

Article

Not peer-reviewed version

Artificial Intelligence and Control Charts: A Big Problem

[Fausto Galetto](#)*

Posted Date: 26 September 2025

doi: 10.20944/preprints202509.2184.v1

Keywords: Control Charts; exponential distribution; TBE; T Charts; Minitab; JMP; Reliability Integral Theory



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Artificial Intelligence and Control Charts: A Big Problem

Fausto Galetto

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

Abstract

We use the data of some published papers to compare those authors findings with ours. From the analysis we get different results: the cause is that they use the Probability Limits of the PI (Probability Interval) as they were the Confidence Limits (Control Limits of the Control Charts, CCs). 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 (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. The same happens for the Confidence Limits of the parameters of the distribution involved in the papers (Weibull, Inverse Weibull, Gamma, Binomial, Maxwell). We will show that the Reliability Integral Theory (RIT) is able to solve these problems and the Sequential way of dealing with data.

Keywords: Control Charts; exponential distribution; TBE; T Charts; Minitab; JMP; Reliability Integral Theory

1. Introduction

In the old days of Statistics almost all the methods have been based on the Normal distribution and their connected ones (t, F, Snedecor, ...); when the data were not "normally distributed" *suitable transformations to Normality* have been used.

Quite recently simulations have become in fashion; several articles in ResearchGate mention simulations [1–14]; Journals do the same [15]. See various statements [from 1-14] in the Excerpt 1.

... Performed a detailed power *simulation* study including distributions of different shapes...

...Such a formulation is direct in principle but difficult to implement in practice because there is no standard form for the integral.... An alternative is to use numerical computations based on Monte-Carlo *simulations*, and we use this approach here...

... In general, the *simulation* results reveal that the OOC run length distributions are also skewed to the right, ...

... *Simulation* studies based on Monte-Carlo also show that the proposed tests are good competitors of the existing tests and have some advantage in certain situations...

... Using *simulation*, the performance of the proposed procedure is compared against the heavily used testing procedure to test ...

... Our *simulation* study shows that the Type I error of both tests are nearly identical,

... Since the sampling distribution of ... does not depend on nuisance parameters (under the null hypothesis), the appropriate p-values can easily be computed by *simulations*....

... There may not be a direct, closed-form formula for calculating the TARL. Due to this

absence, *simulation* techniques are typically employed to estimate TARL...

... according to the Monte-Carlo *simulation*, the corresponding control limit can be found...

... According to the Monte-Carlo *simulation*, these control limits are 1.638, 2.875, 1.450, 0.494, and 0.462, respectively...

...-They mainly rely on Monte-Carlo *simulations* for the performance assessment of their proposed test...

...This section compares the performance of the proposed tests in terms of power and Type-I error rate using numerical *simulations*. ...

... Note that the Monte-Carlo *simulation* is used since there is no closed-form for the distribution of D under H03:...

... Their ARL profiles were evaluated via Monte Carlo *simulation* (25,000 iterations)....

... Exact derivations and computer *simulations* are used to study performance properties of the proposed charts...

... we obtain the control limits through *simulation*. Then for consistency, we used *simulations* in R to compute 16 all of the control limits, and where possible we used the analytical expressions to verify the results. As expected, we find that higher nominal IC ARL values require higher UCLs and lower LCLs. < Table ... Here > Next we compare the performance of the proposed charts. 7. Performance Comparisons The performances of the one-chart and two-chart schemes are evaluated in a *simulation* study...

Excerpt 1. Statements from papers [1–14].

Unfortunately often sound Theory is disregarded [1–23]. In fact, the authors use wrong Control Limits (LCL, UCL) and therefore are wrong in defining the state of control of the Process (IC, In Control, OOC, Out of Control) and the related ARLs (Average Run Lengths).

Since 1989, the author (FG) tried to inform the Scientific Community about the flaws in the use of (“wrong”) *quality methods for making Quality* [24] 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 [25], including the Process Management and Control Charts (CC) for Process Control. Control Charts (CC) use *sequentially* the collected data to assess if a Production or Service process output is to be considered In Control (IC) or Out Of Control (OOC); the decision is very important for taking Corrective Actions (CA), if needed.

To show our Theory we will use some of the data found in the papers [16–18]; we will show the drawbacks of the “theory (wrong)” presented in the papers.

But before that we mention the very interesting the statements in the Excerpt 2:

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 2. From the paper “*Misguided Statistical Process Monitoring Approaches*”.

We agree with the authors in the excerpt 2, but, nevertheless, they did not realise the problem that we are showing here: *wrong Control Limits in CCs for Rare Events, with data exponentially or Weibull*

or Maxwell distributed or free-distributed data. Several papers compute “a-scientific” control limits... See References...

Test of Hypotheses and the Confidence Intervals (CI) are intimately related and so equivalent for decision making. Using the data in [16–18] with good statistical methods [26–74] we give our “reflections on Artificial Intelligence and 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*” [19] and some in the Appendix A) compute in an *a-scientific way* the Control Limits of CCs for “Individual Measures or Exponential, Weibull, Maxwell Gamma and free-distributed data”, indicated as I-CC (Individual Control Charts); we dare to show, to the Scientific Community, how to compute the *True Control Limits (True Confidence 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 [19–25]: this is shown through the suggested published papers. Humbly, given our strong commitment to Quality [19–25,43–74], 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 [19–25,43–74].

To let the reader follow our way of approaching the problem of estimation we will use various figures and some data: this is caused by the fact that there are wrong ideas in the literature.

By reading [19] and other papers, 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 [19]”) 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 [19].*” See the standard mathematical statistical results/methods in the Appendix A and meditate (see the formulae there and in the AI queries)!

Hence, the problem becomes “...the *standard ... methods* as is typical ...”: are those standards typical methods (in the “*The Garden ...*” in [19] and [1–15]) scientific?

To understand the readers need to know “*Some ideas on Hypothesis Test and The Statistical Hypotheses with the related risks*” [26–42], or, in alternative, they can read [43–74].

Due to length limitations of the paper, we must ask the reader to know the following ideas: the definition of *statistical hypothesis* as a statement about a *population parameter* θ (e.g. the “*true*” mean, the “*true*” shape, the “*true*” variance, the “*true*” reliability, the “*true*” failure rate, ...n that we *assume* to exists and has a *value* even though it is *unknown* to us), related to the *statistical model* $F(x|\theta)$ associated with a *random variable* (RV) X . 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 *hypothesized* 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 .

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* C (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*.

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 Θ_0 is a subset of the *parameter space* Θ and Θ_1 a subset disjoint from Θ_0 ; $\Theta_0 \cup \Theta_1 = \Theta$ and $\Theta_0 \cap \Theta_1 = \emptyset$; *before* collecting any data, with H_0 we accept a probability α of wrong decision, while with H_1 we accept a probability β of wrong decision. A hypothesis test of $H_0: \{\theta \in \Theta_0\}$ versus the alternative hypothesis $H_1: \{\theta \in \Theta_1\}$ 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 .

If «*actually* (but we do not know)» $H_0: \{\theta \in \Theta_0\}$ is true and the hypothesis test (the *rule*, the *computed quantity* \underline{S} , in the Figure 1), 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 α . If, on the other hand, «*actually* (but we do not know)» $\theta \in \Theta_1$ 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 β .

These two different situations are depicted in the Table 1 (for simple parametric hypotheses). The framework of a test of hypothesis is depicted in the Figure 1.

Table 1. Statistical Hypotheses and risks.

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

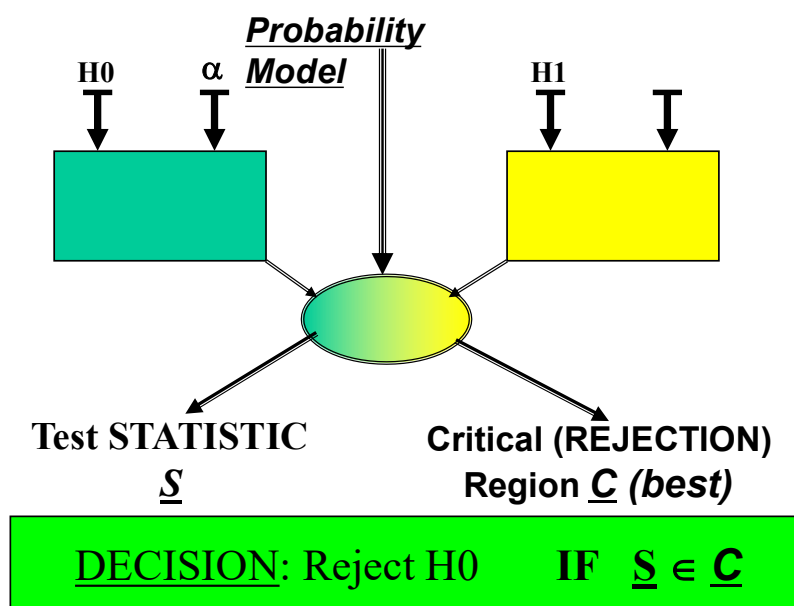


Figure 1. The pictorial framework of a Statistical Hypothesis (based on a Probability Model).

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 evidence to state the contrary”. It is evident that

$$\alpha = P[\text{reject } H_0 | H_0 \text{ true}] \quad \text{and} \quad \beta = P[\text{accept } H_0 | H_0 \text{ false}] \quad (1)$$

A *likelihood ratio test* is any test that has a rejection region of the following form $\{s(D): q(D) \geq c\}$, 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 .

This has great importance for Control Charts, as you can see in the Figure 3.

Suppose C is the “critical” (or rejection) region for a test, based on a «*statistic $s(D)$* » (the formula to elaborate the sampled data D , providing the value $\underline{s(D)}$).

Then for testing $H_0: \{\theta \in \Theta_0\}$, the test makes a mistake if « $\underline{s(D)} \in \underline{C}$ », so that the probability of a Type I Error is $\alpha = P(S(D) \in C)$ [$S(D)$ is the random variable giving the result $\underline{s(D)}$]. It is important the *power of the test* $1 - \beta$, which is the probability of rejecting H_0 when *in reality* H_0 is false

$$1 - \beta = P[\text{reject } H_0 | H_0 \text{ false}] \quad (2)$$

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

To find the RV $S(D)$ and the region C , we use the likelihood function $L(\theta | D = \{x_1, x_2, \dots, x_n\})$

$$L(\theta | D) = \prod_{i=1}^n f(x_i) \quad (3)$$

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 $C = \{s(D): q(n) = L_1/L_0 \geq k_\alpha\}$, where $q(n)$ is the Likelihood Ratio L_1/L_0 and the quantity k_α is chosen in such a way that the Type I Error has a risk (probability) α as in the formula (4), with fixed n (the *sample size*),

$$\int \int \dots \int_{q(n) \geq k_\alpha} L_0 \, dx_1 dx_2 \dots dx_n = \alpha \quad (4)$$

The most powerful critical region C has the highest power $1-\beta(\theta)$.

