Article

Not peer-reviewed version

# Control C Charts with Reflections on Charts for Process Control

Fausto Galetto *

Posted Date: 7 November 2025

doi: 10.20944/preprints202511.0444.v1

Keywords: control charts; poisson distribution; exponential distribution; TBE; T Charts; Minitab; JMP; Reliability Integral Theory

 doi:10.20944/preprints202511.0444.v1



*Article*

# Control C Charts with Reflections on Charts for Process Control

**Fausto Galetto**

Department of Manufacturing Systems & Economics, Politecnico of Turin, Turin, Italy, fausto.galetto@gmail.com

**Abstract**

We start with the ideas in the papers "Chakraborti et al., Properties and performance of the c-chart for attributes data, Journal of Applied Statistics, January 2008" and "Bayesian Control Chart for Number of Defects in Production Quality Control. Mathematics 2024, 12, 1903. https://doi.org/10.3390/math12121903"; then we use the Jarrett (1979) data from "A Note on the Intervals Between Coal-Mining Disasters" and the analysis by Kumar et al., and by Zhang et al. From the analysis of all the papers data we get different results: the cause is that they use the Probability Limits of the PI (Probability Interval) as they were the Control Limits (so they name them) of the Control Charts (CCs): those authors do not extract the complete information from the statistical data of CCs from the data not normally distributed. The Control Limits in the Shewhart CCs are based on the Normal Distribution (Central Limit Theorem, CLT) and are not valid for non-normal distributed data: consequently, the decisions about the "In Control" (IC) and "Out Of Control" (OOC) states of the process are wrong. The Control Limits of the CCs are wrongly computed, due to unsound knowledge of the fundamental concept of Confidence Interval. Minitab and other software (e.g. JMP, SAS) use the "T Charts", claimed to be a good method for dealing with "rare events", but their computed Control Limits of the CCs are wrong. We will show that the Reliability Integral Theory (RIT) is able to solve these problems.

**Keywords:** control charts; poisson distribution; exponential distribution; TBE; T Charts; Minitab; JMP; Reliability Integral Theory

**MSC:** 46N30; 62F15

---

## 1. Introduction

Since 1989, the author (FG) tried to inform the Scientific Community about the flaws in the use of *("wrong") quality methods for making Quality* [1] and in 1999 about the GIQA (Golden Integral Quality Approach) showing how to manage Quality during all the activities of the Product and Process Development in a Company [2], including the Process Management and Control Charts (CC) for Process Control.

First we show how to *deal correctly with c-Control Charts* by analysing a literature case [3], which uses data of "Nonconformities in Printed CircuitBoards (example 7.3)" in the Montgomery book "*Introduction to Statistical Quality Control 8th ed.*" and the ideas in "Bayesian Control Chart for Number of Defects in Production Quality Control", published in Mathematics 2024: it is quite interesting that in a "rejected paper" a Peer Reviewr wrote "*The content of the paper is mainly centered on applied statistics and industrial statistics. Hence, in my opinion the journal Mathematics is not the best venue for this manuscript*." Decision 13 February 2025.

Later, we show how to *deal correctly with I-CC (Individual Control Charts)* by analysing a literature case comprising the famous data-set on coal-mining disasters of Jarrett (1979); these data are considered and analysed in [4,5].

We found very interesting the statements in the Excerpt 1:

 doi:10.20944/preprints202511.0444.v1



> In the recent paper "*Misguided Statistical Process Monitoring Approaches*" by W. Woodall, N. Saleh, M. Mahmoud, V. Tercero-Gómez, and S. Knoth, published in *Advanced Statistical Methods in Process Monitoring, Finance, and Environmental Science*, 2023, We read in the **Abstract**: *Hundreds of papers on flawed statistical process monitoring (SPM) methods have appeared in the literature over the past decade or so. The presence of so many ill-advised methods, and so much incorrect theory, adversely affects the SPM research field. Critiques of some of the various misguided, and/or misrepresented, approaches have been published in the past 2 years in an effort to stem this tide. These critiques are briefly reviewed here.* References…

**Excerpt 1.** From the paper "Misguided Statistical Process Monitoring Approaches".

We agree with the authors in the excerpt 1, but, nevertheless, they did not realise the problem that we are giving here: *wrong Control Limits in CCs for Rare Events, with data exponentially or Weibull distributed*. See References…

Using the data in [3–5] with good statistical methods [6–33] we give our "reflections on Control Charts (CCs)".

We will try to state that *several papers* (that are not cited here, but you can find in the "*Garden of flowers*" [24] and some in the Appendix C) compute in an *a-scientific way* (see the formulae in the Appendix C) the Control Limits of CCs for "Individual Measures or Exponential, Weibull and Gamma distributed data", indicated as I-CC (Individual Control Charts); we dare to show, to the Scientific Community, how to compute the *True Control Limits*. If the author is right, then all the *decisions, taken up today, have been very costly to the Companies using those Control Limits*; therefore, "*Corrective Actions*" are needed, according to the Quality Principles, because NO "*Preventive Actions*" were taken [1,2,27–36]: this is shown through the suggested published papers. Humbly, given our strong commitment to Quality [34–46], we would dare to provide the "truth": *Truth makes you free* [hen ("hic et nunc"=here and now)].

On 22nd of February 2024, we found the paper "*Publishing an applied statistics paper: Guidance and advice from editors*" published in *Quality and Reliability Engineering International* (QREI-2024, 1-17) [by C. M. Anderson-Cook, Lu, R. B. Gramacy, L. A. Jones-Farmer, D. C. Montgomery, W. H. Woodall; the authors have important qualifications and Awards]; since I-CC is a part of "*applied statistics*" we think that their hints will help: the authors' sentence "*Like all decisions made in the face of uncertainty, Type I (good papers rejected) and Type II (flawed papers accepted) errors happen since the peer review process is not infallible.*" is very important for this paper: the interested readers can see [34–46].

By reading [24], the readers are confronted with this type of practical problem: we have a warehouse with two departments

a) in the 1st of them, we have a sample (the "*The Garden of flowers…* in [24]") of "products (*papers*)" produced by various production lines (*authors*)

b) while, in the other, we have some few products produced by the same production line (*same author*)

c) several inspectors (*Peer Reviewers, PRs*) analyse the "quality of the products" in the two departments; the PRs can be the same (but we do not know) for both the departments

d) The final result, according to the judgment of the inspectors (PRs), is the following: the products stored in the 1st dept. are good, while the products in the 2nd dept. are defective. It is a very clear situation, as one can guess by the following statement of a PR: "*Our limits* [in the 1st dept.] *are calculated using* standard mathematical statistical *results/methods* as is typical in the vast literature of similar papers [4,5,24]." See the standard mathematical statistical *results/methods* in the Figures A1, A2, A3, of the Appendix A and meditate (see the formulae there)!

Hence, the problem becomes "…the *standard … methods* as is typical …": are those standards typical methods (in the "*The Garden …* in [24]" and in the Appendix C) scientific?

 doi:10.20944/preprints202511.0444.v1



The practical problem, for TBE data (Exponentially ditsributed)" becomes hence a *Theoretical* one [1–46] (all references and Figure 1): we show here, immediately, *the* ***wrong formulae*** (either using the parameter $\theta = \theta_0$ or its estimate $\bar{t}_0$, with $\alpha =$ 0.0027) in the formula ***(1)***

$$LCL = \theta_0 ln(1-\alpha/2) = 0.00135\,\bar{t}_0 \quad UCL = \theta_0 ln(\alpha/2) = 6.6077\,\bar{t}_0 \qquad (1)$$

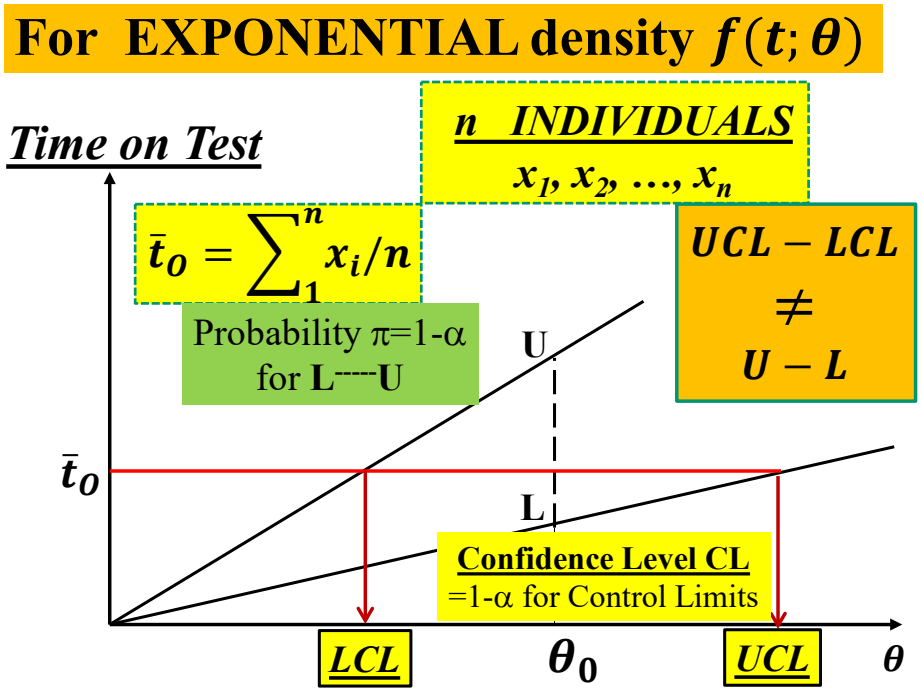



**Figure 1.** Theoretical Difference between L----U and LCL----UCL.

In the formulae (1), in the (named) interval LCL----UCL (Control Interval), the LCL must be L and the UCL must be U, vertical interval L----U (Figure 1); the actual interval LCL----UCL is the horizontal one in the Figure 1, which is not that of the formulae (1). Since ***the errors have been continuing for at least 25 years***, we dare to say that this paper is an *Education Advance* for all the Scholars, for the software sellers and the users: they should study the books and papers in [1–46].

The readers could think that the I-CCs are well known and well dealt in the scientific literature about Quality. We have some doubt about that: we will show that, at least in one field, the I-CC_TBE (with TBE, Time Between Event data) usage, it is not so: there are several published papers, in "scientific magazines and Journals (well appreciated by the Scholars)" with wrong Control Limits; a sample of the involved papers (from 1994 to January 2024) can be found in [23,24]". Therefore, *those authors do not extract the maximum information from the data in the Process Control*. "*The Garden…*" [24] and the excerpts 1, with the Deming's statements, constitute the Literature Review.

> *"Management need to grow-up their knowledge because experience alone, without theory, teaches nothing what to do to make Quality" "Experience alone, without theory, teaches management nothing about what to do to improve quality and competitive position, nor how to do it.), ... understanding of quality requires education. There is no substitute for knowledge. It is a hazard to copy. It is necessary to understand the theory of what one wishes to do or to make..... hundreds of people are learning what is wrong. .... I make this statement on the basis of experience, seeing every day the devastating effects of incompetent teaching and faulty applications. Again, teaching of beginners should be done by a master, not by a hack".*

**Excerpt 2.** Some statements of Deming about Knowledge and Theory (Deming 1986, 1997).

We hope that the Deming statements about knowledge will interest the Readers (Excerpt 2). A preliminary case is shown in Appendix A.

 doi:10.20944/preprints202511.0444.v1



## 2. Materials and Methods

*2.1. A Reduced Background of Statistical Concepts*

This section is essential to understand the "problems related to Control Charts" as we found in the literature. We suggest it for the formulae given and for the difference between the concepts of PI (*Probability Interval*, NOT "Prediction Interval", asd interpreted by a Peer Reviewer!) and CI (*Confidence Interval*): this is overlooked in "*The Garden …* [24]" (a sample of it is in the Appendix C).

See a first case in the appendix A. Therefore, we humbly ask the reader to carefully meditate on the content.

Engineering Analysis is related to the investigation of phenomena underlying products and processes; the analyst can communicate with the phenomena only through the *observed data*, collected with sound experiments (designed for the purpose): any phenomenon, in an experiment, can be considered as a *measurement-generating process* [MGP, a black box that we do not know] that provides us with information about its behaviour through a *measurement process* [MP, known and managed by the experimenter], giving us the observed data (the "*message*").

It is a law of nature that the data are variable, even in conditions considered fixed, due to many unknown causes.

MGP and MP form the Communication Channel from the phenomenon to the experimenter.

The information, necessarily incomplete, contained in the data, has to be extracted using sound statistical methods (the best possible, if we can). To do that, we consider a *statistical model* $F(x|\theta)$ associated with a *random variable* (RV) X giving rise to the measurements, the "*determinations*" $\{x_1, x_2, \ldots, x_n\}$=D of the RV, constituting the "*observed sample*" D; n is the sample size. Notice the function $F(x|\theta)$ [a function of real numbers, whose form we assume we know] with the symbol θ accounting for an unknown quantity (or some unknown quantities) that we want to estimate (assess) by suitably analysing the sample D.

We indicate by $f(x|\theta) = dF(x|\theta)/dx$ the pdf (probability density function) and by $F(x|\theta)$ the Cumulative Function, where $\theta$ is the set of the parameters of the functions.

When $\theta = \{\mu, \sigma^2\}$ we have the *Normal model*, written as $n(x|\mu, \sigma^2)$, with (parameters) mean E[X]=μ and variance Var[X]=$\sigma^2$

$$f(x|\mu,\sigma^2) = n(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} \quad (2)$$

When $\theta = \{\theta\}$ we have *Exponential model*, with (the single parameter) mean E[X]=$\theta = 1/\lambda$ and variance Var[X]=$\theta^2 = 1/\lambda^2$, written in two equivalent ways $f(x|\theta) = (1/\theta)e^{-x/\theta} = \lambda e^{-\lambda x} = f(x|\lambda)$.

When $\theta = \{\mu\}$ we have *Poisson model*, with (the single parameter) mean E[X]=$\mu$ and variance Var[X]=$\mu$, written as $f(x|\mu) = e^{-\mu}\mu^x/x!$, with x=0, 1, 2, …., n, …

When we have the *observed sample* D=$\{x_1, x_2, \ldots, x_n\}$, our general problem is to estimate the value of the parameters of the model (representing the parent population) from the information given by the sample. We define some criteria which we require a "good" estimate to satisfy and see whether there exist any "best" estimates. We assume that the parent population is distributed in a form, the model, which is completely determinated but for the value $\theta_0$ of some parameter, e.g. unidimensional, θ, or bidimensional θ={μ, $\sigma^2$}; we consider only one or two parameters, for easiness.

We seek some function of θ, say τ(θ), named *inference function*, and we see if we can find a RV T which can have the following properties: unbiasedness, sufficiency, efficiency. Statistical Theory allows us the analysis of these properties of the estimators (RVs).

We use the symbols $\bar{X}$ and $S^2$ for the unbiased estimators $T_1$ and $T_2$ of the mean and the variance.

Luckily, we have that $T_1$, in the Exponential model and Poisson model $f(x|\theta)$, is efficient [6–21,25–33], and it extracts the total available information from any random sample, while the couple $T_1$ and $T_2$, in the Normal model, are jointly sufficient statistics for the inference function τ(θ)=(μ, $\sigma^2$), so extracting the maximum possible of the total available information from any random sample. The estimators (which are RVs) have their own "distribution" depending on the parent model F(x|θ) and

 doi:10.20944/preprints202511.0444.v1



on the sample D: we use the symbol $\varphi(t,\theta,n)$ for that "distribution". It is used to assess their properties. For a given (collected) sample D the estimator provides a value t (real number) named the *estimate* of τ(θ), unidimensional.

A way of finding the estimate is to compute the *Likelihood Function* $L(\theta|D)$ [LF] and to maximise it: the solution of the equation $\frac{\partial}{\partial\theta}L(\theta|D)$=0 is termed *Maximum Likelihood Estimate* [MLE].

The LF is important because it allows us finding the MVB (*Minimum Variance Bound*, Cramer-Rao theorem) [1,2,6–16,26–36] of an unbiased RV T [related to the inference function τ(θ)], such that

$$Var(T) \geq \frac{[\tau(\theta)]^2}{E\{[\partial lnL(\theta|D)/\partial\theta]^2\}} = MVB(T) \quad (3)$$

The inverse of the MVB(T) provides a *measure of the total available amount of information* in D, relevant to the inference function τ(θ) and to the *statistical model* F(x|θ).

Naming $I_T$(T) the information extracted by the RV T we have that [6–21,26–36]

$I_T$(T)=1/MVB(T) ⇔ T is an Efficient Estimator.

If T is an *Efficient Estimator* there is no better estimator able to extract more information from D.

The estimates considered before were "*point estimates*" with their properties, looking for the "best" single value of the inference function τ(θ).

We must now introduce the concept of *Confidence Interval* (CI) and Confidence Level (CL) [6–21,26–36]. This point is very important for ***all*** the Control Charts: it was overlooked by thoousands of authors, Peer Reviers and Editors…

The "*interval estimates*" comprise all the values between $\tau_L$ (Lower confidence limit) and $\tau_U$ (Upper confidence limit); the CI is defined by the *numerical* interval CI={$\tau_L$----$\tau_U$}, where $\tau_L$ and $\tau_U$ are two quantities computed from the *observed sample* D: when we make the statement that τ(θ)∈CI, we accept, before any computation, that, doing that, we can be right, in a long run of applications, (1-α)%=CL of the applications, BUT we cannot know IF we are right in the single application (CL=Confidence Level).

We know, before any computation, that we can be wrong α% of the times but we do not know when it happens.

The reader must be very careful to distinguish *between* the *Probability Interval* PI={L----U}, where the endpoints L and U depends on the distribution $\varphi(t,\theta,n)$ of the estimator T (that we decide to use, which *does not depend on the "observed sample" D*) and, on the probability π=1-α (that we fix before any computation), as follows by the probabilistic statement (4) [se the Figure 1 for the exponential density, when n=1]

$$P[L \leq T \leq U] = \int_L^U \varphi(t,\theta,n)dt = 1-\alpha \quad (4)$$

and Confidence Interval CI={$\tau_L$----$\tau_U$} which depends on the "observed sample" *D*.

Notice that the Probability Interval PI={L----U} *does* not *depend* on the data D: L and U are the Probability Limits. Notice that, on the contrary, the Confidence Interval CI={$\tau_L$----$\tau_U$} *does depend* on the data D.

Shewhart identified this approach, L and U, on page 275 of [19] where he states:

> "For the most part, however, we never know $f_\Theta(\Theta,n)$ [this is the symbols of Shewhart for our $\varphi(t,\theta,n)$] in sufficient detail to set up such limit… We usually chose a *symmetrical* range characterised by limits $\overline{\Theta} \pm t\sigma_\Theta$ *symmetrically spaced* in reference to Θ. Tchebycheff's Theorem tells us that the probability P that an observed value of Θ will lie within these *symmetric* limits so long as the quality standard is maintained satisfies the inequality P>1-1/t². We are still faced with the choice of t. Experience indicated that t=3 seems to be an acceptable economic value". See the excerpts 3,…

 doi:10.20944/preprints202511.0444.v1



The Tchebycheff Inequality: IF the RV X is arbitrary with density f(x) and finite variance $\sigma^2$ THEN we have the probability $P[|X-\mu| \geq k\sigma] \leq 1/k^2$, where $\mu = E[X]$. This is a "Probabilistic Theorem".

It can be transferred into *Statistics*. Let's suppose that we want to determine experimentally the unknown mean $\mu$ within a "stated error ε". From the above (Probabilistic) Inequality we have $P[\mu-\varepsilon < X < \mu+\varepsilon\,] \geq 1-\sigma^2/\varepsilon^2$; IF $\sigma \ll \varepsilon$ THEN the event $\{|X-\mu| < \varepsilon\}$ is "very probable" in an experiment: this means that the observed value $x$ of the RV X can be written as $\mu-\varepsilon < x < \mu+\varepsilon$ and hence $x-\varepsilon < \mu < x+\varepsilon$. In other words, using $x$ as an estimate of $\mu$ we commit an error that "most likely" does not exceed $\varepsilon$. IF, on the contrary, $\sigma \not\leq\not\leq \varepsilon$, we need n data in order to write $P[\mu-\varepsilon < \bar{X} < \mu+\varepsilon\,] \geq 1-\sigma^2/(n\varepsilon^2)$, where $\bar{X}$ is the RV "mean"; hence we can derive $\bar{x}-\varepsilon < \mu < \bar{x}+\varepsilon$., where $\bar{x}$ is the "empirical mean" computed from the data. In other words, using $\bar{x}$ as an estimate of $\mu$ we commit an error that "most likely" does not exceed $\varepsilon$. See the excerpts 3, 3a, 3b.

*Notice* that, when we write $\bar{x}-\varepsilon < \mu < \bar{x}+\varepsilon$, we consider the Confidence Interval CI [6–21,25–33], and no longer the Probability Interval PI [6–21,25–33].

These statistical concepts are very important for our purpose when we consider the Control Charts, especially the Individual CCs, I-CC.

*Notice* that the error made by several authors [3–5,24] is generated by *lack of knowledge* of the difference between PI and CI [6–21,25–33]: they think *wrongly* that CI=PI, a diffused disease [3–5,24]! They should study some of the books/papers [6–21,25–33] and remember the Deming statements (excerpt 2).

The Deming statements are important for Quality. Managers, scholars; the professors must learn Logic, Design of Experiments and Statistical Thinking to draw good decisions. The authors must, as well. Quality must be their number one objective: they must learn *Quality methods as well*, using Intellectual Honesty [1,2,6–21,25–33]. Using (4), *those authors do* not *extract the maximum information from the data in the Process Control*. To *extract the maximum information from the data* one needs statistical valid Methods [1,2,6–21,25–33].

