



HAL
open science

AN OPTIMAL CONTROL CHART FOR NON-NORMAL PROCESSES

Emmanuel Duclos, Maurice Pillet

► **To cite this version:**

Emmanuel Duclos, Maurice Pillet. AN OPTIMAL CONTROL CHART FOR NON-NORMAL PROCESSES. 1st IFAC Workshop on Manufacturing Systems: Manufacturing Modelling, Management and Control, 1997, Austria. 12 p. hal-00974314

HAL Id: hal-00974314

<https://hal.science/hal-00974314>

Submitted on 6 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AN OPTIMAL CONTROL CHART FOR NON-NORMAL PROCESSES

Emmanuel DUCLOS, Maurice PILLET

*ODS Europe, 12 rue pré paillard 74940 Annecy le Vieux
LLP/CESALP - BP 806 74016 Annecy Cedex - FRANCE
☎ : (+33) 50.66.60.80 Fax : (+33) 50.66.60.20
Email : duclos@esia.univ-savoie.fr*

Abstract : In this paper we present the application of a control chart for non-normal processes. This chart, which looks like an \bar{X}/S control chart is built with a least-squares L-estimator, which can replace the arithmetic mean and standard deviation usually calculated for Shewhart charts. This estimator has the property to provide a minimum variance estimation of the process position and scattering. This, disregarding data distribution. We focused our attention on « multi-generators » processes, like screw-machines or multi-die holder for injection molding, these processes have the property to generate non-normally distributed pieces.

Keywords : Least-squares estimation, Non-Gaussian processes, Optimal estimation, Process control, Statistics

1. INTRODUCTION

When piloting a process with S.P.C. control charts (\bar{X}/R) we usually assume that the mean is normally distributed. In spite of its robustness, this assumption is far from always being satisfied. In fact, either the \bar{X} distribution is close to a normal distribution or the sample is large enough to satisfy the assumption of normality of the mean. However, large samples are prohibited for economic reasons. Disregarding this assumption can cause problems when the hypothesis of normality is not valid.

First of all, statistical error, type I and II, correspond no longer with those defined for a normal distribution when control limits are placed at $\pm 3\sigma$ from the target (Shilling, 1976).

The second point concerns precision of estimations. Indeed, we can show that the mean is not the optimal estimator in term of variance when the population is non-normal. It means that we can find an estimator without bias which can provide better performances than the average.

So in order to solve these problems, we propose in this article the use of a control chart (L chart) build

with a minimum variance estimator whose performances have been compared to those of the average in term of variance and distribution shape.

We will study this estimator in the case of data incoming from a « Multi-generator » process.

2. MULTI-GENERATOR PROCESSES

Consider a process which consists of several elementary machines. Among these processes, we can quote for example injection presses.

Each criterion produced by an elementary machine is distributed among an elementary characteristic which can be normal or non-normal according to the criterion studied. Then, the global distribution or population is said to be a « mix of probability distributions » which is unlikely to give a normal population.

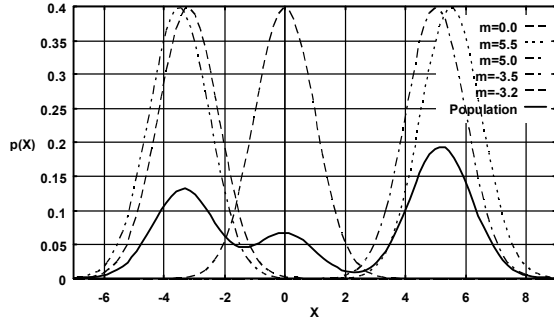


Fig.1. Non-normal population resulting of a mix of elementary distributions.

To illustrate this article we have chosen the following mix of distributions (Figure 1).

Assuming that each elementary distribution has the same probability ($\pi = 1/6$), the population is defined by :

$$f_p(X) = \pi \cdot f_{N_1}(X) + \pi \cdot f_{N_2}(X) + \dots + \pi \cdot f_{N_6}(X)$$

Even when a Multi-generator process is under control, observations are non-normally distributed. Because of its physical characteristics, it is sometimes impossible to make it normal or economically unsuitable.

