

## Introduction

Ceci est la mise au propre d'un cours donné par Virginie Mas-Leroux et Monique Pontier. Il s'appuie sur l'expérience de Madame Leroux et sur la bibliographie donnée en annexe (essentiellement les ouvrages de Baillargeon [2, 3] et Pillet [5]).

Plan :

Capabilité(s)

Cartes de contrôles

Plans d'échantillonnage simple

Plans d'échantillonnage double

Cartes multi-critères

Cas des petites séries

AMDEC

Caractéristiques clés

# 1 Capabilité d'un procédé

## 1.1 Spécifications d'un procédé

Lorsqu'un objet est produit en série au cours d'un procédé (de fabrication), certaines de ses caractéristiques sont numériques : elles doivent être impérativement proches d'un objectif idéal, comprises dans une fourchette "acceptable". Celle-ci est fixée par le bureau d'études.

**Définition 1.1** On appelle "spécifications" un couple de valeurs  $(T_i, T_s)$  entre lesquelles doit se trouver la grandeur mesurable  $X$  concernée. On les appelle aussi "tolérances inférieure et supérieure".

Ce mot de "tolérance" est également utilisé mais dans un autre sens en fiabilité (cf. le cours correspondant du second semestre), où l'on parle aussi de "limites de dispersion".

Ces limites sont établies par l'ingénieur-concepteur pour les valeurs individuelles mesurées, et non pour les moyennes sur des échantillons prélevés.

On se fonde sur l'hypothèse (forte!) que ces mesures suivent une loi gaussienne  $\mathcal{N}(m, \sigma^2)$ ,  $m$  et  $\sigma^2$  devant être estimées sur des échantillons. De fait, dans des cas assez rares, l'hypothèse gaussienne n'est pas utilisée, mais par exemple il arrive parfois que l'on utilise des lois du type  $P(X \geq x) = e^{-e^{-x}}$ .

En général, on voudrait que  $m$  soit l'objectif idéal et  $m = \frac{T_i + T_s}{2}$ . Ceci n'est pas toujours le cas. Le but de ce chapitre est de donner un moyen qui permette de mesurer le risque de produire des objets défectueux (non conformes), i.e.  $X_j$  n'appartient pas à  $[T_i, T_s]$ . D'où l'importance de maîtriser l'usage de la loi gaussienne et des tables afférentes.

## 1.2 Quelques exercices sur la loi gaussienne

- (i)  $X$  est la mesure de la caractéristique étudiée. On suppose que  $X$  suit la loi  $\mathcal{N}(m, \sigma^2)$
- . Calculer  $\mathbb{P}\{X \leq T_i\}$ ,  $\mathbb{P}\{X \geq T_s\}$  en fonction de  $T_i, T_s, m, \sigma^2$ .
  - . Application numérique :  $T_i = 78 - 12, T_s = 78 + 12, m = 78, \sigma^2 = 3.8$ .
  - . Quel est le pourcentage attendu d'objets non conformes ?

- (ii)  $X$  est la mesure de la caractéristique étudiée. On suppose qu'un meilleur réglage de la production permet de diminuer  $\sigma$ . Que doit être  $\sigma$  pour que le pourcentage de défectueux diminue de moitié ?

Soit  $\Phi$  la fonction de répartition de la loi gaussienne centrée ‘réduite :

$$\begin{aligned}\Phi(3.72) &= 10^{-4}, \quad \Phi(4.265) = 10^{-5}, \quad \Phi(4.7534) = 10^{-6}, \\ \Phi(5.1993) &= 10^{-7}, \quad \Phi(5.512) = 10^{-8}, \quad \Phi(5.998) = 10^{-9}.\end{aligned}$$

### 1.3 Détermination de la capabilité d’un procédé stable

**Définition 1.2** *On dit qu’un procédé est “stable” s’il a été suffisamment testé et corrigé pour que la suite des mesures  $(X_i)$  suive une loi gaussienne.*

Au vu de  $n$   $k$ -échantillons préalables, on admet que

- la loi est gaussienne,
- la suite des moyennes  $(\bar{X}_i, i = 1, \dots, n)$  est stable,
- la suite des étendues  $(R_i = \max_j(X_i^j) - \min_j(X_i^j), i = 1, \dots, n)$  ou des écarts-types  $(s_i)$  est stable.

**Question :** quel outil statistique permet-il de tester correctement ces trois hypothèses ?

#### 1.3.1 Estimation de $\sigma$

Deux cas sont distingués :  $k \leq 10$  ou  $k \geq 10$ .

.  $k \leq 10$ ,  $\hat{\sigma} = \frac{1}{d_2(k)} \frac{1}{n} \sum_{i=1}^n R_i$  où  $d_2(k)$  est donné par une table, calculé en sorte que cet estimateur soit sans biais.

.  $k \geq 10$ ,  $\hat{\sigma} = \frac{1}{c_4(k)} \frac{1}{n} \sum_{i=1}^n s_i$  où  $c_4(k)$  est donné par une table, calculé en sorte que cet estimateur soit sans biais. C’est la pratique usuelle, tout à fait contestable, cet estimateur n’a pas beaucoup de propriétés, il vaudrait mieux choisir un estimateur uniformément plus puissant (UPP).

**Question :** que proposeriez vous à la place ?

$k$	2	3	4	5	6	7	8	9	10
$d_2$	1.128	1.693	2.059	2.326	2.534	2.704	2.847	2.970	3.078
$k$	10	12	14	15	16	18	20	22	25
$c_4$	0.9727	0.9776	0.9810	0.9823	0.9835	0.9854	0.9869	0.9882	0.9896

De façon générale,  $c_4(k) = \sqrt{\frac{2}{k-1} \frac{\Gamma(k/2)}{\Gamma((k-1)/2)}}$ .

Se vérifie, exercice, en calculant  $E(s_i)$  où  $(k-1)s_i^2$  suit une loi  $\chi_{k-1}^2$ , c'est à dire une loi  $\Gamma(\frac{k-1}{2}, \frac{1}{2})$ .

Ces estimateurs ne sont pas très bons, on peut suggérer à la place l'estimateur UPP :

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{1}{k-1} \sum_{j=1}^k (X_i^j - \bar{X}_i)^2}.$$

Ce défaut provient de la rigidité des cartes de contrôles qui font des relevés par échantillon et ne mémorise pas le détail des valeurs individuelles.

On pourrait pourtant, puisqu'on calcule d'abord les  $s_i^2$  en faire D'ABORD la moyenne, PUIS extraire la racine carrée, plutôt qu'extraire toutes les racines et d'en faire ENSUITE la moyenne....

### 1.3.2 Estimation de $m$

Classiquement  $\hat{m} = \frac{1}{n} \sum_{i=1}^n \bar{X}_i$ , noté souvent  $\bar{\bar{X}}$ .

### 1.3.3 Estimation de la proportion de non conformes

Cette proportion est estimée par

$$\hat{\mathbb{P}}\{X \leq T_i\} + \hat{\mathbb{P}}\{X \geq T_s\} = \Phi\left(\frac{T_i - \hat{m}}{\hat{\sigma}}\right) + \Phi\left(\frac{-T_s + \hat{m}}{\hat{\sigma}}\right).$$

### 1.3.4 Exemples

*Exercice (Baillargeon page 137) : on considère un procédé de placage en or de circuits imprimés. On dispose de  $n = 25$  échantillons de taille  $k = 10$  de l'épaisseur du placage  $X$ . On obtient  $\sum_{j=1,25} \bar{X}_j = 922.5$  et  $\sum_{j=1,25} s_j = 63.53$ . L'objectif est une épaisseur de 36,  $T_i = 30, T_s = \infty$ . Estimer les paramètres de la variable  $X$  en supposant qu'elle est gaussienne, en déduire le nombre de non conformes sur une production de 10 000 circuits.*

$$\bar{\bar{X}} = 36.9 ; \hat{\sigma} = \frac{63.53}{25c_4(10)} ; \mathbb{P}\{X \leq 30\} = \Phi\left(-\frac{6.9}{63.53}25 \times 0.9727\right) = \Phi(-2.64)$$

où  $\Phi$  est la fonction de répartition de la loi gaussienne standard. Le nombre non conformes attendu est de l'ordre de 410.

## 1.4 Indice de capabilité $C_p$

**Définition 1.3**  $C_p := \frac{T_s - T_i}{6\hat{\sigma}}$ .

Parfois, on appelle "capabilité" la quantité  $6\hat{\sigma}$ .

Interprétation gaussienne : si  $C_p = 1$ , et si l'on suppose  $m = T_0 = \frac{T_i + T_s}{2}$ , alors on obtient la proportion de non conformes suivante :

$$\begin{aligned} p_d &= \mathbb{P}\{X \leq T_i\} + \mathbb{P}\{X \geq T_s\} = \Phi\left(\frac{T_i - m}{\hat{\sigma}}\right) + 1 - \Phi\left(\frac{T_s - m}{\hat{\sigma}}\right) = \\ &= \Phi(-3) + 1 - \Phi(3) = 0.0024. \end{aligned}$$

Si  $C_p > 1$ , on a une proportion de défectueux de moins de 0.0024,

Si  $C_p < 1$ , on a une proportion de défectueux de plus de 0.0024, ce qui n'est pas acceptable.

**Remarque 1.4** Les praticiens aiment bien que  $C_p = 1.33$  ce qui correspond à  $\frac{T_s - m}{\hat{\sigma}} = 4$ , et à une proportion de défectueux de 6 pour 100 000.

Si  $m \neq T_0$ , cette estimation n'est plus très rigoureuse, le nombre effectif de non conformes augmente si l'intervalle  $[T_i, T_s]$  croît ou décroît, le pire étant atteint lorsque  $T_i$  ou  $T_s = m$  ; dans ce cas la proportion de non conformes peut atteindre 50 pour cent. D'où la nécessité de définir d'autres indices.

## 1.5 Autres indices de capabilité

**Définition 1.5**  $C_{pk} := \min\left(\frac{T_s - \bar{X}}{3\hat{\sigma}}, \frac{\bar{X} - T_i}{3\hat{\sigma}}\right)$ ,  $C_{pu} := \frac{T_s - \bar{X}}{3\hat{\sigma}}$ ,  $C_{pl} := \frac{\bar{X} - T_i}{3\hat{\sigma}}$ .

Ces indices pallient bien l'inconvénient que  $m$  diffère de la valeur  $T_0$ .

**Proposition 1.6**  $C_p = C_{pk} \Leftrightarrow \bar{X} = T_0$ . Plus précisément

$$C_p - C_{pk} = \frac{|c_0 - \bar{X}|}{3\hat{\sigma}}.$$

**Preuve** en exercice.

## 1.6 Interprétation probabiliste de $C_{pk}$

Si l'on revient à la définition de  $C_{pk}$  en acceptant l'hypothèse gaussienne avec les paramètres estimés  $(\bar{\bar{X}}, \hat{\sigma})$ , et si  $C_{pk} \geq 2.5$ , alors l'évènement  $\{X \leq T_i\} \cup \{X \geq T_s\}$  est de probabilité évaluée par

$$\mathbb{P}\left\{\frac{X - \bar{\bar{X}}}{\hat{\sigma}} \leq \frac{T_i - \bar{\bar{X}}}{\hat{\sigma}}\right\} + \mathbb{P}\left\{\frac{X - \bar{\bar{X}}}{\hat{\sigma}} \geq \frac{T_s - \bar{\bar{X}}}{\hat{\sigma}}\right\}$$

majorée par  $2(1 - \Phi(3C_{pk}))$  puisque  $\min(\frac{T_s - \bar{\bar{X}}}{\hat{\sigma}}, \frac{\bar{\bar{X}} - T_i}{\hat{\sigma}}) \geq 3C_{pk}$  ce qui majore la proportion de non conformes :

**la proportion de non conformes est majorée par  $2(1 - \Phi(3C_{pk}))$ .**

Exercice : évaluer cette proportion de non-conformes pour  $C_{pk}$  variant de 2.5 à 0.8.