As you can find in any good book or paper [6–21,25–33] there is a strict relationship between CI and Test Of Hypothesis, known also as Null Hypothesis Significance Testing Procedure (NHSTP). In Hypothesis Testing, the experimenter wants to assess if a "thought" value of a parameter of a distribution is confirmed (or rejected) by the collected data: for example, for the mean μ (parameter) of the Normal $n(x|\mu,\sigma^2)$ density, he sets the "null hypothesis" $H_0=\{\mu=\mu_0\}$ and the probability P=α of being wrong if he decides that the "null hypothesis" $H_0$ is true, when actually it is opposite: $H_0$ is wrong. We analyse the *observed sample* $D=\{x_1, x_2, \ldots, x_n\}$ and we compute the *empirical* (observed) *mean* $\bar{x}$ and the *empirical* (observed) *standard deviation* $s$; hence, we define the *Acceptance interval*, which *is the CI*

$$\bar{x} - t_{1-\alpha/2}s/\sqrt{n} < \mu < \bar{x} + t_{1-\alpha/2}s/\sqrt{n} \tag{5}$$

Notice that the following interval (for the Normal model) [see the appendix B]

$$\mu_0 - t_{1-\alpha/2}\sigma_0/\sqrt{n} \text{ ------ } \mu_0 + t_{1-\alpha/2}\sigma_0/\sqrt{n} \tag{6}$$

is the Probability Interval such that $P\left[\mu_0 - t_{1-\alpha/2}\sigma_0/\sqrt{n} < \bar{X} < \mu_0 + t_{1-\alpha/2}\sigma_0/\sqrt{n}\,\right] = 1-\alpha$, and NOT the Confidence Interval and thus NOT the LCL-----UCL (for the CCs).

*A fundamental reflection is in order*: the formulae (5) and (6) tempt the unwise guy to think that he can get the *Acceptance interval*, which is *the CI* [1–23], by substituting the assumed values $\mu_0, \sigma_0$ of the parameters with the *empirical* (observed) *mean* $\bar{x}$ and *standard deviation* $s$. *This trick is valid only for the Normal distribution*.

More ideas about this can be found in [34–46].

In the field of Control Charts, following Shewhart, instead of the formula (5), we use (7)

$$\bar{x} - z_{1-\alpha/2}s/(c_4\sqrt{n}) < \mu < \bar{x} + z_{1-\alpha/2}s/(c_4\sqrt{n}) \tag{7}$$

 doi:10.20944/preprints202511.0444.v1



where the value $t_{1-\alpha/2}$ of the t distribution is substituted by the value $z_{1-\alpha/2}$ of the Normal distribution, actually $z_{1-\alpha/2}$=3, and a coefficient $\boldsymbol{c_4}$ is used to make "unbiased" the estimate of the standard deviation, computed from the information given by the sample.

Actually, Shewhart himself does not use the coefficient $\boldsymbol{c_4}$ is as you can see from page 294 of Shewhart book (1931), where $\bar{X}$ is the "Grand Mean", computed from D [named here *empirical* (observed) *mean* $\bar{x}$], $\sigma$ is "estimated standard of each sample" (named here s, with sample size n=20, in excerpt 3)

$$\bar{X} \pm 3\frac{\sigma}{\sqrt{n}} = 13{,}540 \pm 3\frac{440}{\sqrt{20}}$$

**Excerpt 3.** From Shewhart book (1931), on page 294.

The application of these ideas in the Individual CCs can be seen in the Appendix A, in the Figure A1: the standard deviation is derived from the Mobile Range (which is exponentially distributed as the original UTI data). The formula in the excerpt 3 tells us that the process is OOC (Out Of Control).

### 2.2. *Control Charts for Process Management*

Statistical Process Management (SPM) entails Statistical Theory and tools used for monitoring any type of processes, industrial or not. The Control Charts (CCs) are the tool used for monitoring a process, to assess its two states: the first, when the process, named IC (In Control), operates under the common causes of variation (variation is always naturally present in any phenomenon) and the second, named *OOC* (Out Of Control), when the process operates under some *assignable causes of variation*. The CCs, using the observed data, allow us to decide if the process is IC or OOC. CCs are a statistical test of hypothesis for the process null hypothesis $H_0$={IC} versus the alternative hypothesis $H_1$={OOC}. Control Charts were very considered by Deming [9,10] and Juran [12] after Shewhart invention [19,20].

We start with Shewhart ideas (see the excerpts 3, 3a and 3b).

2. *Statistics to be Used when Quality is Controlled*

When the number $n$ of measurements of some quality $X$ have been made under the conditions of control, we find in general that the function $f$ in (20) can be assumed to be one or the other of the following two forms without introducing practical difficulties:

$$f(x) = \frac{n}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \qquad (22)$$

or

$$f(x) = \frac{n}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \left[1 - \frac{k}{2}\left(\frac{x}{\sigma} - \frac{x^3}{3\sigma^3}\right)\right], \qquad (23)$$

where $x = X - \bar{X}$.

**Excerpt 3a**. From Shewhart book (1931), on page 89.

In the excerpts, $\bar{X}$ is the (experimental) "Grand Mean", computed from D (we, on the contrary, use the symbol $\bar{x}$), $\sigma$ is the (experimental) "estimated standard of each sample" (we, on the

 doi:10.20944/preprints202511.0444.v1



contrary, use the symbol s, with sample size n=20, in excerpts 3a, 3b), $\bar{\sigma}$ is the "estimated mean standard deviation of all the samples" (we, on the contrary, use the symbol $\bar{s}$).

On page 95, he also states that

*"even when nothing is known about the condition under which the distribution was observed, we find that the average and the standard deviation enable us to estimate ... the number of observations lying within any symmetrical range $\bar{X} \pm z\sigma$, where $z>1$. In fact, the proportion of the total number of observed values within any such limits is always greater than $1-1/z^2$. This follows from ... Tchebycheff's Theorem."* He then adds *"we see that no matter what set of n observed values we may have, the number of these values $N_z$ lying within the closed range $\bar{X} \pm z\sigma$ is greater than $n(1-1/z^2)$"*.

$$\bar{X} \pm 3\frac{\sigma}{\sqrt{n}} = 13{,}540 \pm 3\frac{440}{\sqrt{20}}$$

$$\bar{\sigma} \pm 3\frac{\sigma}{\sqrt{2n}} = 423 \pm 3\frac{440}{\sqrt{40}}$$

**Excerpt 3b**. From Shewhart book (1931), on page 294.

So, we clearly see that Shewhart, the inventor of the CCs, *used the data to compute the Control Limits*, LCL (Lower Control Limit) and UCL (Upper Control Limit) both for the mean $\mu_X$ (the 1st parameter of the Normal pdf) and for $\sigma_X$ (the 2nd parameter of the Normal pdf). They are considered the limits comprising 0.9973n of the observed data. Similar ideas can be found in [5–21,25–42] (with Rozanov, 1975, we see the idea that CCs can be viewed as a Stochastic Process).

We invite the readers to consider that *if* one assumes that the process is In Control (IC) and *if* he knows the parameters of the distribution he can test if the assumed known values of the parameters are confirmed or disproved by the data, *then* he does not need the Control Charts; it is sufficient to use NHSTP! (see App. B)

Remember the ideas in the previous section and compare Excerpts 3, 3a, 3b (where *LCL, UCL depend on the data*) with the following Excerpt 4 (where *LCL, UCL depend on the Random Variables*) and appreciate the profound "logic" difference: this is the cause of the many errors in the CCs for TBE [Time Between Events (see [4,5,24]).

Let $X_{ij}$, i=1, 2, …, j=1, 2, …, n be independent and identically distributed (IID) normal $N(\mu_0, \sigma^2_0)$ r.v.'s, where $\mu_0$ and $\sigma^2_0$ are the *specified IC* mean and variance, respectively. The Xij represents the jth observation from the ith rational subgroup (sample) obtained at the i-th sampling stage, and Yi denotes the i-th charting statistic based on the i-th sample. When n≥2, at sampling stage i, $Y_i = \bar{X}_i$ the subgroup mean, is typically used for monitoring the process mean, while in case of individual observations (i.e., n=1) $Y_i=X_{i1}\equiv X_i$. The control limits of the standard Shewhart chart (X chart or the $\bar{X}$ chart) are given by $\mathbf{UCL_1=\mu_Y+k\sigma_Y}$ **and** $\mathbf{LCL_1=\mu_Y-k\sigma_Y}$ (1) where $\mu_Y$ and $\sigma_Y$ are the specified IC mean and standard deviation of the charting statistic Yi, and k denotes the distance of the control limits from the center line (CL) in the units of the standard deviation of the charting statistic. An OOC signal is triggered when for the first time $Y_i \notin [LCL_1, UCL_1]$. It should be mentioned that in the above setup, the control charts are used to monitor the process in real time, by comparing the value of the i-th charting statistic ($\bar{X}_i$ or $X_i$) to the control limits. Therefore, as long as an OOC signal is not triggered, samples are continued to be drawn from the process.

The same type of arguments are used in another paper [4] *JQT*, 2017 where the data are Erlang distributed with $\lambda_0$ is the scale parameter and the Control Limits LCL and UCL are defined [copying Xie et al.] **erroneously** as

$$P[T_r > UCL] = \alpha_0/2 \text{ and } P[T_r < LCL] = \alpha_0/2.$$

 doi:10.20944/preprints202511.0444.v1



**Excerpt 4.** From a paper in the "Garden… [24]". Notice that one of the authors wrote several papers….

The formulae, in the excerpt 4, $LCL_1$ and $UCL_1$ are actually the *Probability Limits* (*L and U*) of the *Probability Interval* PI in the formula (4), when $\varphi(t,\theta,n)$ is the pdf of the Estimator T, related to the Normal model F(x; μ, σ²). Using (4), *those authors do not extract the maximum information from the data in the Process Control*. From the Theory [6–36] we derive that the interval $L=\mu_Y-3\sigma_Y$-----$\mu_Y+3\sigma_Y=U$ is the PI such that the RV $Y=\bar{X}$

$$P[\mu_Y - 3\sigma_Y \leq Y=\bar{X} \leq \mu_Y + 3\sigma_Y] = 0.9973 \quad (7a)$$

and it is not the CI of the mean $\mu=\mu_Y$ [as wrongly said in the Excerpt 4, where actually ($LCL_1$----$UCL_1$)=PI].

The same error is in other books and papers (not shown here but the reader can see in [21–24]).

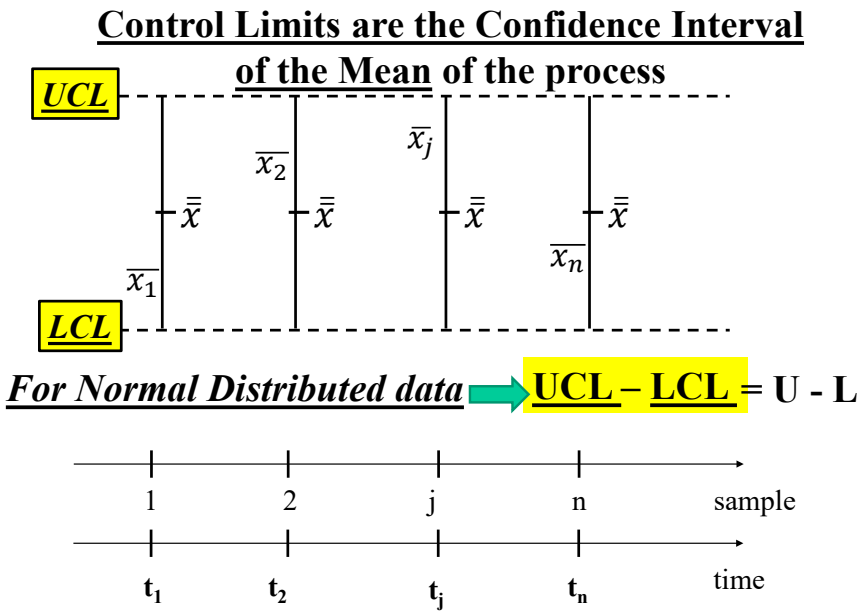



**Figure 2.** Control Limits $LCL_X$---$UCL_X=L$---U (Probability interval), for Normal data (Individuals $x_{ij}$, sample size k) "sample means" $\bar{x}_i$ and "grand mean" $\bar{\bar{x}}$.

The data plotted in the CCs [6–21,25–36] (see the Figure 2) are the means $\bar{x}(t_i)$, determinations of the RVs $\bar{X}(t_i)$, i=1, 2, ..., n (n=number of the samples) computed from the collected data of the i-th sample $D_i=\{x_{ij}$, j=1, 2, ..., k} (k=sample size)}, determinations of the RVs $X(t_{ij})$ at *very close instants* $t_{ij}$, j=1, 2, ..., k. In other applications I-CC (see the Figure 3), the data plotted are the Individual Data $x(t_i)$, determinations of the Individual Random Variables $X(t_i)$, i=1, 2, ..., n (n=number of the collected data), modelling the measurement process (MP) of the "Quality Characteristic" of the product: this model is very general because it is able to consider every distribution of the Random Process $X(t)$, as we can see in the next section. From the excerpts 3, 3a, 3b and formula (5) it is clear that Shewhart was using the Normal distribution, as a consequence of the *Central Limit Theorem (CLT)* [6–20,26–36]. In fact, he wrote on page 289 of his book (1931) "*… we saw that, no matter what the nature of the distribution function of the quality is, the distribution of the arithmetic mean approaches normality rapidly with increase in n* (*his n* is our k), *and in all cases the expected value of means of samples of n* (our k) *is the same as the expected value of the universe*" (CLT in Excerpt 3, 3a, 3b).

 doi:10.20944/preprints202511.0444.v1



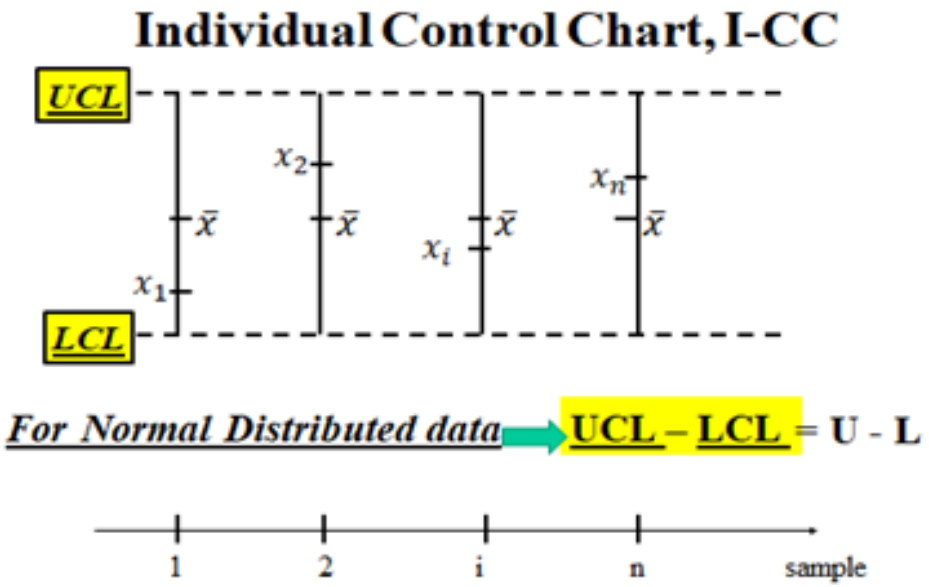



**Figure 3.** Individual Control Chart (sample size k=1). Control Limits LCL┈┈UCL=L┈┈U (Probability interval), for Normal data (Individuals $x_i$) and "grand mean" $\bar{x}$.

Let k be the sample size; the RVs $\bar{X}(t_i)$ are assumed to follow a normal distribution and uncorrelated; $\bar{X}(t_i)$ [i[th] rational subgroup] is the mean of RVs IID $X(t_{ij})$ j=1, 2, ..., k, (k data sampled, at very near times $t_{ij}$).

To show our way of dealing with CCs we consider the process as a "*stand-by system* whose transition times from a state to the subsequent one" are the collected data. The lifetime of "stand-by system" is the sum of the lifetimes of each unit. The process (modelled by a "stand-by …") behaves as a Stochastic Process $X(t)$ [25–33], that we can manage by the Reliability Integral Theory (RIT): see the next section; this method is very general because it is able to consider every distribution of $X(t)$.

If we assume that $X(t)$ is distributed as f(x) [probability density function (pdf) of "transitions from a state to the subsequent state" of a stand-by subsystem] the pdf of the (RV) mean $\bar{X}(t_i)$ is, due the CLT (page 289 of 1931 Shewhart book), $\bar{X}(t_i) \sim N(\mu_{\bar{X}(t_i)}, \sigma^2_{\bar{X}(t_i)})$ [experimental mean $\bar{x}(t_i)$] with mean $\mu_{\bar{X}(t_i)}$ and variance $\sigma^2_{\bar{X}(t_i)}$. $\bar{\bar{X}}$ is the "grand" mean and $\sigma^2_{\bar{\bar{X}}}$ is the "grand" variance: the pdf of the (RV) grand mean $\bar{\bar{X}} \sim N(\mu_{\bar{\bar{X}}}, \sigma^2_{\bar{\bar{X}}})$ [experimental "grand" mean $\bar{\bar{x}}$]. In Figure 2 we show the determinations of the RVs $\bar{X}(t_i)$ and of $\bar{\bar{X}}$.

When the process is Out Of Control (OOC, *assignable causes of variation*, some of the means $\mu_{\bar{X}(t_i)}$, estimated by the experimental means $\bar{x}_i = \bar{x}(t_i)$, are "statistically different)" from the others [6–21,25–36]. We can assess the OOC state of the process via the Confidence Interval (provided by the Control Limits) with CL=0.9973; see the Appendix B. **Remember** the *trick valid **only** for the Normal Distribution …*.; consider the PI, **L**=$\mu_Y$-3$\sigma_Y$┈┈$\mu_Y$+3$\sigma_Y$=**U**; putting $\bar{\bar{x}}$ in place of $\mu_Y$ and $\bar{s}/\sqrt{k}$ in place of $\sigma_Y$ we get the CI of $\mu_{\bar{\bar{X}}}$ when the sample size k is considered for each $\bar{X}(t_i)$, with CL=0.9973. The quantity $\bar{s}$ is the mean of the standard deviations of each sample. This allows us to compare each (subsystem) mean $\mu_{\bar{X}(t_q)}$, q=1,2, …, n, to any other (subsystem) mean $\mu_{\bar{X}(t_r)}$ r=1,2, …, n, and to the (Stand-by system) grand mean $\mu_{\bar{\bar{X}}} = \mu$. If two of them are different, the process is classified as OOC. The quantities $LCL_X = \bar{\bar{x}} - 3\bar{s}/\sqrt{k}$ and $UCL_X = \bar{\bar{x}} + 3\bar{s}/\sqrt{k}$ are the Control Limits of the CC. When the Ranges $R_i$=max($x_{ij}$)-min($x_{ij}$) are considered for each sample we have $LCL_X = \bar{\bar{x}} - A_2\bar{R}$, $UCL_X = \bar{\bar{x}} + A_2\bar{R}$ and $LCL_R = D_3\bar{R}$, $UCL_R = D_4\bar{R}$, where $\bar{R}$ is the "mean range" and the coefficients $A_2$, $D_3$, $D_4$ are tabulated and depend on the sample size k [6–21,25–36].

See the *Appendix B: it is important for understanding our ideas*.

The interval $LCL_X$┈┈$UCL_X$ is the "Confidence Interval" with "Confidence Level" CL=1-α=0.9973 for the unknown mean $\mu_{X(t)}$ of the Stochastic Process X(t) [25–36]. The interval $LCL_R$┈┈┈$UCL_R$ is

 doi:10.20944/preprints202511.0444.v1



the "Confidence Interval" with "Confidence Level" CL=1-α=0.9973 for the unknown Range of the Stochastic Process X(t) [25–36].

***Notice*** that, **ONLY** for normally distributed data, the length of the Control Interval (UCLx-LCLx, which is the Confidence Interval) equals the length of the Probability Interval, PI (U-L): **UCLx-LCLx=U-L**.

The error highlighted, i.e. the *confusion between* the Probability Interval and the Control Limits (Confidence Interval!) has *no consequences* for decisions *when* the data are *Normally distributed*, as considered by Shewhart. On the contrary, it has BIG consequences for decisions ***WHEN the data are Non-Normally distributed*** [4,5,24].

We think that the paper "Quality of Methods for Quality is important", [1] appreciated and mentioned by J. Juran at the plenary session of the EOQC (European Organization for Quality Control) Conference (1989), should be considered and meditated.

