ASPECTS OF STATISTICAL PROCESS CAPABILITY ANALYSIS: REVIEW AND EXTENSIONS

M. El-Moniem Soleha
Ain Shams University
Department of Statis.
Cairo, Egypt

Abstract

This paper is mainly an overview paper that demonstrates some historical issues of the process capability problem. A detailed statistical discussion of the main important early results is launched, by recalling Shewhart's pioneering work, as well as some other productive papers. A representation of some extended views is also provided including an improved numerical new findings. Some concluding remarks are given on which a numerical investigation is conducted to reflect the real state of the process capability of the statistically analyzed data. Suggestions for future researches are also explored by implementing some advanced statistical concepts for refining a set of data, and highlighting a prospective "predictive dimension" for the analysis of the process "future" performance.

Kev words:

Natural tolerances, reference interval, process potential index, location capability index, statistical performance graph, (P, γ) - type tolerances.

1. Introduction

The assessment of the process performance has received a considerable attention since the early forties and fifties. For instance, the essentials of the process capability analysis are dated back to the appearance of the outstanding second book of W..A. Shewhart. The majority of his everlasting concepts have been included in this productive book, which is titled by "statistical Method from the viewpoint of quality control", and published by the U.S. Department of Agriculture in 1939. He proposed that: the so-called "technical specifications" -abbreviated by (specs)- imposed on a certain measurable quality characteristic can be considered as external elements being analogous to the process intended to generate that characteristic. In this case, the process has its own behavioral characteristics which may – or may not- obey some outside requirements. Such behavioral characteristics are referred to as "statistical tolerances" or ,generally, natural tolerances. These have been statistically approximated by $\bar{x} \pm 3s$, where $\bar{x}$ and s are, respectively, the average and the standard deviation of the set of measurements $\{x_1, x_2, \ldots, x_n\}$. Such measurements are obtained from a certain quality characteristic (x)- a measurable one – generated by that process. The previous statistical tolerances had been used (see [4]) to tackle a certain problem – namely to construct analytically a tolerance interval from a "probabilistic" point of view.

Formally, the problem of statistical tolerances of (P, γ) - type can be put as follows: given a continuous random variable X, having the density $f(x;\theta)$ where θ is a vector of parameters which individualizes f, find two statistics $L_{inf}(x_1, x_2, \ldots, x_n)$ and $L_{sup}(x_1, x_2, \ldots, x_n)$ such that a given proportion P ($0 < P < 1$) of the population {x} to be found in the interval [L_i, L_s] with a given probability - say γ $(0 < \gamma < 1)$. This means:

$$\Pr ob\left\{\int_{L_1}^{L_2} f(x;\theta)dx \geq P\right\} = \gamma \qquad (1.1)$$

The above – mentioned interval (1.1)has been considered as a "predictive" one (see [10]). Finally, the proposed concept of "statistical tolerances" as a technique for "statistical control", and the procedure of constructing "control charts" stand for the main contributions to the development of the science of statistical quality control, introduced by Shewart. In this context, Wilks [15] had put up a genuine solution for the following "related" statistical problem.

given a continuous random variable X for which we do not know its density (nor its analytical form, neither its parameters), how large the set of measurements on X be, such that the extreme values of these measurements to be natural tolerances of X, for given P and γ. That is, if one seeks the value of n (sample size) such that $x_{(1)} = \min\{x_1, x_2,..., x_n\}$ and $x_{(n)} = \max\{x_1, x_2,..., x_n\}$ to fulfil the relationship :

$$\Pr ob\left\{\int_{x_{(1)}}^{x_{(n)}} f(x;\theta)dx \geq P\right\} = \gamma \qquad (1.2)$$

where $0 < P, \gamma < 1$

(this is a distribution - free or distribution- space problem since he does not know f (see [14]). Wilks had arrived at his famous equation which gives $n : P^{n-1} - (n-1)P^n = 1 - \gamma$. A solution of the previous problem – being represented by (1.2) – had been given by Wold [16], in which the considered the case of having a normal distribution. His results have been improved (see [7]) to provide more mathematical tractability.

