on trouve ensuite :

$$K' = \sqrt{\frac{3n' - 3}{3n' - 4}} \times 1,42 = 1,49$$

Si donc l'on part du plan ISO 3951 "σ connu"

$$n = 2 \quad K = 1,42$$

et que l'on cherche le plan équivalent (mêmes valeurs de  $p_{95}$  et  $p_{10}$ ) lorsque  $\sigma$  est estimé par la méthode de T. ENKAWA donne les paramètres

$$n' = 5 \quad K' = 1,49$$

alors que ISO 3951 donne :

$$n' = 5 \quad K' = 1,53$$

La démarche adoptée dans ISO 3952 est inverse. On part du plan  $n' = 5$   $K' = 1,53$  pour aboutir au plan  $n = 2$   $K = 1,42$ .

Dans les deux cas, le point d'indifférence est le même :

$p_{50} = 1 - F(1,42) = 7,78\%$  (la norme ISO 3951 donne  $p_{50} = 7,82\%$ )  
Ce qui précède ne constitue nullement une critique de ISO 3951 ! Il est bien évident que la coïncidence des courbes d'efficacité "σ connu" et "σ estimé" ne peut pas être parfaite, en raison du caractère discret (valeur entière) de l'effectif d'échantillon ( $n$  ou  $n'$ ). Mais on trouve plutôt dans cet exemple une confirmation de la méthode de T. ENKAWA, qui a l'avantage d'être applicable à des conditions quelconques ( $\alpha, p_A$ ) ( $\beta, p_R$ ).

Observons finalement que la méthode donnée au paragraphe 2 de cette note donnerait, à partir de

$$n = 2 \quad K = 1,42$$

l'approximation grossière :

$$\begin{cases} n' = 2 \left[ 1 + \frac{(1,42)^2}{2} \right] = 4,016 \text{ (4 ou 5 ?)} \\ K' = 1,42 \end{cases}$$

au lieu de  $n' = 5$   $K' = 1,49$