### 2.3. *Control Charts for Attributes*

We consider here the papers [3] and "Bayesian Control Chart for Number of Defects in Production Quality Control. *Mathematics* 2024. From the papers we read:

> *If individual inspection units are randomly chosen at evenly spaced intervals of time, the number of nonconformities in the $i^{th}$ inspection will adhere to a Poisson distribution with parameter and is given by* $e^{-\mu}\mu^x/x!$....
>
> *The c-chart with standard unknown. There are situations in practice when the true average number of non-conformities in an inspection unit, c, is unknown or unspecified. This … is referred to as the standard unknown case, or simply Case U. In this case, it is common to estimate c from a set of data, typically from m (mutually) independent inspection units, taken when the process is thought to be in-control. Such a set of data is referred to as reference data or Phase I data, and this phase of the analysis is called the retrospective phase or Phase I.*
>
> With different symbols they give the same ideas:
>
> *The establishment of control limits on the c-chart is typically based on the estimation of the average number of nonconformities* $\bar{c}$ *in an initial sample. The control limits can be defined as follows* $LCL = \bar{c} - \sqrt{\bar{c}}$, $UCL = \bar{c} + \sqrt{\bar{c}}$.... *The estimated 3-sigma control limits of the c-chart are then given by lower control limit (if it is negative it is set to zero).*

**Excerpt 5.** From the a. m. papers.

Since we know that for the *Poisson model*, we have mean E[X]=$\mu$ and variance Var[X]=$\mu$, it is clear that the Control Limits $LCL = \bar{c} - \sqrt{\bar{c}}$, $UCL = \bar{c} + \sqrt{\bar{c}}$ are exactly the same for the Normal distribution (seen above, with k=1, $LCL_X = \bar{\bar{x}} - 3\bar{s}/\sqrt{k}$ and $UCL_X = \bar{\bar{x}} + 3\bar{s}/\sqrt{k}$), assuming the validity of the CLT (Central Limit Theorem), so that we could write the probability statement $P[L = \mu - \sqrt{\mu} \leq X \leq \mu + \sqrt{\mu} = U]$… It is easily seen that LCL and UCL are obtained by puggling in $\bar{c}$ for $\mu$ in the previous probability statement.

This is *Theoretically Wrong*. We name tis the "*classical*" method.

As a matter of fact, letting C be the RV "number of Nonconforming in a sample" (Estimator) and $\bar{\mathbf{c}}$ the "estimate (grand mean)" of the mean from the Theory (6-46) we derive the Figure 1b. Thus LCL┈┈UCL ≠ L┈┈U (Probability interval).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 November 2025 doi:10.20944/preprints202511.0444.v1



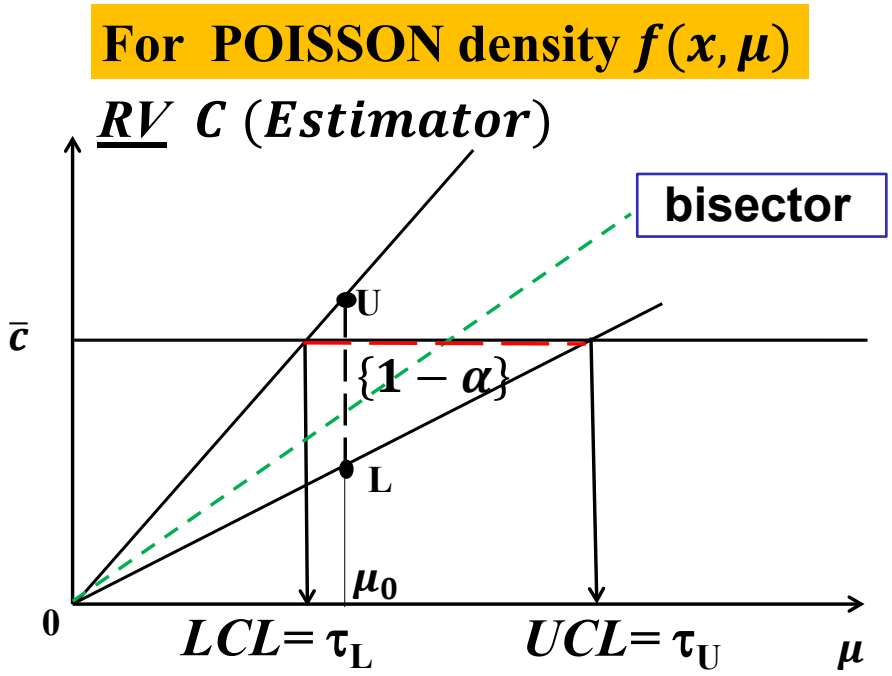



**Figure 1b**. Control Limits for the c Chart. LCL---UCL ≠ L---U (Probability interval) "grand mean" $\bar{c}$.

The error implied in the Chakraborti Control Limits formula is repeated by that author in several other papers on CCs for Exponential distribution [24] (see later).

To illustrate its ideas Chakraborti consider an example (7.3 from Montgomery book 8th ed.). A total of 26 successive inspection samples (consisting of 100 individual units of product) were considered in Phase I to estimate c. Since in this phase 2 samples were OOC they were discarded and then 24 were the samples for the Control Limits. The estimate is $\bar{c} = 19.67$ and using the "*wrong formulae*" LCL1=6.36 and UCL1=32.97; on the contrary, usign the the Theory (Figure 1b) we got LCL2=9.19 and UCL2=36.10; our result, in Figure 4, show the different Control Limits with other 20 new samples were collected for Phase II: Chakraborti and Montgomery conclude "*No lack of control is indicated.*".

Figure 4 says the opposite: the sample 44 is OOC.

Notice that the statitical software JMP does not find the OOC (see the Figure 4b), because it uses the "*wrong formulae*".

We caried out several "10000 simulations" with various μ to check LCL1 and UCL1 versus LCL2 and UCL2; the result is clear: The "*classical*" method *misses real* OOC and *finds non-existent* OOC.

Therefore it is clear that authors, Peer Reviewers and Editors should use the Thoery to analyse the "Scientificness" of papers.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 November 2025 doi:10.20944/preprints202511.0444.v1



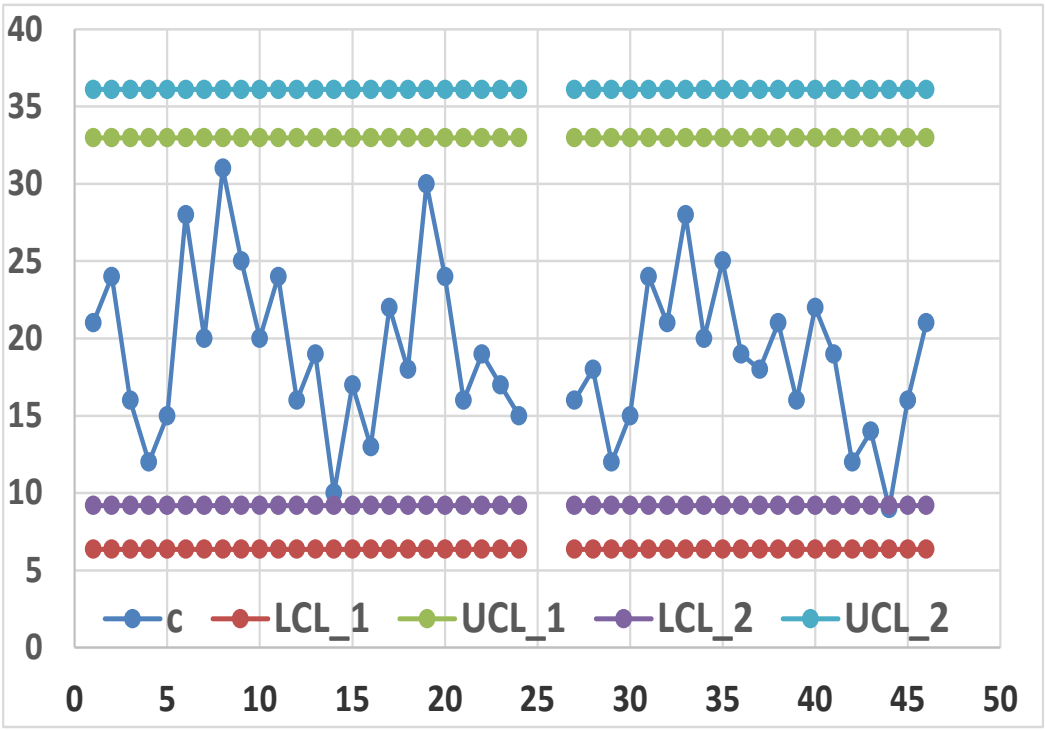



**Figure 4.** Control Limits (by FG) for the c Chart of the Montgomery case.

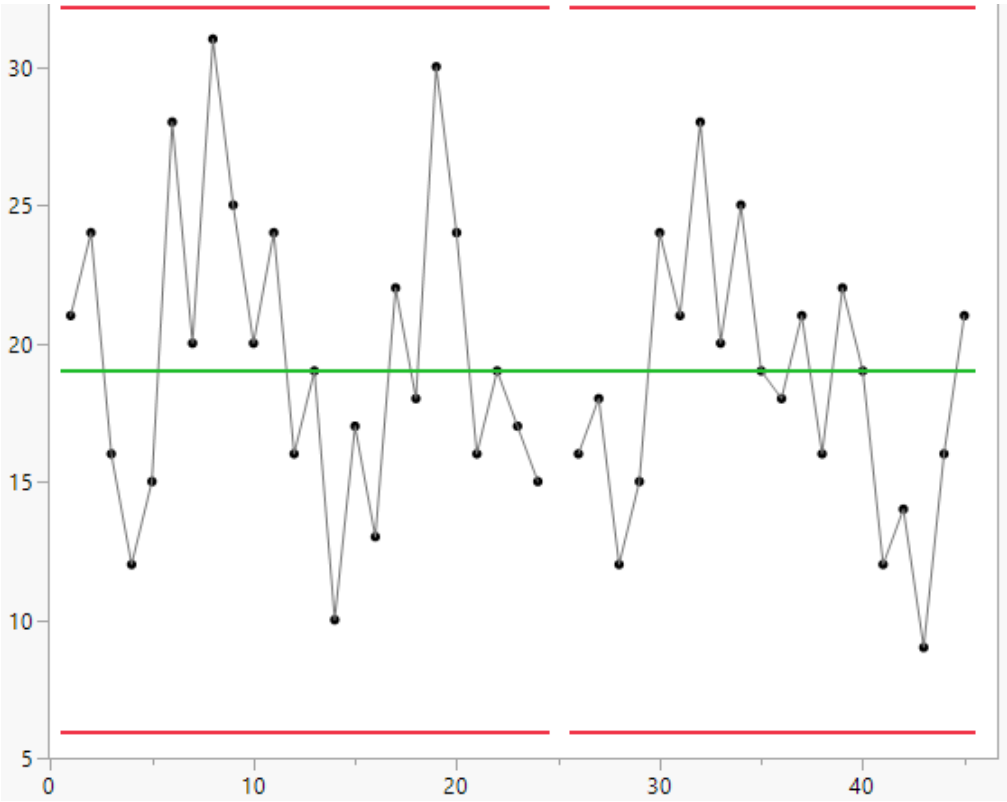



**Figure 4b**. Control Limits (by JMP) for the c Chart of the Montgomery case.

### 2.4. *Statistics and RIT*

We are going to present the fundamental concepts about RIT (Reliability Integral Theory) that we use for computing the Control Limits of CCs. RIT can be found in the author's books…

 doi:10.20944/preprints202511.0444.v1



RIT can be used for parameters estimation and Confidence Intervals (CI), (Galetto 1981, 1982, 1995, 2010, 2015, 2016), in particular for Control Charts (Deming, 1986, 1997, Shewhart 1931, 1936, Galetto 2004, 2006, 2015). In fact, any Statistical or Reliability Test can be depicted by an "Associated Stand-by System" [25–36] whose transitions are ruled by the kernels $b_{k,j}(s)$; we write the *fundamental system of integral equations for the reliability tests*, whose duration t is related to interval 0----t; the collected data $t_j$ can be viewed as the times of the various failures (of the units comprising the System) [$t_0$=0 is the start of the test, t is the end of the test and g is the number of the data (4 in the Figure 4)]

Firstly, we assume that the kernel $b_{j,j+1}(s-t_j)$ is the pdf of the exponential distribution $f(s-t_j\,|\,\mu,\sigma^2)=\lambda e^{-\lambda(s-t_j)}$, where $\lambda$ is the failure rate of each unit and $\lambda=1/\theta$: $\theta$ is the MTTF of each unit. We state that $R_j(t-t_j)$ is the probability that the stand-by system does not enter the state g (5 if we connsider 4 units), at time t, when it starts in the state j (0, 1, …, 4) at time $t_j$, $\bar{W}_j(t-t_j)$ is the probability that the system does not leave the state j, $b_{j,j+1}(s-t_j)ds$ is the probability that the system makes the transition j→j+1, in the interval s----s+ds.

The system reliability $R_0(t)$ is the solution of the mathematical system of the Integral Equations (8)

$$R_j(t-t_j)=\bar{W}_j(t-t_j)+\int_{t_j}^{t} b_{j,j+1}(t-t_j)R_{j+1}(t-s)ds$$
$$\text{for } j=0,1,\ldots,\quad g-1,\quad R_g(t|t_g)=\bar{W}_g(t|t_g) \tag{8}$$

With $\lambda e^{-\lambda(s-t_j)}$ we obtain the solution (see Figure 5, putting the Mean Time To Failure MTTF=θ=123 days, $\lambda=1/\theta$); notice that it is 1-CDF (Poisson) with $\mu=\lambda t$

$$R_0(t)=e^{-\lambda t}[1+\lambda t+(\lambda t)^2/2!+(\lambda t)^3/3!+(\lambda t)^4/4!] \tag{8a}$$

The reliability system (8) can be written in matrix form,

$$R(t-r)=\bar{W}(t-r)+\int_r^t B(s-r)R(s)ds \tag{9}$$

At the end of the reliability test, at time t, we know the data (the times of the transitions $t_j$) and the "*observed*" *empirical sample* D={$x_1$, $x_2$, …, $x_g$}, where $x_j=t_j-t_{j-1}$ is the length between the transitions; the transition instants are $t_j=t_{j-1}+x_j$ giving the "*observed*" *transition sample* D*={$t_1$, $t_2$, …, $t_{g-1}$, $t_g$, t=end of the test} (times of the transitions $t_j$).

We consider now that we want to estimate the unknown MTTF=θ=1/λ of each item comprising the "associated" stand-by system [24–30]: each datum is a measurement from the exponential pdf; we compute the determinant $\det B(s|r;\,\theta,D^*)=(1/\theta)^g\exp[-T(t)]$ of the integral system (9), where $T(t)$ is the "Total Time on Test" $T(t)=\sum_1^g x_i$ [$\boldsymbol{t_0}$ in the Figure 5]: the "Associated Stand-by System" [25–33] in the Statistics books provides the pdf of the sum of the RV $X_i$ of the "*observed*" *empirical sample* D={$x_1$, $x_2$, …, $x_g$}. At the end time t of the test, the integral equations, constrained by the *constraint* D*, provide the equation

$$(\partial lndetB(s|r;\,\theta,D^*))/\partial\theta=\theta/g-T(t)=0 \tag{10}$$

It is important to notice that, in the case of exponential distribution [11–16,25–36], it is exactly the same result as the one provided by the *MLM Maximum Likelihood Method*.

If the kernel $b_{j,j+1}(s-t_j)$ is the pdf $f(s-t_j\,|\,\mu,\sigma^2)=e^{-(s-t_j-\mu)^2/(2\sigma^2)}/(\sqrt{2\pi}\sigma)$ the data are normally distributed, $X\sim N(\mu_X,\sigma_X^2)=e^{-(x-\mu_X)^2/(2\,\sigma_X^2)}/(\sqrt{2\pi}\,\sigma_X)$, with sample size n, then we get the usual estimator $\bar{X}=\sum X_i/n$ such that $E(\bar{X})=\mu_X$.

 doi:10.20944/preprints202511.0444.v1



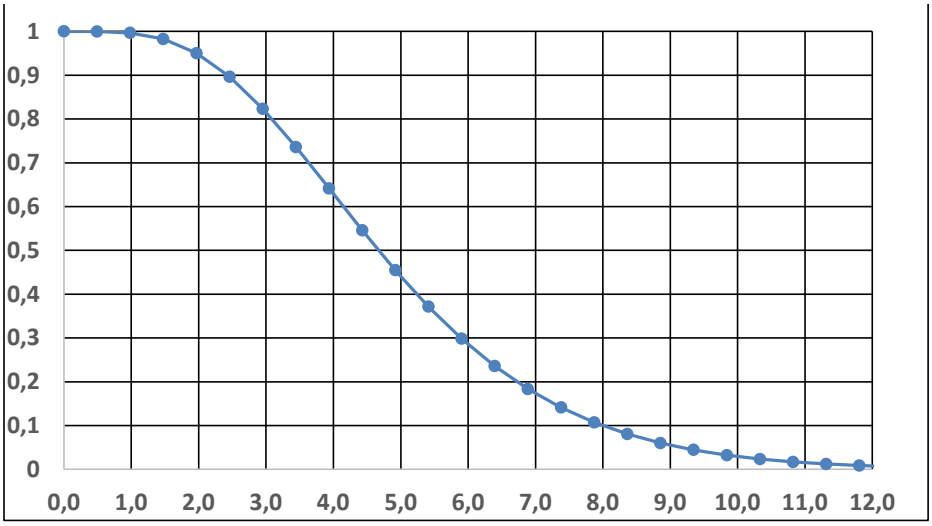



**Figure 5.** Reliability $R_0(\mu = \lambda \mathbf{t_0})$ of a "4 units Stand-by system" with MTTF=θ=123 days; $\mathbf{t_0}$ is the total time on test of the 4 units. To compute the CI (with CL=0.8), find the abscissas of the intersections at $R_0(\mu_L = \lambda_L \mathbf{t_0}) = 0.9$ and $R_0(\mu_U = \lambda_U \mathbf{t_0}) = 0.1$ ….

The same happens with any other distribution provided that we write the kernel $b_{i,i+1}(s)$.

The reliability function $R_0(t|\theta)$, [formula (8)], with the parameter $\theta$, of the "Associated Stand-by System" provides the *Operating Characteristic Curve* (OC Curve, reliability of the system) [6–36] and allows to find the Confidence Limits ($\theta_L$ Lower and $\theta_U$ Upper) of the "unknown" mean $\theta$, to be estimated, for any type of distribution (Exponential, Weibull, Rayleigh, Normal, Gamma, …); by solving, with unknown $\theta$, the two equations $R_0(\boldsymbol{t_0}|\theta) = 1 - \alpha/2$ *and* $R_0(\boldsymbol{t_0}|\theta) = \alpha/2$; we get the two values ($\theta_L$, $\theta_U$) such that

$$R_0(\boldsymbol{t_o}|\theta_L) = \alpha/2 \quad and \quad R_0(\boldsymbol{t_o}|\theta_U) = 1 - \alpha/2 \qquad (11)$$

**where** $\boldsymbol{t_o}$ is the (computed) "total of the length of the transitions $x_i = t_j - t_{j-1}$ data of the *empirical sample* D" and CL=$1 - \alpha$ is the Confidence Level. CI=$\theta_L$------$\theta_U$ is the Confidence Interval: $\theta_L = 1/\lambda_U$ and $\theta_U = 1/\lambda_L$.

For example, with Figure 5, we can derive $\theta_L = 62.5\ days = 1/\lambda_U$ and $\theta_U = 200\ days = 1/\lambda_L$, with CL=0.8. It is quite interesting that the book [14] Meeker et al., "*Statistical Intervals: A Guide for Practitioners and Researchers*", John Wiley & Sons (2017) use the same ideas of FG (shown in the formula 11) for computing the CI; the only difference is that the author FG defined the procedure in 1982 [26], 35 years before Meeker et al.

*2.5. Control Charts for TBE Data. Some Ideas for Phase I Analysis*

Let's consider now TBE (Time Between Event) data, *exponentially or Weibull distributed*. Quite a lot of authors (in the "*Garden …* [24]") *compute wrongly the Control Limits of these CCs*.

The formulae, shown in the section "*Control Charts for Process Management*", are based on the Normal distribution (thanks to the CLT; see the excerpts 3, 3a and 3b); unfortunately, they are used also for NON_normal data (e.g. see formulae (1)): for that, sometimes, the NON_normal data are transformed "with suitable transformations" in order to "produce Normal data" and to apply those formulae (above) [e.g. Montgomery in his book].

Sometimes we have few data and then we use the so called "*Individual Control Charts*" I-CC. The I-CCs are very much *used for exponentially (or Weibull) distributed data*: they are also named "rare events Control Charts for TBE (Time Between Events) data", I-CC_TBE.

The Jarret data (1979) are used also in the paper (*found online, 2024, March 1*) [4] (Kumar, Chakraborti et al. with various presence in the "*Garden …* [24]") who decided to consider the paper [5] (Zhang et al. also present in the "*Garden …*]"): they use the first 30 observed time intervals as phase

 doi:10.20944/preprints202511.0444.v1



1 and start the monitoring at m = 31. You find the original data in the Table 1 in the paper [3]; moreover, for the benefit of the readers we provide them in the section 3 "Results".

It is a very good example for understanding better the problem and the consequences of the difference between PI (Probability Intervals) and the Control Limits, using RIT.

Let's see what the authors say: Kumar, Chakraborti et al. , *Journal of Quality Technology*, 2016, present the case by writing (highlight due to FG):

> *5. An Illustrative Example*
>
> In this Section, we use the data from Table 6 of Zhang et al. (2006) (see also Jarrett 1979) in order to illustrate the application of the proposed charting schemes. The data consist of the time intervals in days between explosions in coal mines (i.e., the events) from 15 March 1981 to 22 March 1962 (*190 observations in total*) in Great Britain. As in Zhang et al. (2006), *we consider* the first m=30 observations to be from the *in-control process*, from which we estimate that $\hat{\lambda}_0 = 0.0081$ (or equivalently, the mean TBE is approximately 123 days). In the sequel, we assume that this is the *true* in-control value $\lambda_0$. Since our numerical analysis showed that r=4 is the best choice we apply the $t_4$*-chart* … Thus, the *remaining 160 observations are first converted by accumulating a set of four consecutive failure times* and the corresponding observations are shown in Table… These are the times until the fourth failure, and are *used for monitoring the process* in order to detect a change in the mean TBE; an increase (which means process improvement) or a decrease (process deterioration).

