Revue de statistique appliquée

Jürgen Sturhahn

Plans d'échantillonnage double par mesures avec une seule limite de tolérance et une caractéristique de qualité normalement distribuée

Revue de statistique appliquée, tome 26, n° 1 (1978), p. 45-57

<http://www.numdam.org/item?id=RSA_1978__26_1_45_0>

PLANS D'ÉCHANTILLONNAGE DOUBLE PAR MESURES AVEC UNE SEULE LIMITE DE TOLÉRANCE ET UNE CARACTÉRISTIQUE DE QUALITÉ NORMALEMENT DISTRIBUÉE

Jürgen STURHAHN

I – INTRODUCTION

Cet article propose une méthode permettant d'obtenir des plans d'échantillonnage double optimaux dans les contrôles par mesures. Dans ces plans, la probabilité d'acceptation doit avoir la valeur $1 - \alpha$ pour une proportion de défectueux p_1, correspondant à la valeur AQL (acceptable quality level), et la valeur β pour une proportion de défectueux p_2 ($p_2 > p_1$), correspondant à la valeur LQ (limiting quality). Ces plans doubles seront comparés aux plans simples qui leur sont équivalents. Par "équivalents" on entend des plans dont les courbes d'efficacité coïncident aux deux points (p_1, $1 - \alpha$) et (p_2, β). Les livraisons, dont la qualité est à contrôler, doivent se présenter sous forme de lots. Chaque lot se compose de n_L éléments, présentant une caractéristique de qualité X, normalement distribuée avec une valeur moyenne μ et une variance σ^2.

Dans ce qui suit on ne considérera que le cas d'une limite de tolérance supérieure T_0 unilatérale donnée (en ce qui concerne le cas d'une limite de tolérance inférieure donnée, ou de deux limites de tolérances données, se référer à [6]).

On dira qu'un élément est "bon", si $x \leqslant T_0$, d'où il résulte que la proportion de défectueux p du lot s'exprime par :

$$p = 1 - \Phi_s\left(\frac{T_0 - \mu}{\sigma}\right), \tag{1.1}$$

où Φ_s désigne la fonction de distribution de la variable normale réduite. La qualité d'un lot peut être caractérisée par le quantile $u_{1-p} = (T_0 - \mu)/\sigma$ correspondant à p. On convient dans ce qui suit que u_{1-p} est l'indice de qualité" du lot.

Mot-clé : **Echantillonnage double par mesures.**

II – PLANS D'ECHANTILLONNAGE DOUBLE ($\overline{x} . \sigma$)

II.1 – Instructions pour le contrôle, et courbe d'efficacité.

Dans un plan d'échantillonnage double on procède d'abord à un premier échantillonnage. Les résultats obtenus sur l'échantillon déterminent si une décision concernant l'acceptation ou le rejet du lot peut déjà être prise à ce stade, ou bien si une décision définitive n'est possible qu'après examen d'un deuxième échantillon (indépendant du premier). Les instructions pour le contrôle sont les suivantes :

1) A partir des n_1 résultats de mesures x_ν ($\nu = 1, \ldots, n_1$) sur un échantillon de n_1 éléments, on calcule la moyenne arithmétique

$$\overline{x}_1 = \frac{1}{n_1} \sum_{\nu=1}^{n_1} x_\nu \tag{2.1}$$

2) On compare la grandeur de contrôle

$$z_1 = \overline{x}_1 + k_a \sigma \tag{2.2}$$

à la limite de tolérance T_0,

Si z_1 $\leqslant T_0$ le lot est accepté. (2.3)

Si z_1 $> T_0$ on calcule une deuxième grandeur de contrôle z'_1 :

$$z'_1 = \overline{x}_1 + k_r \sigma. \tag{2.4}$$

Si z'_1 $\leqslant T_0$ on examine un deuxième échantillon de n_2 éléments, et on met en oeuvre les points 3 et 4

Si z'_1 $> T_0$ le lot est rejeté.

3) A partir des n_2 résultats de mesures x_ν ($\nu = n_1 + 1, \ldots, n_1 + n_2$) du deuxième échantillon de n_2 éléments, on calcule la moyenne arithmétique

$$\overline{x}_2 = \frac{1}{n_2} \sum_{\nu=n_1+1}^{n_1+n_2} x_\nu \tag{2.6}$$

La moyenne globale pour les deux échantillons est

$$\overline{x} = (n_1 \overline{x}_1 + n_2 \overline{x}_2)/n \tag{2.7}$$

avec $n = n_1 + n_2$.