This paper is organized as follows. Section 2 begins with reviewing two main indicators by which the process capability is described. These estimates – usually called indices - are assessing the variability of the process as well as its location. A two-levels procedure is postulated for analyzing the process performance, using the process potentiality and capability indices. Section 3, reports some corrected indices that are capable of measuring, more accurately, the quality characteristic at hand, if it is departing from normality. Section 4, some concluding remarks will be introduced, which take into account the descriptive nature of the analyzed data, and consider it as a "practical criterion" to the best choice among the various process capability indices. In section 5, a numerical example is given to illustrate the systematic procedure for estimating the suitable process capability indices for the data. These estimates will be taken as a basis to validate – or not – the process capability of the data. In section 6, some suggestions are explored for improving the previous estimated indices by implementing some statistical concepts, such as the predictive point of view, the p-value, and the power-transformation for refining a set of data. These concepts could be implemented in the data-analysis to give it a "predictive dimension". Thus, "predictive indices" could be constructed for the process "future" performance. **Finally**, a discussion of an extended case of multi-dimensional process is presented.

2. Process capability analysis

The analysis of the process capability requires a brief background about the process variability. Then, one needs to know the two generations of capability indices. In this context, graphical representation of indices will be useful for decisional purposes. This section ends up with postulating a procedure for the levels of the process performance analysis.

(2.1) : Basic Background

Process capability is conventionally considered to be an equivalent term to its normality. This means that, process capability refers to the normality of that process. Uniformity means a small variability of the underlying characteristic of the generated output. Process variability must behave itself within the specified limits imposed on the characteristic (X) and more than that, it is desirable that measured values be nested around a target (T) of X, considered as the optimal value of that characteristic. In many cases, this target value may be situated in the middle of the specified interval, that is USL - T = T - LSL (USL = Upper specified limit, LSL = Lower specified limit). In such situations T could be compared with $\overline{x}$ (the sample mean) of the measured values.

(2.2) Generations of capability indices:

To describe the process capability – according to the previous background - one needs two indicators: the first is referring to variability:

$$\hat{C}_p = \frac{USL - LSL}{6s} \qquad (2.1)$$

where s is the standard deviation of measurements, and the second is referring to the location of the process:

$$\hat{C}_{pk} = \frac{|USL - \overline{x}|}{3s}, \ \frac{|LSL - \overline{x}|}{3s} \qquad (2.2)$$

These $\hat{C}_p$ and $\hat{C}_{pk}$ (which are in fact the estimates of the theoretical ones, C_p, C_{pk} where E(X) is replaced by $\overline{x}$ end STDEV (X) is replaced by s) are called respectively the **potential index** and the **capability index as regards location**. Shortly, one refers to them as **capability indices ("first generation ones")** - see Kotz and Lovelace [9].

It is important to notice that {$\hat{C}_p$ and $\hat{C}_{pk}$} have an informational value if the process is stable (that is in the state of statistical control) and the quality characteristic at hand obeys the normal law. Otherwise, these indices must be considered as **preliminary** indicators: no important decisions must be taken based on them, until the process is stable and statistical model of the characteristic is identified.

The concept by which the previous indices (2.1) and (2.2) are constructed, has been modified (see[3] by proposing a more refined indices, namely C_{pm} and C_{pmk} using a target value (T) and the well-known idea of quality loss-function (see[4]) :

$$C_{pm}=\frac{USL-LSL}{6\sigma_0}, \qquad (2.3)$$

$$C_{pmk}=\min\left\{\frac{|USL-\mu|}{3\sigma_0},\ \frac{|LSL-\mu|}{3\sigma_0}\right\} \qquad (2.4)$$

Where

$$\sigma_0^2 = E(X-T)^2 = E[(X-\mu)+(\mu-T)]^2$$
$$= E(X-\mu)^2 + (\mu-T)^2$$

Since $E(X)=\mu$ and $Var(X)=\sigma^2$, we have $\sigma_0^2=\sigma^2+(\mu-T)^2$ and hence

$$C_{pm}=\frac{USL-LSL}{6\sqrt{\sigma^2+(\mu-T)^2}}=\frac{\frac{USL-LSL}{6\sigma}}{\sqrt{1+\frac{(\mu-T)^2}{\sigma^2}}}=\frac{C_p}{\sqrt{1+v^2}} \qquad (2.5)$$

where $v=\frac{\mu-T}{\sigma}$

(here $T=(USL+LSL)/2$).

If $T=\mu$, then C_{pm} becomes C_p and $C_{pmk}=C_{pk}$.