Let CR_n be the "Critical Region" found by (4) and β_n be the probability (5), function of n,

$$\beta_n = \int \int \dots \int_{q(n) \leq CR_n} L_1 dx_1 dx_2 \dots dx_n \tag{5}$$

By (4) and (5), increasing n, we arrive to select a final sample size n, such that $\beta_n = \beta$, the desired risk.

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

We will use $\alpha = \beta$ in the following discussion. After the data analysis, we can decide if the data suggest us to "accept (= not reject)" $H_0: \{\theta \in \Theta_0\}$ or "accept" $H_1: \{\theta \in \Theta_1\}$, and after that we can compute the Confidence Interval, $CI = \theta_L \dots \theta_U$, of the parameter θ , with Confidence Level $CL = 1 - \alpha = 1 - (\alpha/2 + \beta/2) = 1 - (\alpha/2 + \alpha/2)$.

When we consider the Control Charts we want to test the two Hypotheses H_0 : {the process is "IC (In Control)"} against H_1 : { the process is "OOC (Out Of Control)"} , and after the data analysis we can compute the Control Interval (which is actually a Confidence Interval), $LCL \dots UCL$.

If we use the Table 4 data (Time between failures data (from "Improved Phase... for Monitoring TBE")) it is easy to see that (as said with the above warehouse example) the practical problem becomes a Theoretical one [19–74] (all references and Figure 21). Since those data are well "exponentially distributed" we anticipate here, immediately, the wrong formulae (either using the parameter $\theta = \theta_0$ or its estimate \bar{t}_0 , with $\alpha = 0.0027$) in the below formula (6) (as you can find in [19])

$LCL =$	$\theta_0 \ln(1 - \alpha/2)$	$0.00135 \bar{t}_0$	$UCL =$	$\theta_0 \ln(\alpha/2)$	$6.6077 \bar{t}_0$	(6)
---------	------------------------------	---------------------	---------	--------------------------	--------------------	-----

The readers should understand clearly the Theoretical and Practical Difference between $L \dots U$ (the Probability Interval) and $LCL \dots UCL$ (the Confidence Interval), pictorially shown in the Figure 2: the two lines L and U depends on the parameter θ (to be estimated) and on the two probabilities α and β , while the two points L and U depends on the assumed value θ_0 of the parameter and on the two chosen probabilities α and β ; after the data analysis, we compute the estimate \bar{t}_0 of the parameter θ and from that the Confidence Interval $LCL \dots UCL$, with Confidence Level $CL = 1 - \alpha$. It is clear now the wrong ideas in the formulae (6).

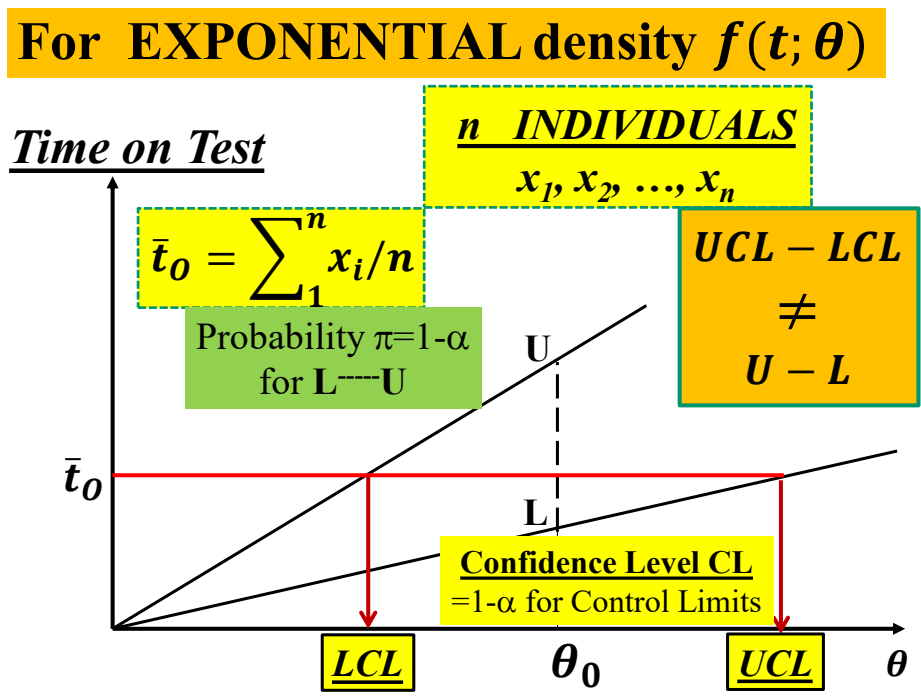


Figure 2. Theoretical and Practical Difference between $L \dots U$ and $LCL \dots UCL$.

In the formulae (6), for the interval LCL—UCL (*named* Control Interval, by the authors [24]), the LCL *actually* must be L and the UCL *actually* must be U, vertical interval L—U (Figure 2); the actual interval LCL—UCL is the horizontal one in the Figure 2, which is not that of the formulae (6). Artificial Intelligence provides various wrong formulae for the interval LCL—UCL, as done by many authors [19]. 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 [19–74].

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 [20–25]”. Therefore, *those authors do not extract the maximum information from the data in the Process Control. “The Garden...”* [19] 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 3. 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).

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 $1-\beta$, they use for the decision (*as per the principles F1 and F2*) [43–74].

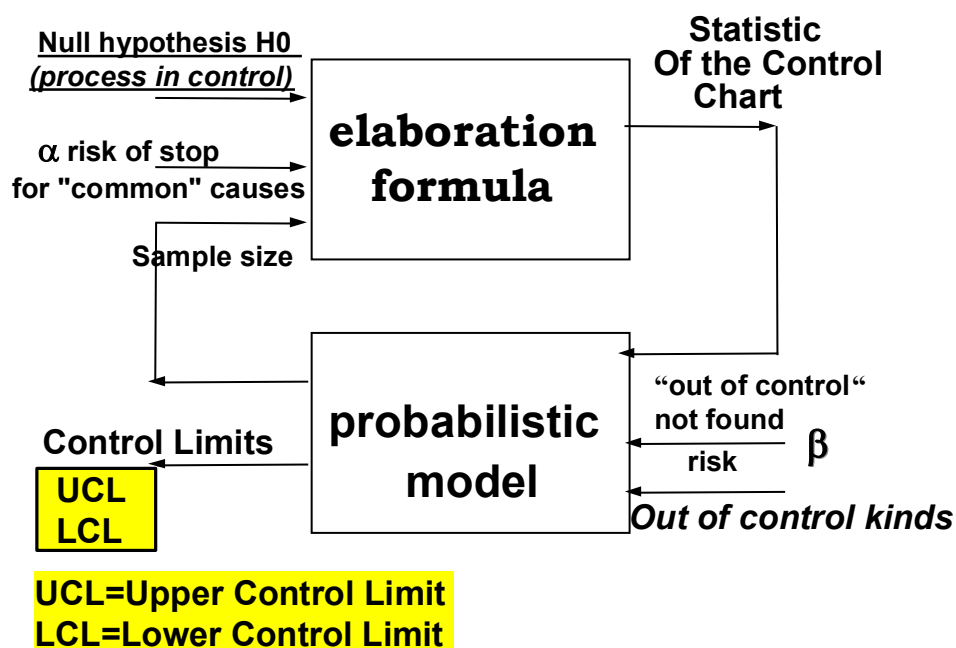


Figure 3. LCL and UCL of Control Charts with their risks.

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 and data are analysed.

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 (Excerpt 2): 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 [43–74]. 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).

2. Materials and Methods

2.1. A Reduced Background of Statistical Concepts

After the ideas given in the Introduction, we provide the following ones essential to understand the “problems related to I-CC” as we found in the literature. We suggest it for the formulae given and for the difference between the concepts of PI (*Probability Interval*) and CI (*Confidence Interval*): this is overlooked in “*The Garden ...* [19]”

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*” $D=\{x_1, x_2, \dots, x_n\}$ 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 θ is the set of the parameters of the functions.

We state in the Table 2 a sample of models where θ is a set of parameters:

Table 2. Some probability models for data analysis.

Name	$F(x \theta)$	Parameters			Symbol
Exponential	$1 - \exp(-x/\theta)$	θ			$E(x \theta)$
Weibull	$1 - \exp[-(x/\eta)^\beta]$	β	η		$W(x \beta,\eta)$
Inverted Weibull	$1 - e^{-(\eta/x)^\beta}$	β	η		$IW(x \beta,\eta)$
Maxwell	$\frac{\sqrt{2/\pi}}{\sigma^3} \int_0^x t^2 e^{-t^2/(2\sigma^2)} dt$		σ^2		$MW(x \sigma)$
Normal	$\frac{1}{\sqrt{2\pi}\sigma} \int_0^x e^{-(t-\mu)^2/(2\sigma^2)} dt$	μ	σ^2		$N(x \mu, \sigma^2)$
Distribution-free	No form of $F(x \theta)$				D-free

Two important models are the Normal and the Exponential, but we consider also the others for comparison. When $\theta = \{\mu, \sigma^2\}$ we have the *Normal model*, written as $N(x|\mu, \sigma^2)$, with (parameters) mean $E[X]=\mu$ and variance $\text{Var}[X]=\sigma^2$ with pdf

$$f(x|\mu, \sigma^2) = n(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} \quad (7)$$

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

When we have the *observed sample* $D=\{x_1, x_2, \dots, x_n\}$, 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 determinate but for the value θ_0 of some parameter, e.g. unidimensional, θ , or bidimensional $\theta=\{\mu, \sigma^2\}$, or $\theta=\{\beta, \eta, \omega\}$ as in the GIW($x|\beta, \eta, \omega$), or $\theta=\{\beta, \eta, \omega, \theta\}$ as in the MPGW($x|\beta, \eta, \omega, \theta$).

We seek some function of θ , say $\tau(\theta)$, 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 $f(x|\theta)$, is efficient [23–54], 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 $\tau(\theta)=(\mu, \sigma^2)$, 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 $F(x|\theta)$ and 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 $\tau(\theta)$, unidimensional.

A way of finding [23–54] the estimate is to compute the *Likelihood Function* $L(\theta|D)$ [LF] and to maximise it: the solution of the equation $\partial L(\theta|D)/\partial\theta=0$ is termed *Maximum Likelihood Estimate* [MLE]. Both are used also for sequential tests.

The LF is important because it allows us finding the MVB (*Minimum Variance Bound*, Cramer-Rao theorem) [23–54] of an unbiased RV T [related to the inference function $\tau(\theta)$], such that

$$\text{Var}(T) \geq \frac{[\tau(\theta)]^2}{E\left\{\left[\frac{\partial \ln L(\theta|D)}{\partial \theta}\right]^2\right\}} = \text{MVB}(T) \quad (8)$$

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

Naming $I_r(T)$ the information extracted by the RV T we have that [23–54]

$$I_r(T)=1/\text{MVB}(T) \Leftrightarrow T \text{ is an Efficient Estimator.}$$

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 $\tau(\theta)$.