4) La grandeur de contrôle

$$z = \overline{x} + k\sigma \tag{2.8}$$

détermine la décision définitive concernant l'acceptation ou le rejet du lot.

$$\text{Si} \quad z \begin{cases} \leqslant T_0 & \text{le lot est accepté} \\ > T_0 & \text{le lot est rejeté} \end{cases} \tag{2.9}$$

La fonction d'efficacité $L_D(p)$ exprime la probabilité que soit accepté un lot dont la proportion de défectueux est égale à p, soit à la suite des résultats de mesures sur le premier échentillon, soit à la suite des résultats globaux de mesures sur le premier et le deuxième échantillons, c'est-à-dire :

$$\begin{aligned} L_D(p) = {} & \Pr(\bar{X}_1 \leqslant T_0 - k_a \sigma) \\ & + \Pr(T_0 - k_a \sigma < \bar{X}_1 \leqslant T_0 \; k_r \sigma, \bar{X} \leqslant T_0 - k\sigma) \end{aligned} \tag{2.10}$$

En standardisant $\bar{X}_1$ et $\bar{X}$ et en introduisant l'indice de qualité u_{1-p}, on aboutit à

$$\begin{aligned} L_D(p) = {} & \Phi_s([u_{1-p} - k_a]\sqrt{n_1}) \\ & + \Psi_s([u_{1-p} - k]\sqrt{n}, [u_{1-p} - k_r]\sqrt{n_1} \; ; \rho) \\ & - \Psi_s([u_{1-p} - k]\sqrt{n}, [u_{1-p} - k_a]\sqrt{n_1} \; ; \rho) \end{aligned} \tag{2.11}$$

Ψ_s désignant la fonction de répartition normale bidimensionnelle réduite. Le coefficient de corrélation est défini par $\rho = \sqrt{n_1/n}$.

II.2 – Calcul des paramètres

$L_D(p)$ est déterminé par les paramètres k_a, k_r, k, n_1 et n_2. La condition posée au début, selon laquelle la courbe d'efficacité de l'échantillonnage double doit passer par les points $(p_1, 1 - \alpha)$ et (p_2, β), ne permet pas de déterminer tous ces paramètres. On peut cependant introduire une condition supplémentaire, en choisissant, parmi l'ensemble des plans qui remplissent les conditions définies plus haut, celui qui se présente comme optimal si on se réfère à certains critères donnés. Comme le nombre d'éléments N_2 du deuxième échantillon (et par conséquent aussi le nombre total d'éléments N) est une variable aléatoire, qui ne peut prendre que les valeurs $N_2 = 0$ ou $N_2 = n_2$, ce qui entraîne respectivement $N = n_1$ ou $N = n = n_1 + n_2$, il est judicieux d'établir l'échantillonnage double de manière à minimiser l'espérance mathématique E_N du nombre total d'éléments des deux échantillons. Evidemment E_N ne dépend pas que des paramètres de l'échantillonnage, mais aussi de l'indice de qualité u_{1-p}, et par là de la proportion de défectueux du lot. On obtient donc les paramètres de l'échantillonnage double en résolvant le problème d'optimisation suivant

$$\min_{\substack{k_a, k_r, k \\ n_1, n_2}} E_N = E(n_1 + N_2)$$

compte tenu de

$$\begin{aligned} L_D(p_1) &= 1 - \alpha \\ L_D(p_2) &= \beta \end{aligned} \tag{2.12}$$

avec

$$\begin{aligned} E_N &= n_1 + n_2 \Pr(T_0 - k_a\sigma < \bar{X}_1 \leqslant T_0 - k_r\sigma) \\ &= n_1 + n_2 \{\Phi_s([u_{1-p} - k_r]\sqrt{n_1}) - \Phi_s([u_{1-p} - k_a]\sqrt{n_1})\}. \end{aligned} \tag{2.13}$$

Il est nécessaire de fixer a priori l'objectif recherché, par exemple si E_N doit être minimisé pour les valeurs AQL ou LQ, si on cherche une solution de Bayes, ou bien par exemple une solution Min-Max, c'est-à-dire la minimisation de E_N pour la proportion de défectueux p pour laquelle E_N est maximal. Dans cet article nous avons opté pour cette dernière possibilité, c'est-à-dire que nous avons pris comme objectif la fonction

$$\underset{\substack{k_a,k_r,k \\ n_1,n_2}}{\mathrm{Min}} \left\{ \underset{p\in[0,1]}{\mathrm{Max}} \; E_N(p) \right\}$$

Dans le cas considéré ici, où la variance σ^2 est connue, $\underset{p\in[0,1]}{\mathrm{Max}} \; E_N(p)$ est facile à déterminer à l'aide de considérations de symétrie.