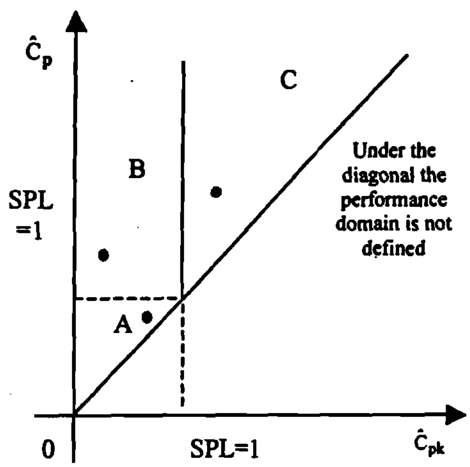
These **"second generation"** of capability indices (C_{pm}, C_{pmk}) illustrate more realisticly the true possibilities of the process to reach its target value. The index C_{pmk} has been studied in detail by Kotz [9].

Obviously, in practice we work with $\hat{C}_{pm}$, $\hat{C}_{pmk}$ - the estimators of those indices.

Considering now the equality $C_{pk}=(1-k)\,C_p$, where $k=2|T-\mu|/(USL-LSL)$ and since $0\le k\le 1$, we have always $C_p\ge C_{pk}$: if $T=\mu$, then $k=0$ and $C_{pk}=C_p$.

(2.3) Graphical representation of indices:

In the above context, graphical techniques can be a useful statistical tool for providing information embedded in the data. If the data represent the outputs of a certain "stable" process the "graphical representation", that connects indices (2.1) and (2.2), will have an important informational value. This means that, it gives a graphical demonstration for the "performance levels" of the process. One can use such graphs in judging the performance of the process for decision-making purposes. By adopting Smith's concept [11], one can choose a suitable graphical method to construct a "decisional graph" for summarizing the outputs of a certain process. This could be done by fixing a "minimal" statistical performance level (SPL) of the process as SPL = 1:



The above graph partition the performance domain into a defined and undefined ones. For the defined domain, it has been partitioned (for decision-making purposes) into three non-overlapping zones A, B and C as follows :

- Zone A: both indices $\hat{C}_p$, $\hat{C}_{pk}$ less then 1, that is unacceptable statistical performance; of the process.
- Zone B: good potential $(\hat{C}_p > 1)$, inadequate location $(\hat{C}_{pk} < 1)$;
- Zone C: it is the zone of the desired capability of the process (both indices greater than 1).

(2.4) Levels of the process performance analysis:

Generally speaking, one can postulate a "two-levels" procedure for "analyzing" the process performance. This analysis will be beneficial in evaluating the "present" state of the process, as well as, it "future" performance. The "first level" is the "potentiality" level in which, we compare the index "Cp" with a specified performance level (SPL). The outcomes of the first level will be: either (a) Cp < SPL, which means the process has "no potentiality", (b) Cp $\geq$ SPL, then the process has "potentiality".

The "second level" is the capability level, and in which "each" outcome of the first level is checked for "capability", by comparing Cpk with (SPL). This will have two resulting outcomes (the process has capability or not) for "each" (a) and "b".

By combining the results of the two levels the process performance will take "one" of the following states:

(1) If the two indices (Cp) and (Cpk) are both "less than" (SPL), the process will have a "bad" performance.

(2) If (Cp < SPL), but (Cpk ≥ SPL), the process has "no potentiality", but it "has capability". This means that the process has "a reduced performance".

(3) If (Cp ≥ SPL), but (Cpk < SPL), then the process has potentiality but with "low performance".

(4) If (Cp ≥ SPL) and (Cpk ≥ SPL), then the process has potentiality as well as capability. This means that it is an "optimal" process with a "highest" (or excellent) level of performance.

3. Limitations and Insights :

It has been widely agreed that the lack of "stability" of the process represents a main obstacle in using the previous indices to judge the performance of the process. This means that if the existing quality characteristic does not conform to the normal law, the "reliability" of the derived indices to judge the process performance is "doubtful".