## 1.7 Capabilités théorique et pratique

Si l'on connaît la **vraie** valeur de  $\sigma$ , la capabilité "théorique" est  $C_p^{th} = \frac{T_s - T_i}{6\sigma}$ . On voudrait que cette quantité soit supérieure à  $C_0$ , valeur "idéale" de capabilité. Mais on ne dispose que de l'estimation  $\hat{\sigma}$ .

(i) Lorsque  $\sigma$  ou plutôt  $\sigma^2$  est estimé à partir des variances expérimentales des  $k$ -échantillons, soit :

$$s_i^2 = \frac{1}{k-1} \sum_{j=1}^k (x_i^j - \bar{x}_i)^2, \quad i = 1, \dots, n,$$

on obtient  $\hat{\sigma}^2 = \frac{1}{n} \sum_i s_i^2 = \frac{1}{n(k-1)} \sum_i \sum_{j=1}^k (x_i^j - \bar{x}_i)^2$  auquel cas  $n(k-1)\hat{\sigma}^2$  suit une loi de khi-deux à  $p = n(k-1)$  degrés de liberté. Ceci donne  $\sigma^2$  dans un intervalle de confiance :

$$\chi_p^2[b_p, \infty[ = 0.95 = \mathbb{P}\left\{p\left(\frac{\hat{\sigma}}{\sigma}\right)^2 \geq b_p\right\},$$

donc, à 95 pour cent, on peut dire que  $\sigma \leq \hat{\sigma} \sqrt{\frac{p}{b_p}}$  soit

$$C_p^{th} = C_p^{exp} \frac{\hat{\sigma}}{\sigma} \geq C_p^{exp} \sqrt{\frac{b_p}{p}} \geq C_0 \text{ dès que } C_p^{exp} \geq \sqrt{\frac{p}{b_p}} C_0.$$

**Proposition 1.7** Soit  $p = n(k-1)$  et  $b_p$  tel que  $\chi_p^2[b_p, \infty[ = 0.95$ , alors  $C_p^{exp} \geq \sqrt{\frac{p}{b_p}} C_0 \Rightarrow C_p^{th} \geq C_0$ .

Il existe des tables pour cela. Mais il vaut mieux se rappeler que pour  $p \geq 30$ , la loi de  $\sqrt{2}\chi_p^2 - \sqrt{2p-1}$  est à peu près la loi gaussienne centrée réduite.

*Exercice* : Si l'on veut une "confiance" de  $1 - \alpha$ , le coefficient correcteur dans le cas  $p \geq 30$  est  $\rho_p = \frac{\sqrt{2p}}{\sqrt{2p-1} + q_\alpha}$  où  $q_\alpha$  est le  $\alpha$ -quantile de la loi gaussienne standard (soit  $\mathbb{P}\{Y \leq q_\alpha\} = \alpha$ ).

**Preuve** : Soit  $Y$  de loi gaussienne standard,

$$\mathbb{P}\{\chi_p^2 \geq b_p\} = \mathbb{P}\{Y \geq \sqrt{2b_p} - \sqrt{2p-1}\} = 1 - \alpha$$

donne  $\sqrt{2b_p} - \sqrt{2p-1} = q_\alpha$  soit  $b_p = \frac{1}{2}(\sqrt{2p-1} + q_\alpha)^2$ . Dans le cas  $p \geq 30$ , le coefficient correcteur

$$\rho_p = \sqrt{\frac{p}{b_p}} = \frac{\sqrt{2p}}{\sqrt{2p-1} + q_\alpha}$$

où  $q_\alpha$  est le  $\alpha$ -quantile de la loi gaussienne standard.

(ii) Par contre, lorsque  $\sigma$  est estimé à partir des étendues des  $k$ -échantillons, soit

$$R_i = \sup_j x_i^j - \inf_j x_i^j, \quad \hat{\sigma} = \frac{1}{nd_2(k)} \sum_{i=1}^n R_i,$$

le problème est de connaître la loi de  $\bar{R}$ . A priori, on sait seulement que  $E[\bar{R}] = \sigma d_2(k)$ , mais la loi de cette variable aléatoire n'est pas facilement accessible, et l'on ne dispose pas de tables permettant d'évaluer dans ce cas la probabilité  $\mathbb{P}\{\frac{\bar{R}}{\sigma d_2(k)} \geq b\}$  et de trouver  $b$  tel que cette probabilité soit  $1 - \alpha$  afin d'obtenir l'intervalle de confiance  $1 - \alpha$  de la forme

$$\left[0, \frac{\bar{R}}{bd_2(k)}\right].$$

•

Remarquer néanmoins que ce calcul d'intervalle de confiance convient dès que le nombre d'observations, ou plutôt le nombre de degrés de liberté  $p = (n-1)k$  est assez grand, soit par exemple  $p \geq 30$ , ce qui est souvent le cas même si  $k \leq 10$ .... Ce qui signifie qu'il est beaucoup plus pertinent d'estimer  $\sigma^2$  de toutes façons par les variances expérimentales que par les étendues.

*Exercice* : calculer  $\rho_p$  pour  $p = 40, 50, 75, 100$  et comparer avec les coefficients donnés par Baillargeon page 142.

Le coefficient correcteur est différent pour  $C_{pk}$ . En effet, il y a en plus une erreur sur

la moyenne estimée par  $\bar{X}$ :

$$C_{pk}^{th} = \min\left(\frac{T_s - m}{3\sigma}, \frac{m - T_i}{3\sigma}\right) = \min\left(\frac{T_s - \bar{X}}{3\sigma} + \frac{\bar{X} - m}{3\sigma}, \frac{\bar{X} - T_i}{3\sigma} + \frac{m - \bar{X}}{3\sigma}\right)$$

c'est à dire  $\frac{\hat{\sigma}}{\sigma} C_{pk}^{exp} \pm E$  où  $E = \frac{\bar{X} - m}{3\sigma}$  suit une loi de Gauss centrée de variance  $\frac{1}{9nk}$  et il faut ajouter l'erreur à 95 pour cent due à  $E$  soit  $\frac{1.645}{3\sqrt{nk}}$ , si  $nk$  est bien sûr le nombre TOTAL d'observations avec lesquelles on a calculé la moyenne  $\bar{X}$ . Par exemple, pour  $nk = 20$  on trouve bien cette erreur de 0.09 qui s'ajoute au coefficient correcteur de  $C_p$  égal à 1.37.



## 2 Cartes de contrôle pour des grandeurs mesurables

### 2.1 Introduction, objectifs

Les objectifs de ces cartes dans le contrôle de la qualité de la production sont :

- contrôle de la stabilité (on produit des choses en série identiques)
- contrôle du nombre de produits non conformes dans chaque lot produit.

Il y a aussi le contrôle des caractéristiques qualitatives, “contrôle de la qualité par attribut”, hors programme. Ici, on cherche à contrôler une grandeur mesurable, i.e. à valeurs réelles, d’où le titre du chapitre.

Préalables à poser avant la mise en oeuvre :

- . quelles caractéristiques veut-on contrôler ?
- . instruments de mesure à utiliser ?
- . taille des échantillons à l’intérieur d’un lot (le plus souvent entre 5 et 10) ?
- . fréquence des contrôles ? (en cas de production continue : tous les jours ? toutes les heures ?)
- . coût du contrôle ? (à mettre en regard avec le coût de produire des objets défectueux...)

De fait, il faut savoir qu’en général, il n’est pas utilisé de bases statistiques réelles.

### 2.2 Définition d’une carte de contrôle, différents types de cartes

Il s’agit d’un relevé des mesures d’une caractéristique numérique en fonction du temps, autour de la moyenne, entre deux lignes horizontales, d’ordonnées appelées LIC et LSC, ou limites provisoires de contrôle, égales à la valeur “cible” (l’objectif idéal)  $\pm 3\hat{\sigma}$ . Remarquer que ces limites coïncident avec les limites de tolérance  $T_i$  et  $T_s$  lorsque  $C_p = 1$ .

*dessin*

Il s’agit de diagnostiquer des écarts à cette bande de “confiance”. Il y a différents types de cartes :

cartes  $\bar{X}$  et  $R$  dans le cas d’une taille d’échantillon  $k \leq 10$ ,

cartes  $\bar{X}$  et  $s$  dans le cas d’une taille d’échantillon  $k \geq 10$ .

Il existe d’autres types de cartes : carte EWMA (cf. [5] pages 248 et sq.) qui utilise des moyennes glissantes ou carte CUSUM par exemple (cf. [2] p. 216 ou [5] pages 256 et sq.) quand on s’intéresse à l’écart à la valeur cible.

## 2.3 Sources de variabilité

La variabilité, qui contredit la stabilité de la production, peut venir des matières premières, de la main d'oeuvre, des outils de fabrication, des instruments de mesure. (cf. le diagramme d'ISHIKAWA vu en master 1 ou le schéma AMDEC qui sera présenté dans un chapitre ultérieur).

**Définition 2.1** *On dit qu'un procédé de fabrication est dans un **état stable** si la répartition expérimentale de la valeur mesurée sur l'échantillon est approximativement gaussienne (on dit que l'on a la **maitrise statistique du procédé**) et à peu près identique d'un échantillon à l'autre, tout en restant dans l'intervalle  $[c_0 - 3(\hat{\sigma}), c_0 + 3(\hat{\sigma})]$ , et sans présenter de "tendance".*

Normalement c'est  $c_0$ , la valeur cible ; de fait, quand le procédé est bien maîtrisé, on remplace  $c_0$  par  $\bar{X}$ .

## 2.4 Carte $\bar{X}$

Cas  $k \geq 10$

La fréquence du prélèvement dépend de la vie de l'entreprise, du coût, etc. Le nombre  $n$  de  $k$ -échantillons doit être supérieur à 20 pour que  $\bar{X}$  et  $s$  soient fiables. On opère les calculs

$$\forall i = 1, \dots, n, \bar{X}_i = \frac{1}{k} \sum_j X_i^j, s_i^2 = \frac{1}{k-1} \sum_j (X_i^j - \bar{X}_i)^2.$$

On obtient un graphique avec le temps en abscisse et les estimations en ordonnée, chaque point représentant un échantillon, si  $k = 1$  il s'agit d'un contrôle exhaustif.

On définit trois zones correspondant à :

$$m \pm \sigma_{\bar{X}} ; m \pm 2\sigma_{\bar{X}} ; m \pm 3\sigma_{\bar{X}},$$

la deuxième est une zone de surveillance, et la troisième une zone de contrôle. Bien sûr,  $m$  est estimée en général par  $\bar{X}$  obtenue avec de précédents échantillons (une fois ceux ci considérés comme bien maîtrisés).

*Exercice :*  $\hat{\sigma}_{\bar{X}} = \frac{1}{\sqrt{k}} \hat{\sigma}$ .

On rappelle que  $\sigma$  est estimé par  $\frac{\bar{s}}{c_4(k)}$ , traditionnellement, le coefficient  $\frac{3}{\sqrt{k}c_4(k)}$  est noté  $A_3(k)$  qui figure dans des tables

$k$	11	12	13	14	15	16	17	18	19	20	22	24	25
$A_3$	0.927	0.886	0.850	0.817	0.789	0.763	0.739	0.718	0.698	0.680	0.647	0.619	0.606
$1/c_4$	1.025	1.023	1.022	1.019	1.018	1.017	1.016	1.015	1.014	1.013	1.011	1.010	1.010

Cas  $k \leq 10$  pour les petits échantillons.