3. PRINCIPAL OF AN L-ESTIMATOR

If a population is normally distributed, we can prove that the variance of the arithmetic mean reaches the Frechet limit. So, the arithmetic mean is said to be the optimal estimator in terms of variance. Unfortunately, this is not the case when the population is non-normally distributed.

Therefore, we have been interested in an estimator based on order statistics known as L-estimators, which have the capability to take into account the population shape, through their coefficients (David, 1981).

The L-estimator proposed by E.H. Lloyd (1952), based on a Least-Squares algorithm, requires no hypothesis on the population shape whereas the method of maximum likelihood supposes a normal parent.

3.1 L-Estimators and L-Statistics

Consider a sample (x_1, x_2, \dots, x_n) of n independent observations sampled at a time k.

$x_{(1)}, x_{(2)}, \dots, x_{(n)}$ are ordered observations of \mathbf{X}_k such as $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$.

$x_{(n)}$ is called n ordered statistics of the ordered vector \mathbf{X}_k .

The linear combination of \mathbf{X}_k 's ordered statistics with a vector \mathbf{C}_j of real coefficient is an L-statistic or L-estimator.

The estimation of θ parameter at a time k is given by :

$$\hat{\theta}_k = {}^t \mathbf{C} \cdot \mathbf{X}_k = \sum_{i=1}^n C_i \cdot x_{(i)} \quad (1)$$

The average is a particular case of L-estimator since all coefficients are equal : $C_i = \frac{1}{n} \forall i \in [1;n]$.

The range is another particular L-estimator for a scale parameter : $C_1 = -1, C_n = 1$ and $C_i = 0 \forall i \neq 1, i \neq n$

The problem is then, to calculate the L-estimator's coefficients in order to obtain the expected performances.

3.2 The Least Square L-estimator

Choice of μ_p and σ_p parameters. The population's parameters of location and scale which can be estimated by the least-squares L-estimator are not necessarily the mean and standard deviation of the population. However, we show that they are the most appropriate parameters, if we want to maximise the process capability indice. (Pillet, 1997).

The purpose of S.P.C. is to improve production quality and therefore to minimise production cost. Taguchi defined the loss function (2) which represents the cost of non-quality :

$$L = K \left(\sigma^2 + (\bar{X} - \text{Target})^2 \right) \quad (2)$$

$$C_{pm} = \frac{IT}{6\sqrt{\sigma^2 + (\bar{X} - \text{Target})^2}} \quad (3)$$

Where K is a constant, \bar{X} is the average of the population and σ its standard deviation.

We notice that minimising Taguchi's loss function is equivalent to maximising the capability indice Cpm (3).

So to maximise the Cpm indice we have to minimise the population standard deviation and the quadratique error between sample mean and the process target. In consequence, we will consider \bar{X} and σ respectively as localisation μ_p and scale σ_p parameters of the population.

Generalised Least-Squares. In this paragraph we describe the construction of the least-squares L-estimator. We also expose some basic results concerning the Generalised Least Square theory.

Suppose the following multiple linear model $\mathbf{X} = \mathbf{W}\mathbf{b} + \mathbf{u}$ where :

- \mathbf{X} is the variable of interest (sample n observations of the process).
- \mathbf{u} is a random vector modelling noise on Y (common causes).

- \mathbf{b} is the vector of parameters that we expect to estimate (localisation and scale of the process).
 - \mathbf{W} is a matrix including non random variables.
- The following conditions are supposed to be satisfied : $E(u) = \vec{0}$ et $V(u) = \sigma^2 \cdot I_n$

Estimating parameter \mathbf{b} by the least-squares consists of minimising the u random variable influence. The Ordinary Least Square estimator is defined by :

$$\hat{\mathbf{b}} = (\mathbf{W}'\mathbf{W})^{-1} \cdot \mathbf{W}'\mathbf{Y}$$

Application on an ordered sample

Assume $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$ are ordered observations of a vector \mathbf{X}_k where $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$