3.1. Adaptation of indices :

An attempt was pursued (see[10]) for the sake of introducing a corrected index to adapt with the departure from normality. This index takes into account the asymmetry of the distribution:

$$C_s = \frac{d - |\mu - T|}{3\sqrt{\sigma^2 + (\mu - T)^2 + |\mu_3/\sigma|}} \qquad (3.1)$$

where T = (LSL + USL)/2, d = (USL - LSL)/2 and $\mu_3 = E(X - \mu)^3$ - the third central moment which describes the departure from symmetry.

Another corrected index was proposed (see[17]) to tackle the case of having skewed data. This index takes the form.

$$C_{pu} = \frac{USL - Median}{3s - Median} \qquad (3.2)$$

in order to describe better a skewed distribution.

In order to cope with a flat distribution of the data, a measure of kurtosis, was suggested [5] for this case. Thus, the potential index was corrected to have the form.

$$\hat{C} = \frac{USL - LSL}{6s\sqrt{1 + |3 - b_2|}} \quad (3.3)$$

where $b_2 = \frac{1}{n}\sum(x_i - \bar{x})^4 / s^4$ is the sample coefficient of kurtosis.

ISO document ISO/DIS 3534-2/2004 (ϕ 1.2 "Statistical process management", pages 11 - 30) introduced the so-called "reference interval" $x_{99.865\%} - x_{0.135\%}$, where $x_{a\%}$ is the a% - fractile of the distribution.

This interval is then used to construct "the process performance index"

$$P_p = \frac{USL - LSL}{x_{99.865\%} - x_{0.135\%}} \quad (3.4)$$

which should be used when the process is not stable (is not in statistical control). Specifically for the normal distribution, the "reference interval" is 6σ or 6s.

3.2. Unification of indices:

An important idea had been realized (see[8]) for getting a general formula to unify some, of these indices. This general index takes the form :

$$C_p(u, v) = \frac{d - u|\mu - M|}{3\sqrt{\sigma^2 + v(\mu - T)^2}} \quad (3.5)$$

where u, v are real numbers, d = (USL - LSL)/2, M = (LSL + USL)/2, T - target value (which may be M sometimes). Hence, we get easily:

1) $C_p(0, 0) = C_p$ (classical potential index);

2) $C_p(0, 1) = C_{pm}$ (Taguchi)

3) $C_p = (1, 1) = C_{pmk}$ (Kotz).

Some other details may be found in the references [2], [9].

4. Concluding Remarks

in this section some remarks will be introduced that are, mainly, **implementing** the idea of considering the (nature) of the data, **in** assessing the process capability indices.

These remarks represent a set of guide lines to choose among the above-mentioned variety of process capability indices.

Depending upon the numerically calculated values of some simple descriptive measures of symmetry, skewness and kurtosis of a set of data, the following proposed recommendations are giving :

1. Start your analysis by testing for normality, choosing the one which is suitable for your sample size. If your test confirms the presence of symmetry, index (2.1) will be suitable for your analysis.
2. If the test shows that the set of data is asymmetric one, then index (3.1) will be suitable for assessing the process capability indices for such set of data.
3. Perform a simple test of skewness. If your set of data is a skewed one, then index (3.2) will be convenient for your analysis.
4. If a test of (kurtosis) is significant for your set of data, then index (3.3) will give more realistic values of the process capability indices.

Finally, a numerical example will be given to demonstrate the use of these recommendations, and in which it is found that normality is prevailing in the exciting set of data.

5. An illustrative numerical example

The aim of this numerical example is to measure the defective fractions P_{inf} and P_{sup} as a basic step to compute the total probable defective fraction p_t. Then by recalling the potential index $\hat{C}_p$ one can validate – or not – the process capability (or the process performance).

This will be beneficial in assessing the natural variability and judging whether it is larger or smaller than the imposed target ones.

A measurable characteristic X has two specifications namely LSL = 5.4 c.u and USL = 5.9 c.u (c.u = conventional units). The target value

T is fixed as (USL + LSL)/2 = 5.65 and hence we can consider the theoretical mean value (μ) as T. A sample of size n = 50 has been collected and the experimental values were as follows:

5.5; 5.6; 5.6; 5.8; 6.0; 5.5; 5.9; 5.8; 5.7; 5.8; 5.4; 5.7; 5.8; 5.4; 5,7; 6.0; 5.8; 5.4; 6.0; 5.6; 5.6; 5.5; 5.6; 5.8; 5.8; 6.0; 5.5; 6.0; 5.8; 5.8; 5.5; 5.6; 5.5; 6.0; 5,8; 5.9; 5.8; 6.1; 5.7; 5.5; 5.5; 5.6; 5.4; 5.9; 5.9; 5.8; 5.6; 6.2; 5.9; 5.5.