**Excerpt 5.** From Kumar, Chakraborti et al., "Journal of Quality Technology", 2016.

In the paper of Zhang et al. (2006) we read:

> The second example presented here uses real data taken from Jarrett (1979). The data set consists of time intervals in days between explosions in coal mines from March 15, 1851 to March 22, 1962. The data are reproduced in Table 6. *We first establish the ARL-unbiased exponential chart with the first 30 observations* using the proposed approach. The phase I analysis is presented in Table …. *Then the established control limits are used to monitor the subsequent data*, i.e., from number 31 to number 190. The resultant ARL-unbiased exponential chart for all the data is plotted in Fig…., where the dotted lines represent the *estimated control limits* which stop updating at point number 30 and the continuing straight lines represent the established control limits. It is revealed, surprisingly, that the mean of the *time intervals between explosions remained constant for a very long period of time (about 40 years)*, and *the accident rate only started to decrease sometime after the 125th explosion*. There is an alarm at data point number 80 (value = 0), which is considered here a false alarm.

**Excerpt 6.** From Zhang et al. (2006), "IIE Transactions".

Notice that both the papers [4,5] are (and were) present in the "*Garden* … [24]".

Zhang et al., 2006, compute the Control Limits from the first 30 data and find LCL=0.268 and UCL=1013.9 (you can see them in their Table 7, that is "our" excerpt 11).

 doi:10.20944/preprints202511.0444.v1



All the data [30+40 $t_4$] are very interesting for our analysis; we recap the two important points, given by the authors (Kumar et al.):

1. *… first m=30 observations to be from the* ***in-control process****, from which we estimate … the mean TBE approximately, 123 days*; we name it $\theta_0$.
2. *… we apply the* ***$t_4$-chart****… Thus, … converted by accumulating a set of four consecutive failure times … the times until the fourth failure, used for monitoring the process to detect a change in the mean TBE.*

The 3 authors (Kumar, Chakraborti et al.) state: "*… the control limits … t4-chart are seen to be equal to LCL=63.95, UCL=1669.28 with CL* (Centre Line)*=451.79.*"

Notice that the authors Zhang et al. and Kumar, Chakraborti et al. find *different Control Limits* to be *used for monitoring the same process*: a very interesting situation; the reason is that they use "different" statistics in Phase II.

The FG findings for the Phase I, using the first 30 data, compute different Control Limits with RIT: RIT solves the I-CC_TBE with *exponentially distributed data* as those of Table 1, considered by Zhang et al. and Kumar et al.

In the previous section, we computed the CI=$\theta_L$------$\theta_U$ of the parameter $\theta$, using the (subsample) "transition times durations": $\boldsymbol{t_0}$="total of the *transition times durations* (length of the transitions $x_i=t_j - t_{j-1}$ data) in the *empirical sample* (subsample with n=4 only, as an example)" and Confidence Level CL=$1-\alpha$.

When we deal with a I-CC_TBE we compute the LCL and UCL of the mean θ through the *empirical mean* $\bar{\boldsymbol{t}}_{\boldsymbol{0}} = \boldsymbol{t_0}/n$ of each transition, for the *n=30 data* (*Phase I of Zhang et al. and Kumar et al.*); we solve the two following equations (12) for the two unknown values LCL and UCL, for $R(\bar{\boldsymbol{t}}_{\boldsymbol{0}}|\theta)$ of each item in the sample, similar to (11)

$$R(\bar{\boldsymbol{t}}_{\boldsymbol{0}}|LCL) = \alpha/2, \qquad R(\bar{\boldsymbol{t}}_{\boldsymbol{0}}|UCL) = 1-\alpha/2 \tag{12}$$

**where now** $\bar{\boldsymbol{t}}_{\boldsymbol{0}} = \boldsymbol{t_0}/n$ is the "mean, to be attributed, to the single lengths of the single transitions $x_i=t_j-t_{j-1}$ data in the *empirical sample* D with the Confidence Level CL=$1-\alpha$: $LCL = 1/\lambda_U$ and $UCL = 1/\lambda_L$.

In the next sections we can see the Scientific Results found by a Scientific Theory (we anticipate them: the Control Limits are LCL=18.0 days and UCL=88039.3 days).

## 3. Results

In these sections we provide the scientific analysis of the Jarret data [3] and compare our result with those of Chakraborty [4,5]: the findings are completely different and the decisions, consequently, should be different, with different costs of wrong decisions.

### *3.1. Control Charts for TBE Data. Phase I Analysis*

The Jarret data are in the Table 1.

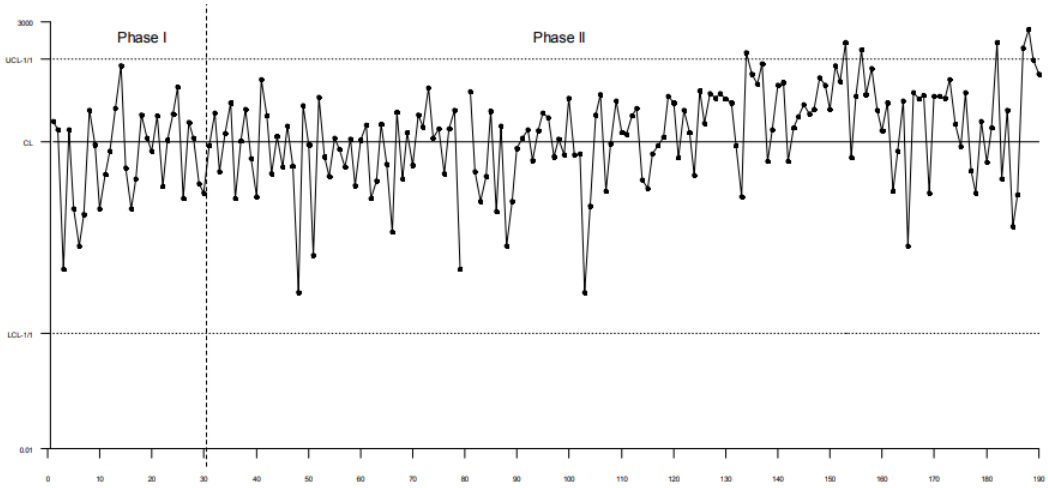
**Table 1.** Data from "A Note on the Intervals Between Coal-Mining Disasters", Biometrika (1979).

 doi:10.20944/preprints202511.0444.v1

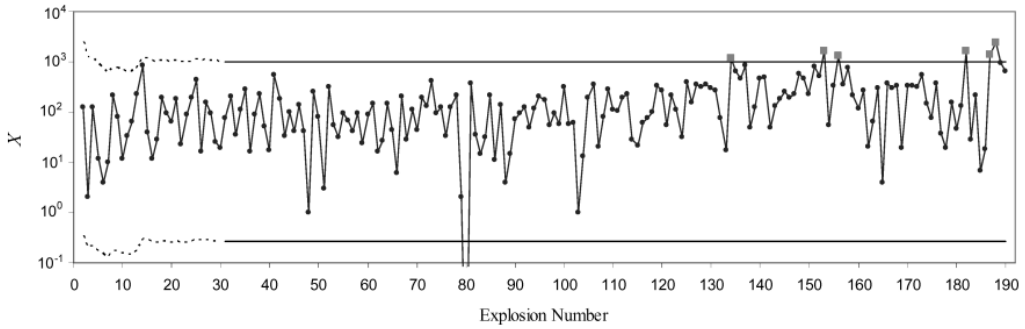


*Time intervals in days between explosions in mines, from 15 March 1851 to 22 March 1962 (to be read down columns)*

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 157 | 65 | 53 | 93 | 127 | 176 | 22 | 1205 | 1643 | 312 |
| 123 | 186 | 17 | 24 | 218 | 55 | 61 | 644 | 54 | 536 |
| 2 | 23 | 538 | 91 | 2 | 93 | 78 | 467 | 326 | 145 |
| 124 | 92 | 187 | 143 | 0 | 59 | 99 | 871 | 1312 | 75 |
| 12 | 197 | 34 | 16 | 378 | 315 | 326 | 48 | 348 | 364 |
| 4 | 431 | 101 | 27 | 36 | 59 | 275 | 123 | 745 | 37 |
| 10 | 16 | 41 | 144 | 15 | 61 | 54 | 456 | 217 | 19 |
| 216 | 154 | 139 | 45 | 31 | 1 | 217 | 498 | 120 | 156 |
| 80 | 95 | 42 | 6 | 215 | 13 | 113 | 49 | 275 | 47 |
| 12 | 25 | 1 | 208 | 11 | 189 | 32 | 131 | 20 | 129 |
| 33 | 19 | 250 | 29 | 137 | 345 | 388 | 182 | 66 | 1630 |
| 66 | 78 | 80 | 112 | 4 | 20 | 151 | 255 | 292 | 29 |
| 232 | 202 | 3 | 43 | 15 | 81 | 361 | 194 | 4 | 217 |
| 826 | 36 | 324 | 193 | 72 | 286 | 312 | 224 | 368 | 7 |
| 40 | 110 | 56 | 134 | 96 | 114 | 354 | 566 | 307 | 18 |
| 12 | 276 | 31 | 420 | 124 | 108 | 307 | 462 | 336 | 1358 |
| 29 | 16 | 96 | 95 | 50 | 188 | 275 | 228 | 19 | 2366 |
| 190 | 88 | 70 | 125 | 120 | 233 | 78 | 806 | 329 | 952 |
| 97 | 225 | 41 | 34 | 203 | 28 | 17 | 517 | 330 | 632 |




**Excerpt 7.** The CC of the 190 data from "Improved Shewhart-Type Charts for Monitoring Times Between Events", Journal of Quality Technology, 2016, (Kumar, Chakraborti et al): the first 30 are used to find the Control Limits for the other 40 $t_4$ (time between "4 failures": 4*40=160).




**Excerpt 8.** The CC of the 190 data [named by the authors " Figure 7. ARL-unbiased exponential chart for the coal mining data"] from "Design of exponential control charts using a sequential sampling scheme", IIE Transactions, (Zhang et al., 2006) [the first 30 data are used to find the Control Limits].

 doi:10.20944/preprints202511.0444.v1



Notice that the authors Zhang et al. and Kumar, Chakraborti et al. find *different Control Limits* to be *used for monitoring the same process*: a very interesting situation; the reason is that they use "different" statistics in Phase II. The results are in the excerpts 7 and 8.

For exponentially distributed data (12) becomes (13) [6–33], k=1, with CL=1 − $\alpha$

$$e^{-[\bar{t}_0/LCL]} = 1 - \alpha/2 \qquad \text{and} \qquad e^{-[\bar{t}_0/UCL]} = \alpha/2 \qquad (13)$$

*The endpoints of the CI=LCL-------UCL are the Control Limits of the I-CC_TBE.*

*This is the right method to extract the "true" complete information contained in the sample (see the Figure 9).*

The Figure 9 is justified by the Theory [6–33] and is related to the formulae [(12), (13) for k=1], for the I-CC_TBE charts.

Remember the book Meeker et al., "*Statistical Intervals: A Guide for Practitioners and Researchers*", John Wiley & Sons (2017): the authors use the same ideas of FG; the only difference is that FG invented 30 years before, at least.

Compare the formulae [(13), for k=1], theoretically derived with a sound Theory [6–33], with the ones in the Excerpt [in the Appendix C (a small sample from the "*Garden …* [24]")] and notice that the two Minitab authors (Santiago&Smith) use the "empirical mean $\bar{t}_0$" in place of the $\theta_0$ in the Figure 1: it is the ***same*** trick of replacing $\bar{\bar{x}}$ to the mean $\mu$ which is valid for the Normal distributed data only; e.g., see the formulae (1)!

*Analysing the first 30 data* of the two articles (Zhang et al. and Kumar, Chakraborti et al.) we get a total $t_0$=3568 days and a mean $\bar{t}_0 = t_0/30$=118.9 days; notice that it is rather different from the value 123 computed by (Kumar, Chakraborti et al.). Fixing α=0.0027, *with RIT we find* the CI={72.6= $\theta_L$-------$\theta_U$ = 220.4} for the parameter $\theta$, and the Control Limits LCL=18.0 days and UCL=88039.3 days.

Compare these with the $LCL_{Zhang}$=0.268 and $UCL_{Zhang}$=1013.9 (Zhang et al., 2006) and the "$LCL_{Kumar}$=*63.95*, $UCL_{Kumar}$=*1669.28 with* $Centre\_Line_{Kumar}$=*451.79*" (Kumar, Chakraborti et al., for *the* $t_4$***-chart***).

Quite big differences… with profound consequences on the decision on the states IC or OOC of the process; the Figure 1, and the "scientific" formula (13) justify our findings.

Now we try to explain why those authors (Zhang et al. and Kumar, Chakraborti et al.) got their results.

In Zhang et al. ("*Design of exponential control charts using a sequential sampling scheme*", *IIE Transactions*); at page 1107, we find the values $L_u$ and $U_u$ versus $LCL_{Zhang}$ and $UCL_{Zhang}$

| | |
|---|---|
| $L_u$=0.286 versus the above value $LCL_{Zhang}$=0.268 | $U_u$=966.6 versus the above value $UCL_{Zhang}$=1013.9 |

We do not know the cause of the "little" difference.

The interesting point is that with these Control Limits the Process "appears" IC (In Control), for the *first m=30 observations*, Phase I; see the excerpt 11.

So, *one is induced to think that* the mean $\bar{t}_0 = t_0/30$=118.9 days can be used to find the $\lambda_0$=1/118.9 for using the Control Limits in the Phase II (the next 160 data) (see the excerpt 12, with the words "***plugging into …***").

Notice that the formulae in excerpt 9 are very similar to those in Appendix C (related to [24])".

This fact generates an IC which is not real. See the Figure 9.

The analysis of the first 30 data show that three possible distributions can be considered, using Minitab 21: Weibull, Gamma, and Exponential; since the Weibull is acceptable, we could use it. See the Table 2.

 doi:10.20944/preprints202511.0444.v1



Given $\lambda_0$ and $\alpha^*$, an exponential chart with LCL, $L_u$, and UCL, $U_u$, determined by Equations (5) and (6) respectively is ARL unbiased, i.e., its ARL curve achieves its maximum value at $\lambda = \lambda_0$:

$$L_u = \frac{-\ln(1-\alpha^*/2)}{\lambda_0}\gamma_{\alpha^*}, \quad (5)$$

$$U_u = \frac{-\ln(\alpha^*/2)}{\lambda_0}\gamma_{\alpha^*}, \quad (6)$$

where $\gamma_{\alpha^*} = \frac{\ln(1-(\alpha^*/2))/\ln(\alpha^*/2)}{\ln(\alpha^*/2)/(1-(\alpha^*/2))}$; see Appendix for a proof.

$\alpha$=0.0027, $\alpha^*$=0.00372, $\gamma_{\alpha^*}$=1.2927

**Excerpt 9.** Formulae for the Control Limits for the first 30 data, IIE Transactions, (Zhang et al., 2006).

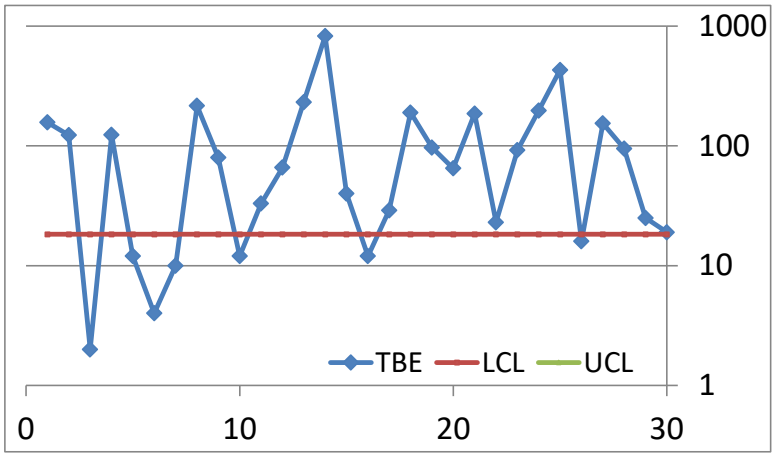



**Figure 6.** The CC of the first m=30 observations used to find the Control Limits (logarithmic scale; only LCL shown). RIT used (exponential distribution, in spite of Table 2…): only LCL is shown because UCL≫1000.

Anyway, for comparison with Zhang et al., 2006, we use the exponential distribution, which is not the "best" to be considered: as you can see the Process is OOC, with 7 points below the LCL.

Therefore, these data should be discarded for the computation of $\lambda_0$.

Hence, the Control Limits (Zhang et al., 2006), based on the estimate "assumed as true" $\lambda_0$=0.0081

| | |
|---|---|
| $L_u$=0.286 and $LCL_{Zhang}$=0.268 | $U_u$=966.6 and $UCL_{Zhang}$=1013.9 |

cannot be used for the next 160 data.

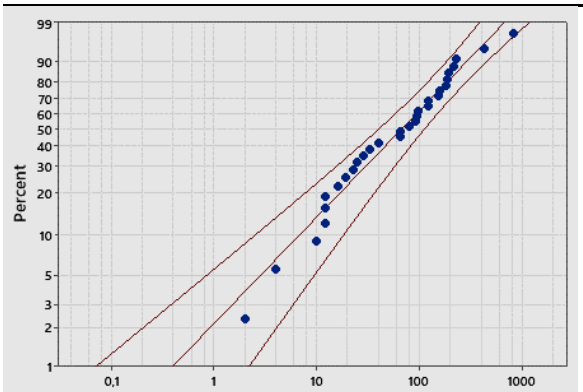
Is the statement (assumption!) "… *first m=30 observations to be from the in-control process…*" sound? NO! The Figure 9 proves [formulae (13)] that the process is OOC.

Using the formulae in the Excerpt 9, *those authors do not extract the maximum information from the data in the Process Control*.

**Table 2.** Estimation of the possible distributions for the first m=30 observations.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 November 2025 doi:10.20944/preprints202511.0444.v1

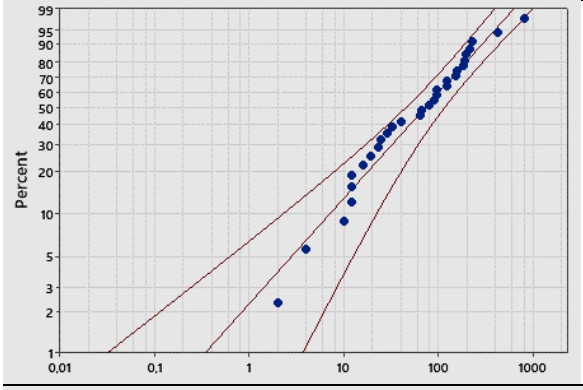





Weibull (95%)

| | |
|---|---|
| Shape | 0,8215 |
| Scale | 105,9 |
| N | 30 |
| AD | 0,290 |
| P-Value | >0,250 |

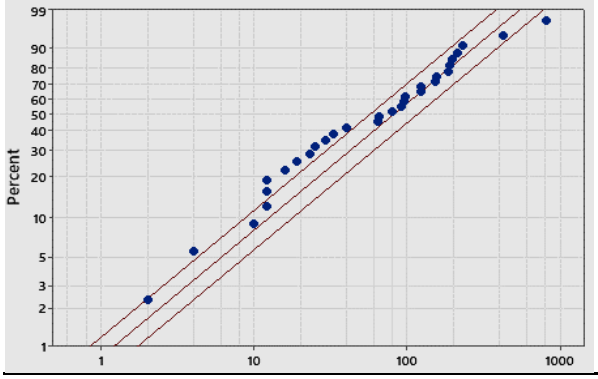
Scale: 105.94; Shape: 0.82




Gamma (95%)

| | |
|---|---|
| Shape | 0,7650 |
| Scale | 155,5 |
| N | 30 |
| AD | 0,356 |
| P-Value | >0,250 |

Scale: 155.47; Shape: 0.77

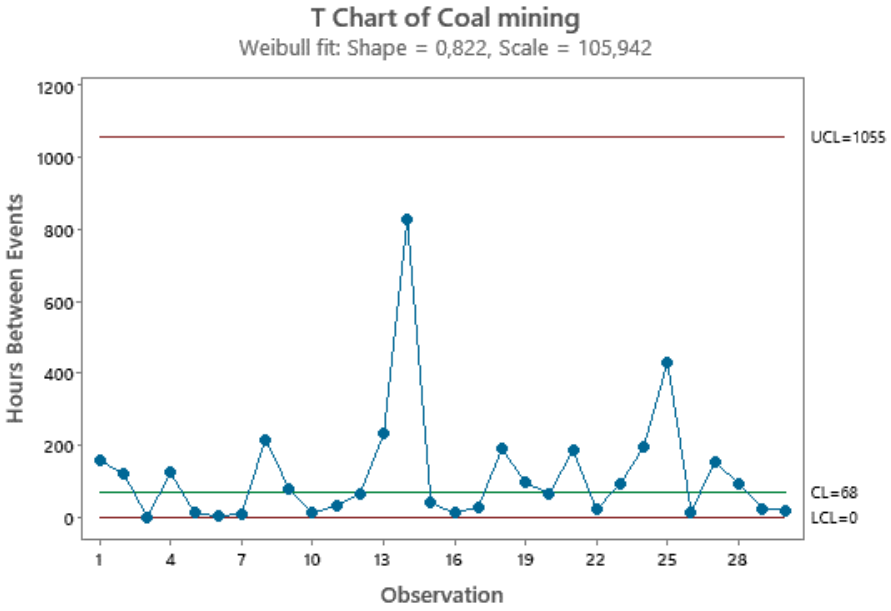



Exponential (95%)

| | |
|---|---|
| Mean | 118,9 |
| N | 30 |
| AD | 0,902 |
| P-Value | 0,148 |

Scale: 118.93, Shape: 1

Before ending this section, let's see what MINITAB, which use the ideas of Santiago & Smith, provides us in Phase I (Figure 7a)



 doi:10.20944/preprints202511.0444.v1



**Figure 7a**. MINITAB CC of the first m=30 observations used to find the Control Limits (Minitab uses the formulae in the Appendix C, applied to Weibull distribution); process IC due to wrong Control Limits.