$U_{(r)}$ is a standardized variable :

$$U_{(r)} = \frac{(x_{(r)} - \mu_p)}{\sigma_p}$$

and moments of the order statistic $U_{(r)}$ are :

$$E[U_{(r)}] = \alpha_r \quad Var[U_{(r)}] = \omega_{rr} \quad Cov[U_{(r)}, U_{(s)}] = \omega_{rs}$$

Lets resume now the problem exposed by Lloyd. The moment of order statistics $X(r)$ can be written under vectorial form :

$$E[X(r)] = \mu_p \cdot \mathbf{e} + \sigma_p \alpha$$

where α is a vector of α_r and \mathbf{e} is a unit vector

$$\mathbf{X} = \begin{pmatrix} x(1) \\ \vdots \\ x(n) \end{pmatrix} \quad \mathbf{e} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad \mathbf{a} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

The preceding equation can be written :

$$E[X] = A \cdot \theta$$

where $\theta = \begin{pmatrix} \mu \\ \sigma \end{pmatrix}$ and $A = \begin{bmatrix} 1 & \alpha_1 \\ \vdots & \vdots \\ 1 & \alpha_n \end{bmatrix}$

The variance-covariance matrix of \mathbf{X} is : $\sigma^2 \cdot \Omega$ where Ω is an $(n \times n)$ matrix of ω_{rs} elements.

Aitken (1935) proved that such a problem could be solved by applying a least-squares algorithm on ordered statistics.

Since $Var(u) = Var(X) = \sigma^2 \Omega$, the model is general. The solution of this model derives from the ordinary least-squares. In fact, it is noticeable that Ω is a positive matrix, so a $M(n, n)$ regular matrix exists such as: $M^t M = \Omega^{-1}$

A Generalised Least Square Estimation of θ is defined as (4) :

$$\hat{\theta} = \begin{pmatrix} \hat{\mu} \\ \hat{\sigma} \end{pmatrix} = (A' \cdot \Omega^{-1} \cdot A)^{-1} \cdot A' \cdot \Omega^{-1} \cdot X \quad (4)$$

Variance of both estimators of location and scale is (5) :

$$\begin{aligned} Var[\hat{\mu}] &= \sigma_p^2 \cdot \frac{\alpha' \cdot \Omega^{-1} \cdot \alpha}{\det(A' \Omega^{-1} A)} \\ Var[\hat{\sigma}] &= \sigma_p^2 \cdot \frac{e' \cdot \Omega^{-1} \cdot e}{\det(A' \Omega^{-1} A)} \end{aligned} \quad (5)$$

This theoretical description of the Least-Squares L-estimator brings to light, that this estimator can't be used without knowing moments of the ordered observations. This can cause some problems of application. The next paragraph will present an application of this estimator for Statistical Control of a multi-generator process.

4. L-CHARTS PERFORMANCES

4.1 Calculation of the variance-covariance matrix.

To be able to appreciate improvement brought by the L-estimator compared to the arithmetic mean, we achieved computer simulations. In order to determine the L-estimator coefficients, it is necessary to calculate the matrix of variances-covariances of standardized ordered statistics. These calculations were made with random data generated according to a known distribution. Each coefficient of the matrix is calculated with the relation (6)

$$\omega_{ij} = \frac{1}{m-1} \sum_m \left[(x_{(i)} - \bar{x}_{(i)}) (x_{(j)} - \bar{x}_{(j)}) \right] \quad (6)$$

Where m is the number of samples.

Coefficients of variance-covariance matrix can also be calculated from the theoretical expression :

$$Cov[X_{r:n}, X_{s:n}] = \int_{-\infty}^{\infty} \int_{-\infty}^y (x - \mu_{r:n}) (y - \mu_{s:n}) \cdot f_{rs}(x, y) \cdot dx dy \quad (7)$$

with

$$f_{rs}(x, y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} \cdot F^{r-1}(x) \cdot f(x) (F(y) - F(x))^{s-r-1} \cdot p(y) [1 - P(y)]^{n-s}$$

This expression necessitates powerful means of calculation and even more, if the sample size is large and the distribution is non-normal. Further more, most practical applications deal with unknown populations, so that, this calculus can't be achieved. Although theoretical calculus was made to reach a great precision, computer simulations were preferred because of their similarity with the practical situation.