$$\begin{aligned} \underset{p\in[0,1]}{\mathrm{Max}} \; E_N(p) &= n_1 + n_2 \,.\, \underset{p\in[0,1]}{\mathrm{Max}} \; \{\Phi_s[(u_{1-p} - k_r]\sqrt{n_1}) \\ &\quad - \Phi_s([u_{1-p} - k_a]\sqrt{n_1})\} \\ &= n_1 + n_2 \,.\, \left\{ 2\,\Phi_s\left(\frac{k_a - k_r}{2}\sqrt{n_1}\right) - 1 \right\} \end{aligned} \tag{2.14}$$

Pour des raisons faciles à comprendre, il est raisonnable de limiter l'espace de résolution en considérant les conditions k_a, k_r, $k \geqslant 0$, ainsi que $n_1, n_2 \in \mathbf{N}$. La solution numérique du problème d'optimisation non linéaire à plusieurs variables entières ainsi obtenu se trouve cependant être encore extrêmement pénible. On peut toutefois mettre en évidence que, en choisissant judicieusement n_1 et n_2, et par conséquent en réduisant le problème à la détermination des paramètres k_a, k_r et k, on peut considérablement réduire les calculs, sans modifier de façon importante les valeurs de E_N (cf. [2]). Pour des raisons d'organisation, dans le cadre de l'activité de l'entreprise, il est recommandé de fixer $n_2 = n_1$ (ou $n_2 = 2n_1$) (cf. [3], contrôle par attributs). Si l'on fixe ainsi le rapport des nombres d'éléments des deux échantillons il est raisonnable de donner à n_1 des valeurs telles que (cf. [6], [2])

$$n_1 = [n_e/2] + 1 \qquad n_2 = n_1 \tag{2.15a}$$

ou bien

$$n_1 = [n_e/3] + 1 \qquad n_2 = 2n_1 \tag{2.15b}$$

n_e représentant le nombre d'éléments de l'échantillonnage simple équivalent.

Pour la solution numérique du problème d'optimisation non linéaire résultant de ces hypothèses, la référence [6] indique un moyen approprié qui permet de déterminer les paramètres k_a, k_r et k.

Si, par exemple, on recherche un plan d'échantillonnage double par variable équivalent à un plan donné du "Military-Standard 105 D" (contrôle par attributs), on obtient, pour les paramètres n_1, n_2, k_a, k_r et k, le rapport E_N/n_e (pour déterminer n_e, on calcule évidemment l'échantillonnage simple équivalent dans un contrôle par variables). Les tableaux 1 et 2 donnent les résultats obtenus dans quelques cas particuliers : valeur de $(E_N/n_e)_{max}$, et valeur maximale Δ_{max} de la valeur absolue de l'écart entre les courbes d'efficacité des échantillonnages simple et double.

L'examen des valeurs de Δ_{max} montre que la concordance entre les courbes d'efficacité des échantillonnages double et simple peut être considérée comme tout à fait acceptable. Les valeurs obtenues pour $(E_N/n_e)_{max}$ permettent de s'attendre, à long terme, à une économie importante sur les dépenses si l'on choisit un plan d'échantillonnage double de préférence à un plan simple, surtout si l'on tient compte de ce que, lorsque $p = p_1$ ou $p = p_2$ les valeurs de E_N/n_e sont nettement inférieures à celles de $(E_N/n_e)_{max}$. En comparant les valeurs des tableaux 1 et 2, on peut, de plus, constater que les valeurs de $(E_N/n_e)_{max}$ pour $n_2 = 2n_1$ sont supérieures à celles pour $n_2 = n_1$. D'un autre côté, il faut s'attendre à ce que, si $p \ll p_1$ ou $p \gg p_2$, les valeurs de E_N/n_e pour $n_2 = 2n_1$ soient inférieures à celles pour $n_2 = n_1$ (cf. [6]), car

$$\begin{aligned} \lim_{p\to 1} E_N = \lim_{p\to 0} E_N = \lim_{u_{1-p}\to\pm\infty} & [n_1 + n_2 \{\Phi_s([u_{1-p} - k_r]\sqrt{n_1}) \\ & - \Phi_s([u_{1-p} - k_a]\sqrt{n_1})\}] \\ & = n_1 \end{aligned}$$

Avec

$$n_1 = [n_e/2] + 1 \quad ; \quad n_2 = n_1$$
$$\Rightarrow \quad E_N/n_e \approx 0{,}5$$

Avec

$$n_1 = [n_e/3] + 1 \quad ; \quad n_2 = 2n_1$$
$$\Rightarrow \quad E_N/n_2 \approx 0{,}33$$

III. PLANS D'ECHANTILLONNAGE DOUBLE ($\bar{x}$; s)

Les instructions pour le contrôle concernant les plans d'échantillonnage double ($\bar{x}$; s) ne diffèrent pas, dans leurs bases fondamentales, de celles qui ont été détaillées pour les plans d'échantillonnage ($\bar{x}$; σ) ; le calcul supplémentaire des variances d'échantillonnage est nécessaire, et entraîne des modifications des valeurs de contrôle, la variance σ^2, inconnue, étant remplacée par sa valeur estimée.