On relève l'étendue  $R_i$  pour chaque échantillon  $i$ , on fait leur moyenne  $\bar{R}$  et on rappelle que  $\sigma$  est estimé par  $\hat{\sigma} = \frac{\bar{R}}{d_2(k)}$ . Quant à  $\sigma_{\bar{X}}$  elle est estimée par  $\frac{\bar{R}}{\sqrt{k}d_2(k)}$ , traditionnellement, le coefficient  $\frac{3}{\sqrt{k}d_2(k)}$  est noté  $A_2(k)$  qui figure dans des tables.

$k$	2	3	4	5	6	7	8	9	10
$A_2$	1.880	1.023	0.729	0.577	0.483	0.419	0.373	0.337	0.308

Limites provisoires : au départ, on prend comme limites  $c_0 \pm 3\hat{\sigma}_{\bar{X}}$ , puis on passe à  $\bar{\bar{X}} \pm 3\hat{\sigma}_{\bar{X}}$  lorsque l'on est sûr que la moyenne est la cible et que l'on a enlevé les points aberrants (hors limites de contrôle) qui augmentent induement l'écart-type. Ensuite, on actualise  $\bar{\bar{X}}$  et  $\hat{\sigma}$  au fil des échantillons observés, jusqu'à stabilisation ; on obtient alors les limites de contrôle, soit  $\bar{\bar{X}} \pm i\hat{\sigma}_{\bar{X}}$  avec  $i = 1, 2, 3$ .

## 2.5 Carte $s$ , cas $k \geq 10$

On introduit de même des limites pour  $s$  qui proviennent de quantiles de la loi de khi-deux. Là aussi, on fournit des tables donnant les coefficients  $B_3(k)$  et  $B_4(k)$  tel que  $LIC_s = B_3(k)\bar{s}$ ,  $LSC_s = B_4(k)\bar{s}$ . On peut montrer  $B_i(k) = 1 \pm \frac{3\sqrt{1-c_4^2(k)}}{c_4(k)}$ .

$k$	11	12	13	14	15	16	17	18	19	20	22	24	25
$B_3$	0.321	0.354	0.382	0.405	0.428	0.448	0.446	0.482	0.497	0.510	0.534	0.555	0.565
$B_4$	1.679	1.646	1.618	1.594	1.572	1.552	1.534	1.518	1.503	1.490	1.466	1.445	1.435

## 2.6 Carte $R$ , cas $k \leq 10$

On relève l'étendue  $R_i$  pour chaque échantillon  $i$ , on fait leur moyenne  $\bar{R}$ . On calcule les limites de contrôle de  $R$  avec des coefficients figurant dans des tables :  $LIC_R = D_3(k)\bar{R}$ ,  $LsC_R = D_4(k)\bar{R}$ .

$k$	2	3	4	5	6	7	8	9	10
$D_3$	0	0	0	0	0	0.076	0.136	0.184	0.223
$D_4$	3.267	2.575	2.282	2.115	2.004	1.924	1.864	1.816	1.777

Ces coefficients viennent d'une estimation d'intervalles de confiance déduits de la loi de  $\bar{R}$ , moyenne des étendues d'un échantillon gaussien.

## 2.7 Diagnostics

On recherche les points à l'extérieur des limites, les points particuliers. On appelle "causes communes" la variabilité "normale" au sens gaussien, qui donnent des points distribués selon la loi normale. Ainsi, "normalement", on devrait avoir dans la première zone,  $[\bar{X} - \hat{\sigma}_{\bar{X}}, \bar{X} + \hat{\sigma}_{\bar{X}}]$  68,26 pour cent d'observations ; dans chacune des deux zones suivantes : 13.6 pour cent ; dans chacune des deux zones les plus proches des limites : 2.145 pour cent.

On appelle "causes spéciales" la variabilité exogène, due éventuellement à un dysfonctionnement de la machine, ces points hors limite sont une alerte qui permet de conduire à un réglage.

Il existe, en plus des points hors limites, une série de tests automatiques (7 tests) pas forcément tous fondés statistiquement. Par exemple tendance, périodicité, points alignés horizontalement..... mais il vaut mieux privilégier l'aspect statistique.

## 2.8 Ouverture sur une meilleure estimation de $\sigma$ , pour $k \geq 10$

Pour chaque échantillon  $s_i^2 = \frac{1}{k-1} \sum_{j=1}^k (X_i^j - \bar{X}_i)^2$ . On propose ici l'estimateur sans biais et uniformément plus puissant de  $\sigma^2$  :

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n s_i^2 = \frac{1}{n(k-1)} \sum_{i=1}^n \sum_{j=1}^k (X_i^j - \bar{X}_i)^2.$$

Notons que  $\bar{X}_i = \frac{1}{k} \sum_{j=1}^k X_i^j$  est de variance  $\frac{1}{k} \sigma^2$ , et on peut estimer la variance  $\sigma_{\bar{X}}^2$  par  $\hat{\sigma}_{\bar{X}}^2 = \frac{1}{k} \frac{\sum_{i=1}^n s_i^2}{n}$ . Avec cet estimateur, **pour les cartes  $s$**  on peut éviter  $A_3(k)$  et les limites de contrôle pour la suite des  $\bar{X}_i$  sont alors  $\bar{X} \pm l \sqrt{\frac{\sum_{i=1}^n s_i^2}{nk}}$ ,  $l = 1, 2, 3$ .

Par ailleurs, les  $s_i^2$  sont des variables aléatoires indépendantes de même loi et  $\hat{\sigma}_{\bar{X}}^2$  converge presque sûrement et dans  $L^2$  vers  $\sigma^2/k$ . C'est un estimateur "consistant", de maximum de vraisemblance et uniformément plus puissant.

L'utilisation de cet estimateur permet de s'affranchir également des coefficients  $B_4$ ,  $B_3$  pour les cartes  $s$ . En effet, la loi de  $\frac{n(k-1)\hat{\sigma}^2}{\sigma^2}$  est  $(\chi_{n(k-1)}^2)$  ce qui permet des intervalles de confiance. On cherche  $a$  et  $b$  tels que  $\mathbb{P}(a \leq \frac{n(k-1)\hat{\sigma}^2}{\sigma^2} \leq b) = \Phi(3) - \Phi(-3) = 0.9973$  ( $\Phi$  fonction de répartition de la gaussienne standard), soit

$$\sigma^2 \in \left[ \frac{n(k-1)\hat{\sigma}^2}{b}, \frac{n(k-1)\hat{\sigma}^2}{a} \right] = \left[ \frac{k-1}{b} \sum_i s_i^2, \frac{k-1}{a} \sum_i s_i^2 \right]$$

avec une très grande probabilité. Ces limites doivent être calculées sur des séries précédentes pour l'estimation  $\hat{\sigma}^2$ . Les  $s_i^2$  surveillés ensuite doivent se trouver dans les bandes de confiance évaluées sur les observations précédentes :

$$LIC_{s^2} = \frac{n(k-1)}{b} \hat{\sigma}^2 ; LSC_{s^2} = \frac{n(k-1)}{a} \hat{\sigma}^2$$

avec le  $\hat{\sigma}^2$  estimé précédemment sur  $n$   $k$ -échantillons, et  $a$  et  $b$  tels que respectivement

$$\mathbb{P}\{\chi_{n(k-1)} \leq a\} = \Phi(-j), \quad \mathbb{P}\{\chi_{n(k-1)} \leq b\} = \Phi(j), \quad j = 1, 2, 3.$$

*Exercice : Comparer l'exigence des deux types de limites*

- pour les cartes  $\bar{X}$  selon que l'on utilise  $\bar{s}$  ou  $\sqrt{\bar{s}^2}$

- cartes  $s^2$  contre cartes  $s$ .

**Preuve :** On remarquera que  $\bar{s} = \frac{1}{n} \sum_i s_i < \sqrt{\frac{1}{n} \sum_i s_i^2}$ , ce qui veut dire qu'a priori l'estimateur sous-estime  $\sigma$ . Reste à vérifier que les coefficients  $A_3$  et  $B_4$  compensent suffisamment : on peut calculer l'espérance de  $\bar{s}$  car

$$E[\bar{s}] = E[s_i] = E\sqrt{\frac{\sum_{i=1}^{k-1} (x_j^i - \bar{x}_i)^2}{k-1}}$$

et  $\sigma^2 \sum_{i=1}^{k-1} (x_j^i - \bar{x}_i)^2$  suit une loi de chi-deux à  $k-1$  degrés de liberté :

$$E[\bar{s}] = \frac{\sigma}{\sqrt{k-1}} \int_0^\infty \sqrt{x} \frac{x^{(k-3)/2} e^{-x/2}}{2^{(k-1)/2} \Gamma((k-1)/2)} dx$$

et par conséquent pour obtenir l'estimateur sans biais  $\frac{\bar{s}}{c_4(k)}$  on pose

$$c_4(k) = \sqrt{\frac{2}{k-1} \frac{\Gamma(k/2)}{\Gamma((k-1)/2)}}$$

qui est un nombre croissant vers 1.

Calculons sur un exemple les deux estimations de  $\sigma$  :

-  $\bar{s} = 0.2686, c_4(10) = 0.9727, \hat{\sigma}_1 = 0.27614,$

-  $\frac{\sum_{i=1}^{25} s_i^2}{25} = 0.06687, \hat{\sigma}_2 = 0.2586.$

Le premier estimateur généralisé utilisé sur-estime l'écart-type et de ce fait, sous-estime la capabilité, et donne des limites de contrôle trop larges.

## 2.9 Carte EWMA

Exponentially Weighted Moving Average, c'est à dire moyenne mobile à pondération exponentielle. Ce type de cartes commencent à être utilisées dans les entreprises car beaucoup plus performantes pour déceler de faibles dérives. De plus, elles sont faciles à interpréter. Leur intérêt est qu'elles tiennent compte du passé, la  $i$ ème valeur  $M_i$  est calculée par récurrence

$$M_i = \lambda \bar{x}_i + (1 - \lambda)M_{i-1} \quad \lambda \in ]0, 1[,$$

avec comme valeur initiale  $M_0 = c_0$ , la valeur "cible". Pour calculer les limites, on doit calculer la variance des variables  $M_i$  qui dépendent de  $i$ . Admettons que tous les échantillons de taille  $k$  sont de variance  $\sigma^2$ , alors la variance de  $\bar{x}_i$  est comme précédemment  $\frac{\sigma^2}{k}$ . De plus, le  $i$ ème échantillon est indépendant du passé,  $\bar{x}_i$  et  $M_{i-1}$  sont indépendantes :

$$Var(M_i) = \lambda^2 \frac{\sigma^2}{k} + (1 - \lambda)^2 Var(M_{i-1}).$$

Puisque  $M_0$  est une constante, sa variance est nulle, la résolution de la récurrence donne:

$$Var(M_i) = \lambda^2 \frac{\sigma^2}{k} \frac{1 - (1 - \lambda)^{2i}}{1 - (1 - \lambda)^2}$$

qui se simplifie en

$$Var(M_i) = \frac{\lambda \sigma^2}{k} \frac{1 - (1 - \lambda)^{2i}}{2 - \lambda}$$

dont la racine est une suite croissante vers  $\frac{\lambda \sigma}{\sqrt{k} 2 - \lambda}$ . Les limites de contrôle sont donc pour chaque point  $M_i$

$$\left[ c_0 - 3\sigma \sqrt{\frac{\lambda}{k} \frac{1 - (1 - \lambda)^{2i}}{2 - \lambda}}, c_0 + 3\sigma \sqrt{\frac{\lambda}{k} \frac{1 - (1 - \lambda)^{2i}}{2 - \lambda}} \right].$$

Bien sûr, il faut estimer  $\sigma$  à l'aide des données dont on dispose.

*Exercice* : sur les données de l'exercice 1 feuille 2, regarder si une telle carte ne permet pas de détecter plus tôt l'anomalie, c'est à dire avant les deux points hors contrôle décelés par la carte Shewart.