Since processes are not necessarily well known and operating conditions are always evolving, building a model of the population's distribution does not seem realistic.

4.2 Computer Simulations

Since a control chart is a set of two estimations : Punctual estimation and confidence interval estimation, we studied the L-estimator performances in terms of variance and shape. Since the method is non parametric, performances of the L-estimator are different according to the distribution of the population.

The more different from a normal law the distribution is, the more efficient the L-estimator. In consequence we studied a case where the population is significantly non-normal (figure 1) to appreciate the L-estimator efficiency compared to the mean.

Variance and bias of estimations. Results in table 2, show that the variance of the L-estimator of location is always much lower than the average. The relative efficiency of the L-estimator compared to the mean is (8) :

Table 1 Variance of both location estimators and relative efficiency of the L-estimator compared to the arithmetic mean

n	3	4	5	6	7	8	9	10
$Var[\hat{\mu}]$	4.29	2.45	1.52	1.01	0.72	0.55	0.44	0.36
$Var[\bar{X}]$	5.32	3.99	3.19	2.66	2.28	2.00	1.77	1.60
<i>eff</i>	0.80	0.62	0.48	0.38	0.32	0.28	0.25	0.23

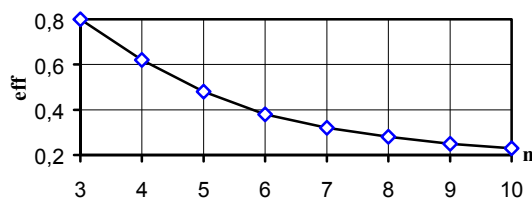


Fig. 2. Relative efficiency of the L-estimator compared to the arithmetic mean

$$eff = Var[\hat{\mu}] / Var[\bar{X}] \quad (8)$$

To provide the same performances in term of variance as the L-estimator when n=4, the average requires a sample size n=7. The benefits of using an L-estimator can be either the sample size reduction, or the improvement of commandability when sample size is maintained and the variance of the estimation is reduced.

In opposition, the L-estimator for the scale parameter brings no improvement compared to the standard deviation, in terms of variance.

However, it gives a non biased estimation of the population dispersion even when the population is non-normal (Figure 3, Table 2). Whereas the empiric standard deviation S_n , is a biased estimator when sample size is small. In fact, the coefficients c_4 usually used to correct this bias is unsuitable when the population is non-normal.

A relative bias for estimators of a scale parameter can be defined as (9) :

$$b_{\hat{\sigma}} = (\hat{\sigma} - \sigma_p) / \sigma_p \quad (9)$$

Of course these values are specific to this example, but similar results can be found for other significantly non-normal populations.

Symmetry of distributions. The distribution shape was studied to determine how to apply confidence interval tests for control charts.

The setting of limits on \bar{X} control charts are based on the assumption of normality, justified by the central limit theorem.

Table 2 Relative bias of the L-estimator of dispersion and standard deviation

n	3	4	5	6	7	8
$b_{\hat{\sigma}}$	-9.2%	-5.2%	-3.4%	-2.4%	-1.7%	-1.4%
b_{S_n}	0.1%	0.03%	0.06%	0.03%	0.03%	0.01%

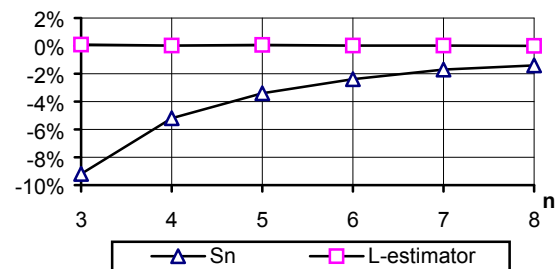


Fig. 3. Relative bias of the L-estimator of dispersion.