TABLEAU 1

Plans d'échantillonnage double (σ connu, $n_2 = n_1$) équivalents aux plans "Military-Standard 105 D" (contrôle normal)

Plan de référence 105 D Lettre-code et AQL	n_1	k_a	k_r	k	E_N/n_e 1. $p = p_1$ 2. $p = p_2$	$(E_N/n_e)_{max}$	Δ_{max}
F 10	6	1,039	0,246	0,586	1. 0,752 2. 0,761	0,876	0,00252
J 2.5	11	1,823	1,137	1,484	1. 0,737 2. 0,788	0,857	0,00405
N 0.65	24	2,425	1,944	2,172	1. 0,763 2. 0,791	0,890	0,00385
Q 0.65	50	2,431	2,100	2,274	1. 0,746 2. 0,814	0,895	0,00829

TABLEAU 2

Plans d'échantillonnage double (σ connu, $n_2 = 2n_1$) équivalents aux plans "Military-Standard 105 D" (contrôle normal)

Plan de référence 105 D Lettre-code et AQL	n_1	k_a	k_r	k	E_N/n_e 1. $p = p_1$ 2. $p = p_2$	$(E_N/n_e)_{max}$	Δ_{max}
F 10	4	1,313	0,011	0.585	1. 0,811 2. 0,800	0,915	0,00255
J 2.5	8	1,817	0,993	1,499	1. 0,708 2. 0,865	0,916	0,00476
N 0.65	16	2,617	1,856	2,171	1. 0,858 2. 0,804	0,924	0,00785
Q 0.65	33	2,550	1,976	2,271	1. 0,831 2. 0,886	0,941	0,00206

Soient

$$s_1^2 = \frac{1}{n_1 - 1} \sum_{\nu=1}^{n_1} (x_\nu - \bar{x}_1)^2 \tag{3.1}$$

$$s_2^2 = \frac{1}{n_2 - 1} \sum_{\nu=n_1+1}^{n_1+n_2} (x_\nu - \bar{x}_2)^2 \tag{3.2}$$

En posant :

$$f_1 = n_1 - 1, f_2 = n_2 - 1, \ f = f_1 + f_2$$
$$s^2 = (f_1 s_1^2 + f_2 s_2^2)/f \tag{3.3}$$

est l'estimation de plus faible variance de σ^2. Selon les formules (2.2), (2.4) et (2.8)

$$z_1 = \bar{x}_1 + k_a s_1 \tag{3.4}$$

$$z_1' = \bar{x}_1 + k_r s_1 \tag{3.5}$$

$$z = \bar{x} + k s. \tag{3.6}$$

Pour la fonction d'efficacité, on a alors

$$L_D(p) = \underbrace{\Pr(\bar{X}_1 \leqslant T_0 - k_a S_1)}_{P_{a_1}} + \underbrace{\Pr(T_0 - k_a S_1 < \bar{X}_1 \leqslant T_0 - k_r S_1 , \bar{X} \leqslant T_0 - k S)}_{P_{a_2}}. \tag{3.7}$$

En standardisant $\bar{X}_1$ et $\bar{X}$ suivant $U = (\bar{X}_1 - \mu)\sqrt{n_1}/\sigma$ et $V = (\bar{X} - \mu)\sqrt{n}/\sigma$, et en introduisant

$$\begin{aligned} \delta_1 &= \sqrt{n_1}\,(T_0 - \mu)/\sigma \\ \delta &= \sqrt{n}\,(T_0 - \mu)/\sigma \end{aligned} \tag{3.8}$$

ainsi que

$$\begin{aligned} W_1 &= f_1\, S_1^2/\sigma^2 \\ W &= f\, S^2/\sigma^2 \end{aligned} \tag{3.9}$$