Notice that JMP (using the ideas of Santiago & Smith), provides us in Phase I the same type of information (Figure 7b)

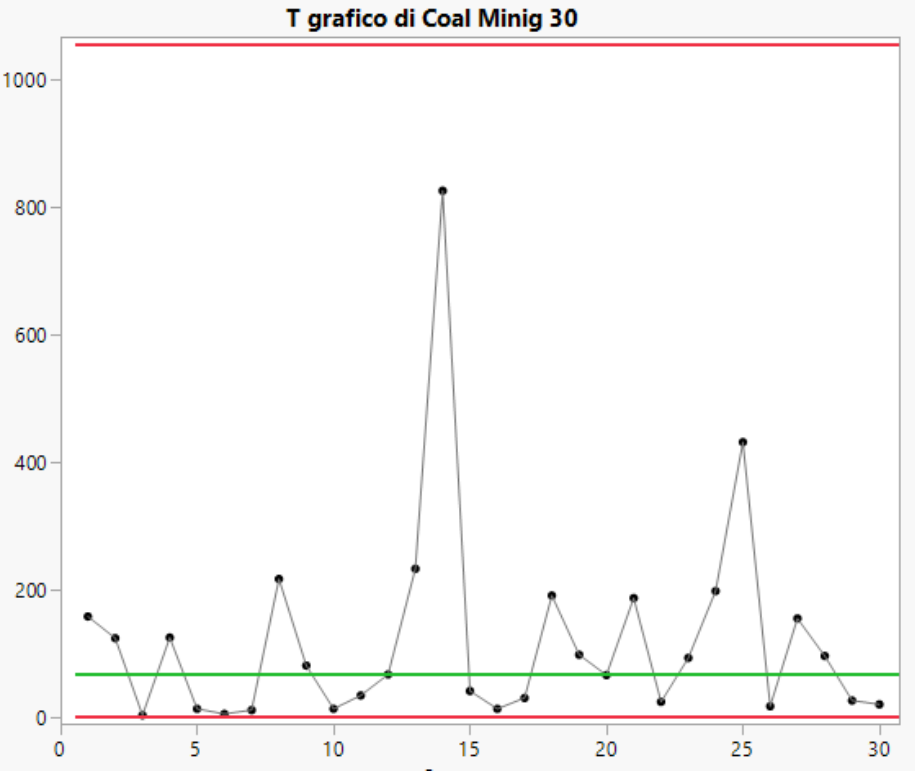



**Figure 7b**. JMP CC of the first m=30 observations used to find the Control Limits (JMP uses the formulae in the Appendix C, applied to Weibull distribution); process IC due to wrong Control Limits.

For the software Minitab, the process is IC, the same as Zhang et al. and Kumar et al.; same result could have been found by JMP (Appendix B) and SAS, and all the authors in the "*Garden* [24]"…

### 3.2. *Control Charts for TBE Data. Phase II Analysis*

We saw in the previous section what usually it is done during the Phase I of the application of CCs: estimation of the mean and standard deviation; later, their values are assumed as "true known" parameters of the data distribution, in view of the Phase II.

In particular, for TBE individual data the exponential distribution is assumed with a known parameter $\lambda_0$ or $\theta_0$.

We consider now what it is done during the Phase II of the application of CCs for TBE data individual exponentially distributed.

We go on with the paper "Improved Shewhart-Type Charts for ….", *Journal of Quality Technology*, 2016, (Kumar, Chakraborti et al. with various presence in the "*Garden* …") who analysed the Jarret data. In their paper we read:

 doi:10.20944/preprints202511.0444.v1



Thus we focus on Shewhart-type TBE charts and try to improve their performance. The process is said to be in-control (IC) when $\lambda=\lambda_0$, where $\lambda_0$ is the given (known) or specified value of the failure rate. It should be noted here that when $\lambda_0$ is not known from previous knowledge, it has to be *estimated from a preliminary IC sample*. In the sequel, *we assume that* $\lambda_0$ is known or that it has been estimated from a (sufficiently) large Phase I sample. Omissis….. Therefore, following Kumar and Chakraborti (2014), the UCL and LCL given in (2) can be more conveniently rewritten

$$UCL = \frac{\chi^2_{2r,1-\alpha_0/2}}{2\lambda_0} \text{ and } LCL = \frac{\chi^2_{2r,\alpha_0/2}}{2\lambda_0}, \quad (3)$$

**Excerpt 10.** From "Improved … Monitoring Times Between Events", J. Quality Technology, '16.

They combining 4 data to generate a $\mathbf{t_4}$ chart giving the formulae in the excerpt 10 (with r in place of 4; *notice the authors mentioned…* ).

*Notice* the formulae: the mentioned authors provide *their LCL and UCL* which are actually the Probability Limits L and U of the Probability Interval (PI) and NOT the Control Limits of the Control Chart, as it is easily seen By using the Theory of CIs (Figures 1 and 5).

All the Jarret data [30+40 $t_4$] are very interesting for our analysis; we recap the two important points, given by the authors (Chakraborti et al.):

1. *… first m=30 observations to be from the in-control process, from which we estimate … the mean TBE approximately, 123 days*; we name it $\theta_0$.
2. *… we apply the* ***$t_4$-chart****… Thus, … converted by accumulating a set of four consecutive failure times … the times until the fourth failure, used for monitoring the process to detect a change in the mean TBE.*

The 3 authors (Chakraborti et al.) state: "*… the control limits … t4-chart are seen to be equal to LCL=63.95, UCL=1669.28 with CL* (Centre Line)*=451.79*", named by them "*ARL-unbiased {1/1, 1/1}*".

The 3 authors (Chakraborti et al.) state also: "*… the control limits … t4-chart are seen to be equal to LCL=217.13, UCL=852.92 with CL* (Centre Line)*=451.79*", named by them "*ARL-unbiased {M:3/4, M:3/4}*".

*Dropping out the OOC data* (*from the first 30 observations*), in Phase I, with RIT, we find that now the process is IC: the distribution fitting the remaining data is the Weibull with parameters $\eta$=140.6 days and $\beta$=1.39; since the CI of the shape parameter is 0.98---2.15, with CL=90%, we can assume $\beta$=1 (exponential with $\theta$=127.9); therefore, we have the "true" LCL=18.6, quite different from the LCLs of the authors (Chakraborti et al.).

Considering the **40** $\mathbf{t_4}$ data, the distribution fitting the data is the Weibull with parameters $\eta$=990.2 days and $\beta$=1.18; since the $1\in$CI of the shape parameter, with CL=90%, we can assume $\beta$=1 (exponential with $\theta$=924.5); therefore, we have the "true" LCL=72.9 and UCL=1987, quite different from the Control Limits of the authors (Kumar, Chakraborti et al.): hence there is a profound consequence on the analysis of the *t4-chart*; see the Figure 12.

 doi:10.20944/preprints202511.0444.v1



Thus, we consider the (modified) $t_r$-chart control limits

$$UCL = \frac{\chi^2_{2r,1-p^*}}{2\lambda_0} \text{ and } LCL = \frac{\chi^2_{2r,\gamma p^*}}{2\lambda_0}, \quad (5)$$

where $\gamma \geq 1$ and $p^* \in (0,1)$. The $CL$ is as in (4). Thus, the probability of a charting statistic $t_r$ plotting above the $UCL$ is $p^*$ and plotting below the $LCL$ is $\gamma p^*$, respectively. The constant $\gamma$ may be viewed as an adjustment factor. If $\gamma = 1$, we have the original $t_r$-chart with equal tail probability limits, with the probability in each of the two tails beyond the control limits (i.e., above

with $p^* + \gamma p^* = (1+\gamma)p^* = \alpha_0$.

**Excerpt 11.** From "Improved … Monitoring Times Between Events", J. Quality Technology, '16.

Considering, on the contrary, the value θ=127.9 from the *first* [<30] *INDIVIDUAL observations of the IC process*, transformed into the one for the $\mathbf{t_4}$ chart, we have the "true" LCL=40.38 and UCL=1100.37; so, we have 4 LCLs and 4 UCLs (see Figure 12).

The 3 authors (Kumar, Chakraborti et al.) show the formulae for the Control Limits (excerpt 11):

Now we face a problem: in which way could the 3 authors (Kumar, Chakraborti et al.) compute *their* Control Limits from the individual *first m=30 observations*?

FG (in spite of excerpt 11) did not find the way: is $\theta_0=1/\lambda_0$ the value estimated from the *first m=30 observations*? He *suspects* that they used a trick (not shown): they use the fist 30 data to find $\theta_0$=123 days (from the total $\mathbf{t_0}$ of days in the Phase I) and then, they consider the total as though it were $\mathbf{4t_0}$, i.e. 30 $t_4$, to find LCL and UCL for the ***$t_4$-chart*** … We did it, but we could not find them…

Doing that, Chakraborti et al. missed the fact that the process, in the individual *first m=30 observations* was OOC and they should not use the "$\boldsymbol{\lambda_0}$ specified value of the failure rate, that has to be *estimated from a preliminary IC sample*."

Since the data are assumed (by the 3 authors) Exponentially distributed it follows from the Theory that "… *by accumulating a set of four consecutive failure times (times until the fourth failure)…*" we should find the $t_4$ data (determinations of the RV $T_4$) Erlang distributed. Since k=4 we could expect that the CLT could apply and find that the 40 $t_4$ data (determinations of the 40 RV $T_4$) follow "approximately" the Normal distribution.

Considering the 40 $t_4$ data and searching for the distribution, we are disillusioned: the distributions Normal, Lognormal, Exponential, Gamma, Smallest Extreme, Largest Extreme, Logistic, Loglogistic do not fit the data; only the Weibull, and the Box-Cox and Johnson transformations seem adequate… See the Gamma (Erlang) fitting in the Figure 8.

 doi:10.20944/preprints202511.0444.v1



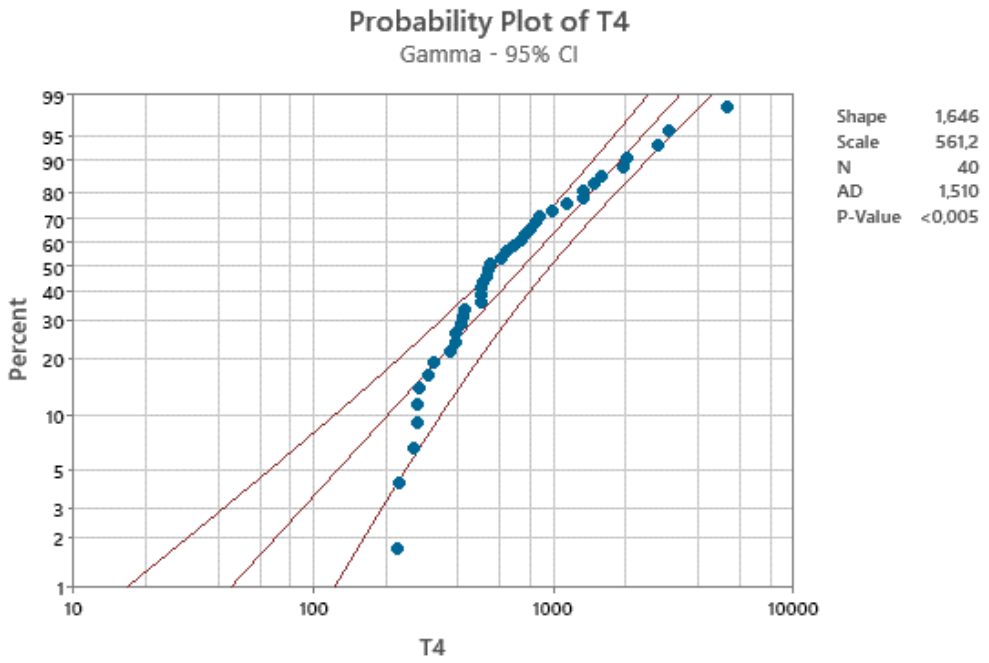



**Figure 8.** Fitting of the Gamma distribution to the **40 $t_4$** data: the Erlang expected distribution is not applicable to the **40 $t_4$** data.

Therefore, the authors formulae for the Control Limits are inadequate in three ways, because

a) the gamma (Erlang) distribution does not apply, with CL=95%

b) then, the formulae in the excerpt 11 cannot be applied.

c) the formulae, in their paper, $\boldsymbol{P[T_r > UCL] = \alpha_0/2}$ **and** $\boldsymbol{P[T_r < LCL] = \alpha_0/2}$ are generated by the confusion (of the authors) between LCL and L and UCL and U, as you can see in the Figure 9, based on the non-applicable Gamma distribution; you see the vertical line intercepting the two probability lines in the points L and U such that $\boldsymbol{P[T_r > U] = \alpha_0/2}$ **and** $\boldsymbol{P[T_r < L] = \alpha_0/2}$ versus the horizontal line, at $\boldsymbol{\bar{t}_0}$, intercepting the two lines at LCL and UCL.

It is clear that the two intervals, L----U and LCL----UCL, are *different* and have *different definitions, meaning and length*, through the Theory [6–8,11–16,25–36]. Notice, in the Figure 9, the logarithmic scale for both axes (to have readable intervals).

It should be noted that we drew two horizontal lines, one at $\boldsymbol{\bar{t}_0}$ (the experimental mean) and the other at $\boldsymbol{\theta_0}$ (the assumed known mean) to show the BIG difference between the interception points: we showed the TRUE LCL and UCL; the reader can guess that the WRONG Control Limits (not shown in the figure) have quite different values from the TRUE Control Limits.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 November 2025 doi:10.20944/preprints202511.0444.v1



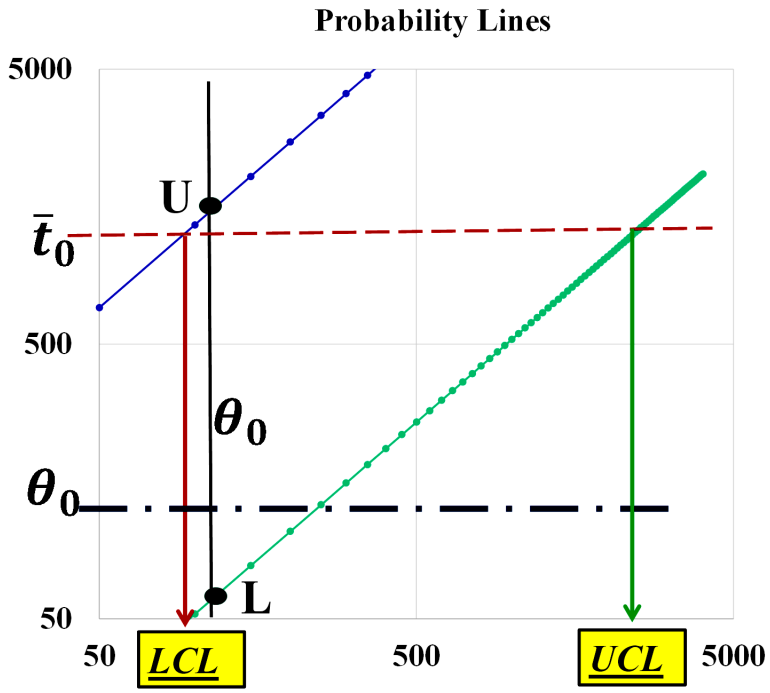



**Figure 9.** The two intervals L---U and LCL---UCL are different (different definitions and meaning). Axes logarithmic (to have readable intervals). Notice the two horizontal lines at $\bar{t}_0$ and at $\theta_0$ ….

The reader is humbly asked to be very attentive in the analysis of the Figure 9: FG thanks him!

Using the original 160 data, divided in 40 samples of size 4, we could compare the estimates with those found by the 40 $t_4$ (in the paper "*Improved … for Monitoring TBE*", *Journal of Quality Technology*, 2016"); you see them in the Table 3.

**Table 3.** Estimation of the possible distributions for the 40 $t_4$ and the Phase II 160 observations.

| *Notice that only the Weibull … is Yes* | | Using 40 $t_4$ data | | Using 160 data | |
|---|---|---|---|---|---|
| Exponential | NO | | 924.02 | | 231.13 |
| *Weibull* | *Yes* | *1.18* | *989.44* | *0.795* | *201.35* |
| Gamma | NO | 1.65 | 561.21 | 0.718 | 322.12 |
| Normal | NO | | | | |
| Normal | NO | | | | |

Notice that the Figure 10a shows the same behaviour as the Phase I figure; it would be interesting to understand IF, with that, the authors would have been able to find an ARL=370: the process is Out Of Control both considering the *first* individual 30 data and the last individual 160 data ….

How many OOC are in the Jarret data?

Analysing the 40 t4 data with Minitab (which uses the Santiago&Smith formulae, Appendix C) we get Figure 11a, and with JMP we get Figure 11b;

 doi:10.20944/preprints202511.0444.v1



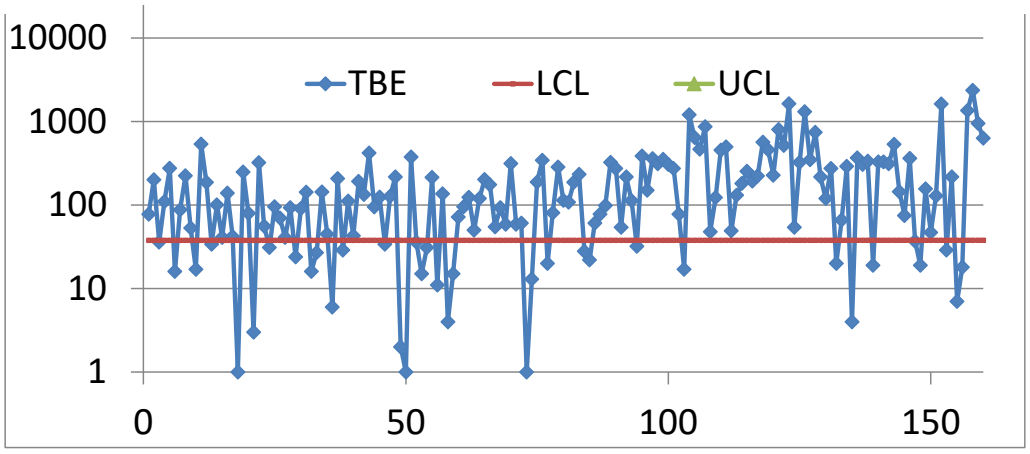



**Figure 10a**. Control Charts and LCL from the last 160 data. Process OOC: 35 points below the LCL. Vertical axis logarithmic. RIT used.