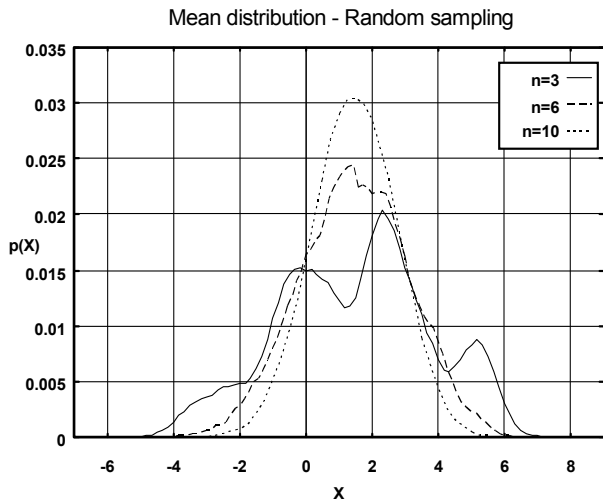


Fig. 4.

The theorem essentially states, that under general conditions, the distribution of sample means approaches normality. Nevertheless (Burr, 1967) showed that for significantly non-normal distributions, sample mean was far from being normal. Yourstone (1992) proposed recently modified control limits to keep risks close to 0.27% in the case of skewed population.

Figure 4 underlines the fact that when the sample size is small, normality is far from being satisfied. Coefficients of Kurtosis approaches 3 by inferior value as n increases. It means that the distribution has heavier tails than the normal law.

On the contrary we notice that the distribution of the L-estimator doesn't converge to a normal law (Figure 5), but a narrower one. Coefficients of Kurtosis are actually around 6. Since coefficients of the Means-Squares L-estimator are non zero for extreme values, this estimator doesn't reject outliers. As a result distribution tails are as heavy as the average ones, hence control limits are kept at $3\sigma/\sqrt{n}$ to minimise type I false alarms.

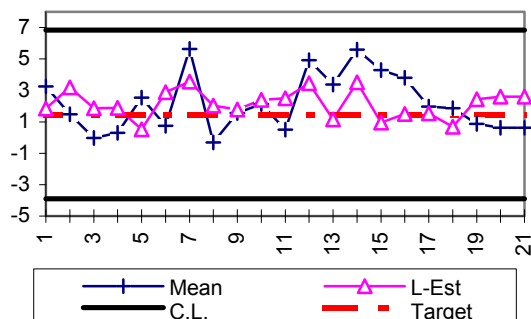


Fig. 8 L-chart of location compared to an \bar{X} chart.

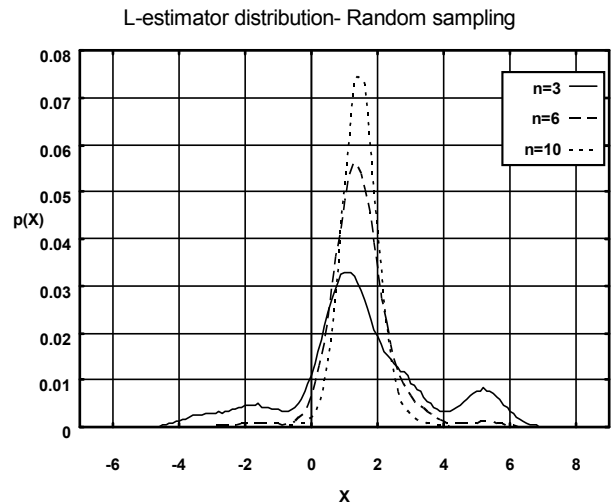


Fig. 5.

5. EXAMPLE OF APPLICATION

5.1 The L-chart

The interest for such an estimator is evident for an industrial application since the sample size can be reduced without loss of efficiency compared to the mean. The application of this chart requires a preliminary run to calculate the variance-covariance matrix. However this run isn't restricting as it was established by simulation that 40 samples were sufficient to calculate Ω precisely. Such a preliminary run is then short enough for an industrial application. This procedure is equivalent to using a preliminary control chart to determinate the mean and the standard deviation of the process.

When using standard control charts, the process which is under control, is always supposed to be stationary. Statistical Process Control aim is to keep the process under control, which means evolution of the process position or dispersion. In consequence, each time an out of control state is detected, the model (matrix Ω and vector α) has to be computed again to take into account the evolution of the process.