(3.7) devient

$$P_{a_1} = \Pr(U \leqslant \delta_1 - k_a \sqrt{n_1/f_1}\,\sqrt{W_1}) \tag{3.10}$$

et
$$P_{a_2} = \Pr(\delta_1 - k_a \sqrt{n_1/f_1}\,\sqrt{W_1} \leqslant U \leqslant \delta_1 - k_r \sqrt{n_1/f_1}\,\sqrt{W_1}, \; V \leqslant \delta - k\sqrt{n/f}\sqrt{W}). \tag{3.11}$$

Les variables aléatoires W_1 et W suivent des répartitions χ^2 avec respectivement $f_1 = n_1 - 1$ et $f = n_1 + n_2 - 2$ degrés de liberté. U et W_1, ainsi que V et W sont indépendants. En posant $k_a' = k_a \sqrt{n_1/f_1}$, et en tenant compte de ce que la fonction de densité de probabilité de W_1 est :

$$g_1(w_1) = \frac{w^{f_1/2 - 1} \exp\{-w_1/2\}}{2^{f_1/2}\,\Gamma(f_1/2)} \tag{3.12}$$

on obtient, si $w_1 = 2t_1$ et $k_1 = 1/\Gamma(f_1/2)$

$$P_{a_1} = k_1 \int_0^{\infty} \Phi_S(\delta_1 - k_a' \sqrt{2t_1})\, \underbrace{t_1^{f_1/2 - 1} \exp\{-t_1\}\, dt_1}_{f_{p_1}(t_1)}. \tag{3.13}$$

Pour le calcul numérique de la probabilité d'acceptation il convient de préférer cette fonction P_{a_1} à une forme également envisageable obtenue en se basant sur la fonction de distribution de la variable t non centrée.

La probabilité P_{a_2} se détermine de façon analogue. On obtient

$$P_{a_2} = \underbrace{\Pr(U \leqslant \delta_1 - k'_r\sqrt{W_1}\ ,\ V \leqslant \delta - k'\sqrt{W})}_{P'_{a_2}}$$

$$- \underbrace{\Pr(U \leqslant \delta_1 - k'_a\sqrt{W_1}\ ,\ V \leqslant \delta - k'\sqrt{W})}_{P''_{a_2}} \qquad (3.14)$$

où $$k'_r = k_r\sqrt{n_1/f_1}\ ,\ k' = k\sqrt{n/f}.$$

Comme les vecteurs aléatoires $(U\ ,V)$ et $(W_1\ ,W)$ sont indépendants des fonctions de densité de probabilité correspondantes $\Psi_s(u\ ,v\ ;\rho)$ et $g(w_1\ ,w)$, on obtient

$$P'_{a_2} = \int_0^\infty \int_0^w \int_{-\infty}^{\delta_1 - k'_r\sqrt{w_1}} \int_{-\infty}^{\delta - k'\sqrt{w}} \Psi_s(u\ ,v\ ;\rho)\, g(w_1\ ,w)\, du\, dv\, dw_1\, dw. \qquad (3.15)$$

avec $0 < w_1 < w < \infty$

Le coefficient de corrélation correspond à celui des plans $(\overline{x}\ ;\sigma)$: $\rho = \sqrt{n_1/n}$. $g(w_1\ ,w)$ devient

$$g(w_1\ ,w) = \frac{w^{f_1/2 - 1}(w - w_1)^{f_2/2 - 1} \exp\{-w/2\}}{2^{(f_1+f_2)/2}\,\Gamma(f_1/2)\,\Gamma(f_2/2)} \qquad (3.16)$$

(cf [1], [6]).

Si, pour simplifier, on pose $w = 2t$ ainsi que $w_1/w = y$, on obtient, à partir de (3.15)

$$P'_{a_2} = k_2 \int_{t=0}^{\infty} \int_{y=0}^{1} \Psi_s(\delta_1 - k'_r\sqrt{2\,ty}\ ,\delta - k'\sqrt{2t}\ ;\rho)\ \underbrace{t^{(f_1+f_2)/2 - 1}\, y^{f_1/2 - 1}\, (1-y)^{f_2/2 - 1} \exp\{-t\}}_{f_{p_2}(t\ ,y)}\, dy\, dt \qquad (3.17)$$

où $$k_2 = \frac{1}{\Gamma(f_1/2)\,.\,\Gamma(f_2/2)}$$

Les mêmes calculs s' appliquent à P'_{a_2}, en remplaçant systématiquement k'_r par k'_a. On obtient alors comme fonction d'efficacité du plan double $(\overline{x}\ ;s)$:

$$L_D(p) = k_1 \int_0^\infty \Phi_s(\delta_1 - k'_a\sqrt{2t_1})\, f_{p_1}(t_1)\, dt_1$$

$$+ k_2 \int_{t=0}^{\infty} \int_{y=0}^{1} [\Psi_s(\delta_1 - k'_r\sqrt{2\,ty}, \delta - k'\sqrt{2t}\ ;\rho)$$

$$- \Psi_s(\delta_1 - k'_a\sqrt{2\,ty}\ , \delta - k'\sqrt{2t}\ ;\rho)]\, f_{p_2}(t\ ,y)\, dy\, dt. \qquad (3.18)$$