## 2.10 Carte CUSUM

C'est à dire CUmulative SUMs : sommes cumulées. De fait, ce qui est cumulé ce sont les écarts à la cible  $c_0$ . Il en existe de deux types "type tableau" et "avec masque en V", on ne présente ici que le type tableau. A chaque échantillon on fait correspondre par récurrence deux valeurs:

$$\begin{aligned} (1) \quad S_{H_i} &= \sup(0, \bar{X}_i - c_0 - 0.5\sigma_{\bar{X}} + S_{H_{i-1}}, \\ (2) \quad S_{L_i} &= \sup(0, c_0 - \bar{X}_i - 0.5\sigma_{\bar{X}} + S_{L_{i-1}}, \end{aligned}$$

avec les deux conditions initiales nulles. Les limites de contrôles proposées dans ce cas sont bien sûr centrées autour de 0, (il s'agit d'écarts):

$$-h\sigma_{\bar{X}}, +h\sigma_{\bar{X}}, \quad h = 4 \text{ ou } 5.$$

## 3 Plans d'échantillonnage simples

cf. [3] chapitre 1.

### 3.1 Introduction

En plus de la maîtrise statistique des procédés et des contrôles de qualité, un autre point de vue est de **vérifier la conformité** des produits aux normes (AFNOR par exemple). On ne va pas vérifier tous les produits (ce qu'on appelle un contrôle exhaustif) pour plusieurs raisons :

- le contrôle peut être destructif (endommager le produit),
- cela coûte cher,
- c'est physiquement impossible,
- finalement, ce n'est pas forcément efficace.

Bien sûr, cela ne vaut pas pour les industries de luxe (séries très limitées) ou celles faisant intervenir des problèmes de sécurité. Sinon, on fait un tirage aléatoire : "contrôle par prélèvement". Il y a deux types de contrôles :

- produit conforme/non conforme (contrôle par "attribut"),
- décompte de défauts ou nombre de "non-conformités".

On n'aborde dans ce chapitre que le premier type de contrôle.

Il y a deux étapes du contrôle :

- avant de vendre (produit fini),
- avant d'acheter (matières premières).

### 3.2 Définition, modèle probabiliste

**Définition 3.1** *Un plan de contrôle est une modalité de prélèvement qui permet de définir une règle de décision pour accepter ou rejeter un lot.*

*Il est déterminé par un couple  $(n, c)$  où  $n$  est la taille du prélèvement et  $c$  est le seuil tel que si le nombre de non conformes  $X \leq c$  le lot est accepté, sinon, il est rejeté.*

Un plan est d'autant plus satisfaisant qu'il permet de mieux discriminer les lots de bonne qualité des lots de mauvaise qualité, c'est à dire lorsque la qualité est bonne (faible pourcentage de pièces défectueuses), la probabilité d'acceptation doit être élevée.

Ceci est le plan simple ; il y a aussi le plan double, le plan multiple, le plan progressif.

Soit un lot de  $N$  pièces, on en prélève  $n$  au hasard. On suppose  $p$  la proportion de non conformes (la "qualité" du lot),  $p \in [0, 1]$ , et  $n$  très petit devant  $N$ . Cette hypothèse permet d'assurer que le tirage sans remise n'affecte pas la proportion  $p$ .



Exemple :  $N = 400$  fiches électriques. On propose le plan  $(50, 1)$ , on tire 50 fiches au hasard, on accepte le lot si  $X \leq 1$ , on le rejette sinon.

*Exercice* : (i) Donner la loi de  $X$ , et la loi approchée puisque  $n$  est petit devant  $N$  (plus précisément  $N > 100$ ,  $N > 10n$ .)

(ii) Donner une autre approximation si  $n$  tend vers l'infini,  $p$  tend vers 0 et  $np$  tend vers  $\lambda$  constante strictement positive.

(iii) L'approximation gaussienne  $(np, np(1-p))$  avec  $np, n(1-p) \geq 10$  de fait ne convient pas.

### 3.3 Courbe d'efficacité

Que se passe-t-il puisqu'en fait on ne connaît pas  $p$  ? On étudie la fonction  $P_a : p \mapsto \mathbb{P}_p\{X \leq c\}$ . Le graphe de cette fonction s'appelle la "courbe d'efficacité" à tracer à  $n$  et  $c$  fixés. Selon l'approximation choisie, on obtient l'une des deux fonctions :

$$P_a : p \mapsto \sum_{k=0}^c C_n^k p^k (1-p)^{n-k} ; \quad \sum_{k=0}^c e^{-np} \frac{(np)^k}{k!}.$$

Le paramètre  $p$  s'appelle la "qualité effective" du lot.

**Proposition 3.2** La fonction  $P_a$  vaut 1 en 0, est nulle en 1, est décroissante pour tout  $(n, c)$ .

**Preuve** en exercice, exo 4 feuille 3 ; indication : calculer la dérivée  $P'_a(p)$  dans les deux cas d'approximation.

$$P'_a(p) = -nC_{n-1}^c p^c (1-p)^{n-c-1} \text{ ou } -ne^{-np} \frac{(np)^c}{c!}.$$

Ces courbes s'interprètent de la façon suivante : le tracé de la courbe indique la probabilité d'acceptation du lot en fonction du "vrai"  $p$  inconnu. Cela veut dire que sur un grand nombre de lots de qualité  $p$ , on va en accepter une proportion  $P_a(p)$ , et une proportion  $1 - P_a(p)$  sera rejetée. Pour mieux séparer les deux décisions, on aimerait

$P_a(p)$  "grand" quand  $p$  est "petit" ;  $P_a(p)$  "petit" quand  $p$  est "grand".

**Proposition 3.3** *Soit  $(p, n)$  fixés, alors si  $c$  augmente,  $P_a(p)$  augmente aussi, le contrôle est moins sélectif.*

**Preuve** en exercice.

**Proposition 3.4** *Soit  $(p, c)$  fixés, alors si  $n$  augmente,  $P'_a(p)$  est de plus en plus négative et  $P_a$  décroît plus rapidement.*

*Plus précisément, au point d'inflexion  $c/(n-1)$  ou  $c/n$ , qui se rapproche de 0 quand  $n$  augmente, la pente vaut  $-nC_{n-1}^c (\frac{c}{n-1})^c (1 - \frac{c}{n-1})^{n-c-1}$  ou  $-ne^{-c\frac{c}{c!}}$  est de plus en plus négative et la discrimination est meilleure.*

**Preuve** en exercice.

En conclusion, en augmentant  $n$  on obtient une meilleure discrimination entre les deux décisions.

### 3.4 Niveaux de qualité acceptable, toléré, NQA, NQL

**Définition 3.5** *NQA, "niveau de qualité acceptable", est le sup des  $p$  "acceptables" pour un risque fournisseur  $\alpha$  donné :*

$$\forall p \leq NQA, \mathbb{P}\{X > c/p\} \leq \alpha.$$

**Proposition 3.6**

$$\sup_{p \leq NQA} [1 - P_a(p)] = 1 - P_a(NQA).$$

La condition à vérifier par le fournisseur sur le plan d'échantillonnage est que  $(n, c)$  doit être tel que

$$1 - P_a(NQA) \leq \alpha.$$

Remarquons que l'on peut ne pas avoir l'égalité car la variable aléatoire  $X$  est à valeurs entières.

**Preuve** : puisque  $P_a$  est décroissante,  $1 - P_a$  est croissante et le sup est bien atteint en NQA. •

Interprétation : c'est le risque de refuser un lot alors qu'il est de qualité acceptable (cf. le risque de première espèce pour le test de  $p$  petit contre  $p$  grand).

**Définition 3.7**  $NQL$ , "niveau de qualité limite", est l'inf des  $p$  "tolérés" pour un risque client  $\beta$  :

$$\forall p \geq NQL, \mathbb{P}\{X \leq c/p\} \leq \beta.$$

**Proposition 3.8**

$$\sup_{p \geq NQL} [P_a(p)] = P_a(NQL).$$

La condition à vérifier par le client sur le plan d'échantillonnage est que  $(n, c)$  doit être tel que

$$P_a(NQL) \leq \beta.$$

Remarquons que, ici non plus, on peut ne pas avoir l'égalité car la variable aléatoire  $X$  est à valeurs entières.

**Preuve** : puisque  $P_a$  est décroissante l'inf est bien atteint en NQL. •

Interprétation : c'est le risque pour le client d'accepter un lot alors qu'il est de qualité médiocre (cf. le risque de seconde espèce).

L'objectif est d'établir un "bon" plan d'échantillonnage  $(n, c)$ , les deux couples  $(\alpha, NQA)$  et  $(\beta, NQL)$  étant fixés selon un accord entre le fournisseur et le client. Ceci se ferait facilement avec l'approximation gaussienne (mais on a vu que celle-ci est illusoire !) Pour les lois binomiales et de Poisson, il existe les tables de J.M. CAMERON (jointes).

Graphiquement, considérons par exemple la courbe d'efficacité du plan  $(n, c)$ ,  $P_a(p) = (1 - p)^n + np(1 - p)^{n-1} + \dots$ , il s'agit de trouver  $n$  et  $c$  tels que cette courbe passe par les points  $(NQA, \alpha)$  et  $(NQL, \beta)$ .

$$1 - \alpha \leq \sum_{k=0}^c e^{-n.NQA} \frac{(n.NQA)^k}{k!} ; \beta \geq \sum_{k=0}^c e^{-n.NQL} \frac{(n.NQL)^k}{k!}.$$

Ce problème (système de deux inéquations à deux inconnues) n'est pas résoluble analytiquement, d'où l'usage de tables.

#### Liste des abréviations et sigles pour les plans d'échantillonnage.

$\alpha$ =risque fournisseur

$\beta$ =risque client

$c$ =critère d'acceptation (si  $X \leq c$  on accepte le lot),  $N$  taille du lot,  $n$  taille de l'échantillon

$(n, c)$ =plan d'échantillonnage simple.

NQA= niveau de qualité acceptable (AQL en anglais),

NQL= niveau de qualité toléré (LTPD en anglais).

### 3.5

Remarquons que la prise de décision ne dépend pas de la taille du lot. Néanmoins, on introduit la notion de **criticité**. Et selon la criticité requise (niveau de contrôle plus ou moins sévère) et selon la taille des lots, des lettres codes sont proposées qui recommande la taille de l'échantillonnage eu égard aux NQA et NQL fixés (cf. ci-dessous 3.5.1 et tableaux joints).

**Définition 3.9** *Non conformité critique : non conformité qui, d'après le jugement et l'expérience, est susceptible de conduire à un manque de sécurité ou à des risques d'accidents pour les utilisateurs, le personnel d'entretien, ou ceux qui dépendent du produit en question, ou bien une non conformité qui, d'après le jugement et l'expérience, pourrait empêcher l'accomplissement de la fonction du produit. NQA = 0.01.*

*Non conformité majeure : non conformité qui, sans être critique, risque de provoquer une défaillance ou de réduire de façon importante la possibilité d'utilisation du produit pour le but qui lui est assigné. NQA = 0.025.*

*Non conformité mineure : non conformité qui ne réduira vraisemblablement pas beaucoup la possibilité d'utilisation du produit pour le but qui lui est assigné ou qui traduit, par rapport aux normes établies, une divergence n'entraînant pas de conséquences appréciables sur l'utilisation ou le fonctionnement efficace du produit ; ces non conformités mineures sont souvent associées à l'apparence du produit. NQA = 0.04.*

### 3.5.1 Exemples de plans d'échantillonnage proposés pour la spécification selon la taille du lot

$N$	$n$	$c$
$\leq 2500$	4	0
(2500, 10000)	5	0
(10000, 50000)	10	0
(50000, 100000)	15	0
$\geq 100000$	27	1

Dans les quatre premiers cas, la fonction d'acceptation est  $p \mapsto (1 - p)^n$ ,  $F'_a(p) = n(1 - p)^{n-1}$ . Tracer le graphique, donner risque fournisseur et risque client pour  $NQA = 0.01$ ,  $NQL = 0.1$ .