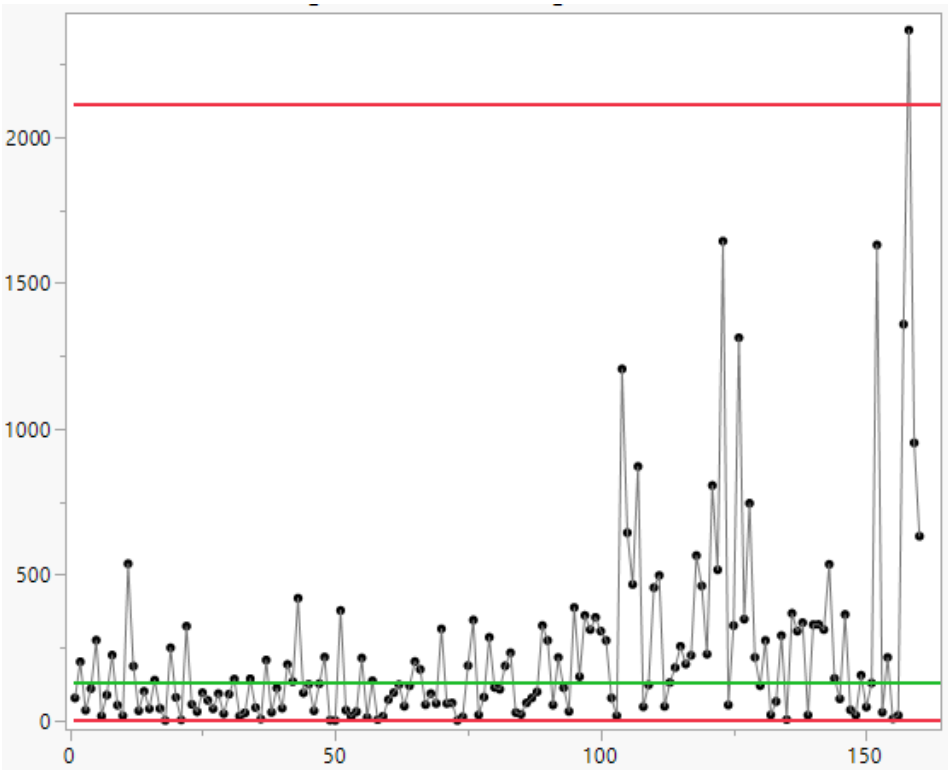



**Figure 10b.** T Chart computed by JMP for the last 160 data. Process IC: opposite to Figure 10 a.

 doi:10.20944/preprints202511.0444.v1



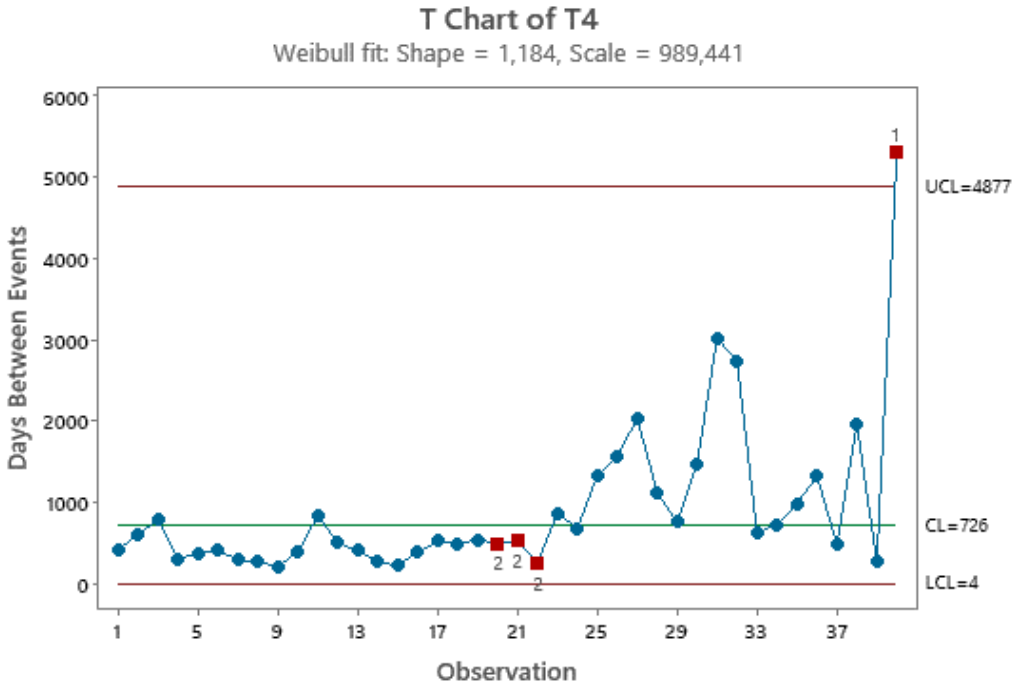



**Figure 11a.** T Chart for the **40 t4** data. Minitab T-Chart (by Santiago & Smith).

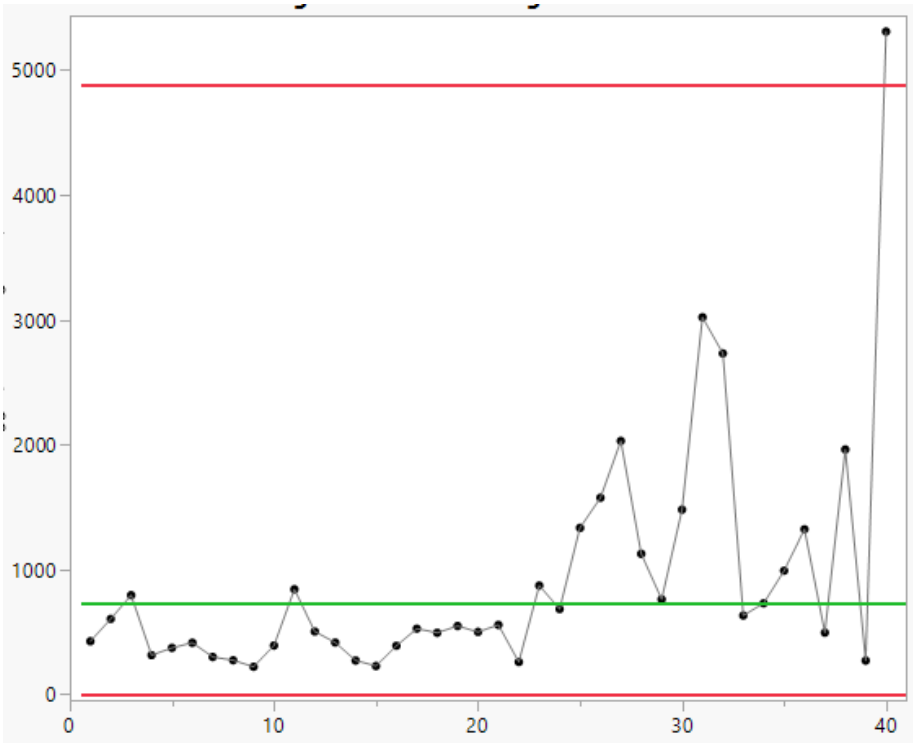



**Figure 11b**. T Chart for the **40 t4** data, computed with JMP; compare with Figure 11a.

See the authors *Acknowledgement* (in the paper)….

Notice that the OOC points in the Figure 10 disappear when we plot the **40 t4**, as you can see in the Figure 12, where we draw 4 LCLs and 4 UCLs, from Kumar, Chakraborti et al. analysis and FG analysis (with RIT). We have the Figure 12.

| LCL_K1 | LCL_K2 | UCL_K1 | UCL_K2 | LCL_G1 | LCL_G2 | UCL_G1 | UCL_G2 |
|---|---|---|---|---|---|---|---|
| 63.95 | 217.13 | 1669.28 | 852.92 | 40.38 | 72.91 | 1100.37 | 1986.96 |

 doi:10.20944/preprints202511.0444.v1



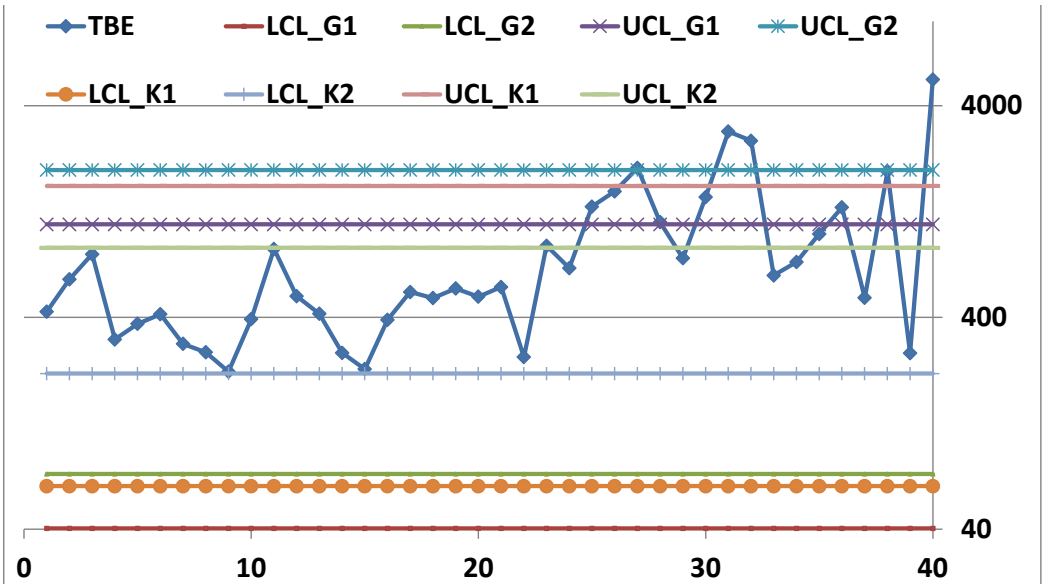



**Figure 12.** Control Charts and LCL from the **40 t4**, data. Notice: we draw 4 LCLs and 4 UCLs.

Notice the Figure 11a,b. Compare it with Figure 12. What can we deduce from this analysis? That the Method is fundamental to draw sound decisions.

Which is the best Method? Only the one which is *Scientific, based on a sound Theory*. Which one?

From the Theory [6–8,11–16,25–36] we know that we must assess the "true (to be used in the Phase II)" Control Limits from the data of an IC Process in the Phase I: therefore, only LCL_G1=40.38 and UCL_G1=1100.37 are the Scientific Control Limits; you can compare with the others in Figure 12.

From Figure 12 we see that the first **20 $t_4$** have a mean lower than the last **20 $t_4$**: the mean time between events increased with calendar time. We can assess that by computing the two mean values: their ratio is 3.18 and we can see if it is significant and we find that it is so with CL=0.9973 (0.0027 probability of being wrong).

See also Figure 13a,b.

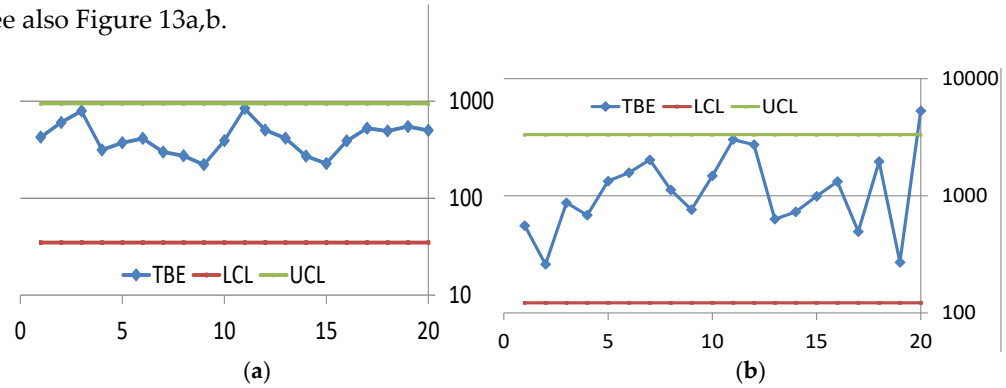



**Figure 13.** (**a**) First 20 t4 data. (**b**) Second 20 t4 data.

From Figures 11a,b, 12 and 13ab we saw that the mean time between explosions was changing with time: it became larger (improved); a method that should show better this behaviour is the EWMA Control Chart. We do not analyse this point; also, for this chart there is the problem of the confusion between the intervals L------U and LCL------UCL.

## 4. Discussion

We first considered the c Chart (paper [3]) and saw that a method exists for better computation of LCL------UCL. Later, we decided to use the Jarrett (1979) data in Table 1 and the analysis by Kumar, Chakraborti, Rakitzis, (2017) in the Journal of Quality Technology [4] and that by Zhang, Xie, Goh

 doi:10.20944/preprints202511.0444.v1



(2006) in the IIE Transactions [5], (papers that you can find also in the "*Garden of flowers*" [24] and in the Appendix C).

We got different results from all those authors: the cause is that they use the Probability Limits of the PI (Probability Interval) as they were the Control Limits (so named by them) of the Control Charts.

The proof of the confusion between the intervals L------U (Probability Interval) and LCL------UCL (Confidence Interval) in the domain of Control Charts (for Process Management) highlight the importance and novelty of these ideas in the Statistical Theory and in the applications.

For the "location" parameter in the CCs, from the Theory, we know that two mean $\mu_{\bar{X}(t_q)}$ (parameter), q=1,2, …, n, and any other mean $\mu_{\bar{X}(t_r)}$ (parameter), r=1,2, …, n, are different, with risk α, if their estimates are not both included in their common Confidence Interval as the CI of the grand mean $\mu_{\bar{\bar{X}}} = \mu$ (parameter) is.

Let's consider the formula (4) and apply it to a "Normal model" (due to CLT, and assuming known variance), sequentially we can write the "real" fixed interval L----U comprising the RV $\bar{\bar{X}}$ (vertical interval) and the Random Interval comprising the unknown mean $\mu$ (horizontal interval) (Figure 14)

$$P\left[L = \mu - \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{k}} \leq \bar{\bar{X}} \leq \mu + \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{k}} = U\right] = P\left[\bar{\bar{X}} - \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{k}} \leq \mu \leq \bar{\bar{X}} + \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{k}}\right] \qquad (14)$$

When the RV $\bar{\bar{X}}$ assume its determination (numerical value) $\bar{\bar{x}}$ (grand mean) the Random Interval becomes the Confidence Interval for the parameter μ, with CL=1-α: risk α that the horizontal line does not comprise the "mean" μ.

This is particularly important for the Individual Control Charts for Exponential, Weibull and Gamma distributed data: this is what Deming calls "*Profound Knowledge* (understanding variation)" [9,10]. In this case, the Figure 14 looks like the Figure 1, where you see the Confidence Interval, the realisation of the horizontal Random Interval.

The case we considered shows clearly that the analyses, in the Process Management, taken so far have been wrong and the decisions have been misleading, when the collected data follow a Non-Normal distribution [24].

Since a lot of papers (related to Exponential, Weibull and Gamma distributions), *with the same problem* as that of "*The garden of flowers*" [24], are published in reputed Journals we think that the title "*History is written by the winners. Reflections on Control Charts for Process Control*" is suitable for this paper: *the authors of the wrong papers* [24] *are the winners*.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 November 2025 doi:10.20944/preprints202511.0444.v1



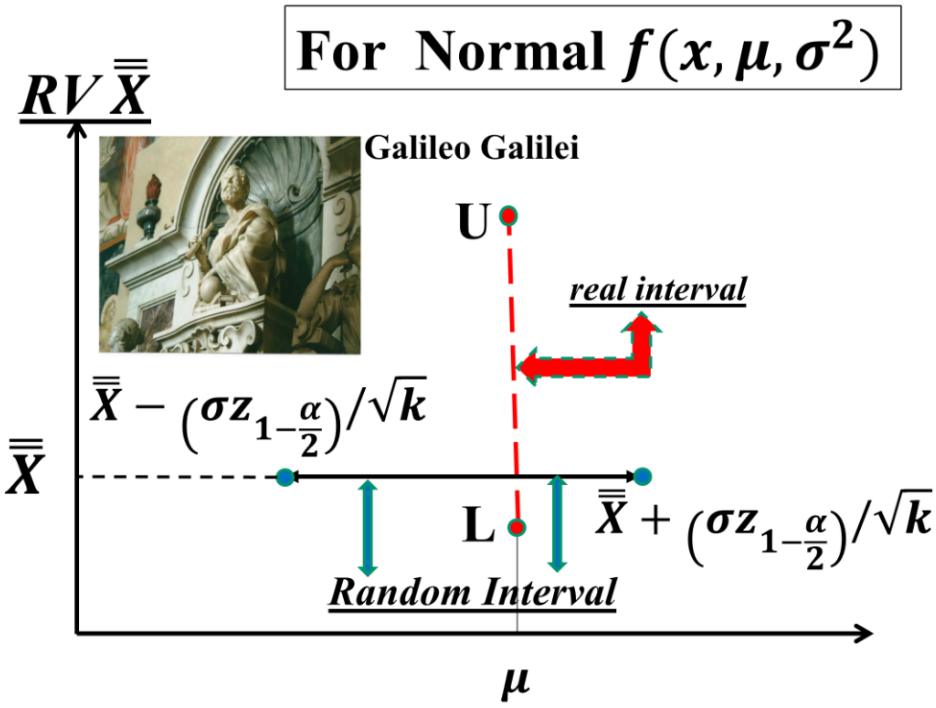



**Figure 14.** Probability Interval L--U (vertical line) versus Random Intervals comprising the "mean" μ (horizontal random variable lines), for Normally distributed RVs $\bar{X} \sim N(\mu, \sigma^2)$.

Our study is limited to the Individual Control Charts with Exponentially, Weibull and Gamma distributed data.

Further studies should consider other distributions which cannot be transformed into the three above distributions: Exponential, Weibull and Gamma.

## 5. Conclusions

With our figures (and the Appendix C, that is a short extract from the "*Garden …* [24]") we humbly ask the readers to look at the references [1–46] and find how much the author has been fond of Quality and Scientificness in the Quality (Statistics, Mathematics, Thermodynamics, …) Fields.

The errors, in the "*Garden …* [24]", are caused by the lack of knowledge of sound statistical concepts about the properties of the parameters of the parent distribution generating the data, and the related Confidence Intervals. For the I-CC_TBE the computed Control Limits (which are actually the Confidence Intervals), in the literature are wrong due to lack of knowledge of the difference between Probability Intervals (PI) and Confidence Intervals (CI); see the Figure 17 (remembering also the Figures 1 and 14). Therefore, the consequent decisions about Process IC and OOC are wrong.

We saw that RIT is able to solve various problems in the estimation (and Confidence Interval evaluation) of the parameters of distributions for Control Charts. The basics of RIT have been given.

We could have shown many other cases (from papers not mentioned here, that you can find in [22–24]) where errors were present due to the lack of knowledge of RIT and sound statistical ideas.

Following the scientific ideas of Galileo Galilei, the author many times tried to compel several scholars to be scientific (Galetto 1981-2025). Only Juran appreciated the author's ideas when he mentioned the paper "Quality of methods for quality is important" at the plenary session of EOQC Conference, Vienna. [1]

For the control charts, it came out that RIT proved that the T Charts, for rare events and TBE (Time Between Events), used in the software Minitab, SixPack, JMP or SAS are wrong. So doing the author increased the h-index of the mentioned authors who published wrong papers.

RIT allows the scholars (managers, students, professors) to find sound methods also for the ideas shown by Wheeler in Quality Digest documents.

 doi:10.20944/preprints202511.0444.v1



We informed the authors and the Journals who published wrong papers by writing various letters to the Editors…: no "Corrective Action", a basic activity for Quality has been carried out by them so far. The same happened for Minitab Management. We attended a JMP forum in the JMP User Community and informed them that their "Control Charts for Rare Events" were wrong: they preferred to stop the discussion, instead to acknowledge the JMP faults [46].

So, dis-quality continues to be diffused people and people continue taking wrong decisions…

Deficiencies in products and methods generate huge cost of Dis-quality (poor quality) as highlighted by Deming and Juran. Any book and paper are products (providing methods): their wrong ideas and methods generate huge cost for the Companies using them. The methods given here provide the way to avoid such costs, especially when RIT gives the right way to deal with Preventive Maintenance (risks and costs), Spare Parts Management (cost of unavailability of systems and production losses), Inventory Management, cost of wrong analyses and decisions.

We think that we provided the readers with the belief that *Quality of Methods for Quality is important*.

The reader should remember the Deming's statements and the ideas in [6–46].

Unfortunately, many authors do not know Scientifically the role (concept) of Confidence Intervals (Appendix B) for Hypothesis Testing.

Therefore, *they do not extract the maximum information form the data in the Process Control*.

Control Charts are a means to test the hypothesis about the process states, $H_0$={Process In Control} versus $H_1$={Process Out Of Control}, with stated risk $\alpha$=0.0027.

We have a big problem about Knowledge: sound Education is needed.

We think that the Figure 16 conveys the fundamental ideas about the need of Theory for devising sound Methods, to be used in real applications in order to avoid the Dis-quality Vicious Circle.

Humbly, given our commitment to Quality, Education, Mathematics anf Physics (versus wrong methods: Fuzzy, Taguchi, Six-Sigma, Control Charts, …) and our long-life love for them [1–46], we would venture to quote Voltaire:

"*It is dangerous to be right in matters on which the established men are wrong.*" because "Many are destined to reason wrongly; others, not to reason at all; and others, to persecute those who do reason." So, "*The more often a stupidity is repeated, the more it gets the appearance of wisdom.*" and "It is difficult to free fools from the chains they revere."

Let's hope that Logic and Truth prevail and allow our message to be understood (Figures 15 and 16).

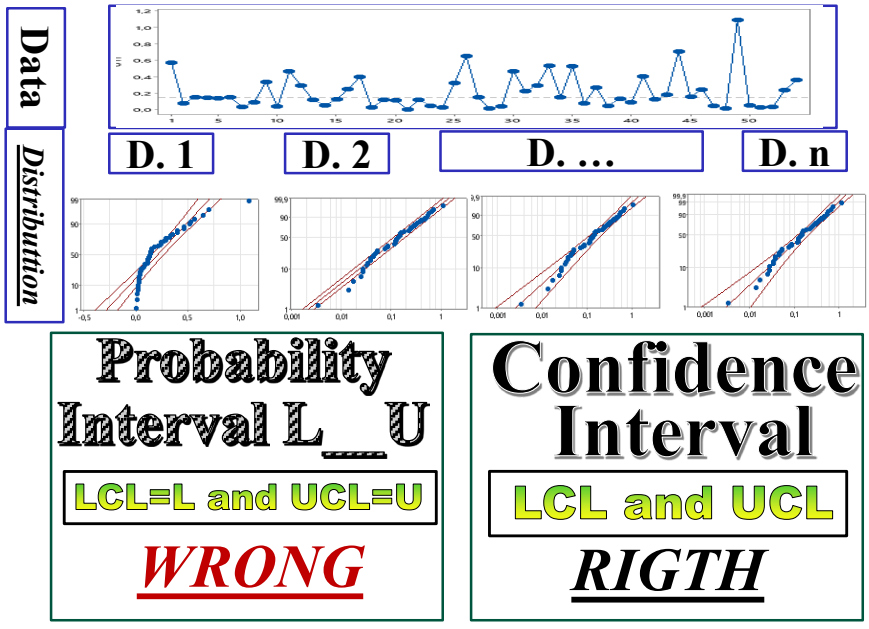



**Figure 15.** Probability Intervals L----U versus Confidence Intervals LCL----UCL in Control Charts.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 November 2025 doi:10.20944/preprints202511.0444.v1



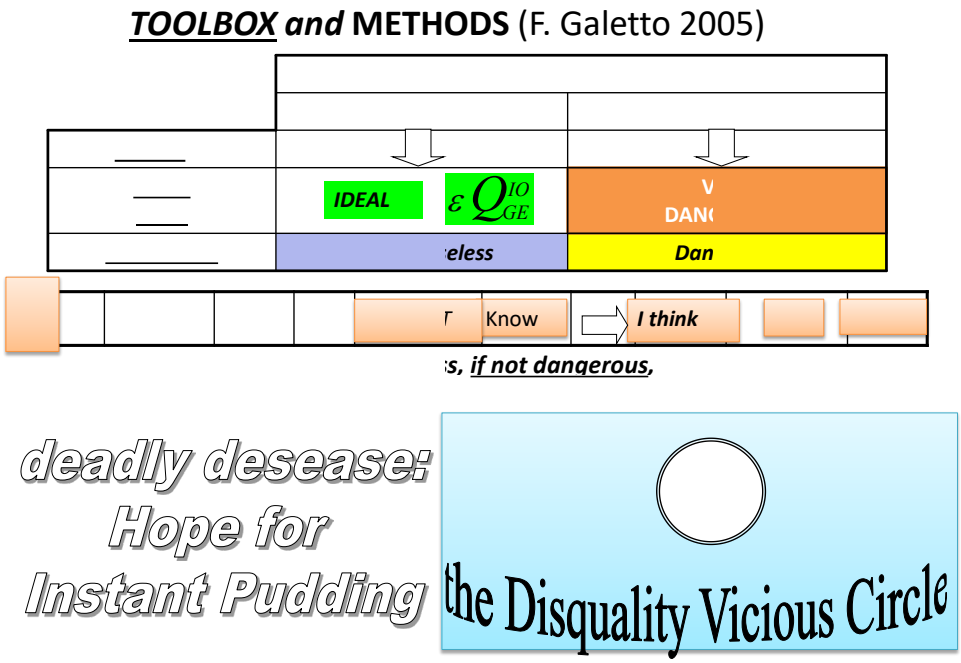



**Figure 16.** Knowledge versus Ignorance, in Tools and Methods.

The objective of collecting and analysing data is to take the right action. The computations are merely a means to characterize the process behaviour. However, it is important to use the right Control Limits take the right action about the process states, i.e., In Control versus Out Of Control.