We recap the very important concept of *Confidence Interval* (CI) and *Confidence Level* (CL) [23–54].

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

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

The reader must be very careful to distinguish *between* the *Probability Interval* $PI=\{L\text{---}U\}$, 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 $\pi=1-\alpha$ (that we fix before any computation), as follows by the probabilistic statement (9) [see the Figure 2 for the exponential density, when $n=1$]

$$P[L \leq T \leq U] = \int_L^U \varphi(t, \theta, n) dt = 1 - \alpha \quad (9)$$

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

Notice that the *Probability Interval* $PI=\{L\text{---}U\}$, given in the formula (9), *does not depend* on the data D, as you can pictorially see in Figure 2: L and U are the *Probability Limits*. Notice that, on the contrary, the *Confidence Interval* $CI=\{\tau_L\text{---}\tau_U\}$ *does depend* on the data D, pictorially seen in Figure 2. This point is essential for all the papers in the References.

Shewhart identified this approach, L and U, on page 275 of [40] 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 $\bar{\theta} \pm t\sigma_\theta$ *symmetrically spaced* in reference to $\bar{\theta}$. Tchebycheff’s Theorem tells us that the probability P that an observed value of $\bar{\theta}$ will lie within these *symmetric* limits so long as the quality standard is maintained satisfies the inequality $P > 1-1/t^2$. 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,...

The Tchebycheff Inequality: IF the RV X is arbitrary with density $f(x)$ and finite variance σ^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 μ 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 μ we commit an error that “most likely” does not exceed ε . IF, on the contrary, $\sigma \not\ll \varepsilon$, 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 μ we commit an error that “most likely” does not exceed ε . See the excerpts 3, 3a, 3b.

Notice that, when we write $\bar{x} - \varepsilon < \mu < \bar{x} + \varepsilon$, we consider the Confidence Interval CI [23–54], and no longer the Probability Interval PI [23–54].

These statistical concepts are very important for our purpose when we consider the Sequential tests and the Control Charts, especially with Individual data.

Notice that the error made by several authors [1–19] is generated by *lack of knowledge* of the difference between PI and CI [23–54]: they think *wrongly* that CI=PI, a diffused disease [1–19]! They should study some of the books/papers [23–54] 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 (9), *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 [23–54] there is a strict relationship between CI and Test Of Hypothesis, known also as Null Hypothesis Significance Testing Procedure (NHSTP). In Hypothesis Testing (see the Introduction), 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 $P=\alpha$ of being wrong if he decides that the “null hypothesis” H_0 is true, when actually it is opposite: H_0 is wrong. When we analyse, at once, the *observed sample* $D=\{x_1, x_2, \dots, x_n\}$ and we compute the *empirical* (observed) *mean* \bar{x} and the *empirical* (observed) *standard deviation* s , we define the *Acceptance interval*, which is the CI

$$LCL = \bar{x} - t_{1-\alpha/2}s/\sqrt{n} < \mu < \bar{x} + t_{1-\alpha/2}s/\sqrt{n} = UCL \quad (10)$$

Notice that the interval (for the Normal model, μ'' assumed)

$$\mu'' - t_{1-\alpha/2}\sigma/\sqrt{n} \text{-----} \mu'' - t_{1-\alpha/2}\sigma/\sqrt{n} \quad (11)$$

is the Probability Interval such that $P[\mu'' - t_{1-\alpha/2}\sigma/\sqrt{n} < \bar{X} < \mu'' - t_{1-\alpha/2}\sigma/\sqrt{n}] = 1 - \alpha$.

A fundamental reflection is in order: the formulae (10) and (11) tempt the unwise guy to think that he can get the *Acceptance interval*, which is the CI [23–54], by substituting the assumed values μ_0, σ_0 of the parameters with the *empirical* (observed) *mean* \bar{x} and *standard deviation* s . This trick is valid only for the Normal distribution.

The formulae (10) can be used sequentially to test $H_0=\{\mu=\mu_0\}$ versus $H_1=\{\mu=\mu_1<\mu_0\}$; for any value $2 < k \leq n$; we obtain $n-2$ CIs, decreasing in length; we can continue until either $\mu_1 < LCL$ or $UCL < \mu_0$, or both (verify) $\mu_1 < LCL$ and $UCL < \mu_0$.

More ideas about these points can be found in [23–54].

In the field of Control Charts, with Shewhart, instead of the formula (10), we use (12)

$$\bar{x} - \frac{z_{1-\alpha/2}S}{c_4\sqrt{n}} < \mu < \bar{x} + \frac{z_{1-\alpha/2}S}{c_4\sqrt{n}} \quad (12)$$

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

Actually, Shewhart does not use the coefficient 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}], σ 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.

2.2. Control Limits by AI Versus Sound Theory

In the first part of this section we provide the ideas of the Statistical Theory, while in the second one we see what AI tells us.

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 [29,30] and Juran [32] after Shewhart invention [40,41].

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

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

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:

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

or

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

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

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

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, which is the Lower *Confidence* Limit) and UCL (Upper Control Limit, the Upper *Confidence* Limit) both for the mean μ_X (1st parameter of the Normal pdf) and for σ_X (2nd parameter of the Normal pdf). They are considered the limits comprising 0.9973n of the observed data. Similar ideas can be found in [23–54] (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 Shewhart Control Charts; it is sufficient to use NHSTP or the Sequential Test Theory!

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 [19,42–54]).

Let X_{ij} , $i=1, 2, \dots, j=1, 2, \dots, n$ be independent and identically distributed (IID) normal $N(\mu_0, \sigma^2_0)$ r.v.'s, where μ_0 and σ^2_0 are the *specified IC* mean and variance, respectively. The X_{ij} represents the j th observation from the i th rational subgroup (sample) obtained at the i -th sampling stage, and Y_i denotes the i -th charting statistic based on the i -th sample. When $n \geq 2$, 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., $n=1$) $Y_i = X_{i1} = X_i$. The control limits of the standard Shewhart chart (X chart or the \bar{X} chart) are given by $UCL_1 = \mu_Y + k\sigma_Y$ and $LCL_1 = \mu_Y - k\sigma_Y$ (1) where μ_Y and σ_Y are the *specified IC* mean and standard deviation of the charting statistic Y_i , 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 on *JQT*, 2017 where the data are Erlang distributed with λ_0 is the scale parameter and the Control Limits LCL and UCL are defined [copying Xie et al.] **erroneously** as

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

Excerpt 4. From a paper in the “Garden... [19]”. 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 (9), when $\varphi(t, \theta, n)$ is the pdf of the Estimator T , related to the Normal model $F(x; \mu, \sigma^2)$. Using (9), *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 $Y = \bar{X}$

$$P[\mu_Y - 3\sigma_Y \leq Y = \bar{X} \leq \mu_Y + 3\sigma_Y] = 0.9973 \quad (12a)$$

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 [19–23]).

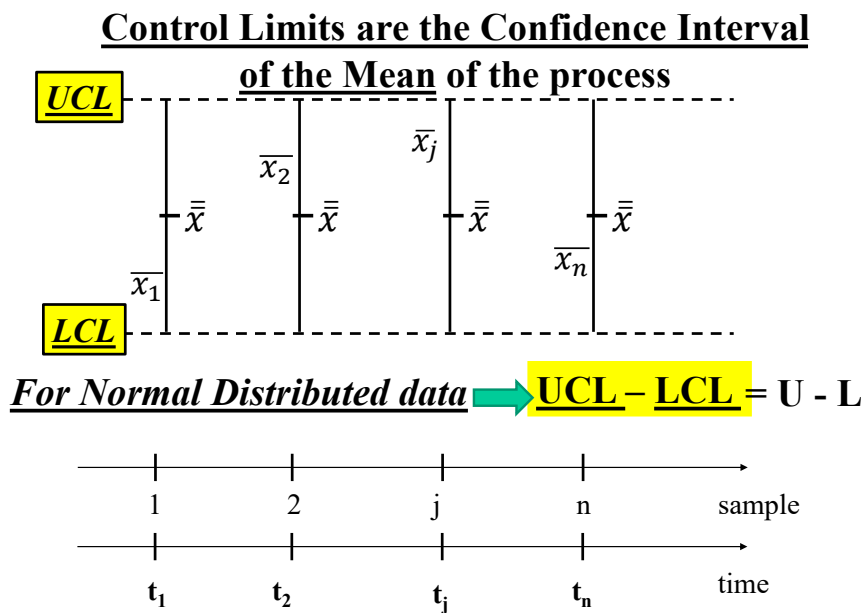


Figure 3. Control Limits $LCL \leq \bar{x} \leq UCL$ (Probability interval), for Normal data (Individuals x_{ij} , sample size k) “sample means” \bar{x}_i and “grand mean” \bar{x} .

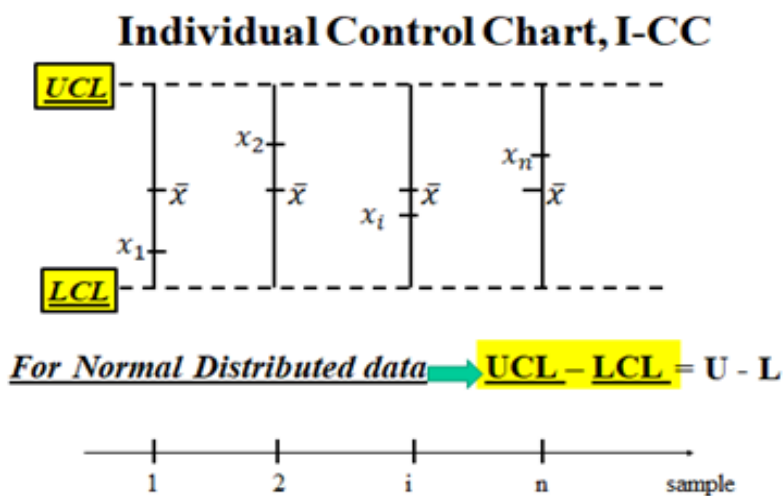


Figure 4. Individual Control Chart (sample size $k=1$). Control Limits $LCL \leq x_i \leq UCL$ (Probability interval), for Normal data (Individuals x_i) and “grand mean” \bar{x} .

The data plotted in the CCs [23–54] (see the Figure 3-4) are the means $\bar{x}(t_i)$, determinations of the RVs $\bar{X}(t_i)$, $i=1, 2, \dots, n$ (n =number of the samples) computed from the *sequentially* collected data of the i -th sample $D_i=\{x_{ij}, j=1, 2, \dots, k\}$ (k =sample size)}, determinations of the RVs $X(t_{ij})$ at *very close instants* $t_{ij}, j=1, 2, \dots, k$. 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 $X(t_i)$, $i=1, 2, \dots, n$ (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 (10) it is clear that Shewhart was using the Normal distribution, as a consequence of the Central Limit Theorem (CLT) [23–54]. 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).