A test for normality is performed (using [6]) that confirms the symmetry of the above set of data.

Thus, by estimating the potential index $\hat{C}_p$ one can have a correct insight about whether the estimated value $\hat{C}_p$ validates (and to what extent) the process capability.

Firstly, concerning the lower and upper values it is noticed that there are no values less than LSL (5.4) but there are eight values above USL = 5.9. Thus one can obtain easily $\bar{x} \approx 5.724$ and $s \approx 0.103$.

Secondly, By performing a test of normality for moderate sample size (see [6]) the presence of this law is confirmed (90% confidence). Thus, the defective fractions are computed as follows :

$$p_{inf} = \mathrm{Prob}\{x < T_i\} = \mathrm{Prob}\left\{\frac{x-\bar{x}}{s} < \frac{T_i-\bar{x}}{s}\right\} \approx 0.08\% \quad (5.1)$$

(which is a very low percentage)

$$p_{sup} = \mathrm{Prob}\{x > T_s\} = 1-\mathrm{Prob}\{x \le T_s\} \approx 4.36\% \quad (5.2)$$

(we denoted USL = T_s and LSL = T_i). Therefore **the total probable** defective fraction is

$$p_t = p_{inf} + p_{sup} \approx 4.44\% \quad (5.3)$$

The conclusion, derived from the above numerical results, assures that the natural variability is larger than the imposed one. Thus "the potential index" is therefore

$$\hat{C}_p = (5.9-5.4)/0.103 \times 6 = 0.5/0.618 \approx 0.81$$

which **invalidates** process capability.

6- Suggestions for future researches

Firstly, a predictive framework could be suggested for solving relation (1.2) which is viewed as a distribution-space problem (see [14]), where the distribution f(x; θ) is unknown.

The predictive point of view demands using a "data - oriented" method for estimating $\hat{f}$ (see [12]). Thus the relation (1.2) will have a predictive property by which its solution will be more realistic. Also, estimating $\hat{f}$ in (1.2) could be considered as initial step to explore the possibility of constructing "predictive" indices for assessing the process "future performance". This means that, the proposed idea is one of constructing an index as a "predictor" of a certain future variable outcomes.

Secondly: The second proposed idea aims at improving the construction of the process performance index by suggesting a specific "reference interval". Recalling the previous index (3.4), one would resort to estimating the "extreme centiles" of $\hat{f}$ and the corresponding data (or reference) interval. A centile could be derived by converting the observation to the probability scale by using a probability integral transform (see [13]). In this context, the implementation of the p-value approach – as a measure of extremeness – is a promising device to be considered in constructing the reference intervals, upon which an "informative" updated indices can be constructed. These indices will reflect the "real state" of the process performance, because they are based on exact probability statements rather than an asymptotic approximations.

Thirdly: This suggested idea is concerned with ways of overcoming "the skewness of the data" used to construct the above process performance index (3.2). Thus, the initial step of the process capability analysis should be devoted to "refining" the given set of data for the sake of moving towards normality and getting rid of skewness. Consequently, it would be convenient to use the shifted power transformation g(X) (see [1]) defined by

$$g(X)=\frac{(X-\alpha)^{\lambda}-1}{\lambda}, \qquad (6.1),$$

as a flexible way of achieving such "data refinement". The distribution of the random variable (X) given by the transformation (6.1) is distributed as $N(\theta_1, \theta_2^2)$ and is usually called the shifted power normal distribution, in which skewness has been ironed out.

Also, one can notice that as λ approaches zero in (6.1) then the power transformation of the observed value (x_λ) tends to log (x). Thus, by following the third suggestion, we have successfully implemented the important "log-transformation" in our data-refinement process.

In conclusion : The "ultimate" goal of the previous set of proposals is to practically explore distributions ($\hat{f}$) based on transformations towards "normality" as flexible models for the real data and in which estimation of "Reference intervals" is appropriate for providing a probabilistically "updated future" indices.