La détermination des paramètres k_a, k_r et k du plan double $(\bar{x}\ ; s)$, se fait alors comme pour le plan double $(\bar{x}\ ; \sigma)$, en résolvant le problème d'optimisation non linéaire

$$\begin{aligned}\underset{k_a, k_r, k}{\mathrm{Min}} \ \underset{p \in [0,1]}{\mathrm{Max}} \ E_N &= E(n_1 + N_2) \\ &= n_1 + n_2 k_1 \int_0^\infty [\Phi_s(\delta_1 - k_r' \sqrt{2t_1}) \\ &\quad - \Phi_s(\delta_1 - k_a' \sqrt{2t_1})] f_{p_1}(t_1)\, dt_1\end{aligned}$$

de façon que

$$\begin{aligned}L_D(p_1) &= 1 - \alpha \\ L_D(p_2) &= \ \beta \\ k_a, k_r, k &\geqslant 0,\end{aligned} \tag{3.19}$$

n_1 et n_2 sont, comme pour le plan double $(\bar{x}\ ; \sigma)$ donnés a priori ; en ce qui concerne n_e, on considère le nombre d'éléments dans un plan simple $(\bar{x}\ ; s)$ défini par $(p_1\ , 1 - \alpha)$ et $(p_2\ , \beta)$.

La résolution numérique de (3.19) exige, même avec des ordinateurs modernes, une durée importante de calcul, en particulier lorsque le maximum minimisé de l'espérance mathématique E_N ne peut être déterminé que numériquement (cf [4]). Pour simplifier le problème, on se sert du fait que les valeurs de contrôle Z_1, Z_1' et Z ont une distribution approximativement normale, si n_1 et $n \geqslant 5$ (cf [7]).

En posant

$$\begin{aligned}E\{Z_1\} &\approx \mu + k_a \sigma \\ V\{Z_1\} &\approx \sigma^2 A^2(n_1\ , k_a) \\ E\{Z_1'\} &\approx \mu + k_r \sigma \\ V\{Z_1'\} &\approx \sigma^2 A^2(n_1\ , k_r),\end{aligned} \tag{3.20}$$

où $\qquad A(n_1\ , k_{a,r}) = 1/\sqrt{1/n_1 + k_{a,r}^2/2(n_1 - 1)}$,

on peut déterminer de façon approximative l'espérance mathématique du nombre d'éléments de l'échantillonnage de la façon suivante

$$\begin{aligned}E_N \approx n_1 + n_2 \{&\Phi_s([u_{1-p} - k_r] A(n_1\ , k_r)) \\ &- \Phi_s([u_{1-p} - k_a] A(n_1\ , k_a))\},\end{aligned} \tag{3.21}$$

Le maximum par rapport à p est obtenu pour

$$u_{1-p} = -a + \sqrt{a^2 - b} \tag{3.22}$$

où $\qquad a = \dfrac{k_a A(n_1\ , k_a)^2 - k_r A(n_1\ , k_r)^2}{A(n_1\ , k_r)^2 - A(n_1\ , k_a)^2}$

et $\quad b = \dfrac{k_r^2\, A(n_1, k_r)^2 - k_a^2\, A(n_1, k_a)^2 + 2\ln(A(n_1, k_a)/A(n_1, k_r))}{A(n_1, k_r)^2 - A(n_1, k_a)^2}$.

Des comparaisons numériques permettent de mettre en évidence que la position du maximum ainsi obtenu est en très bon accord avec la position du maximum exact même pour de très petits nombres d'éléments de l'échantillonnage $n_1 < 5$ (cf [6]). Il s' en suit que, pour la résolution de (3.19), on peut se passer de la détermination numérique de $\max\limits_{p \in [0,1]} E_N(p)$. La fonction – objectif minimisante se réduit à

$$E_{N_{max}} = n_1 + n_2 k_1 \int_0^{\infty} \{\Phi_s(\sqrt{n_1}\,[-a + \sqrt{a^2 - b}] - k_r' \sqrt{2t_1}) \tag{3.23}$$

$$- \Phi_s(\sqrt{n_1}\,[-a + \sqrt{a^2 - b}] - k_a' \sqrt{2t_1})\} f_{p_1}(t_1)\, dt_1$$

où a et b sont définis par (3.22).