5.2 L-chart for Start-up processes

In order to reduce to a minimum the period of reference and then control the process with the only values μ_p and σ_p fixed by the user's knowledge, we propose a start-up procedure for the L-chart by exponentially weighted moving average (EWMA) of the L-estimator's coefficients.

Coefficients of the estimator (C_{EWMA}) evolve from an initial value ($C_{initial}$) to the optimal coefficients ($C_{L-estimator}$) by using relationship (10).

$$C_{EWMA} = \frac{k}{k_{max}} \cdot C_{L-Estimator} + \left(1 - \frac{k}{k_{max}}\right) \cdot C_{initial} \quad (10)$$

Where k is the number of a samples and k_{max} is the number of samples required to calculate the matrix of variance-covariance with precision. The coefficient k_{max} should be larger than 40.

Coefficients of the vector $C_{initial}$ depend on the parameter to estimate. In order to estimate the parameter of location μ_p , coefficient of $C_{initial}$ are equal to the mean ($1/n$). As a result, the estimator will provide the mean of samples at the beginning of the run. On the contrary, as k increases, more weight is given to the coefficients of the L-estimator $C_{L-estimator}$ which are more and more reliable. In order to estimate the parameter of scale σ_p , the vector $C_{initial}$ defined as : $C_{initial}(1) = -1/d2$, $C_{initial}(2) = 1/d2$ and $C_{initial}(i) = 0 \quad \forall i \neq 1, i \neq 2$ gives an estimation of population's standard deviation with $R/d2$.

Thanks to this step, variance of estimation for the location parameter, decreases from σ_p^2/n to $\sigma_{L-estimator}^2$.

6. CONCLUSION

Through this short example, we have shown that it is interesting to replace the traditional estimators of Statistical Process Control (the average and the standard deviation), by the Least-Squares L-estimator. This estimator has the advantage of providing non biased estimations with minimum variance. These two characteristics are essential to minimise the cost of non quality. In fact, it is shown that Taguchi's loss function is minimised when observations are centred on the target and their variance is minimal. In addition, the application of this L-estimator is not limited to Shewhart control charts. It can easily be extended to CUSUM charts.

On the other hand, the use of the Least-Squares L-estimator is facilitated by its systematic method of calculation whatever the distribution of the population. In order to make this method transparent

for a manufacturer, we have proposed a procedure to launch production without any preliminary run.

Finally, for the construction of L charts, we have systematically placed control limits to $Target \pm 3\sigma/\sqrt{n}$. This specification allows us to reduce α risks (as compared to the average), however it seems unsatisfactory because, there are unquantifiable risks on both sides of the distribution. L-estimators are non non-normally distributed in spite of the fact that they are asymptotically normally distributed. We are presently working on a systematic method of calculation for control limits that would make L charts more efficient in term of Average Run Length (A.R.L.).

REFERENCES

- Aitken, A. C. « *On least squares and linear combinations of observations.* » Proc. Roy. Soc. Edimb. 55,42, 1935
- Burr, I. W. « *The effect of Non-Normality on Constants for \bar{X} and R Charts* » Industrial Quality Control, Vol. 3, No. 11, May 1967, pp563-568.
- David, H. A. « *Order Statistics* », New York, Willey, 1981
- Downton, F. « *Least Square Estimates* », Annals of Mathematical Statistics, Vol 25, pp 303-316, 1954
- Duclos, E. « *Optimisation de la Maîtrise Statistique des procédés par une méthode de filtrage d'ordre* », Revue de Statistiques Appliquées (RSA), N° XLIV Vol 2, p 61-79 - 1996
- Lloyde, E. H. « *Least Square Estimation of Location and Scale Parameters using Order Statistics* », Biometrika, Vol 39, pp 88-95 - 1952
- Pillet, M., Rochon, S., Duclos, E « *SPC - Generalisation of capability index Cpm. Case of unilateral tolerances* » Quality Engineering, 1997.
- Shilling, E.G, Nelson, P. R. « *Effect of Non-Normality on the Control Limits of \bar{X} Charts* », Journal of Quality Technology Vol 8, N° 4, October 1976.
- Yourstone, S. « *Non-normality and design of Control Charts for Average* », Decision Science, Vol 23, pp 1099-1113 - 1992