Finally, a discussion of an extended case of multi-dimensional process is presented as follows. If the process consists of X_i (i = 1, ..., r) characteristics, one has to summarize the "individual" performance of all process characteristics to evaluate the aggregate performance of the "r-dimensional" process. Thus, one can suggest an "extension" for determining the "performance indices" for such multivariate process. That is, according to the multiplicative law, the multivariate "potential" index may, presumably, have the form :

$$C_p(\text{mul}) = \left[\prod_{j=1}^{r} C_{pj}\right]^{\frac{1}{r}} \quad \ldots\ldots\ldots\ldots (6.2)$$

Also, by the some argument, the r-dimensional "capability" index may take the form.

$$C_{pk}(\text{mul}) = \left[\prod_{j=1}^{r} C_{pkj}\right]^{\frac{1}{r}} \quad \ldots\ldots\ldots (6.3)$$

These performance indices, represented by (6.2) and (6.3), have to be compared with a certain or (imposed) specified performance level (SPL) to determine the domain of the multivariate process performance.

In fact, equations (6.2) and (6.3) stand for the coordinate axes of the Cartesian product "C_p (mul) x C_{pk} (mul)", which is determined by (SPL). Thus, one can plot the point (C_p (mul), C_{pk} (mul)) – which is a pair of the calculated performance indices, and proceed by followign the steps of the above-mentioned "decisional-graph" for the "individual" characteristic X of the univariate process. Thus, the above-mentioned equations can be considered as a devised formulas to "reduce" the "r-dimensional" process into a "two-dimensional" space one.

References

1. Atkinson, A., Pericchi, L. and Smith, R.L (1991). Grouped likelihood for the shifted power transformation, Journal of the Royal Statistical Society, Series B 53: 473-482.

2. Bârsan-Pipu, N., Isaic-Maniu, Al., and Vodă, V.Gh. (1998): Decision procedures in statistical process control. Econ. Comp. Econ. Cyb. Stud. Res., nr. 1 - 4, P. 49 -58.

3. Boyles, R.A. (1991): The Taguchi capability index. Journal of Quality Technology, vol. 23, No. 1, P. 17 - 26.

4. Caulcutt, R. (1995): Achieving Quality Improvement. A Practical Guide. Chapman and Hall, London.

5. Crişan, D.M. (2002); On the contribution of capability indices : Theory and practice. Doctoral dissertation, University of Bucurest, Faculty of IMST, Department TCM.

6. D'Agostino, R.B. (1971): An omnibus test of normality for moderate large size sample. Biometrika, vol. 58, nr. 2, P. 341 - 348.

7. Howe, W.G. (1969): Two-sided tolerance limits for normal populations: some improvements. J. Amer. Statist. Assoc., vol 64, nr. 326, P. 610 - 620.

8. Kerstin, V. (1993): A unified approach to capability indices. Research Report No. 3/1993, May, Högskalan i Lulia University, Sweden, IES/ Division of Quality Technology.

9. Kotz, S. and Lovelace, Cynthia (1998): Process Capability Indices in Theory and Practice. ARNOLD, (a member of the Hodder Headline group), London.

10. Montgomery, D.C. (1996): Introduction to SQC. Third Edition. John Wiley and Sons, Inc. New York.

11. Smith, A, Vounatsou, P (1994). Graphical methods for simulation – based inference. Technical report (94-12), Imperial College of Science, Technology and Medicine, London,. U.K.

12. Soleha, M. (1988): On the estimation of the Mixing Density Function in the Mixture of Exponentials. Doctoral dissertation, City University, London, U.K.

13. Soleha, M. (1993). On the integral – transform procedure for the Bayesian predictive approach to model selection., Economic and Business Review, Vol.2, 149-169. Ain Shams University, Cairo, Egypt.

14. Soleha, M. (1995): The role of the distribution–space estimation method in improving model selection procedures. The seventh annual conference on statistics and computer modelling in human and social sciences, Cairo University, Cairo, Egypt.

15. Wilks, S.S. (1941): Determination of sample size for setting tolerance limits. Annals of Math. Statist., vol. 12. P. 91 - 96.

16. Wold, A and Wolfowitz, J. (1946): Tolerance limits for a normal distributicns. Annals of Math. statist., vol. 17, P. 208 -215.

17. Wright, A.P. (1995): A process capability index sensitive to skewness. Journal of Statistical Computation and Simulation, vol. 52, P. 195 - 203.