## 4 Plans d'échantillonnage doubles

### 4.1 Introduction, cf. [3] pages 43-86

Le plan d'échantillonnage simple a l'inconvénient d'exiger le plus souvent une taille d'échantillon assez élevé. Dans le cas d'un très bon lot ( $X \leq c$ ), il faut examiner les  $n$  pièces avant de conclure à l'acceptation. L'échantillonnage double permet de prendre une décision plus rapidement dans les deux cas suivants :

- lot très bon (procédé maîtrisé),
- lot très mauvais (procédé non maîtrisé).

Evidemment, l'effectif  $n_1 + n_2$  des deux échantillons est plus élevé que l'effectif  $n$  du plan simple, mais on peut voir que, en moyenne, on a moins d'objets à examiner (sauf dans les cas "moyens", entre très bons et très mauvais).

**Définition 4.1** *Un plan d'échantillonnage double est la donnée de deux plans  $(n_1, c_1)$  pour le premier prélèvement et  $(n_2, c_2)$  pour le second.*

La méthode consiste à prélever un  $n_1$ -échantillon, on note  $X_1$  le nombre d'objets défectueux,

- si  $X_1 \leq c_1$ , le lot est accepté,
- si  $X_1 > c_2$ , le lot est rejeté,
- si  $c_1 < X_1 \leq c_2$ , on prélève un  $n_2$ -échantillon dont  $X_2$  pièces sont non conformes ;
  - si  $X_1 + X_2 \leq c_2$ , le lot est accepté,
  - si  $X_1 + X_2 > c_2$ , le lot est rejeté.

Exemple : plan simple  $P_1 = (82, 2)$ , plan double  $P_2 = ((50, 0), (50, 3))$ .

$X_1$	0	1	2	3	$\geq 4$
$P_1$	A	A	A	R	R
$P_2$	A	2ème prél.		R	R
$X_2 = 2$	$\emptyset$	A	R	R	$\emptyset$
$X_2 = 1$	$\emptyset$	A	A	R	$\emptyset$
$X_2 = 0$	$\emptyset$	A	A	A	$\emptyset$

### 4.2 Probabilité d'acceptation en fonction de la qualité effective

On suppose ici légitime d'approcher la loi binomiale  $B(n_i, p)$  par la loi de Poisson  $P(n_i p)$ ,  $i = 1, 2$ . L'évènement  $A$ , accepter le lot, est

$$A = \{X_1 \leq c_1\} \cup \{X_1 > c_1, X_1 + X_2 \leq c_2\},$$

de probabilité

$$P_a(p) = P_a^1(p) + P_a^2(p) = \mathbb{P}\{X_1 \leq c_1\} + \sum_{k=c_1+1}^{c_2} \mathbb{P}\{X_1 = k\} \mathbb{P}\{X_2 = c_2 - k\} =$$

$$= \sum_{k=0}^{c_1} e^{-n_1 p} \frac{(n_1 p)^k}{k!} + \sum_{k=c_1+1}^{c_2} e^{-n_1 p} \frac{(n_1 p)^k}{k!} \sum_{j=0}^{c_2-k} e^{-n_2 p} \frac{(n_2 p)^j}{j!}.$$

On utilise des tables de loi de Poisson (ou logiciel stat) pour tracer le graphe de cette application, la courbe d'efficacité,  $P_a : p \mapsto P_a(p)$ .

### 4.3 Courbe d'efficacité d'un plan double

C'est toujours le graphe de la fonction  $P_a$ . L'étude théorique est terriblement technique. On remarque quand même que comme pour le plan simple,  $P_a(0) = 1, P_a(1) = 0$ . La fonction  $P_a^1$  est décroissante et la fonction  $P_a^2$ , elle aussi décroissante, peut s'écrire (cas  $n_1 = n_2 = n$ ) :

$$P_a^2(p) = e^{-2np} \sum_{k=c_1+1}^{c_2} \frac{(np)^k}{k!} \sum_{j=0}^{c_2-k} \frac{k!}{j!(k-j)!}.$$

*Exercice 1, feuille 4* : poser  $x = np$  et dériver plutôt en  $x$ .

### 4.4 Qualité moyenne après contrôle, AOQ

**Définition 4.2** *AOQ (en anglais Average Outgoing Quality) est l'espérance de la variable aléatoire égale à la proportion de pièces défectueuses dans les lots acceptés lorsque, après contrôle de l'échantillon, on a réparé ou remplacé les pièces non conformes. (On dit que l'on a effectué une inspection rectifiante).*

*Exercice 2, feuille 4, indication* : on note  $Y$  la variable aléatoire égale au nombre de défectueux restant après cette inspection ; on note  $A$  l'évènement "lot accepté",  $X$  la variable aléatoire égale au nombre de défectueux restant dans un lot accepté et calculer l'espérance de  $Y$  en remarquant que  $Y = X \cdot 1_A$ . AOQ est l'espérance de  $Y/N$ .

### 4.5 Nombre moyen de pièces prélevées, ASN

ASN= average sample number, en anglais.

Il est clair que lors d'un plan simple  $(n, c)$  on prélève  $n$  pièces dans le lot. En revanche, dans

le cas d'un plan double  $(n_1, c_1), (n_2, c_2)$  ce nombre de pièces prélevées devient aléatoire, c'est alors une variable aléatoire  $Y$  dont on peut calculer l'espérance, d'où ce nombre moyen annoncé :

on prélève a priori au moins  $n_1$  pièces,

ensuite on en prélève  $n_2$  sur l'évènement  $\{X_1 \in [c_1 + 1, c_2]\}$ .

Soit le nombre de pièces prélevées :  $Y = n_1 + n_2 \mathbf{1}_{\{X_1 \in [c_1 + 1, c_2]\}}$ ,  
et  $ASN = n_1 + n_2 \mathbb{P}_p\{X_1 \in [c_1 + 1, c_2]\}$ .

*Exercice 3, feuille 4.*

Comme annoncé, le plan double est moins intéressant pour les qualités "moyennes".

## 4.6 Quantité moyenne contrôlée à long terme

Il s'agit d'évaluer la charge de travail possible dans le cas d'un plan double pour le contrôle de lots d'effectif  $N$  lorsque les lots refusés sont contrôlés à 100 pour cent.

$$I(p) = n_1 P_a^1(p) + (n_1 + n_2) P_a^2(p) + N(1 - P_a(p))$$

où  $P_a(p) = P_a^1(p) + P_a^2(p)$ .

*Exercice : justifier ce calcul et montrer que l'on a également*

$$I(p) = n_1 P_a(p) + n_2 P_a^2(p) + N(1 - P_a(p)) = n_1 + n_2(1 - P_a^1(p)) + (N - n_1 - n_2)(1 - P_a(p)).$$

•



## 5 Cartes multi-critères

### 5.1 Système multigénérateur

Il s'agit d'un système qui produit plusieurs pièces élémentaires différentes, soit  $l$  types différents.

**Hypothèse** : on suppose que chaque mesure de produit élémentaire suit une loi de Gauss.

Alors, la répartition cumulée de ces différentes lois gaussiennes n'est plus nécessairement unimodale.

Questions à se poser :

- . quelle est la taille optimale du prélèvement ?
- . faut-il prélever systématiquement une pièce de chaque type ?
- . comment adapter les cartes classiques à ce nouveau problème ?
- . comment améliorer l'efficacité du contrôle tout en diminuant la taille du prélèvement ?

Mode de prélèvement : il vaut mieux prélever une pièce de chaque type et étudier alors la moyenne  $\bar{X} = \frac{1}{l} \sum_{i=1}^l X_i$  qui suit une loi de Gauss.

Calcul des limites de contrôle : on observe une trentaine d'échantillons sur laquelle on estime la variance de  $\bar{X}$  notée  $\sigma(\bar{X})$ , la cible  $c_0$  est la moyenne des cibles, les limites sont alors  $c_0 \pm 3\hat{\sigma}_{\bar{X}}$ , où l'estimation  $\hat{\sigma}_{\bar{X}}$  est faite comme dans le chapitre 2 mais sur des échantillons de moyennes des  $l$  mesures.

**Attention !** si  $l$  est trop grand, le prélèvement systématique devient trop onéreux, on prélève alors seulement un sous-ensemble des  $l$  types, mais toujours les mêmes.

### 5.2 Cartes à caractéristiques multiples, un exemple

Le problème est de suivre la production d'un type de pièces ayant **deux** caractéristiques (ou plus), par exemple le diamètre et la circularité, c'est à dire que le rayon est constant tout autour de la circonférence.

Le protocole conseillé dans un tel cas est de relever pour chaque  $k$ -échantillon outre le diamètre, l'écart entre le max et le min du diamètre quand on fait tourner la pièce autour de son axe ; on note  $E$  la moyenne de ces écarts. Les limites de contrôle proposées sont alors  $c_0 \pm (3\hat{\sigma}_{\bar{X}} + \frac{1}{2}E)$  :

$\pm 3\hat{\sigma}_{\bar{X}}$  donne un intervalle de confiance pour le diamètre,

$\pm \frac{1}{2}E$  en donne un pour la circularité.

### 5.3 Cartes multicritères

La variable à contrôler est ici à valeurs dans  $\mathbb{R}^p$ ,  $X = (X_j, j = 1, \dots, p)$ . On prélève un  $k$ -échantillon qui donne une moyenne  $\bar{x}$  à valeurs dans  $\mathbb{R}^p$ ,  $\bar{x}_j = \frac{1}{k} \sum_{i=1}^k x_j^i$ ,  $j = 1, \dots, p$ , et

la covariance observée est pour l'échantillon  $e$   $S_{j,l}^2(e) = \frac{1}{k-1} \sum_{i=1}^k (x_j^i - \bar{x}_j)(x_l^i - \bar{x}_l)$ ,  $j$  et  $l$  varient de 1 à  $p$ .

On prélève  $n$   $k$ -échantillons, d'où le vecteur  $\bar{X} = \frac{1}{n} \sum_{e=1}^n \bar{x}_e$  dans  $\mathbb{R}^p$  de matrice de variance estimée par  $\frac{1}{k} S^2$  où

$$S_{jl}^2 = \frac{1}{n} \sum_{e=1}^n S_{jl}^2(e).$$

Alors, si l'on suppose que le vecteur  $X$  dont on prélève des échantillons suit une loi normale  $(m, \Sigma)$ ,  $m \in \mathbb{R}^p$  et  $\Sigma$  matrice de variance, le théorème de Cochran dit que

$$n(k-1)S_{jl}^2 = \sum_{e=1}^n \sum_{i=1}^k (x_j^i - \bar{x}_j)(x_l^i - \bar{x}_l),$$

suit une loi de khi-deux à  $n(k-1)$  degrés de liberté fois  $\Sigma_{jl}$ .

Le problème est donc le suivant : étant donné un vecteur gaussien de moyenne  $m \in \mathbb{R}^p$  et de matrice de variance  $\Sigma$ , au lieu d'un intervalle de confiance pour chaque caractéristique, trouver une région de confiance pour le paramètre  $m$  dans  $\mathbb{R}^p$  qui tienne compte des corrélations. De fait, on construit un *ellipsoïde de confiance* fondé sur la statistique ( $T^2$  généralisé)

$$T^2 = nk(\bar{X} - m)' S^{-2} (\bar{X} - m).$$

Si  $N = nk$  (nombre d'observations) est très grand, la loi limite de  $T^2$  est celle d'un khi-deux à  $p$  degrés de liberté. De fait, on a le résultat plus précis (cf. [1] chapitre 5, corollaire 5.2.1) :

**Proposition 5.1** *Soit un  $N$  échantillon d'un vecteur gaussien de moyenne  $m \in \mathbb{R}^p$  et de matrice de variance  $\Sigma$ ,  $(X^i, i = 1, \dots, N)$ , alors si l'on pose le vecteur  $\bar{X} = \sum_i X^i / N$  et  $S^2$  la matrice de terme général*

$$S_{jl}^2 = \frac{1}{N-1} \sum_{i=1}^N (X_j^i - \bar{X}_j)(X_l^i - \bar{X}_l), \quad j, l = 1, \dots, p,$$

la loi de

$$\frac{N-p}{(N-1)p} N(\bar{X} - m)' S^{-2} (\bar{X} - m)$$

est une loi de Fisher-Snédecor à  $(p, N-p)$  degrés de liberté.