Let k be the sample size; the RVs $\bar{X}(t_i)$ are assumed to follow a normal distribution and uncorrelated; $\bar{X}(t_i)$ [ith rational subgroup] is the mean of RVs IID $X(t_{ij})$ $j=1, 2, \dots, k$, (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), $\bar{X}(t_i) \sim N(\mu_{\bar{X}(t_i)}, \sigma_{\bar{X}(t_i)}^2)$ [experimental mean $\bar{x}(t_i)$] with mean $\mu_{\bar{X}(t_i)}$ and variance $\sigma_{\bar{X}(t_i)}^2$. $\bar{\bar{X}}$ is the "grand" mean and $\sigma_{\bar{\bar{X}}}^2$ is the "grand" variance: the pdf of the (RV) grand mean $\bar{\bar{X}} \sim N(\mu_{\bar{\bar{X}}}, \sigma_{\bar{\bar{X}}}^2)$ [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 [23–54]. We can assess the OOC state of the process via the Confidence Intervals (by the Control Limits) with $CL=0.9973$. **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{x} in place of μ_Y and \bar{s}/\sqrt{k} in place of σ_Y we get the CI of $\mu_{\bar{x}}$ 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 $\mu_{\bar{X}(t_q)}$, $q=1, 2, \dots, n$, to any other (subsystem) mean $\mu_{\bar{X}(t_r)}$, $r=1, 2, \dots, n$, 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, which are the Confidence Limits. 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 [23–54].

We stress that the interval LCL_X — UCL_X is the "Confidence Interval" with "Confidence Level" $CL=1-\alpha=0.9973$ for the unknown mean $\mu_{X(t)}$ of the Stochastic Process $X(t)$ [23–54]. The interval LCL_R — UCL_R is the "Confidence Interval" with "Confidence Level" $CL=1-\alpha=0.9973$ for the unknown Range of the Stochastic Process $X(t)$ [23–54].

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

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** [23–74].

We think that the paper "Quality of Methods for Quality is important", [24] 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.

===== Now we see what AI answered to our queries. =====

First Query: [NOTICE that the answer is like from the Theory!]

How can I compute LCL and UCL in Control Charts for individuals Normally distributed?

ChatGPT says:

To compute the Lower Control Limit (LCL) and Upper Control Limit (UCL) in Control Charts for Individuals (also called Individuals or X Charts), especially when the data is normally distributed, follow this step-by-step process:

1. Understanding Control Charts for Individuals

When you're dealing with individual measurements (i.e. sample size = 1), you usually use:

- An Individuals (X) chart to monitor the process mean.
- A Moving Range (MR) chart to estimate the process variability.

You don't have subgroups like in X-bar and R charts — you're monitoring single data points.

2. Basic Definitions

- \bar{X} : Mean of individual measurements
- MR: Moving range (difference between consecutive individual values)
- \overline{MR} : Average of the moving ranges
- d_2 : A constant depending on subgroup size (for $n=2$, $d_2 \approx 1.128$)

3. Formulae for Control Limits

Individuals (X) Chart:

To estimate the process standard deviation: $\hat{\sigma} = \overline{MR}/d_2$

Then: Center Line (CL): \bar{X} , UCL: $\bar{X} + 3\hat{\sigma}$, LCL: $\bar{X} - 3\hat{\sigma}$

Substitute for $\hat{\sigma}$: UCL = $\bar{X} + 3\overline{MR}/d_2$, LCL = $\bar{X} - 3\overline{MR}/d_2$

Moving Range (MR) Chart: CL: \overline{MR} , UCL: $D_4 \cdot \overline{MR}$, LCL: $D_3 \cdot \overline{MR}$

For moving range of 2 (i.e., $MR = |X_i - X_{i-1}|$): $D_4 = 3.267$, $D_3 = 0$

4. Step-by-Step Example

Suppose you have this data: $X = [10, 12, 11, 13, 12]$

Step 1: Compute \bar{X} , $\bar{X} = (10 + 12 + 11 + 13 + 12)/5$

Step 2: Compute Moving Ranges, $MR = \{|12 - 10|, |11 - 12|, |13 - 11|, |12 - 13|\} = [2, 1, 2, 1]$

$\overline{MR} = (2 + 1 + 2 + 1)/4 = 1.5$

Step 3: Estimate standard deviation $\hat{\sigma} = \overline{MR}/d_2 = 1.5/1.128 \approx 1.33$

Step 4: Compute Control Limits

- UCL = $\bar{X} + 3\hat{\sigma} = 11.6 + 3(1.33) = 11.6 + 3.99 = 15.59$
- LCL = $\bar{X} - 3\hat{\sigma} = 11.6 - 3.99 = 7.61$

So your X Chart limits are: UCL = 15.59, CL = 11.6, LCL = 7.61

Summary of Constants for $n = 2$:

Constant	Value
d_2	1.128
D_3	0
D_4	3.267

Notes

- This method assumes normality of the data and independence of observations.
- If there is autocorrelation, the control chart may not work well.
- Avoid using standard deviation directly unless the population SD is known.

Second Query: [NOTICE that the answers are all wrong!]

How can I compute LCL and UCL in Control Charts for individuals Exponentially distributed?

ChatGPT says:

To compute the Lower Control Limit (LCL) and Upper Control Limit (UCL) for individual control charts (I-Charts) when the data follow an exponential distribution, you need to take into account that the usual assumptions for I-Charts (normality, constant variance, etc.) do not hold. However, there are appropriate methods adapted for exponential data.

Overview of the Problem

- You have individual observations (i.e., no subgroups).
- The data follow an Exponential distribution (e.g., lifetimes, waiting times).
- Traditional I-chart (X chart) limits assuming normality may not be appropriate

Approach to Control Charts for Exponentially Distributed Data

The exponential distribution is often used to model the time between events in a Poisson process. It has a probability density function (PDF) given by: $f(x|\lambda) = \lambda e^{-\lambda x}, x \geq 0$ where λ is the rate parameter, which is the inverse of the mean (i.e., $\lambda = 1/\mu$).

The exponential distribution is not symmetric (it's positively skewed), and has:

- Mean: $\mu = 1/\lambda$, Variance: $\sigma^2 = 1/\lambda^2$

In a SECOND TIME query BEGINNING=====

Control Charts for Individual Measurements:

• In an individual control chart for exponentially distributed data, each data point represents the individual measurement, and we are concerned with monitoring the location (mean) of the process.

For individual measurements from an exponential distribution, you can use Shewhart control charts adapted for exponential data, typically using a transformation or using a non-normal chart like the exponential control chart.

UCL and LCL Formula:

For an exponentially distributed process, the formulas for the UCL and LCL are typically derived from the properties of the exponential distribution.

• Center Line (CL): The center line is the expected value (mean) of the exponentially distributed data, i.e., μ .

• UCL and LCL are typically calculated as a function of the standard deviation (or sometimes the range) of the individual data points and a constant multiplier (commonly denoted as -3 $+3$ for a 3-sigma control chart).

Steps to Calculate:

1. Determine the Average (Mean) μ of the individual measurements.

○ If you have a sample of data points x_1, x_2, \dots, x_n , the mean is: $\mu = \sum_1^n x_i / n$

2. Determine the Control Limits:

For an exponentially distributed process, the standard deviation of the individual data points is also μ , which is the same as the mean (since for an exponential distribution, the mean and standard deviation are equal); UCL (Upper Control Limit): $UCL = \mu + 3\mu = 4\mu$ and LCL (Lower Control Limit): $LCL = \mu - 3\mu = -2\mu$

However, since you can't have negative values for exponentially distributed data (because the data cannot go below zero), if the LCL becomes negative, it should be set to 0.

Therefore, the LCL is: $LCL = \max[0, -2\mu]$

3. Plotting the Control Chart:

Once you have the UCL, LCL, and mean, you can plot the individual data points and compare them against the control limits. The control chart will signal out-of-control points when they fall outside the UCL or LCL.

Example: Suppose you have the following data points from an exponentially distributed process: 2.1, 3.0, 1.8, 2.5, 3.3, 2.9

1. Calculate the mean: $\mu = (2.1 + 3.0 + 1.8 + 2.5 + 3.3 + 2.9) / 6 = 2.67$

2. Calculate the UCL and LCL:

○ UCL: $UCL = 4 \times 2.67 = 10.68$, LCL: $LCL = \max(0, -4.01) = 0$

3. So, for this dataset, the UCL is 10.68 and the LCL is 0.

This is the basic process for computing control limits for individual exponentially distributed data in a control chart. Would you like to go over an example in more detail or dive into other aspects of control charts?

SECOND TIME query END=====

◇ Method 1: Control Chart Using Transformation (Log Transformation)

You can transform exponential data to approximate normality:

1. Let $X_i \sim \text{Exponential}(\lambda)$

2. Then $Y_i = \ln(X_i)$ is approximately Gumbel distributed, which may still not be normal, but may behave better.

3. Construct the chart based on Y_i , then back-transform control limits if needed.

However, this is approximate and not always reliable.

◇ Method 2: Use of Quantiles – Exact Method for Exponential

Since exponential data is skewed, control limits can be set using quantiles of the exponential distribution. Given:

• You assume the exponential distribution is the correct model.

• You want the control limits to capture a certain percentage of the in-control process.

Then the LCL and UCL are set using percentiles from the exponential distribution:

✳ Control Limits for Individuals (based on quantiles):

If you want to set the control limits at, say, 3-sigma equivalent (which covers ~99.73% of normal data), then instead of standard deviation, use percentiles:

Let's say your exponential distribution has rate λ , i.e. $X \sim \text{Exp}(\lambda)$, with mean $\mu = 1/\lambda$.

Then: $LCL = [F_{(\alpha/2)}]^{(-1)} = -(1/\lambda) \ln(1 - \alpha/2)$ and $UCL = [F_{1 - (\alpha/2)}]^{(-1)} = -(1/\lambda) \ln(\alpha/2)$

For example, if you want the control limits to cover 99.73% (like $\pm 3\sigma$ in normal distribution), set $\alpha = 1 - 0.9973 = 0.0027$

So: $LCL = -(1/\lambda)\ln(1-0.00135)$ and $UCL = -(1/\lambda)\ln(0.00135)$

This gives you asymmetric control limits appropriate for exponential data.

◇ Method 3: Use of Moving Range (I-MR Chart Alternative)

If you still want to use the Individuals-Moving Range (I-MR) chart:

1. Use mean of individual observations as the center line.
2. Estimate the average moving range \overline{MR}
3. Estimate process standard deviation: $\hat{\sigma} = d_2\overline{MR}$, where $d_2 \approx 1.128$ for moving range of 2.
4. Then control limits: $UCL = \bar{X} + 3\hat{\sigma}$, and $LCL = \bar{X} - 3\hat{\sigma}$

⚠ Caution: This assumes approximate normality, which does not hold for exponential data. Thus, this is not strictly valid unless the exponential is close to symmetric (i.e., large mean).

☑ Recommended: Use Exponential Distribution Quantiles (Method 2)

If you're confident the process follows an exponential distribution (e.g., lifetimes, waiting times, times between failures), the quantile-based control limits (Method 2) are the most appropriate.