On July-August 2024 we *again* verified (through **Six?????** *new* downloaded papers) that the *Pandemic Disease about the (wrong) Control Limits, that are actually the Probability Limits of the PI* is still present (notice the Journals):

1. Zameer Abbas et al., (30 June 2024): "Efficient and distribution-free charts for monitoring the process location for individual observations", Journal of Statistical Computation and Simulation,
2. Marcus B. Perry (June 2024) [University of Alabama 674 Citations] "Joint monitoring of location and scale for modern univariate processes", Journal of Quality Technology.
3. E. Afuecheta et al., (2023) "A compound exponential distribution with application to control charts", Journal of Computational and Applied Mathematics [the authors use data of Santiago&Smith (Appendix C) and erroneously find that the UTI process IC].
4. N. Kumar (2019), "Conditional analysis of Phase II exponential chart for monitoring times to an event", Quality Technology & Quantitative Management
5. N. Kumar (2021), "Statistical design of phase II exponential chart with estimated parameters under the unconditional and conditional perspectives using exact distribution of median run length", Quality Technology & Quantitative Management
6. S. Chakraborti et al. (2021), "Phase II exponential charts for monitoring time between events data: performance analysis using exact conditional average time to signal distribution", Journal of Statistical Computation and Simulation
7. S. Chakraborti et al. (2025), "Dynamic Risk-Adjusted Monitoring of Time Between Events: Applications of NHPP in Pipeline Accident Surveillance", downloaded from RG

Other papers with the same problem have been downloaded from June 2024 to October 2025…

There will be any chance that the *Pandemic Disease* ends? See the Excerpt 12: *notice* the (ignorant) words "***plugging into …***". The only way out is Knowledge… (Figure 16): *Deming's* [7,8] *Profound Knowledge, Metanoia, Theory*.

 doi:10.20944/preprints202511.0444.v1



As mentioned before, when the parameter is unknown, it is estimated from a phase I sample of size $m$, say, $Y_1, \ldots, Y_m$ which is collected from an IC process. Recently, Kumar and Jaiswal (2020) used an estimator which is a function of the sample median to show the effect of the presence of outliers in the phase I sample. However, the minimum variance unbiased estimator (MVUE) is a commonly used estimator of the rate parameter which is given by $\hat{\lambda}_0 = \frac{m-1}{T}$, where $T = \sum_{i=1}^{m} Y_i$, the sum of all phase I observations. Thus, the phase II (estimated) control limits of the exponential control chart are given by plugging $\hat{\lambda}$ \(\hat{\lambda}_0\) into the control limits in Equation (1) as follows.

$$\widehat{\mathrm{LCL}} = \frac{A_1}{\hat{\lambda}_0} = \frac{A_1 T}{m-1} \quad \text{and} \quad \widehat{\mathrm{UCL}} = \frac{A_2}{\hat{\lambda}_0} = \frac{A_2 T}{m-1} \tag{4}$$

These control limits are known as the conditional control limits conditioned on a given phase I sample (or a given estimated value of $\lambda_0$). In order to examine the conditional performance of the estimated control chart, we consider the CRL conditioned on a given phase I sample which follows a geometric distribution with parameter $\hat{\beta}(\delta) = P[X < \widehat{\mathrm{LCL}}|\lambda_1] + P[X > \widehat{\mathrm{UCL}}|\lambda_1]$.

**Excerpt 12.** From "Conditional analysis of Phase II exponential chart… an event", Q. Tech. & Quantitative Mgt, '19.

**Funding:** This research received no external funding.

**Data Availability Statement:** "MDPI Research Data Policies" at https://www.mdpi.com/ethics.

**Acknowledgments:** In this section, you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

**Conflicts of Interest:** "The author declares no conflicts of interest."

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| LCL, UCL | Control Limits of the Control Charts (CCs) |
| L, U | Probability Limits related to a probability 1-α |
| θ | Parameter of the Exponential Distribution |
| $\theta_L$-----$\theta_U$ | Confidence Interval of the parameter θ |
| RIT | Reliability Integral Theory |

## Appendix A

*A Very Illuminating Case*

We consider a case found in the paper (with 148 mentions) "*Control Charts based on the Exponential distribution*", Quality Engineering, March 2013, of Santiago&Smith, two experts of Minitab Inc. at that time. You find it mentioned in the "*Garden…*" [24] and in the Appendix C.

This is important because we analysed the data with Minitab software and JMP software and we found astonishing results: the cause are the formulae $LCL = \theta_0 ln(1-\alpha/2) = .00135\ \bar{t}_0$, $UCL = \theta_0 ln(\alpha/2) = 6.6077\ \bar{t}_0$.

The author knew that Minitab computes wrongly the Control Limits of the *Individual Control Chart*. He wanted to assess how the JMP Student Version would deal with them using the following 54 data analysed by Santiago&Smith in their paper; they are "Urinary Tract Infection (UTI) data collected in a hospital"; the distribution of the data is the Exponential.

**Table A1.** UTI data ("Control Charts based on the Exponential distribution").

| | UTI | | UTI | | UTI | | UTI | | UTI | | UTI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.46014 | 11 | 0.46530 | 21 | 0.00347 | 31 | 0.22222 | 41 | 0.40347 | 51 | 0.02778 |

 doi:10.20944/preprints202511.0444.v1



| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 0.07431 | 12 | 0.29514 | 22 | 0.12014 | 32 | 0.29514 | 42 | 0.12639 | 52 | 0.03472 |
| 3 | 0.15278 | 13 | 0.11944 | 23 | 0.04861 | 33 | 0.53472 | 43 | 0.18403 | 53 | 0.23611 |
| 4 | 0.14583 | 14 | 0.05208 | 24 | 0.02778 | 34 | 0.15139 | 44 | 0.70833 | 54 | 0.35972 |
| 5 | 0.13889 | 15 | 0.12500 | 25 | 0.32639 | 35 | 0.52569 | 45 | 0.15625 | | |
| 6 | 0.14931 | 16 | 0.25000 | 26 | 0.64931 | 36 | 0.07986 | 46 | 0.24653 | | |
| 7 | 0.03333 | 17 | 0.40069 | 27 | 0.14931 | 37 | 0.27083 | 47 | 0.04514 | | |
| 8 | 0.08681 | 18 | 0.02500 | 28 | 0.01389 | 38 | 0.04514 | 48 | 0.01736 | | |
| 9 | 0.33681 | 19 | 0.12014 | 29 | 0.03819 | 39 | 0.13542 | 49 | 1.08889 | | |
| 10 | 0.03819 | 20 | 0.11458 | 30 | 0.46806 | 40 | 0.08681 | 50 | 0.05208 | | |

The analysis with JMP software, using the Rare Events Profiler, is in the Figure A1.

NOTICE that JMP, for Rare Events, Exponentially distributed, in the Figure A1, uses the Normal distribution! NONSENSE

It finds the UTI process OOC: both the charts, Individuals and Mobile Range are OOC.

The author informed the JMP User Community.

After various discussions, a member of the Staff (using the Exponential Distribution) provided the Figure A2.

You see that, now (Figure A2), the UTI process is IC: both the charts, Individuals and Mobile Range are IC; opposite decision than before (Figure A1), by the same JMP software (but with two different methods: the first is the standard method, while the second was devised by a JMP Staff member).

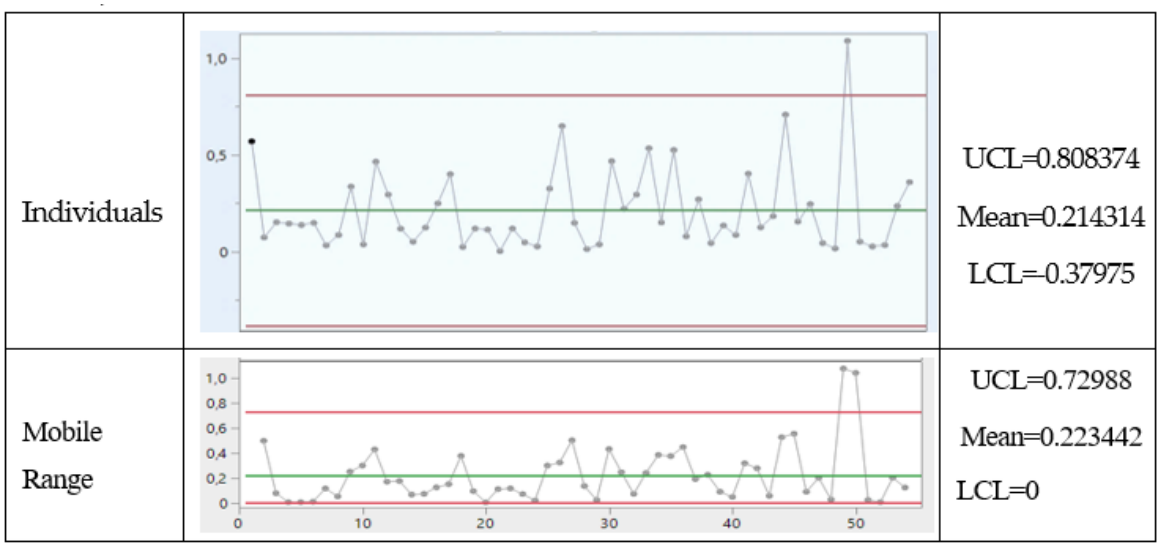



**Figure A1.** First Control Chart by JMP.

 doi:10.20944/preprints202511.0444.v1



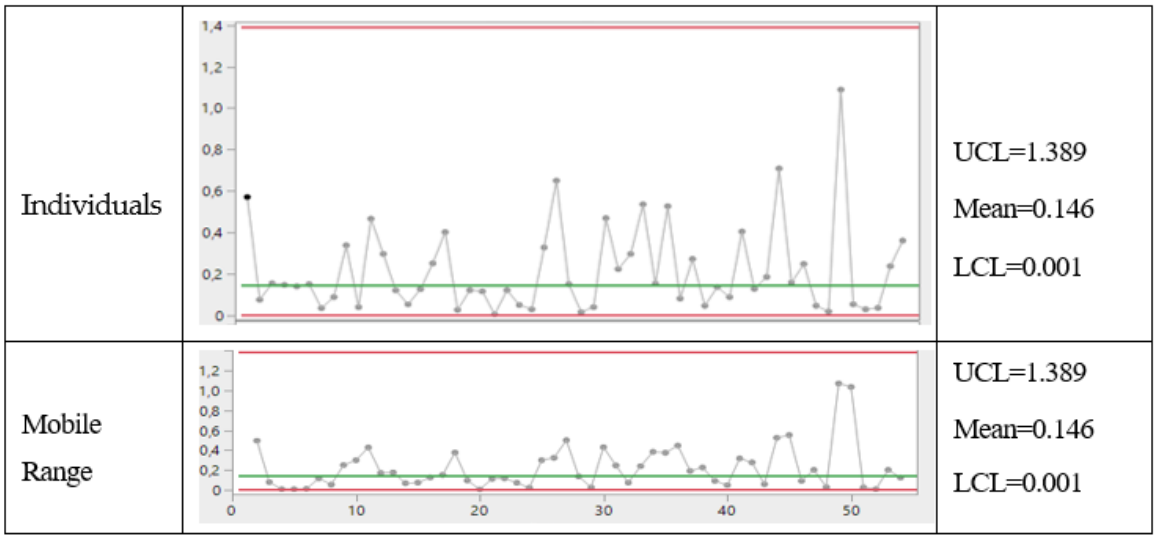



**Figure A2.** Second Control Chart by a member of the Staff of JMP. Notice the numbers (LCL and UCL)!

Notice the LCL, the Mean and the UCL of both charts.

Compute the mean of all the data and you find a different value: therefore, the mean in the charts is not the mean of the process!

If one analyses the data with Minitab, he finds the Figure A3.

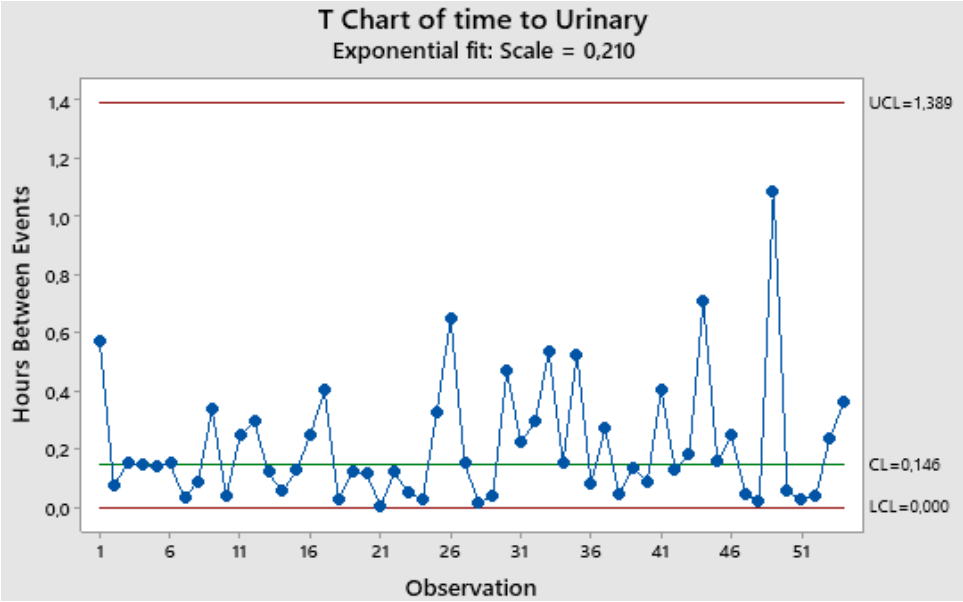



**Figure A3.** Individual Control Chart by Minitab.

You see that now the UTI process is IC: notice the LCL, Mean and UCL.

A natural question arises: which of the three figures is correct?

Actually, they all are wrong, as you can see from the Figure A4:

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 November 2025 doi:10.20944/preprints202511.0444.v1



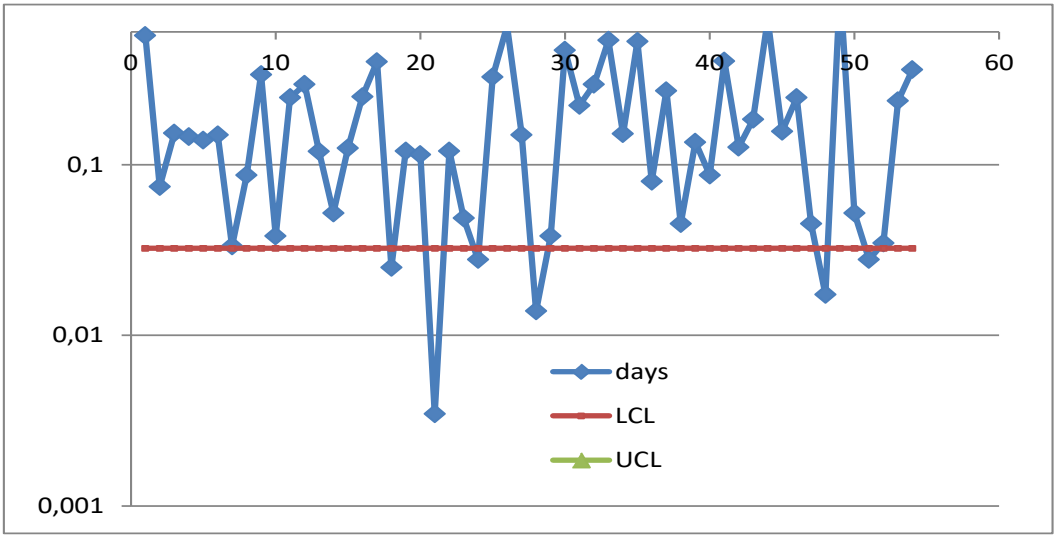



**Figure A4.** Individual Control Chart by FG, using RIT: UTI process OOC.

The author offered JMP to become a better statistical software provider by solving the flaw according to JMP advertising:

> *Our Purpose*
> *Our purpose is to empower scientists and engineers via our statistical discovery software. That's pretty straightforward, and it's never wavered. Sometimes we work directly with the data explorers themselves, sometimes with their companies, and other times with colleges and universities so that the next generation of scientists and engineers will count on JMP by the time they enter the workforce.*
> *Although our purpose is clearly connected to our software, it doesn't end there. As an employer, JMP purposefully creates and maintains a culture of camaraderie that allows the personalities of our diverse and inclusive employee base to shine through. JMP is an equal opportunity employer.*
> *And as inhabitants of this planet, we intentionally measure and work to mitigate our impact on the world. Explore our Data for Green program to see how we're working to empower other organizations to make a difference too.*
> *The catchphrase "corporate social responsibility" could be used for a lot of what we do for our employees, communities, education and the Earth. But starting with JMP founder John Sall, we just try to do the right thing.*

No reaction … and therefore **NO Corrective Action**.

## Appendix B

*The Statistical Hypotheses and the Related Risks*

We define as *statistical hypothesis* a statement about a population parameter (e.g. the "*true*" mean, the "*true*" shape, the "*true*" variance, the "*true*" reliability, the "*true*" failure rate, …). The set of all the possible values of the parameter is called the parameter space Θ. The goal of a *hypothesis test* is to decide, based on a sample drawn from the population, which value *hypothesed* for the population parameter of the parameter space Θ can be accepted as *true*. Remember: nobody knows the truth…

Generally, two competitive hypotheses are defined, the *null hypothesis* $H_0$ and the *alternative hypothesis* $H_1$.

 doi:10.20944/preprints202511.0444.v1



If θ denotes the population parameter, the general form of the null hypothesis is $H_0$: $\theta \in \Theta_0$ versus the alternative hypothesis $H_1$: $\theta \in \Theta_1$, where $\Theta_0$ is a subset of the parameter space $\Theta$ and $\Theta_1$ a subset disjoint from $\Theta_0$. If the set $\Theta_0=\{\theta_0$, a *single value*$\}$ the null hypothesis $H_0$ is called *simple*; on the contrary, the null hypothesis $H_0$ is called *composite*. If the set $\Theta_1=\{\theta_1$, a *single value*$\}$ the alternative hypothesis $H_1$ is called *simple*; on the contrary, the alternative hypothesis $H_1$ is called *composite*.

In a *hypothesis testing problem*, after observing the sample (and getting the empirical sample of the data D) the experimenter (the Manager, the Researcher, the Scholar) must decide either to «accept» $H_0$ as true or to reject $H_0$ as false and then decide, on the opposite, that $H_1$ is true.

Let's make an example: let the reliability goal be $\theta_0$ [θ being the MTTF]; we ask the data D, from the reliability test to confirm the goal we set. Nobody knows the reality; otherwise, there would be no need of any test.

The test data D are the determinations of the random variables related to the items under test; it can happen then that the data, after their elaboration, provide us with an estimate far from $\theta_0$ (and therefore they induce us to decide that the goal has not been achieved).

Generally, in the case of reliability test, the reliability goal to be achieved is called **null hypothesis** $H_0 = \{\theta = \theta_0\}$.

The hypotheses are classified in various manners, such as (and some suitable combinations)

1. Simple Hypothesis: it specifies completely the distribution (probabilistic model) and the values of the parameters of the distribution of the Random Variable under consideration
2. Composite Hypothesis: it specifies completely the distribution (probabilistic model) BUT NOT the values of the parameters of the distribution of the Random Variable under consideration
   a. Parametric Hypothesis: it *specifies completely the distribution* (probabilistic model) *and the values (some or all) of the parameters of the distribution of the Random Variable under consideration*
   b. Non-parametric Hypothesis: it does not *specify the distribution* (probabilistic model) *of the Random Variable under consideration*

A *hypothesis testing procedure* (or simply a *hypothesis test*) is a *rule (decision criterion)* that specifies

1. for which sample values the decision is made to «accept» $H_0$ as true,
2. for which sample values $H_0$ is rejected and then $H_1$ is accepted as true.

based on managerial/Statistics which defines

- the test statistic (a formula to analyse the data)
- the critical region R (rejection region)

to be used for decisions, with the stated risks: *decision criterion*.

The subset of the sample space for which $H_0$ will be rejected is called *rejection region* (or *critical region*). The complement of the rejection region is called the *acceptance region*.

A hypothesis test of $H_0$: $\theta \in \Theta_0$ versus $H_1$: $\theta \in \Theta_1$, ($\Theta_0 \cap \Theta_1 = \varnothing$) might make one of two types of errors, traditionally named Type I Error and Type II Error; their probabilities are indicated as α and β.