### 5.3.1 Cas $p = 1$ , $n$ échantillons de taille $k$

On dispose de  $n$  moyennes  $\overline{x(i)}, i = 1, \dots, n$ , la variance estimée est

$$\hat{\sigma}^2(i) = \frac{1}{k-1} \sum_{j=1}^k (x_j^i - \overline{x(i)})^2,$$

on estime  $\sigma^2$  par

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \hat{\sigma}^2(i)$$

qui converge presque sûrement vers  $\sigma^2$  pour  $n$  infini (loi des grands nombres).

Notons que  $\frac{n(k-1)\hat{\sigma}^2}{\sigma^2} = \sum_{i=1}^n \sum_{j=1}^k (x_j^i - \bar{x}(i))^2 \sigma^{-2}$  suit une loi de khi-deux à  $n(k-1)$  degrés de liberté.

Cette statistique est indépendante de la moyenne  $\bar{X}$  et  $nk \frac{(\bar{X}-m)^2}{\sigma^2}$  suit une loi de khi-deux à 1 degré de liberté, le quotient

$$\frac{k}{k-1} T^2 = \frac{k(\bar{X} - m)^2}{(k-1)\hat{\sigma}^2}$$

suit donc une loi de Fisher  $(1, n(k-1))$ . On a alors un intervalle de confiance pour le paramètre  $m$  : soit  $q_{1-\alpha}$  le quantile de la loi  $F(1, n(k-1))$ , alors

$$\mathbb{P}\left\{\frac{k(\bar{X} - m)^2}{(k-1)\hat{\sigma}^2} \leq q_{1-\alpha}\right\} = 1 - \alpha$$

dont on déduit l'intervalle de confiance pour le paramètre  $m$  :

$$[\bar{X} - \sqrt{q_{1-\alpha}(k-1)\hat{\sigma}^2/k}, \bar{X} + \sqrt{q_{1-\alpha}(k-1)\hat{\sigma}^2/k}].$$

Les limites de contrôle sont alors déterminées par les quantiles correspondant à  $1 - \alpha = \mathbb{P}\{|Y| \leq 3\} = 0.0027$  avec  $Y$  gaussienne centrée réduite.

### 5.3.2 Cas $p > 1, k = 1$

On dispose de  $N$  réalisations de la gaussienne,  $(X_i, i = 1, \dots, N)$ . Là encore on a deux statistique indépendantes : la moyenne  $\bar{X}$  qui est un vecteur, et la matrice de variance estimée  $S^2$ . On utilise la statistique  $T^2$  sous la forme

$$T^2 = N(\bar{X} - m)' S^{-2} (\bar{X} - m).$$

Puisque ici  $k = 1$ ,  $S^2$  est la matrice  $\frac{1}{N-1} \sum_i (X_i - \bar{X})(X_i - \bar{X})'$ , alors en application de la proposition 5.1, la loi de

$$\frac{N-p}{(N-1)p} N(\bar{X} - m)' S^{-2} (\bar{X} - m)$$

est une loi de Fisher à  $(p, N - p)$  degrés de liberté.

La preuve de la proposition 5.1 se trouve dans T.W. ANDERSON, An Introduction to Multivariate Analysis, elle est longue et difficile.

*Exercice:* Proposer un ellipsoïde de confiance pour le paramètre  $m \in \mathbb{R}^p$  en s'inspirant de l'intervalle de confiance donné dans le cas  $p = 1$ .

$$\{m \in \mathbb{R}^p : (\bar{X} - m)' S^{-2} (\bar{X} - m) \leq q_{1-\alpha} \frac{(N-1)p}{(N-p)N}\}$$

où  $q_{1-\alpha}$  est le  $1 - \alpha$ -quantile de la loi de Fisher. Pour courrir le même risque que dans le cas classique, là aussi on choisit  $1 - \alpha = 0.0027$ .

## 6 Cas des petites séries

(cf. [5] pages 287-338).

C'est par exemple la situation de démarrage d'un procédé : on jette les premières pièces fabriquées, mais on en profite pour faire les premières estimations et limites provisoires de contrôle.

### 6.1 Introduction

De plus en plus, il y a production de séries courtes (problème de stocks, production "à flux tendu", diversification de la production, etc.) Dans ce cas, l'hypothèse gaussienne tombe. Mais on peut néanmoins essayer de tirer le maximum de la pratique de la maîtrise statistique des procédés. Voici les objections formulées par les industriels :

- on contrôle à cent pour cent, on règle au fur et à mesure (ceci a posteriori, la MSP est inutile),
- il est trop cher de mettre en place des cartes de contrôle "pour si peu",
- les écarts-types sont dénués de sens.

On va voir qu'il est quand même possible d'intervenir pour améliorer l'efficacité du pilotage (i.e. re-régler quand on pressent que la machine se dérègle) en s'adaptant aux situations les plus diverses telles :

- . produits très chers et peu nombreux,
- . travail sur commande de faibles effectifs (cf. faïencerie de Gien par exemple),
- . production de petits lots avec changement de la production plusieurs fois par jour,
- . forte production mais dont les mesures coûtent cher.

Deux exemples de contrôle à cent pour cent et de pilotage à courte vue :

(i) les limites de tolérance sont  $[-5, +5]$ , la 7ème pièce en sort, l'opérateur règle à  $-6$  ; les mesures suivantes sont  $-2, -3, -4, \dots$  de fait, il a tout dérégulé !

pièces	1	2	3	4	5	6	7	8	9	10
mesures	2	1	4	3	2	2	6	-2	-3	-4

(ii) avec la même machine, on observe :

pièces	1	2	3	4	5	6	7	8	9	10
mesures	1	3	-2	-4	2	1	0	-1	2	3

L'opérateur a réglé à  $-3$  après la 2ème, et  $+4$  après la 4ème.

Dans les deux cas, il semble que le réglage "à vue" en cours de fabrication n'a pas été forcément judicieux. La statistique doit pouvoir permettre d'anticiper et d'intervenir en bonne connaissance de cause. Ici, l'opérateur s'est fondé sur l'observation de la dernière mesure (modèle markovien) alors que **tout** le passé compte. Une telle démarche (contrôle à cent pour cent et pilotage à vue) est coûteuse, inopportune et surtout subjective. Il est important de considérer l'intervalle de confiance plutôt que l'intervalle de tolérance.

## 6.2 Le suivi par valeurs individuelles

(cf. [5] p. 335, section 9.2)

Sur l'exemple ci-dessus on ne s'aperçoit qu'assez tard d'un dérèglement. Soit un exemple avec un produit par jour, le problème est d'estimer l'écart-type. On propose

$$R_i = |x_i - x_{i+1}|.$$

On calcule  $R_i$  à chaque étape et  $\hat{\sigma}$  correspondant (sauf en cas de réglage). On obtient des limites de contrôle à chaque étape  $c_0 \pm 3\hat{\sigma}$  et on a des alertes chaque fois que l'on sort de ces limites, plus restreintes en général que les limites de tolérance. De plus, on contrôle également l'étendue avec une carte  $R$  :

$$[D_3(2)R, D_4(2)R] = [R - 3d_3(2)\frac{R}{d_2(2)}, R + 3d_3(2)\frac{R}{d_2(2)}].$$

mesures	$x_1$	$x_2$	$\dots$	$x_i$
$R_i$	0	$ x_1 - x_2 $	$\dots$	$ x_i - x_{i-1} $
$\hat{\sigma}$	$\emptyset$	$\frac{\sqrt{\pi}R_2}{2}$	$\dots$	$\frac{\sqrt{\pi}R_i}{2}$

et à chaque mesure on vérifie que  $x_i$  est ou non dans les limites  $c_0 \pm 3\hat{\sigma}$ .

*Pillet propose une amélioration avec des cartes EWMA ( $M_i = \lambda\bar{x}_i + (1 - \lambda)M_{i-1}$ ) ou CUSUM, cf. [5] chapitre 6.*

*Exercice: Montrer que  $\hat{\sigma} = \frac{\sqrt{\pi}R}{2}$  est sans biais. En déduire  $d_2(2) = \frac{2}{\sqrt{\pi}} = 1.128$ .*

**Preuve :**  $E(R) = E|X - Y|$  où  $X$  et  $Y$  sont de loi normale  $(m, \sigma^2)$  et indépendantes, donc la loi de  $X - Y$  est normale  $(0, 2\sigma^2)$  et

$$E(R) = 2\sigma\sqrt{2} \int_0^\infty x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \frac{2\sigma}{\sqrt{\pi}}.$$

•

## 6.3 Cartes de contrôle “petites séries”

(cf. [5] p. 353, section 9.4)

Objectif : aide au pilotage des toutes premières pièces ; application de la MSP dans le cas des séries de moins de 20 pièces.

La question est : qu'est ce qui permet de dire que le procédé est mal réglé ? On essaye de faire en sorte que ces cartes ressemblent aux cartes  $(\bar{X}, \bar{R})$  pour ne pas perturber les opérateurs.

Principe de remplissage de ces cartes : on fait les mesures ( $x_i, i \geq 1$ ) et on note à chaque étape :

mesures	$x_1$	$x_2$	$\dots$	$x_i$
$\bar{x}$	$x_1$	$\frac{x_1+x_2}{2}$	$\dots$	$\bar{x}_i$
$R_i$	0	$ x_1 - x_2 $	$\dots$	$R_i$
$\hat{\sigma}$	$\emptyset$	$\frac{\sqrt{\pi}R_2}{2}$	$\dots$	$\frac{R_i}{d_2(i)}$
$\hat{\sigma}_{\bar{X}}$	$\emptyset$	$\frac{\hat{\sigma}}{\sqrt{2}}$	$\dots$	$\frac{\hat{\sigma}}{\sqrt{i}}$

et on vérifie si  $x_i$  est ou non dans les limites  $[c_0 \pm A_2(i)R_i]$  avec  $A_2(i) = \frac{3}{\sqrt{id_2(i)}}$ .

On a de même des limites pour les étendues. :

$$LIC_R = D_5(i)\hat{\sigma}, D_3(i)R_i ; LSC_R = D_6(i)\hat{\sigma}, D_4(i)R_i.$$

vérifier la cohérence de ces divers coefficients, de fait voir que  $D_6/D_4 = D_5/D_3 = d_2 \dots$

Avec ces limites, on peut vérifier sur l'exemple qu'il aurait fallu intervenir dès la 3ème pièce.

mesures	2	1	4	3	2	2	6	-2	-3	-4
$\bar{x}$										
$R$	0									
$\hat{\sigma}$	$\emptyset$									
$\pm 3\hat{\sigma}$	$\emptyset$									

On remarque que la mesure 3 n'est pas dans les limites de contrôle : c'est dès cet instant qu'il aurait fallu intervenir.

Attention : si on fait un réglage, on démarre un nouvel échantillon, donc une nouvelle carte "petite série".

## 6.4 Retrouver un effet de série

cf. Pillet, page 342, section 9.3.

C'est le cas lorsque l'on produit des pièces différentes, cas multiproduit. L'idée est de suivre les écarts à la cible :  $x_i - c_0$ , plutôt que les mesures brutes.

Si a priori, l'écart-type est le même pour tous les types de pièces, on est ramené au cas des grandes séries avec  $X_i - c_0$  sur des cartes  $\bar{X}$ .