📄 Example: Assume $\lambda = 0.5$, so $\mu = 2$

Using $\alpha = 0.0027$, $LCL \approx -2\ln(0.99865) \approx 0.0027$, and $UCL = -2\ln(0.00135) \approx 13.21$

So, your control limits would be: $CL = 2$, $LCL \approx 0.0027$, $UCL \approx 13.21$

🔗 Summary

Method	Use when	Notes
Transformation	Data not too skewed	Approximate, may not always help
Quantiles (recommended)	Exponential model is appropriate	Exact, statistically valid
I-MR Chart	Simpler method needed	Assumes normality; use with caution

If you'd like, I can help generate control limits for your actual data or simulate an example. Just upload your dataset or specify the λ (or sample of data).

Third Query: [NOTICE that the answers are all wrong!]

How can I compute LCL and UCL in Control Charts for individuals Exponentially distributed?

ChatGPT says:

To compute the Lower Control Limit (LCL) and Upper Control Limit (UCL) for Control Charts for individuals when the data is exponentially distributed, we need to acknowledge that:

- The exponential distribution is not symmetric, unlike the normal distribution assumed in traditional Shewhart Individuals (I) charts.
- Standard I-MR (Individuals and Moving Range) charts assume normality, so applying them directly to exponential data can lead to incorrect control limits.
- However, with appropriate transformation or modified control limits, you can still monitor such processes.

🔗 Option 1: Transform the data to normality

If your data $X \sim \text{Exponential}(\lambda)$, then:

- The natural log transformation: $Y = \ln(X)$ results in Y having a distribution that is closer to Extreme Value Type I (Gumbel), but still more manageable than raw exponential.

However, since control charts often require normality, it's common to apply log transformation and then use regular Individuals charts on the transformed data. Here's how:

☑ Steps for Log-Transformed I-chart:

1. Take natural log of all your data points: $Y_i = \ln(X_i)$

□ 2 Compute:

- \bar{Y} : mean of Y_i
- $MR_i = |Y_i - Y_{i-1}|$: moving ranges of logged data
- \overline{MR} : average of moving ranges

□ 3 Estimate the standard deviation using: $\sigma_Y = \overline{MR} / 1.128$

□ 4 Compute UCL and LCL for the I-chart of logged data:

$$UCL_Y = \bar{Y} + 3\sigma_Y \quad LCL_Y = \bar{Y} - 3\sigma_Y$$

□ 5 To get control limits in the original scale, exponentiate the limits:

$$UCL_X = e^{UCL_Y} \quad LCL_X = e^{LCL_Y}$$

This approach makes the data approximately normal, so standard control chart techniques become valid.

🔗 Option 2: Use control charts designed for exponential data

Some practitioners and researchers propose specialized control charts for exponential distributions, like:

1. Exponential EWMA (Exponentially Weighted Moving Average)
 - EWMA is robust to non-normality and can be tailored for exponential data.
2. Cumulative Sum (CUSUM) control charts
 - Particularly useful when dealing with non-normal distributions.
3. Exact UCL/LCL for exponential data

If you insist on not transforming the data, here's a rough approximation based on percentiles: Let $X \sim \text{Exponential}(\lambda)$, then: Mean: $\mu=1/\lambda$, Standard deviation: $\sigma=1/\lambda$

If you approximate 3-sigma limits like a normal distribution (though not ideal), then:

$$UCL=\mu+3\sigma=1/\lambda+3/\lambda=4/\lambda \quad LCL=\mu-3\sigma=1/\lambda-3/\lambda=-2/\lambda$$

But LCL would be negative, which is not possible for exponential data (since it's strictly positive).

So we either:

- Set LCL = 0 (common in some practices), or
- Use percentile-based control limits, like:

$$LCL=F^{-1}(p1), \quad UCL=F^{-1}(p2)$$

where F^{-1} is the inverse CDF of the exponential distribution, and $p1, p2$ are lower and upper percentiles (e.g., 0.0015 and 0.9985 for 3-sigma equivalent).

$$F^{-1}(p)=-\ln(1-p)/\lambda$$

Recommended Approach

Use the log-transformed individuals chart, which works well in practice and maintains simplicity. Here's a summary:

Step	Action
1	Take $Y_i=\ln(X_i)$
2	Compute \bar{Y} , MR_i , and \overline{MR}
3	Estimate $\sigma_Y = \overline{MR}/1.128$
4	Get UCL and LCL in log scale: $UCL_Y = \bar{Y} + 3\sigma_Y$ $LCL_Y = \bar{Y} - 3\sigma_Y$
5	Exponentiate limits: $UCL_X = e^{UCL_Y}$ $LCL_X = e^{LCL_Y}$

Let me know if you'd like a worked example or Excel template for this.

===== End of answers of AI to our queries. =====

NOTICE that the AI answers are wrong for Exponentially distributed data: AI is actually AIG (Artificial Ignorance).

The following section explains why.

2.3. Statistics and Reliability Integral Theory (RIT)

We are going to present the fundamental concepts about RIT (Reliability Integral Theory) that we use for computing the Control Limits (*Confidence* Limits) of CCs. RIT is the natural way for Sequential Tests, because the transitions happen *sequentially*; to provide the ideas, we use a "4 units Stand-by system", depicted by 5 states (Figure 5): 0 is the state with all units not-failed; 1 is the state with the first unit failed; 2 is the state with the second unit failed; and so on, until the system enters the state 5 where all the 4 units are failed (down state, in yellow): any transition provides a datum to be used for the computations. RIT can be found in the author's books...

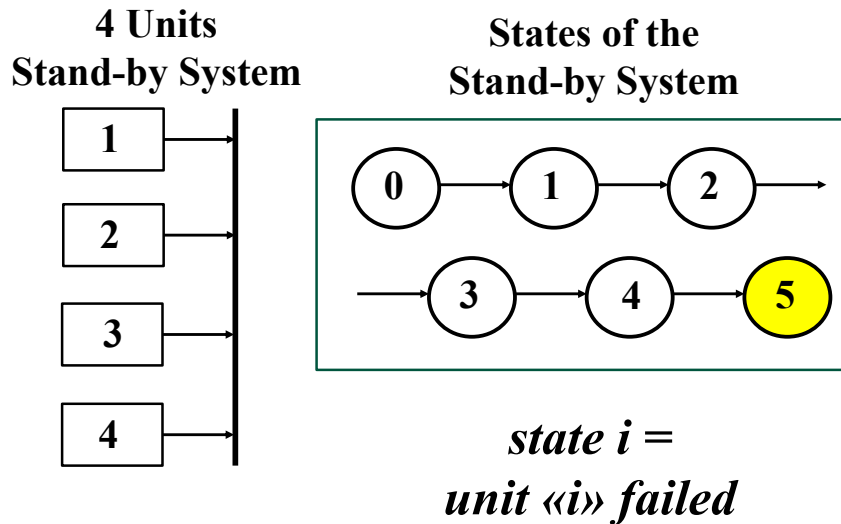


Figure 5. A “4 units Stand-by system” and its states.

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 $0 \rightarrow t$; 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 5)]

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 λ is the failure rate of each unit and $\lambda = 1/\theta$: θ 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 in Figure 5), 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 $j \rightarrow j+1$, in the interval $s \rightarrow s+ds$.

The system reliability $R_0(t)$ is the solution of the mathematical system of the Integral Equations (13)

$$R_j(t - t_j) = \bar{W}_j(t - t_j) + \int_{t_j}^t b_{j,j+1}(t - t_j) R_{j+1}(t - s) ds \quad (13)$$

for $j = 0, 1, \dots, g - 1$, $R_g(t|t_g) = \bar{W}_g(t|t_g)$

With $\lambda e^{-\lambda(s-t_j)}$ we obtain the solution (see Figure 5, putting the Mean Time To Failure MTTF of each unit $= \theta$, $\lambda = 1/\theta$) (see the Figure 6)

$$R_0(t) = e^{-\lambda t} \left[1 + \lambda t + \frac{(\lambda t)^2}{2!} + \frac{(\lambda t)^3}{3!} + \frac{(\lambda t)^4}{4!} \right] \quad (13a)$$

The reliability system (13) can be written in matrix form,

$$R(t - r) = \bar{W}(t - r) + \int_r^t B(s - r) R(s) ds \quad (14)$$

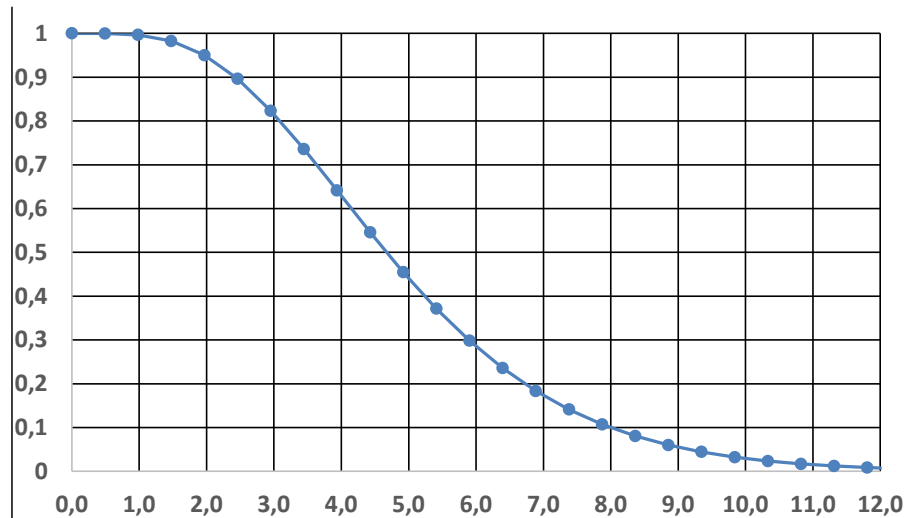


Figure 6. Example of Reliability $R_0(\lambda t_0)$ of a “4 units Stand-by system” with $MTTF=\theta=123$ days; t_0 is the total time on test of the 4 units. To compute the CI (with $CL=0.8$), find the abscissas of the intersections at $R_0(\lambda_U t_0) = 0.9$ and $R_0(\lambda_L t_0) = 0.1$

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 $D=\{x_1, x_2, \dots, x_g\}$, 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 $D^*=\{t_1, t_2, \dots, t_{g-1}, t_g, t=\text{end of the test}\}$ (times of the transitions t_j).

We consider now that we want to estimate the unknown $MTTF=\theta=1/\lambda$ of each item comprising the “associated” stand-by system [19–30]: 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 (14), where $T(t)$ is the “Total Time on Test” $T(t) = \sum_1^g x_i [t_0]$ 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 $D=\{x_1, x_2, \dots, x_g\}$. At the end time t of the test, the integral equations, constrained by the constraint D^* , provide the equation

$$\partial \ln \det B(s|r; \theta, D^*) / \partial \theta = \theta/g - T(t) = 0 \quad (15)$$

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

The same happens with any other distribution (e.g. see the Table 2) provided that we write the kernel $b_{i,i+1}(s)$.