Si le nombre d'éléments est plus grand, le calcul numérique de $L_D(p)$ selon (3.18) conduit à des difficultés importantes. C'est pourquoi nous proposons ci-dessous une approximation, qui tient également compte de ce que les valeurs de contrôle suivent approximativement une distribution normale pour $n_1, n \geqslant 5$.

L'application des égalités (3.7) et (3.20) conduit, en ce qui concerne $L_D(p) = P_{a_1} + P_{a_2}$, à

$$\begin{aligned} P_{a_1} &= \Pr(Z_1 = \overline{X}_1 + k_a S_1 \leqslant T_0) \\ &\approx \Pr(U_A \leqslant [u_{1-p} - k_a]\, A(n_1, k_a)) \\ &= \Phi_s([u_{1-p} - k_a]\, A(n_1, k_a)) \end{aligned} \tag{3.24}$$

où $\quad U_A = (Z_1 - E\{Z_1\})/\sqrt{V\{Z_1\}}$

De même :

$$\begin{aligned} P_{a_2} &= \Pr(Z_1' = \overline{X}_1 + k_r S_1 \leqslant T_0 < \overline{X}_1 + k_a S_1 = Z_1\,,\ Z = \overline{X} + kS \leqslant T_0) \\ &\approx \Pr(U_r \leqslant [u_{1-p} - k_r]\, A(n_1, k_r)\,,\ V \leqslant [u_{1-p} - k]\, A(n, k)) \\ &\quad - \Pr(U_a \leqslant [u_{1-p} - k_a]\, A(n_1, k_a)\,,\ V \leqslant [u_{1-p} - k]\, A(n, k)) \\ &= \Psi_s([u_{1-p} - k_r]\, A(n_1, k_r)\,, [u_{1-p} - k]\, A(n, k)\,; \rho_r^*) \\ &\quad - \Psi_s([u_{1-p} - k_a]\, A(n_1, k_a)\,, [u_{1-p} - k]\, A(n, k)\,; \rho_a^*) \end{aligned} \tag{3.25}$$

où $\quad A(n, k) = 1/\sqrt{1/n + k^2/2(n-2)}$

et $\quad \rho_{a,r}^* = \mathrm{cov}\,\{U_{a,r}, V\}$

$$\approx \sqrt{n_1/n}\, \frac{1 + \dfrac{k_{a,r}\, k}{2} \cdot \dfrac{n}{f}}{\sqrt{\left(1 + \dfrac{k_{a,r}^2}{2} \cdot \dfrac{n_1}{f_1}\right)\left(1 + \dfrac{k^2}{2} \cdot \dfrac{n}{f}\right)}}.$$

Pour déterminer $\rho^*_{a,r}$, on a tenu compte dans les calculs ci-dessus des espérances $E\{Z_1\}$, $E\{Z'_1\}$ et $E\{Z\}$, et non des valeurs approchées résultant de (3.20), car ces dernières conduisent à des résultats trop évidemment inexacts. Les valeurs exactes utilisées ont pour expressions

$$\begin{aligned} E\{Z_1\} &= \mu + k_a \sigma a_1 \\ E\{Z'_1\} &= \mu + k_r \sigma a_1 \\ E\{Z\} &= \mu + k \sigma a_n \end{aligned} \tag{3.26}$$

où

$$a_1 = \sqrt{2/f_1}\, \Gamma((f_1 + 1)/2)/\Gamma(f_1/2)$$

$$a_n = \sqrt{2/f'}\, \Gamma((f + 1)/2)/\Gamma(f/2).$$

Après des calculs assez longs et en utilisant les relations [5] :

$$\begin{aligned} 1 - a_n^2 &\approx 1/(2f) \\ a_1/a_n &\approx 1 \end{aligned} \tag{3.27}$$

on obtient, pour de "grandes" valeurs de f, les expressions ci-dessus de ρ^*_a, ρ^*_r.

La référence [6] indique une méthode de calcul numérique appropriée pour résoudre les équations (3.19), en considérant pour $L_D(p)$ les relations (3.18) ou (3.24) et en tenant compte de (3.25). Les tableaux 3 et 4 donnent les valeurs des paramètres k_a, k_r et k, les valeurs de E_N/n_e, et de Δ_{max}, pour différents plans équivalents de "Military Standard 105 D". Ces valeurs peuvent également être utilisées pour les plans doubles $(\bar{x}\ ;\sigma)$ en tant que plans de référence.