Les limites de contrôle deviennent alors

$$\pm A_2(n)\bar{R}_n = \pm 3\hat{\sigma}/\sqrt{n} ; \pm A_3(n)\bar{s}_n = \pm 3\hat{\sigma}/\sqrt{n}$$

et on fait aussi une carte  $R$  (respectivement  $s$ ) pour s'assurer de la stabilité de l'écart-type.

Sinon, si l'on ne peut considérer l'écart-type comme constant, mais si on le connaît pour chaque type  $i$  de pièces, on effectue le changement de variables

$$X_i \mapsto \frac{X_i - cible_i}{\sigma_i}$$

sur des échantillons tous de même taille.

Si la taille des échantillons varie, les limites de contrôle doivent changer sur la carte selon la taille des échantillons.

Par exemple, une entreprise produit des objets de type A, B ou C. On prélève des 3-échantillons, les écarts types sont a priori connus,  $\sigma_A$ ,  $\sigma_B$ ,  $\sigma_C$ . La règle de contrôle est de suivre les quantités

$$\left| \frac{\bar{X}_i - c_i}{\sigma_i} \right| \leq \sqrt{3}, \quad i = A, B, C.$$

Inconvénients : il y a plus de calculs, mais ceci est mineur avec l'aide de l'informatique ; certains opérateurs sont perturbés de devoir inscrire des écarts au lieu des valeurs observées.

## 6.5 Carte de pré-contrôle

(cf. Pillet [5], page 364, section 9.5.)

Trop souvent, le contrôle et le réglage sont effectués sans qu'il soit laissé de trace écrite, les opérateurs trouvant cela trop lourd. Les palliatifs sont les suivants :

- . si la capabilité est faible ( $< 2$ ), on prend comme limites de contrôle  $[c_0 \pm \frac{T_s - T_i}{6}]$  et on contrôle à cent pour cent,
- . lorsque  $2 \leq C_p < 2.5$ ,  $\hat{\sigma} < \frac{T_s - T_i}{12}$ , on fixe les limites indépendamment de  $\hat{\sigma}$  à  $c_0 \pm \frac{T_s - T_i}{6}$ , et on contrôle par échantillonnage,
- . dès que la capabilité augmente ( $C_p \geq 2.5$ ,  $\hat{\sigma}$  diminue) on prend comme limites de contrôle  $[c_0 \pm \frac{T_s - T_i}{4}]$  et on contrôle par échantillonnage.

Donc, il faut calculer  $C_p$  à partir de  $\hat{\sigma}$  et on peut le faire dès que la 2ème pièce est prélevée (cf. [5] page 365 et sq.).

*Exercice* : On rappelle que  $C_p = \frac{T_s - T_i}{6\hat{\sigma}}$ . On suppose que  $C_p = 2$ ,  $LC = c_0 \pm \frac{T_s - T_i}{6}$ .

(i) Calculer dans ce cas la proportion de hors limite en supposant le procédé centré ( $m = c_0$ ), puis la proportion de défectueux.

(ii) Recalculer ces proportions dans le pire des cas c'est à dire lorsque le procédé est tout à fait décalé à droite ou à gauche :  $m = c_0 \pm \frac{T_s - T_i}{6}$ .

**Preuve** :  $C_p = 2$  donc  $T_s - T_i = 12\hat{\sigma}$ .

$$(i) \quad 1 - \mathbb{P}\left\{X \in \left[c_0 \pm \frac{T_s - T_i}{6}\right]\right\} = 2\mathbb{P}\left\{X - c_0 \geq \frac{T_s - T_i}{6} = 2\hat{\sigma}\right\} = 2(1 - \Phi(2)) = 0.0456.$$



$$\mathbb{P}\{|X - c_0| \geq \frac{T_s - T_i}{2}\} =$$

(ii) Dans ce cas  $m = c_0 + 2\hat{\sigma}$ , la limite est  $m$  et  $\mathbb{P}\{X \geq m\} = 0.5$  ce qui est très probable mais la proportion de défectueux est donnée par

$$\mathbb{P}\{X \geq T_s\} = \mathbb{P}\{X - c_0 - 2\hat{\sigma} \geq T_s - c_0 - 2\hat{\sigma} = 4\hat{\sigma}\} = 1 - \Phi(4) = 3 \times 10^{-5}$$

ce qui reste acceptable.

*Exercice : On suppose que  $C_p = 3$ , soit  $\hat{\sigma} = \frac{T_s - T_i}{6C_p} \sim (T_s - T_i)/18$ . Alors les limites de contrôle sont  $c_0 \pm \frac{T_s - T_i}{4} = c_0 \pm 4.5\hat{\sigma}$ .*

(i) Si le procédé est centré, calculer la proportion de hors limite et de non conformes.

(ii) Si le procédé est décentré, avec  $m = c_0 \pm 4.5\hat{\sigma}$ , quelle est alors la proportion hors limite ? de défectueux ?

## 6.6 Utiliser l'échantillonnage

Tagushi caractérise les paramètres non maîtrisables par différents bruits :

- . bruits "intérieurs" (variation du système de production, usure, dérèglement...)
- . bruits "extérieurs" (température, vibrations...)
- . variations des matières premières.

Ces bruits forment les causes communes révélées par la dispersion gaussienne, s'il n'y a pas de causes "spéciales". Deux erreurs sont possibles :

- . ne pas corriger un procédé dérèglé,
- . régler à tort une machine bien réglée.

D'où l'utilisation de l'échantillonnage et de la réduction de variance avec  $\hat{\sigma}_{\bar{X}} = \frac{1}{\sqrt{k}}\hat{\sigma}$  pour un  $k$ -échantillon, qui permettent de mieux isoler les causes spéciales. En effet, on diminue alors la dispersion due aux bruits "normaux". Cela sert en cas de production

unitaire mais effectuée avec la même machine outil. Par exemple, la machine produit  $k$  objets, il y a  $k$  mesures  $(x_1, \dots, x_k)$  ; on note les écarts aux cibles  $e_j = x_j - c_0^j$  et pour chaque prélèvement  $i$  de  $k$  mesures, soit  $k$  écarts, on note la moyenne des écarts  $\bar{e}_i$ , l'étendue  $R_i$ . D'où une estimation  $\hat{\sigma}$  pour la variable "écart", et on peut réaliser une carte  $(\bar{X}, R)$  et l'on aura une meilleure précision grâce au coefficient réducteur  $\frac{1}{\sqrt{k}}$  :  $\hat{\sigma}_{\bar{e}} = \frac{1}{\sqrt{k}}\hat{\sigma}$ .

## 6.7 Analyse a posteriori

Cette analyse est utilisée dans l'industrie aéronautique aux EU : il y a de très nombreuses pièces à réaliser avec beaucoup de mesures à contrôler sur chacune, et dont les tolérances sont très étroites. Par exemple, on produit 150 pièces différentes, 15 à 25 de chaque type. Dans ce cas, on établit une carte par procédé de fabrication et non par lot (cf. KOONS and LUNER, *Use in low volume manufacturing environment. Quality and reliability*, ASQC Quality Press, 1991). Ces auteurs proposent les étapes suivantes :

- collecte des données, repérage des causes spéciales,
- étude des causes communes de dispersion,
- étude de la capabilité.

### 6.7.1 Etape 1 : collecte des données, repérage des causes spéciales

A l'intérieur d'un procédé de fabrication, on fait des lots homogènes (même stock d'origine, même outillage). Pour chacun de ces sous-lots on fait des relevés (soit à cent pour cent, soit par  $k_i$ -échantillon) des écarts à la cible, d'où

$$(y_1, \dots, y_k), \bar{y}, \hat{\sigma}$$

avec  $\hat{\sigma}_i^2 = \frac{1}{k_i-1} \sum_{j=1}^{k_i} (y_j^i - \bar{y}^i)^2$  qui est meilleur que l'étendue, en tout cas pour repérer les cause spéciales.

Cette fois, on fait un contrôle pour le paramètre  $s^2$  :

$$\forall i = 1, \dots, n, (k_i - 1)\hat{\sigma}_i^2 = \sum_{j=1}^{k_i} (y_j^i - \bar{y}^i)^2$$

pour ne pas avoir de biais. De plus, la statistique  $\frac{1}{\hat{\sigma}^2} \sum_{i=1}^n \sum_{j=1}^{k_i} (y_j^i - \bar{y}^i)^2$  suit une loi de khi-deux à  $\sum_{i=1}^n (k_i - 1)$  degrés de liberté.

La variance estimée de la population totale est

$$\bar{S}^2 = \frac{1}{\sum_{i=1}^n (k_i - 1)} \sum_{i=1}^n (k_i - 1)\hat{\sigma}_i^2$$

et  $\sum_{i=1}^n (k_i - 1)\bar{S}^2$  suit une loi de khi-deux à  $\sum_{i=1}^n k_i - n$  degrés de liberté. D'où les limites de contrôle pour la **variance** (et non plus pour l'écart-type) :

$$\mathbb{P}\{a\sigma^2 \leq (N - n)\bar{S}^2 \leq b\sigma^2\} = 1 - 2 \times 10^{-3}$$

et

$$LIC_{\sigma^2} = \frac{(N-n)\bar{S}^2}{\chi_{0.999}^2}, \quad LSC_{\sigma^2} = \frac{(N-n)\bar{S}^2}{\chi_{0.001}^2}.$$

Dès que la variance estimée  $\hat{\sigma}_i^2$  sort de ces limites (il s'agit d'une analyse *a posteriori*), c'est qu'il existe un risque de causes spéciales.

On enlève le lot incriminé, et on recommence jusqu'à ce que toutes les  $\hat{\sigma}_i^2$  soient dans les limites.

Puis on fait la carte des moyennes ( $\bar{x}_i$ ), moins celles des lots retirés, avec les limites de contrôles (cas  $k_i = k$  pour tout  $i$ ) :

$$\bar{X} \pm 3\sqrt{\frac{\bar{S}^2}{k}}.$$

On peut là encore détecter des causes spéciales.

### 6.7.2 Causes communes

KOONS et LUNER proposent une régression logistique de  $\log \hat{\sigma}_i^2$  avec pour variables explicatives les variables qualitatives pouvant "expliquer" les causes communes, par exemple le numéro de la machine, la matière usinée, le type de dimension mesurée...

### 6.7.3 Capabilité

Dans le cas où l'on n'a pas trouvé de variables qualitatives particulièrement significatives, on peut utiliser  $\hat{\sigma}_i^2$  pour évaluer la capabilité des lots,  $\frac{T_s - T_i}{6\hat{\sigma}}$ .

### 6.7.4 Critiques de la méthode

Les industriels peu cultivés (PME) trouvent cette démarche lourde et compliquée ; la rendre automatique par l'informatisation leur paraît difficile et dangereux ; ils veulent pouvoir garder leur libre arbitre à chaque instant. De plus, la méthode n'aide pas pour les premières pièces.

Néanmoins, il faut souligner que la méthode est plus rigoureuse dans le cas de lots de faible importance. Elle améliore l'étude de la dispersion, donc la capabilité est mieux évaluée.

## 6.8 Capabilité

### 6.8.1 A court terme

Elle se calcule avec l'estimation de  $\sigma$  faite sur l'ensemble des lots dont on aura éliminé ceux sujets aux causes spéciales :  $\hat{\sigma} = \frac{\bar{R}}{d_2(k)}$ , si  $k \leq 10$  pour des  $k$ -échantillon. Si les échantillons sont de taille différentes  $(k_1, \dots, k_m)$  le mieux est de passer par l'estimateur de variance :

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (k_i - 1) \hat{\sigma}_i^2}{\sum_{i=1}^n (k_i - 1)}$$

qui est un estimateur sans biais, et alors  $C_p = \frac{T_s - T_i}{6\sqrt{\hat{\sigma}^2}}$ .