The reliability function $R_0(t|\theta)$, [formula (13)], with the parameter θ , 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 (θ_L Lower and θ_U Upper) of the “unknown” mean θ , to be estimated, for any type of distribution (Exponential, Weibull, Rayleigh, Normal, Gamma, Inverted Weibull, General Inverted Weibull, ...); by solving, with (a general) unknown (indicated as) θ , the two equations $R_0(t_0|\theta) = 1 - \alpha/2$ and $R_0(t_0|\theta) = \alpha/2$; we get the two values (θ_L, θ_U) such that

$$R_0(t_0|\theta_L) = \alpha/2 \text{ and } R_0(t_0|\theta_U) = 1 - \alpha/2 \quad (16)$$

where t_0 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. $CI=\theta_L \text{-----} \theta_U$ is the Confidence Interval: $\theta_L = 1/\lambda_U$ and $\theta_U = 1/\lambda_L$.

For example, with Figure 6, we can derive $\theta_L = 62.5 \text{ days} = 1/\lambda_U$ and $\theta_U = 200 \text{ days} = 1/\lambda_L$, 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 16) for computing the CI; the only difference is that the author FG defined the procedure in 1982 [44], 35 years before Meeker et al.

As said before, we can use RIT for the Sequential Tests; we have only to consider the various transitions and the Total Time on Test to the last transition we want to consider.

2.4. Control Charts for TBE Data: Some Ideas for Phase I Analysis

Let's consider now TBE (Time Between Event, time between transitions) data, *exponentially or Weibull distributed*. Quite a lot of authors (in the "Garden ... [19]") *compute wrongly the Control Limits* (which are the Confidence 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 (6)): 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.

In the previous section, we computed the $CI = \theta_L \dots \theta_U$ of the parameter θ , using the (subsample) "transition times durations": $t_0 =$ "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* $\bar{t}_0 = t_0/n$ of each transition, for the... ; we solve the two following equations (17) for the two unknown values LCL and UCL, for $R(\bar{t}_0 | \theta)$ of each item in the sample, similar to (16)

$$R(\bar{t}_0 | LCL) = \alpha/2, \quad R(\bar{t}_0 | UCL) = 1 - \alpha/2 \quad (17)$$

where now $\bar{t}_0 = t_0/n$ 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 $UCL = 1/\lambda_L$.

For *exponentially distributed data* (17) becomes (18) [6-33], $k=1$, with $CL = 1 - \alpha$

$$e^{-[\bar{t}_0/LCL]} = 1 - \alpha/2 \quad \text{and} \quad e^{-[\bar{t}_0/UCL]} = \alpha/2 \quad (18)$$

The endpoints of the $CI = LCL \dots UCL$ 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 figs. 8, 9, 10). The figures are justified by the Theory [6-33] and are 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 [(18), for $k=1$], theoretically derived with a sound Theory [6-33], with the ones in the Excerpt [in the Appendix (a small sample from the "Garden ... [19]")] and notice that the two Minitab authors (Santiago&Smith) use the "empirical mean \bar{t}_0 " in place of the θ_0 in the Figure 1: it is the same trick of replacing \bar{x} to the mean μ which is valid for the Normal distributed data only; e.g., see the formulae (1)!

In the next sections we can see the Scientific Results found by a Scientific Theory.

3. Results

In this section, firstly, we provide the scientific analysis of the "Piston Rings" data [16] and compare our result with those of the authors: the findings are completely different and the decisions, consequently, should be different, with different costs of wrong decisions.

3.1. Control Charts for Piston Rings data. Phase I Analysis

The Inside Diameter Measurements for Piston Rings data are in the Table 3.

D. C. Montgomery uses a \bar{x}, s Charts to assess the “state of control” of the production process, with Control Limits (see (12))

$$LCL_x = \bar{x} - 3\bar{s}(c_4\sqrt{5}), UCL_x = \bar{x} + 3\bar{s}(c_4\sqrt{5}) \text{ and } LCL_s = 0, UCL_s = 1.964\bar{s}$$

He concludes “There is no indication that the process is out of control, so those limits could be adopted for phase II monitoring of the process.”

We agree with his conclusion.

On the contrary, the authors of [16] decide the opposite, with their “Distribution-free Control Chart for Joint Monitoring of Unknown Location and Scale Parameters”!

To understand we analysed the type of distribution and found that the Normal Distribution with mean $\mu=74.001$ and $\sigma=0.0114$; the CIs are $73.998 < \mu < 74.004$ and $0.0098 < \sigma < 0.0137$, both with CL=99%.

In this case the AI formulae (for normal data) provide the right answer.

On the contrary, the authors of [16] find the opposite.

Table 3. Inside Diameter Measurements (mm) for Automobile Piston Rings (Table 6.3 of the Montgomery Book, Introduction to Statistical Quality Control 8th Ed.).

Sample Number	Observations					\bar{x}_i	s_i
1	74.030	74.002	74.019	73.992	74.008	74.010	0.0148
2	73.995	73.992	74.001	74.011	74.004	74.001	0.0075
3	73.988	74.024	74.021	74.005	74.002	74.008	0.0147
4	74.002	73.996	73.993	74.015	74.009	74.003	0.0091
5	73.992	74.007	74.015	73.989	74.014	74.003	0.0122
6	74.009	73.994	73.997	73.985	73.993	73.996	0.0087
7	73.995	74.006	73.994	74.000	74.005	74.000	0.0055
8	73.985	74.003	73.993	74.015	73.988	73.997	0.0123
9	74.008	73.995	74.009	74.005	74.004	74.004	0.0055
10	73.998	74.000	73.990	74.007	73.995	73.998	0.0063
11	73.994	73.998	73.994	73.995	73.990	73.994	0.0029
12	74.004	74.000	74.007	74.000	73.996	74.001	0.0042
13	73.983	74.002	73.998	73.997	74.012	73.998	0.0105
14	74.006	73.967	73.994	74.000	73.984	73.990	0.0153
15	74.012	74.014	73.998	73.999	74.007	74.006	0.0073
16	74.000	73.984	74.005	73.998	73.996	73.997	0.0078
17	73.994	74.012	73.986	74.005	74.007	74.001	0.0106
18	74.006	74.010	74.018	74.003	74.000	74.007	0.0070
19	73.984	74.002	74.003	74.005	73.997	73.998	0.0085
20	74.000	74.010	74.013	74.020	74.003	74.009	0.0080
21	73.982	74.001	74.015	74.005	73.996	74.000	0.0122
22	74.004	73.999	73.990	74.006	74.009	74.002	0.0074
23	74.010	73.989	73.990	74.009	74.014	74.002	0.0119
24	74.015	74.008	73.993	74.000	74.010	74.005	0.0087
25	73.982	73.984	73.995	74.017	74.013	73.998	0.0162
						$\Sigma = 1,850.028$	0.2351
						$\bar{\bar{x}} = 74.001$	$\bar{\bar{s}} = 0.0094$

The authors of [16] decide that the process is OOC, with their “Distribution-free Control Chart for Joint Monitoring of Unknown Location and Scale Parameters” and write:

In this section, we illustrate the proposed distribution-free SC chart using the well-known piston ring data discussed in Montgomery18 (Table V.1 and V.2, respectively; Figure 3). Piston rings for an automotive engine are produced by a metallic workshop process. The aim is to maintain statistical control of the inside diameters of the rings manufactured by this process. Twenty-five samples each of size 5 (shown in Table V.1 of Montgomery18) are assumed to have been observed a-priori. A Phase-I analysis in Montgomery concluded that one may consider the data set with 125 observations as a reference sample. Therefore, in the present context, $m=125$. Moreover, in Montgomery, 18 15 Phase-II samples (test samples) each of size 5 are also provided. This, in turn, indicates that $n=5$. Through Monte-Carlo simulations, we find $H=6.1$ for a target ARL0 of 500 and for $m=125$ and $n=5$. Values of the 15 SC plotting statistics are obtained as: 2.29, 0.068, 2.17, 0.39, 1.19, 0.76, 0.57, 1.28, 2.46, 2.59, 0.25, 7.53, 8.24, 12.34 and 2.54, respectively. These are displayed in Figure 2 along with the UCL 6.1. The control chart shows that the process remains IC for the first 11 test samples and goes OOC for the first time at sample number 12. In fact, all three consecutive test samples 12, 13 and 14, seem to come from an OOC process, indicating possibly a shift in location, or scale, or both. In this context, it is worth mentioning that the SL chart as in Mukherjee and Chakraborti also declared a shift at the same time point.

Excerpt 5. From “Distribution-free Control Chart for Joint Monitoring of Unknown Location and Scale Parameters”.

3.2. Control Charts for “Time Intervals Between Consecutive Earthquakes ... in the Mount St. Helens Region of the Washington State, After January 1, 1978”. Phase I Analysis

We continue our route for Quality by considering the papers “ARL-Unbiased Exponential Control Charts With Estimated Parameters: Statistical Design and Implementation” (Quality and Reliability Engineering International 2025), “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 2021), and “Improved Phase I Control Charts for Monitoring Times Between Events” (Quality and Reliability Engineering International 2014).

There are, as well, several drawbacks in many papers dealing with Control Charts (CCs) related to “Time Between Events Exponentially distributed data”. As far as 2000 you can find many papers with wrong methods for computing the Control Limits of the CCs; several are listed in the *Garden of Flowers* [19], in Academia.edu (where you find often the two above authors).

The problem is, as well, wrongly solved in Minitab 19&20&21 and JMP, in spite that the Companies Management was informed about it.

Both the papers use the same data (taken form a paper of Santiago & Smith) about the about the recorded time intervals between consecutive earthquakes of magnitude 1.0 or higher in the Mount St. Helens region of the Washington State, after January 1, 1978.

The “theory” given there about the CCs with exponentially distributed data is almost the same; we show it through some excerpts copied from the papers.

Assuming a known value of the “constant event rate” [notice the two different symbols θ and λ] we read the Excerpt 1, for the years 2025 and 2021:

consider a situation where the quality characteristic (e.g., time to an event, TBE) X follows an exponential distribution with probability density function $f(x) = \theta \exp(-\theta x)$, for $x > 0$, where $\theta > 0$ is the event rate. Let θ_0 denote IC rate parameter. For monitoring X (or θ_0) the popular exponential chart (or the t -chart, see [15]) is the two-sided	the time between two consecutive events in a homogeneous Poisson process with constant rate, say, λ . Then, X follows an exponential distribution with mean $1/\lambda$ and has the density function $f(x) = \lambda \exp(-\lambda x)$; $x > 0; \lambda > 0$. The process is said to be IC when $\lambda = \lambda_0$, otherwise it is called OOC when
---	---

<p>equal-tails Shewhart-type chart for a specified FAR=α_0, where the (probability) control limits are obtained from setting $P[X < LCL \theta_0] = P[X > UCL \theta_0] = \alpha_0/2$. These yield $LCL = \theta^{-1}A_1$ and $UCL = \theta^{-1}A_2$ (1) where the LCL and UCL denote the lower and upper control limit, respectively, with $A_1 = -\ln(1 - \alpha_0/2)$ and $A_2 = -\ln(\alpha_0/2)$.</p>	<p>$\lambda = \lambda_1, (\lambda_0 \neq \lambda_1)$ For a nominal false alarm rate (FAR), α_0, the equal-tailed limits LCL and UCL can be obtained from $P[X < LCL IC] = P[X > UCL IC] = \alpha_0/2$ as follows. $LCL = -\ln(1 - \alpha_0/2)/\lambda_0 = A_1/\lambda_0$ and $UCL = -\ln(\alpha_0/2)/\lambda_0 = A_2/\lambda_0$</p>
QREI 2025	QT & QM 2021

Excerpt 6. Statements from the two papers published by *Quality and Reliability Engineering International* and *Quality Technology & Quantitative Management*.