**Table B1.** Statistical Hypotheses and risks.

| | **Decision taken due the Reliability test data (sample)** | |
|---|---|---|
| | **Accept** $H_0$ | **Reject** $H_0$ |
| ***UNKNOWN REALITY*** | | |
| $H_0$ ***True*** | **correct decision probability 1-α** | ***Type I error risk α*** |
| $H_0$ ***False, i.e.*** $H_1$ ***True*** | ***Type II error risk β*** | **correct decision probability 1-β** |

 doi:10.20944/preprints202511.0444.v1



If «*actually*» $H_0$: $\theta \in \Theta_0$ is true and the hypothesis test (the *rule*), due to the collected *data*, incorrectly decides to reject $H_0$ then the test (and the Experimenter, the Manager, the Researcher, the Scholar who follow the rule) makes a Type I Error, whose probability is $\alpha$. If, on the other hand, «*actually*» $\theta \in \Theta_1$ but the test (the *rule*), due to the collected *data*, incorrectly decides to accept $H_0$ then the test (and the Experimenter, the Manager, the Researcher, the Scholar who follow the rule) makes a Type II Error, whose probability is $\beta$.

These two different situations are depicted in the previous table (for simple parametric hypotheses).

Notice that when we decide to "*accept the null hypothesis*" in reality we *use a short-hand statement saying that we do not have enough elements to state the contrary*.

*It is evident that*

$$\alpha = P[reject\ H_0 | H_0\ true] \qquad \text{and} \qquad \beta = P[accept\ H_0 | H_0\ false] \qquad \text{(B1)}$$

Suppose R is the rejection region for a test, based on a «*statistic s(D)*» (the formula to elaborate the sampled data D).

Then for $H_0$: $\theta \in \Theta_0$, the test makes a mistake if «$s(D) \in R$», so that the probability of a Type I Error is $\alpha = P(S(D) \in R)$ [S(D) is the random variable giving the result s(D)].

It is important the *power of the test 1-β*, which is the probability of rejecting $H_0$ when *in reality* $H_0$ is false

$$1 - \beta = P[reject\ H_0 | H_0\ false] \qquad \text{(B2)}$$

Therefore, the *power function* of a hypothesis test with rejection region R is the function of $\theta$ defined by $\beta(\theta) = P(S(D) \in R)$. The function *1-power function* is often named the *Operating Characteristic curve* [*OC curve*].

A good test has power function near 1 for most $\theta \notin \Theta_0$ and, on the other hand, near 0 and for most $\theta \in \Theta_0$.

From a managerial point of view, it is sound using powerful tests: *a powerful test (finds the reality and) rejects what must be rejected*.

It is obvious that we want that the test be the most powerful and therefore one must seek for the statistics which have the maximum power; it's absolutely analogous to the search of efficient estimators.

*We know* that the competition of simple hypotheses can have a good property: the most powerful critical region [i.e. the rejection region found has the highest power $1-\beta(\theta) = P(S(D) \notin R)$ of $H_1$ against $H_0$, for any $\alpha$ (*$\alpha$* sometimes is called *size of the critical region*)]; a theorem regarding the *likelihood ratio* proves that.

Let's define the *likelihood ratio* tests; let $\Theta$ denote the entire parameter space; the *likelihood ratio test statistic* for testing $H_0$: $\theta \in \Theta_0$ versus $\theta \in \Theta_1$ is the ratio [which uses the Likelihood function $L(\theta|D)$]

$$\lambda(D) = \frac{sup_{\Theta_0} L(\theta|D)}{sup_{\Theta} L(\theta|D)} \qquad \text{(B3)}$$

A *likelihood ratio test* is any test that has a rejection region that has the form $\{s(D): \lambda(D) \leq c\}$, where c is any number satisfying $0 \leq c \leq 1$ and s(D) is the "statistic" by which we elaborate the data of the empirical sample D. This test is a measure of how much the evidence, provided by the data D, supports $H_0$.

The previous criterion is very simple if the two competing hypotheses are *both simple*: $H_0$: $\theta = \theta_0$ versus $\theta = \theta_1$.

Let $L_0$ be the Likelihood function $L(\theta_0|D)$ and $L_1$ be the Likelihood function $L(\theta_1|D)$: the most powerful test is the one that has the most powerful critical region $C = \{s(D): L_1/L_0 \geq k_\alpha\}$, where the quantity $k_\alpha$ is chosen in such a way that the Type I Error has a risk (probability) $\alpha$. The most powerful critical region C has the highest power $1-\beta(\theta)$.

Usually when an efficient estimator exists, this provides then a powerful statistic, giving the most powerful test.

 doi:10.20944/preprints202511.0444.v1



For the *Normal model*

$$n(x|\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/(2\sigma^2)} \quad \text{(B4)}$$

the test about H0: $\theta\in\Theta_0=\{\mu,\sigma^2:\mu=\mu_0\ ;\ 0<\sigma^2<\infty^2\}$ where $\mu_0$ is a given number, we get

$$\lambda(D)=\left[\frac{1}{1+t^2/(n-1)}\right]^{n/2} \quad \text{(B5)}$$

where $\boldsymbol{t}$ has the t distribution with n-1 degrees of freedom when H0 is true.

After some algebra, the test of H0 may be performed as follows: we compute the quantity $t_c=\left[\sqrt{n(n-1)}\frac{(\bar{x}-\mu_0)}{\sqrt{\sum(x_i-\bar{x})^2}}\right)$ and if

$$-t_{1-\alpha/2}<t_c<t_{1-\alpha/2} \quad \text{(B6)}$$

H0 is accepted; otherwise H0 is rejected.

It is worthwhile to observe that the Confidence Interval for $\mu'$ is

$$CI=\{\bar{x}-t_{1-\alpha/2}s/\sqrt{n}<\mu'<\bar{x}+t_{1-\alpha/2}s/\sqrt{n}\} \quad \text{(B7)}$$

Hence, the test of H0 is equivalent to the following points, *for any distribution of the data*:

1) Construct a confidence interval for the population mean
2) *IF* $\mu_0\in CI$ *THEN* Accept H0; otherwise H0 is rejected.

This has great importance for Control Charts, as you can see in the figure

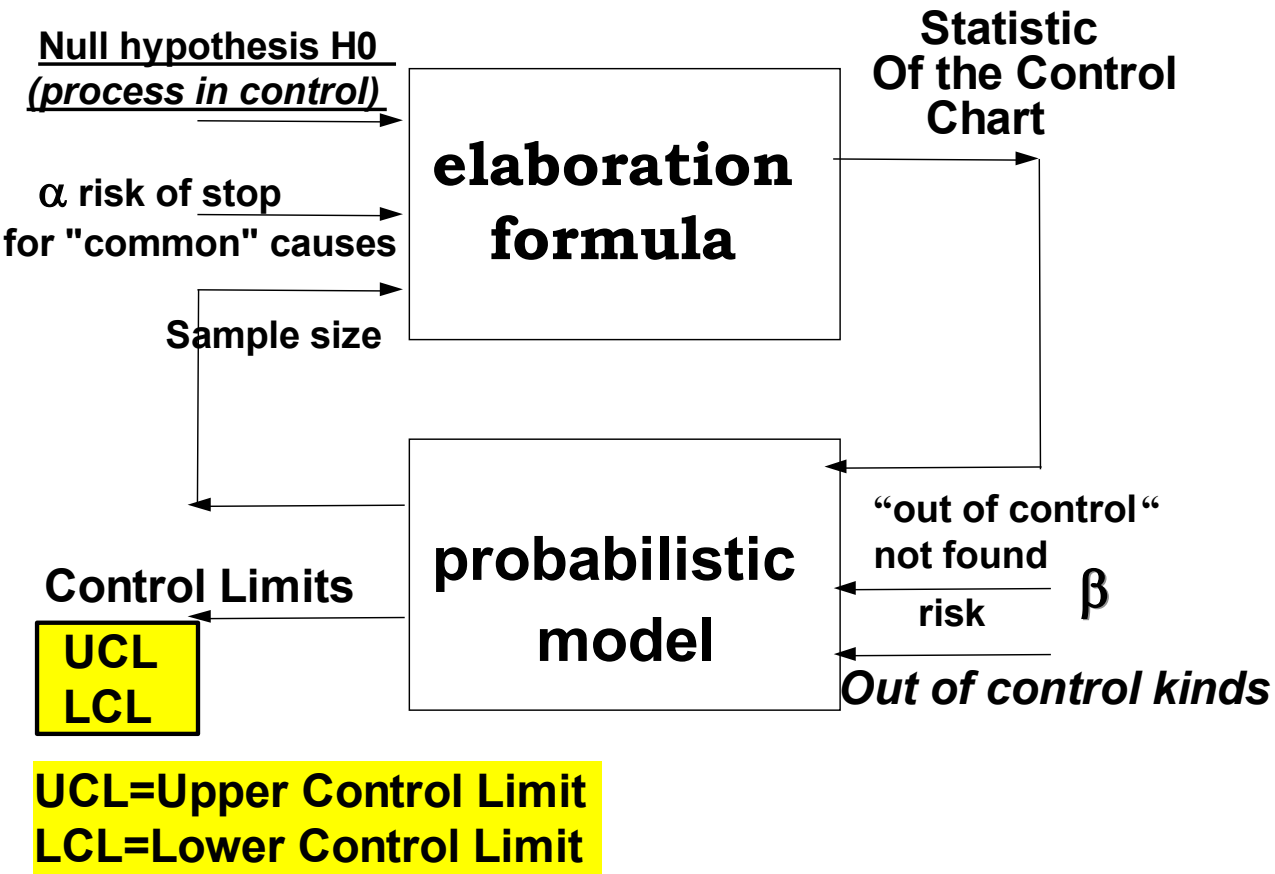



**Figure B1.** LCL and UCL of Control Charts with their risks.

 doi:10.20944/preprints202511.0444.v1



The good Managers, Researchers, Scholars do not forget that the two risks always are present and therefore they must take care of the power of the test 1-β, they use for the decision (*as per the principles F1 and F2)* [24–30].

Such Managers, Researchers, Scholars use the Scientific Method.

It is important to state immediately and in an explicit way that

⇒ the *risks must be stated,*

⇒ *together with the goals (the hypotheses),*

⇒ *BEFORE any statistical (reliability) test is carried out.*

For demonstration of reliability characteristics, with reliability tests, Managers, Students, Researchers and Scholars must take into account, according the F1 principle, the very great importance of W. E. Deming statements

- *A figure without a theory tells nothing.*
- *There is no substitute for knowledge.*
- *There is widespread resistance of knowledge.*
- *Knowledge is a scarce national resource.*
- *Why waste Knowledge?*
- *Management need to grow-up their knowledge because experience alone, without theory, teaches nothing what to do to make Quality*
- *Anyone that engages teaching by hacks deserves to be rooked.*

From these, unfortunately for Quality, for the Customers, for the Users and for the Society, this devastating result

➢ *The result is that hundreds of people are learning what is wrong. I make this statement on the basis of experience, seeing every day the devastating effects of incompetent teaching and faulty applications.*

In many occasions and several *Conferences on Total Quality Management for Higher Education Institutions*, [Toulon (1998), Verona (1999), Derby (2000), Mons (2001), Lisbon (2002), Oviedo (2003), Palermo (2005), Paisley (2006), Florence (2008), Verona (2009)] the author (FG) showed many real cases, found in books and magazines specialized on Quality related to concepts, methods and applications wrong, linked to Quality [21–61]. All the very many documents published (more than 250) by F. Galetto show the profound truth that

*facts and figures are useless, if not dangerous, without a sound theory* (F. Galetto, 2000),

Brain is the most important asset: let's not forget it. (F. Galetto, 2003),

All that is particularly important for the analysis of any type of data (quality or reliability).

## Appendix C

**Typical statement by ALL ...**

 doi:10.20944/preprints202511.0444.v1



A uniform model the exponential TBE charts is that the occurrence of events is modelled by a Poisson process, and the time between events $X_i$ (i=1, 2, …) re independent and identically distributed random variables with pdf $f(x) = \theta^{-1} \exp(-x/\theta)\, for\ x \geq 0$, 0 otherwise, where $\theta$ is the "mean time between events".

The Control Chart plots the quantity produced before observing an event; The Control Limits can be calculated as

$$LCL = \theta \ln(1 - \alpha/2), \qquad UCL = \theta \ln(\alpha/2)$$

**Liu J., Xie M., Sharma P., "A Comparative Study of Exponential Time Between Event Charts", *Quality Technology & Quantitative Management*, 2006 – Issue 3, pp. 347-359**

**ACTUALLY LCL=L and UCL=U**

To construct a t chart, we determine the control limits based on a false alarm rate (α) of 0.0027, equaling that of an individual chart of normal data, and use the median as the centreline". Whenever historical estimates are not available, the scale parameter θ can be estimated using maximum likelihood. …. because both control limits and the centerline are functions of solely θ, by the invariance property of MLEs the estimates are 0.00135 $\bar{t}$, 6.60773 $\bar{t}$, and log(2) $\bar{t}$." .

$$LCL_T = 0.00135\ \bar{t}, \qquad UCL_T = 6.60773\ \bar{t}$$

**E. Santiago, J. Smith, Control charts based on the Exponential Distribution, Quality Engineering, Vol. 25, Issue 2, 85-96**

**ACTUALLY LCL=L and UCL=U**

 doi:10.20944/preprints202511.0444.v1



In a subsequent paper "*Improved Shewhart-Type Charts for Monitoring Times Between Events*", *Journal of Quality Technology*, 2016 (*found online, 2024, March*), we find again the *same error [formula (2)]*:

probability control limits *UCL* and *LCL* of a two-sided $t_r$-chart are respectively given by (see Xie et al. 2002b)

$$P[T_r > UCL] = \alpha_0/2 \text{ and } P[T_r < LCL] = \alpha_0/2.$$

**ACTUALLY LCL=L and UCL=U**

In another paper we found

Suppose LCL and UCL denote the lower and upper control limits of the Phase II $t_r$-chart respectively. Then for a given false alarm rate (FAR) $\alpha_0$, they can be obtained from $P(T_r < LCL|IC) = P(T_r > UCL|IC) = \alpha_0/2$ according to the equal tail probabilities approach. Thus, we have (see also Kumar and Baranwal (2019))

$$\text{LCL} = \frac{\chi^2_{2r,\,\alpha_0/2}}{2\lambda_0} = \frac{A_1}{\lambda_0} \text{ and } \text{UCL} = \frac{\chi^2_{2r,\,1-\alpha_0/2}}{2\lambda_0} = \frac{A_2}{\lambda_0} \quad (1)$$

where $A_1 = \frac{\chi^2_{2r,\,\alpha_0/2}}{2}$, $A_2 = \frac{\chi^2_{2r,\,1-\alpha_0/2}}{2}$ are the design constants and $\lambda_0$ is the known or specified IC rate parameter value. The $\chi^2_{2r,\,a}$ denotes the $a$-th quantile of the chi-square distribution with $2r$ degrees of freedom. The center line (CL) of the $t_r$-chart can be considered as the median of the IC distribution of $T_r$ and it is given by $\text{CL} = \frac{\chi^2_{2r,\,0.5}}{2\lambda_0}$.

**TBE data (exponential distribution, r=1)**

**ACTUALLY LCL=L and UCL=U**

Chakraborti et. al. (with several papers…)

**Excerpt C1.** Typical statements in the "Garden …[24]" where the authors name LCL and UCL what **actually** are the Probability Limits L and U. See the Figure 9 and the Excerpt 12.

Many other cases, with the same errors, can be found in the "*Garden …*[24], and the Conclusions" where *the authors name LCL and UCL what* ***actually*** *are the Probability Limits L and U*.

## References


1. Galetto, F., Quality of methods for quality is important. European Organisation for Quality Control Conference, Vienna. 1989
2. Galetto, F., GIQA, the Golden Integral Quality Approach: from Management of Quality to Quality of Management. Total Quality Management (TQM), Vol. 10, No. 1; 1999
3. Chakraborti et al., Properties and performance of the c-chart for attributes data, Journal of Applied Statistics, January 2008
4. Kumar, N., Chakraborti, S., Rakitzis, A. C. et al., Improved Shewhart-Type Charts for Monitoring Time Between Events. Journal of Quality Technology, 49(3), pp. 278–296. 2017
5. Zhang , C. W., Xie, M., Goh, T. N., Design of exponential control charts using a sequential sampling scheme. IIE Transactions, 38(12), pp. 1105–1116. 2006
6. Belz, M. Statistical Methods in the Process Industry: McMillan; 1973.
7. Casella, Berger, Statistical Inference, 2nd edition: Duxbury Advanced Series; 2002.

 doi:10.20944/preprints202511.0444.v1




8. Cramer, H. Mathematical Methods of Statistics: Princeton University Press; 1961.
9. Deming W. E., Out of the Crisis, Cambridge University Press; 1986.
10. Deming W. E., The new economics for industry, government, education: Cambridge University Press; 1997.
11. Dore, P., Introduzione al Calcolo delle Probabilità e alle sue applicazioni ingegneristiche, Casa Editrice Pàtron, Bologna; 1962.
12. Juran, J., Quality Control Handbook, 4th, 5th ed.: McGraw-Hill, New York: 1988-98.
13. Kendall, Stuart, (1961) The advanced Theory of Statistics, Volume 2, Inference and Relationship:, Hafner Publishing Company; 1961.
14. Meeker, W. Q., Hahn, G. J., Escobar, L. A. Statistical Intervals: A Guide for Practitioners and Researchers. John Wiley & Sons. 2017
15. Mood, Graybill, Introduction to the Theory of Statistics, 2nd ed.: McGraw Hill; 1963.
16. Rao, C. R., Linear Statistical Inference and its Applications: Wiley & Sons; 1965.
17. Rozanov, Y., Processus Aleatoire, Editions MIR: Moscow, (traduit du russe); 1975.
18. Ryan, T. P., Statistical Methods for Quality Improvement: Wiley & Sons; 1989.
19. Shewhart W. A., Economic Control of Quality of Manufactured Products: D. Van Nostrand Company; 1931.
20. Shewhart W.A., Statistical Method from the Viewpoint of Quality Control: Graduate School, Washington; 1936.
21. D. J. Wheeler, Various posts, Online available from Quality Digest
22. Galetto, F., (2014), Papers, and Documents of FG, Research Gate
23. Galetto, F., (2015-2025), Papers, and Documents of FG, Academia.edu
24. Galetto, F., (2024), The garden of flowers, Academia.edu
25. Galetto, F., Affidabilità Teoria e Metodi di calcolo: CLEUP editore, Padova (Italy); 1981-94.
26. Galetto, F., Affidabilità Prove di affidabilità: distribuzione incognita, distribuzione esponenziale: CLEUP editore, Padova (Italy); 1982, 85, 94.
27. Galetto, F., Qualità. Alcuni metodi statistici da Manager: CLUT, Torino (Italy; 1995-2010).
28. Galetto, F., Gestione Manageriale della Affidabilità: CLUT, Torino (Italy); 2010.
29. Galetto, F., Manutenzione e Affidabilità: CLUT, Torino (Italy); 2015.
30. Galetto, F., Reliability and Maintenance, Scientific Methods, Practical Approach, Vol-1: www.morebooks.de.; 2016.
31. Galetto, F., Reliability and Maintenance, Scientific Methods, Practical Approach, Vol-2: www.morebooks.de.; 2016.
32. Galetto, F., Statistical Process Management, ELIVA press ISBN 9781636482897; 2019
33. Galetto F., Affidabilità per la manutenzione, Manutenzione per la disponibilità: tab edizioni, Roma (Italy), ISBN 978-88-92-95-435-9, www.tabedizioni.it; 2022
34. Galetto, F., (2015) Hope for the Future: Overcoming the DEEP Ignorance on the CI (Confidence Intervals) and on the DOE (Design of Experiments), Science J. Applied Mathematics and Statistics, Vol. 3, No. 3, pp. 99-123, doi: 10.11648/j.sjams.20150303.14
35. Galetto, F., (2015) Management Versus Science: Peer-Reviewers do not Know the Subject They Have to Analyse, Journal of Investment and Management. Vol. 4, No. 6, pp. 319-329, doi: 10.11648/j.jim.20150406.15
36. Galetto, F., (2015) The first step to Science Innovation: Down to the Basics., Journal of Investment and Management. Vol. 4, No. 6, pp. 319-329, doi: 10.11648/j.jim.20150406.15
37. Galetto F., (2021) Minitab T charts and quality decisions, Journal of Statistics and Management Systems, DOI: 10.1080/09720510.2021.1873257
38. Galetto, F., (2021) Control Charts for TBE and Quality Decisions, Academia.edu
39. Galetto F. (2021) ASSURE: Adopting Statistical Significance for Understanding Research and Engineering, Journal of Engineering and Applied Sciences Technology, ISSN: 2634 – 8853, 2021 SRC/JEAST-128. DOI: https: //doi.org/ 10.47363/ JEAST/ 2021(3)118
40. Galetto, F., (2006) Does Peer Review assure Quality of papers and Education? 8th Conference on TQM for HEI, Paisley (Scotland)
41. Galetto, F., (2001), Looking for Quality in "quality books", 4th Conference on TQM for HEI, Mons (Belgium)
42. Galetto, F., (2001), Quality QFD and control charts, Conference ATA, Florence (Italy)

 doi:10.20944/preprints202511.0444.v1



43. Galetto, F., (2002), Fuzzy Logic and Control Charts, 3rd ICME Conference, Ischia (Italy)
44. Galetto, F., (2010),The Pentalogy Beyond, 9th Conference on TQM for HEI, Verona (Italy)
45. Galetto, F., (2024), News on Control Charts for JMP, Academia.edu
46. Galetto, F., (2024), JMP and Minitab betray Quality, Academia.edu