### 6.8.2 A long terme

Alors on estime  $\sigma$  sur **toutes** les valeurs individuelles disponibles :

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2$$

et  $C_{pk} = \frac{(T_s - \bar{x}) \wedge (\bar{x} - T_i)}{3\sqrt{\hat{\sigma}^2}}$ .

## 7 Processus de la mise en place de la MSP

Diagramme d'ISHIKAWA

↓

AMDEC

↓

Caractéristiques clés

↓

Etude du procédé (capabilités)

↓

Cartes de contrôle

↓

Plans d'échantillonnage

Les 4 premières étapes permettent l'*état des lieux*, les deux dernières sont la surveillance du procédé qui permettent de remonter aux premières étapes dans une *boucle d'amélioration continue* dite **PDCA** (Plan Do Check act).

## 8 Méthode AMDEC

Analyse de Modes de Défaillance, de leurs Effets et de leur Criticité.

C'est un outil qui permet d'analyser les modes de défaillances, de quantifier leurs effets, de détecter leurs causes et de mettre en place des dispositifs d'assurance-qualité pour limiter les effets des défaillances. Il permet également de mesurer la gravité des conséquences des défaillances, et de classer ces défaillances selon leur "criticité". Cette méthode comporte 5 étapes détaillées ci-dessous. Le principe de base est une technique d'analyse exhaustive et rigoureuse de travail en groupe, pour faire ressortir les actions correctives à mettre en place :

- initialisation
- analyse fonctionnelle
- analyse des défaillances
- cotation des défaillances
- actions correctives à mener.

### 8.1 Initialisation

Il faut d'abord préciser l'objectif recherché, délimiter l'étude, préciser les acteurs en présence :

- le demandeur qui suscite l'analyse et choisit l'étude
- le décideur, le responsable dans l'entreprise du sujet étudié,
- l'animateur (organisateur, meneur des réunions, secrétaire des séances...)

et constituer un groupe de travail de 2 à 5 personnes responsables et compétentes, selon l'objet de l'étude.

On élabore alors une fiche de synthèse.

## 8.2 Analyse fonctionnelle

**Définition 8.1** *Pour analyser les défaillances d'un système, il est nécessaire de bien identifier toutes les fonctions que ce système doit assurer (fonctionnement et/ou stockage) : l'analyse fonctionnelle est donc l'identification de toutes ces fonctions.*

Cette analyse permet d'aboutir à une synthèse rigoureuse donnant une solution fiable sur le plan technique et économique.

*Exemple* Soit un surligneur. Quelles sont ses fonctions ? à qui sert-il ? sur quoi agit-il ? quels sont les milieux extérieurs ?

Cette étape aboutit à la production d'un dossier sur les fonctions à étudier, l'environnement, les objectifs de qualité et de fiabilité, le plan de maintenance, le conditionnement....

## 8.3 Analyse des défaillances

A partir de l'analyse fonctionnelle, on recherche :

- les modes de défaillances possibles,
- leurs effets,
- leurs causes,

- on mesure leur criticité (cotation, pas quantification des défaillances).

. Quelques exemples de modes de défaillances : une fonction n'opère pas (fuite, connexion défaite, court circuit....) ou bien elle est dégradée : perturbations, usure prématurée...

.. Les causes de la défaillance peuvent se trouver dans le milieu, chez les opérateurs, dans la documentation, l'organisation, la technique. Elles influent sur la sécurité du système.

Les effets sont d'importance graduées de 1 à 4 selon 4 critères :

- sécurité : 4= danger jusqu'à risque de mort,
- fiabilité : 1=mineur, 4= catastrophe,
- disponibilité : selon la durée de l'arrêt de fonctionnement,
- maintenabilité : c'est à dire possibilité ou non d'une maintenance régulière et fiable.

Tout ceci est résumé dans une *grille AMDEC* à sept colonnes : le nom de l'élément, sa fonction, causes, mode de défaillance, détection, effets, cotation de la criticité.

Cette notion est détaillée ci-dessous (cf. aussi ci-dessus page 13).

## 8.4 Notion de criticité

**Définition 8.2** *La criticité est évaluée à partir de la fréquence de la défaillance et de sa gravité. Elle détermine le choix des actions correctives, fixe la priorité entre ces actions. C'est un critère pour le suivi de la fiabilité prévisionnelle de l'équipement. Ainsi obtient-on une hiérarchie entre les différentes défaillances.*

**Définition 8.3** *La cotation est le nombre*

$$C = G.F.N,$$

où  $G$  est l'indice de gravité,  $F$  celui de la fréquence,  $N$  de la non détection, Chaque entreprise définit sa cotation pour ces trois indices, de 1 à 4 (voire plus, selon les entreprises).

Exemples :

Indice de 1 pour  $G, F, N$  : arrêt moins de 12h ; moins de 1 par an ; détection efficace,  
Indice de 2 pour  $G, F, N$  : arrêt de 13 à 24h ; moins de 1 par mois ; risque de détection non efficace,

Indice de 3 pour  $G, F, N$  : arrêt moins d'une semaine ; moins de 1 par semaine ; détection peu fiable,

Indice de 4 pour  $G, F, N$  : arrêt > semaine ; plus de 1 par semaine ; détection absente.



Ou encore on peut indiquer les niveaux suivants :

- 1 pour  $G$  minime,  $F$  rare,  $D$  avant montage,
- 2 pour  $G$  mineur,  $F$  faible,  $D$  aux essais,
- 3 pour  $G$  panne probable,  $F$  modérée,  $D$  à l'expédition,
- 4 pour  $G$  panne et détérioration,  $F$  fréquente,  $D$  possible par le client,
- 5 pour  $G$  risque d'accident,  $F$  systématique,  $D$  systématique par le client.

On établit un **seuil de décision** qui permet de mettre en évidence les défaillances sur lesquelles agir en priorité. Ce seuil est déterminé par le goupe de travail, c'est un **seuil de gravité**, limite au dessus de laquelle l'action (préventive ou corrective) est systématique.

## 8.5 Actions

Une fois les risques de défaillance mis en évidence, il s'agit de mettre en place des actions correctives ou préventives. On essaye de diminuer la criticité en établissant une maintenance :

- corrective,
- préventive,
- améliorative.

## 9 Caractéristiques clés

Une caractéristique clé, à la différence des caractéristiques contrôlées par les cartes, est une caractéristique dont la variation a une influence significative sur un procédé ou la réalisation d'un produit : il peut s'agir d'une caractéristique clé propre au produit ou bien propre au procédé de fabrication.

Une caractéristique clé 'procédé' doit être surveillée le plus en amont possible afin d'éviter les retours des produits, surtout ceux à forte valeur ajoutée.

### 9.1 Définition

Une caractéristique clé est une caractéristique d'un matériel ou d'une partie de matériel dont la variation a une influence significative sur

- . le montage du produit,
- . sa performance,
- . sa durée de vie,

qu'il s'agisse du producteur ou du client. Pour le producteur, l'impact de la caractéristique sur la réalisation du produit peut la rendre 'caractéristique clé'. Une fois les caractéristiques clé définie, on peut mettre en place les mesures et s'y tenir :

- journal des états,
- graphiques (suivi des valeurs par pointage, les cartes de contrôle par exemple).

### 9.2 Sources

Quelles sont les informations permettant de définir les caractéristiques clé ? Voici quelques exemples : les spécifications donnent une cotation fonctionnelle et des cotes clé à tenir absolument. Via la description du procédé, on analyse les risques et délimite les contrôles à effectuer. L'historique des 'non-qualité' peut mener à l'analyse des causes de défaillances et de non qualité. Enfin, il est des règles 'métier', des consignes, à respecter absolument.

De façon générale les caractéristiques clé sont définies tout autant par les clients que par les 'spécificateurs' que sont les autorités, les intervenants des différents métiers (aussi bien méthode que qualité, laboratoire, production, etc.) et le bureau d'étude. Le système d'information permet de se reposer sur des éléments factuels recensés a priori (incidents, gamme de fabrication, cycle de production, postes de travail....) afin de déterminer en connaissance de cause quelles sont les caractéristiques clés.

### 9.3 Caractéristique majeure ou non ?

Une caractéristique majeure pour le client doit l'être pour le fournisseur ! D'autres critères de majorité pour une caractéristique clé sont :

- . une variation dommageable et coûteuse,
- . sa mesurabilité sur le produit ou sur des paramètres qui conditionnent son élaboration.

Enfin, les caractéristiques clés doivent être en nombre restreint, sinon, on court le risque d'une conception ne permettant pas une bonne reproductibilité.

## 9.4 Préparer la mesure

Il s'agit de déterminer par avance où et quand au cours du procédé sera prise la mesure. Plus cette mesure sera prise en amont, si c'est possible, moins les problèmes auront de conséquence.

Qui va faire la mesure ? l'idéal serait une mesure automatisée, sinon, l'intégrer dans le procédé et qu'elle soit effectuée par les membres de la ligne.

On fixe bien sûr la taille et la fréquence de la mesure, (penser aux plans d'échantillonnage) : il s'agit d'optimiser le coût des mesures au regard du bénéfice apporté par la surveillance.

Enfin, il faut déterminer :

- . les moyens de mesure et de stockage des données recueillies,
- . l'exploitation des résultats
- . le mode opératoire.

## 9.5 Expérimenter

Voir si les caractéristiques clé sont mesurables de manière pratique, à un coût raisonnable. De plus, il faut effectuer une analyse 'R et R', répétabilité-reproductibilité, pour être assuré de la qualité des mesures.

## 9.6 Exploiter les données

- . mesurer régulièrement et rigoureusement,
- . enregistrer les événements importants qui peuvent expliquer des variations (changement de procédé, maintenance des outils, nouveaux moyens, nouveau personnel...)

- . mettre en place les graphiques (suivi des valeurs par pointage, les cartes de contrôle par exemple),
- . expliquer aux membres de la ligne (ligne=ensemble de postes groupés les uns à la suite des autres dédié à une mission) comment lire les graphiques et quelle attitude avoir à leur égard,
- . présenter les graphiques et commenter régulièrement les tendances au sein de la ligne,
- . monter des actions ou projets d'amélioration ciblés.

C'est là qu'intervient le schéma d'amélioration continue (PDCA) :

planifier-faire-vérifier-agir,  
et ceci en boucle.

En effet, le but est d'obtenir une inversion des pourcentages suivants : souvent, le travail de fond occupe 20 pour cent du temps, contre 80 pour du "réparage", plus ou moins bricolé (on dit qu'on joue les "pompiers"!), alors qu'il serait souhaitable que 80 pour cent du temps soit consacré au travail de fond et seulement 20 pour cent au "réparage" bricolé après coup. On parle parfois de "loi de Pareto", loi qui est évoquée pour dire que 20 pour cent de la population possède 80 pour cent de la richesse...

## References

- [1] T.W. ANDERSON, “An Introduction to Multivariate Analysis”, Wiley Press, 3d edition, 2003.
- [2] G. BAILLARGEON, “Maîtrise statistique des procédés”, Editions SMG, Québec, 1995.
- [3] G. BAILLARGEON, “Plans d’échantillonnage en contrôle de la qualité”, Editions SMG, Québec, 1995.
- [4] P. BARBE et M. LEDOUX, “Probabilité”, Belin, Paris, 1998.
- [5] Maurice PILLET, “Appliquer la maîtrise statistique des procédés, MSP/SPC”, Editions d’organisation, Paris, 4<sup>è</sup>m édition, 2005.
- [6] P. JAFFARD, “Méthode de la statistique et du calcul des probabilités”, Masson, paris, 1996.
- [7] K. VO KHAC : “Théorie des probabilités”, Ellipses, Paris, 1984.
- [8] Page web de Maurice Pillet  
<http://www.qlio.univ-savoie.fr/pillet/>
- [9] Logiciel SIMDI tour, TP de MSP  
<http://www.oqp.univ-savoie.com/>