Notice immediately that the formulae in the Excerpt 6 provide the Probability Limits L and U (vertical segment) of the Random Variable X (*exponentially distributed*) NOT the Control Limits (LCL and UCL, horizontal segment) of the Control Chart [9–25], as the reader can see in the Figure 1 [θ_0 is the assumed known value of the parameter θ which is the mean $\theta = 1/\lambda$ of the exponential pdf $f(x) = \lambda \exp(-\lambda x)$, as written in the 2021 paper (Excerpt 6)].

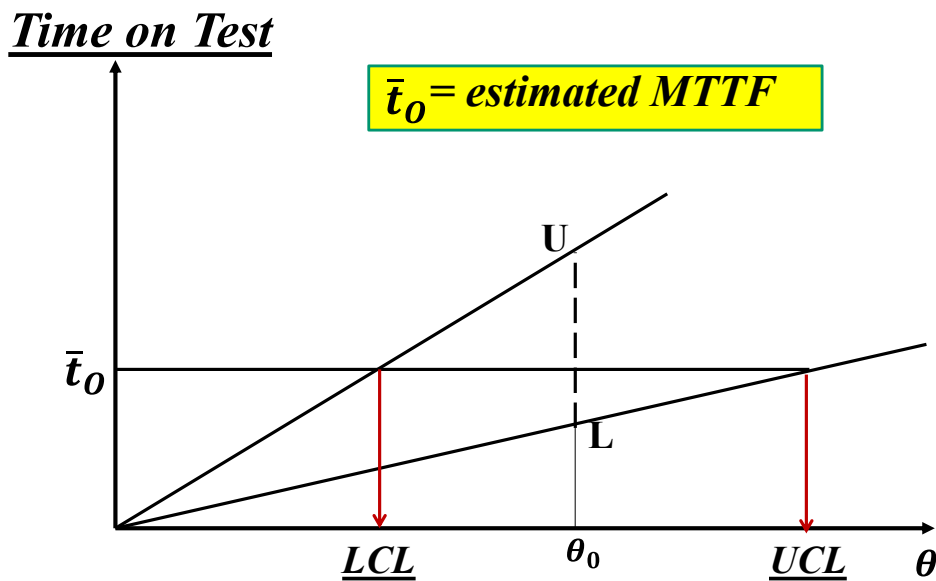
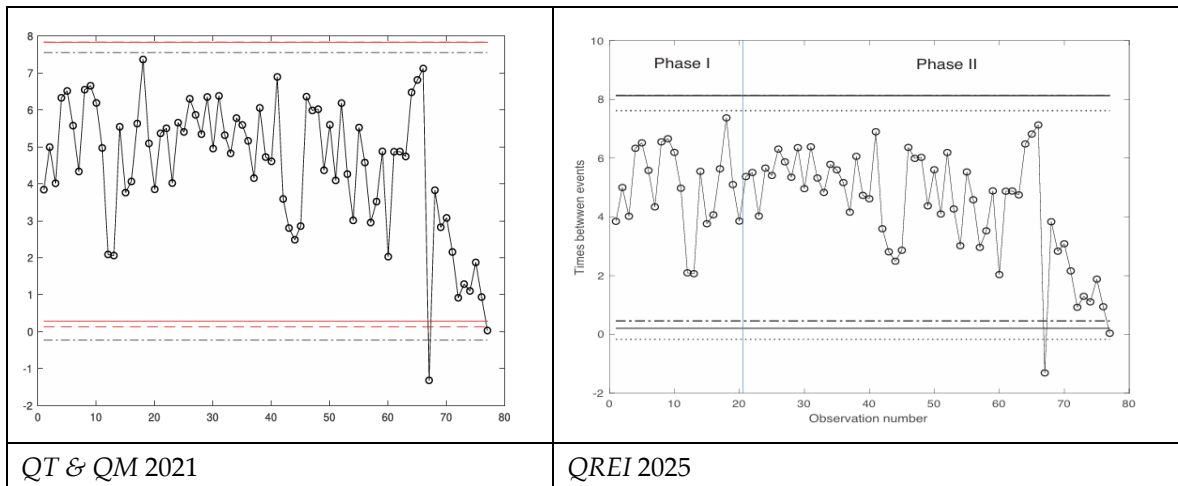


Figure 7. LCL and UCL for the Control Chart (exponentially distributed data).

The two authors analyse the recorded time intervals between consecutive earthquakes in the Mount St. Helens region and provide the figures in the Excerpt 7.

Notice that the data are shown as the natural logarithm of the original time intervals between consecutive earthquakes.

We present immediately in the Figure 8 the scientific analysis of time intervals between consecutive earthquakes, where the Control Limits are derived by the first 20 data as done in the 2025 *QREI* paper.



Excerpt 7. Control Charts from the two papers published by QT & QM and QREI (time intervals between consecutive earthquakes, In vertical scale).

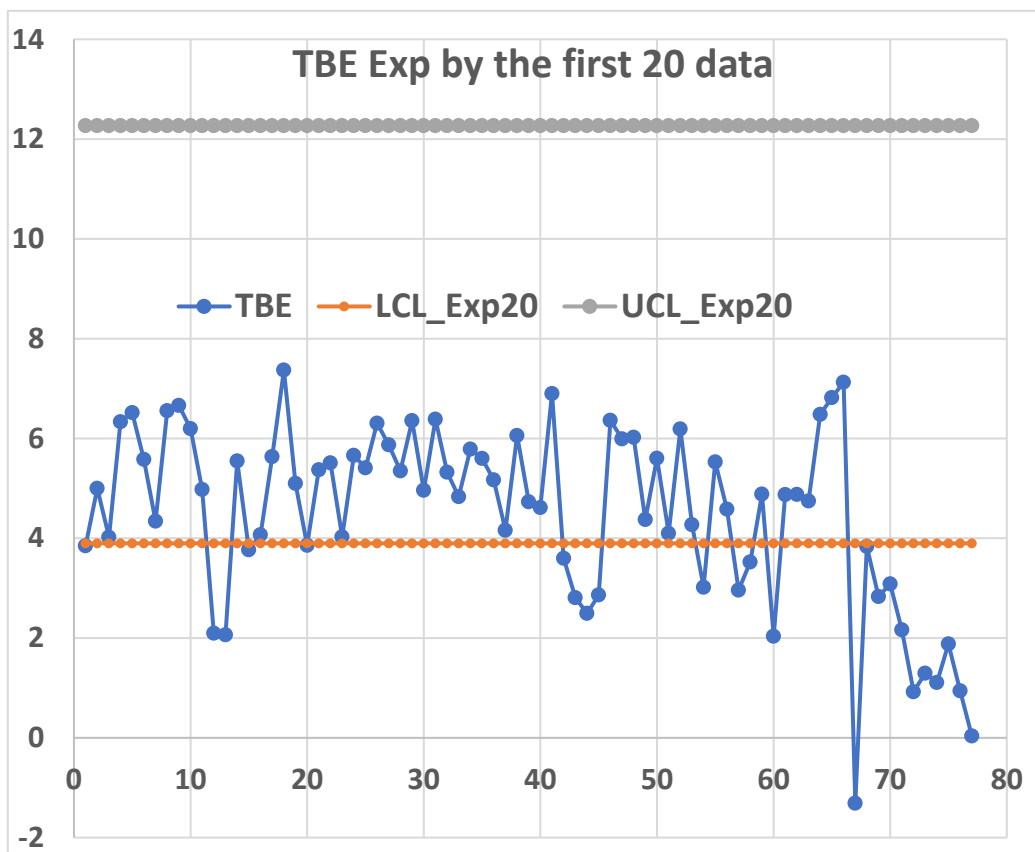


Figure 8. Scientific LCL and UCL, computed with the first 20 data, of the Control Chart (exponentially distributed data [9–36]) of the time intervals between consecutive earthquakes (In vertical scale), to be compared with the Excerpt 7.

It is clear that the results are completely different: the 2021 and 2025 papers have “questionable theory”.

Figure 9 shows the result provided by JMP (left side) while the right side provides the result by Santiago & Smith.

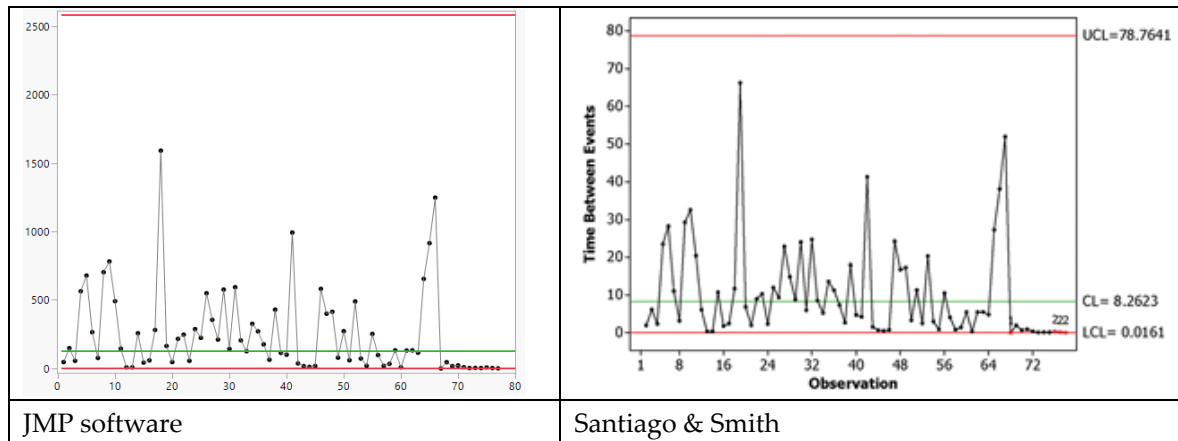


Figure 9. Control Charts of the time intervals between consecutive earthquakes to be compared with the Excerpt 7 and the Figure 7, for finding the IC and OOC situations.