Les valeurs de Δ_{max} montrent que lors du calcul concret des plans d'échantillonnage double $(\bar{x}\ ; s)$, il est beaucoup plus difficile d'arriver à un bon accord entre les courbes d'efficacité des plans doubles et simples. Par exemple, en ce qui concerne toute la série de plans du "Military Standard 105 D", qui ne sont pas cités ici, on n'aboutit lors des calculs qu'à des valeurs $\Delta_{max} > 0{,}02$ (atteignant même 0,04). Il est évident que dans ce cas on aboutit à un résultat décevant. De plus les économies atteintes ne sont (relativement) pas aussi importantes que pour les plans doubles $(\bar{x}\ ;\sigma)$. Avant de prendre une décision définitive il faut cependant tenir compte du fait que, pour les plans $(\bar{x}\ ; s)$ les échantillons sont beaucoup plus importants que pour les plans $(\bar{x}\ ;\sigma)$, si bien que, en valeur absolue, les économies à réaliser sont en règle générale supérieures à celles qui sont obtenues dans les plans $(\bar{x}\ ;\sigma)$.

Un autre fait parle en faveur d'une mise en œuvre effective des plans d'échantillonnage double $(\bar{x}\ ; s)$: les valeurs des rapports E_N/n_e sont pour les proportions de défectueux $p = p_1$ et $p = p_2$ (les plus intéressantes en règle générale) considérablement inférieures aux valeurs maximales. Pour des proportions de défectueux $p \ll p_1$ ou $p \gg p_2$ les considérations concernant les systèmes $(\bar{x}\ ;\sigma)$ (cf. paragraphe 2), peuvent être appliquées.

TABLEAU 3

Plans d'échantillonnage double (σ inconnu, $n_2 = n_1$) équivalents aux plans "Military-Standard 105 D" (contrôle normal)

Plan de référence 105 D Lettre-code et AQL	n_1	k_a	k_r	k	E_N/n_e 1. $p = p_1$ 2. $p = p_2$	$(E_N/n_e)_{max}$	Δ_{max}
F 10	7	1,628	0,303	0,610	1. 0,920 2. 0,736	0,943	0,00180
J 2.5	24	2,036	1,129	1,505	1. 0,876 2. 0,849	0,965	0,01380
N 0.65	80	2,052	1,947	2,171	1. 0,826 2. 0,786	0,909	0,00303
Q 0.65	176	2,476	2,103	2,271	1. 0,798 2. 0,801	0,904	0,00254

TABLEAU 4

Plans d'échantillonnage double (σ inconnu, $n_2 = 2n_1$) équivalents aux plans "Military-Standard 105 D" (contrôle normal)

Plan de référence 105 D Lettre-code et AQL	n_1	k_a	k_r	k	E_N/n_e 1. $p = p_1$ 2. $p = p_2$	$(E_N/n_e)_{max}$	Δ_{max}
F 10	5	1,826	0,263	0,587	1. 0,936 2. 0,709	0,953	0,00942
J 2.5	16	2,277	1,176	1,498	1. 0,900 2. 0,714	0,922	0,00699
N 0.65	54	2,620	1,844	2,173	1. 0,852 2. 0,872	0.927	0,00349
Q 0.65	118	2,609	2,025	2,271	1. 0,878 2. 0,845	0,941	0,00265

REFERENCES

[1] BULGREN W.G., DYKSTRA R.L., HEWETT J.E. – A bivariate t-Distribution with Applications, *Jasa*, Vol. 69, 1974, pp. 525-532.

[2] HALD A. – Optimum Double Sampling Tests of Given Strength. I. The Normal Distribution, *Jasa*, Vol. 70, 1975, pp. 451-456.

[3] Office of the Assistant Secretary of Defense : Sampling Procedures and Tables for Inspection by Attributes, U.S. Government Printing Office, Washington, 1963.

[4] SPURRIER J.D., HEWETT J.E. – Double Sample Tests for the Mean of a Normal Population, *Jasa*, Vol. 70, 1975, pp. 448-451.

[5] STANGE K. – Angewandte Statistik, 1. Teil, Eindimensionale Probleme, Springer-Verlag, Berlin, Heidelberg, New York, 1970.

[6] STURHAHN J. – Doppelstichprobenpläne für messende Prüfung, Diss. RWTH Aachen, 1976.

[7] WILRICH P.-Th. – Nomogramme zur Ermittlung von Stichprobenplänen für messende Prüfung bei einer einseitig vorgeschriebenen Toleranzgrenze, Teil 2: Pläne bei unbekannter Varianz der Fertigung, QZ 15 (1970), S. 181-187.