The same authors wrote a paper “Improved Phase I Control Charts for Monitoring Times Between Events” [17], using a completely different “theory” from the 2021/2025 papers. In the Abstract, they say

*In many situations, the times between certain events are observed and monitored instead of the number of events particularly when the events occur rarely. In this case, it is common to assume that the times between events follow an exponential distribution. Control charts are one of the main tools of statistical process control and monitoring. Control charts are used in phase I to assist operating personnel in bringing the process into a state of statistical control. In this paper, phase I control charts are considered for the observations from an exponential distribution with an and out-of-control performance of the proposed chart. It is seen that **the proposed charts are considerably more in-control robust than** two competing charts and have comparable out of control properties. Copyright © 2014 John Wiley & Sons, Ltd.*

Notice the wrong statement: “It is seen that **the proposed charts are considerably more in-control robust than** two competing ...”. That is the problem: K&C method shows In Control (IC) situations when actually the process is Out Of Control (OOC); the method is wrong.

The authors use the following symbols:

- $\bar{X} = \sum_1^n X_i/n$ denoting the sample average of the $X_i, i=1, 2, \dots, n$,
- α_0 denoting the overall false alarm rate,
- $X_{(l)}, X_{(m)}, X_{(u)}$ denoting the first, the second, and the third quartile, respectively, of the ordered data.

They go on by writing

*... we investigate the IC robustness of the one-sided control charts to the assumption of the underlying exponential distribution via simulation... then carried out for the two-sided charts. The IC robustness is an important attribute of a control chart and should be investigated thoroughly, because in practice, the underlying distribution may not be exactly exponential. **The more IC robust the control chart, the more confidence the user will have on the advertised false alarm rate.***

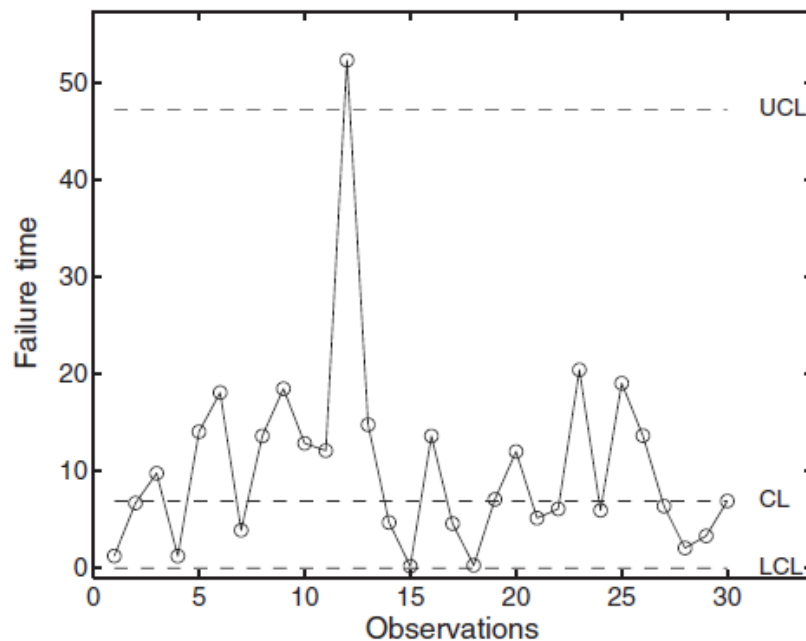
The K&C method shows In Control (IC) situations when actually the process is Out Of Control (OOC); the method is wrong!

The data, given in the paper, are in the Table 4.

Table 4. Time between failures data (“Improved Phase... for Monitoring TBE”. [17]).

1.24	6.69	9.77	1.23	14.03	18.07	3.90	13.61	18.47	12.85
52.32	14.75	4.69	0.18	13.61	4.57	0.28	7.08	12.00	5.15
6.09	20.41	5.93	19.03	13.65	6.37	2.06	3.30	6.91	12.08

The K&C analysis, with its *OOC (Out Of Control)*, is in the Figure 4:

**Figure 10.** CC from “Improved Phase... for Monitoring TBE”. Excerpt 6 from [17].

The authors say:

Table ... shows a set of 30 failure time data generated from a Poisson distribution with a mean of 0.1. For these data, $n = 30$, $l=8$, $m= 15$, and $u = 23$. We monitor these data with the proposed phase I chart. The center line for the proposed two-sided control chart is $CL=X(15)=6.91$, and the lower and UCL are given by $LCL=-53.9213$ and $UCL=47.2320$. Because $LCL<0$, we set the LCL as $LCL=0$. It can be seen from Figure (our 5) that **the eleventh observation 52.32 plots outside the UCL**, which indicates an OOC situation that needs further investigation. Note that for these data, neither the Dovoedo and Chakraborti, nor the Jones and Champ control chart indicates any OOC situation.

On the contrary, the scientific solution is in Figure 5 (vertical axe logarithmic): UCL is >100.

Using RIT (devised by F. Galetto) we find that the Process is OOC (Out Of Control) for the opposite reason stated by the two authors! See Figure 11.

Comparing the figures 10 and 11 it becomes clear that the Control Chart from “Improved Phase... for Monitoring TBE” [17] presents 5 errors about OOC: there is NO OOC for the value (the eleventh observation 52.32) and there are 4 OOC values below the LCL.

The *Garden of Flowers* [19] can be considered as our literature review...

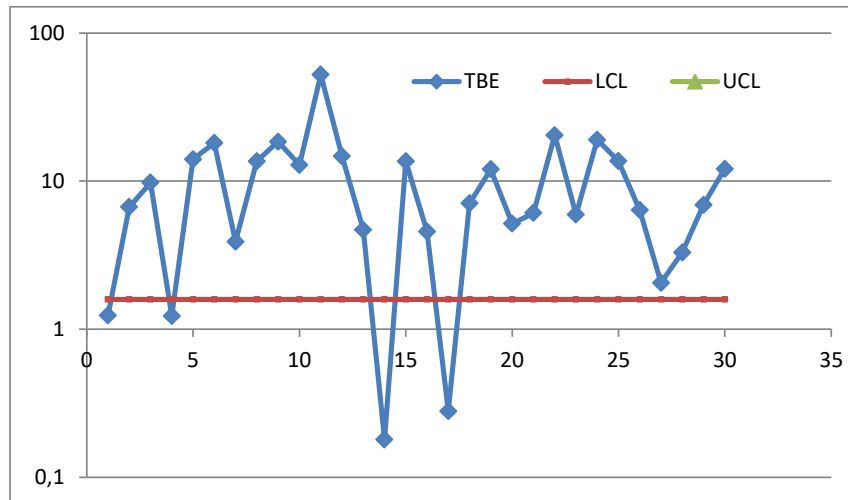


Figure 11. F. Galetto Control Chart of the data from “Improved Phase... for Monitoring TBE”. [17]; vertical axe logarithmic; UCL is >100.(RIT used). Process OOC.

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.

Now we apply RIT to the data (“time intervals between consecutive earthquakes ...” figured) in the Excerpt 7, where the authors applied the “theory” in the Excerpt 6.

We consider the paper *QREI 2025*, where the model (as written by the authors) is “the quality characteristic (e.g., time to an event, TBE) X follows an exponential distribution with probability density function $f(x) = \theta \exp(-\theta x)$, for $x > 0$, where $\theta > 0$ is the event rate.” It is a very strange notation because, in the literature, the event rate is λ (as in the *QT & QM 2021*, related to the Poisson distribution), and $\theta = 1/\lambda$ is the mean of the exponential pdf).

The two authors K&C (with a number of citations 304 and 8,466 respectively) state that the analysis should be based on the Excerpt 8:

is the two-sided equal-tails Shewhart-type chart for a specified FAR α_0 , where the (probability) control limits are obtained from setting $P[X < LCL | \theta_0] = P[X > UCL | \theta_0] = \alpha_0/2$. These yield

$$LCL = \theta_0^{-1} A_1 \text{ and } UCL = \theta_0^{-1} A_2 \quad (1)$$

where the LCL and UCL denote the lower and upper control limit, respectively, with $A_1 = -\ln(1 - \alpha_0/2)$ and $A_2 = -\ln(\alpha_0/2)$. Let

Excerpt 8. From *QREI 2025*, named by the authors “case K” when the “rate” θ_0 is known.

Note all the words in the Excerpt 9 and in particular “plugging-in” for the “case U” when the “rate” θ_0 is unknown and the “estimator $\hat{\theta}_0$ ” is “plugged-in”...

note that in case U, when θ_0 is unknown, a set of control limits may be obtained simply by substituting (plugging-in) an estimator $\hat{\theta}_0$ in the control limits for case K. Thus, a set of intuitive

Excerpt 9. From *QREI 2025*, named “case U” when the “rate” θ_0 is unknown.

This “plugging-in” provides wrong Control Limits LCL and UCL for any distribution not-Normal, while it has no consequences (see the formulae (7)) for the Normal Distribution (in spite of *being theoretically wrong*). The THEORY in the section 2 shows that.

The first thing to do when we analyse data “supposed exponentially distributed” is to assess if the exponential pdf (1) fits “well” the data.

The two authors K&C did not care about that: they “assumed data exponentially distributed”!

A suitable alternative is the Weibull pdf of the TTF (Time To Failure) Random Variable, with parameters β (shape) and η (scale); the mean MTTF depends on both:

$$f(x|\eta, \beta) = (1/\eta)(x/\eta)^{\beta-1} \exp [-(x/\eta)^\beta] \quad (8)$$

The solution of the two equations $\det [B(s|r; \eta, \beta, D) = 0$, for the 77 data is $\hat{\beta} = 0.727$, and Confidence Interval $CI_\beta = [\hat{\beta}_L = 0.589, \hat{\beta}_U = 0.922]$, with Confidence Level 95%, and $\hat{\eta} = 204.23$, Confidence Interval $CI_\eta = [\hat{\eta}_L = 152.72, \hat{\eta}_U = 282.80]$, with Confidence Level 95%.

Since $1 \notin CI_\beta$ we *cannot* assume, with Confidence Level 95%, that the data are exponentially distributed. Our analysis is in the Figure 12.

The Figure 13 shows the difference (graphical) between the distribution of the data and the Exponential: it is clear that ... as shown by the confidence intervals.

Transforming the Weibull data into Exponential data we find the CC in the Figure 14.

Since the authors K&C (in the 2025 paper) used the first 20 data, as Phase I, to find the Control Limits we did the same: we searched for the estimates of the Weibull distribution and found the solution of the two equations $\det [B(s|r; \eta, \beta, D) = 0$, for the first 20 data (Phase I); we got $\hat{\beta}^{Phase I} = 0.823$, and Confidence Interval $CI_\beta = [\hat{\beta}_L^{Phase I} = 0.555, \hat{\beta}_U^{Phase I} = 1.347]$, with CL=95%, and $\hat{\eta}^{Phase I} = 287.75$, and with CL=95% we computed the Confidence Interval $CI_\eta = [\hat{\eta}_L^{Phase I} = 178.18, \hat{\eta}_U^{Phase I} = 523.78]$.

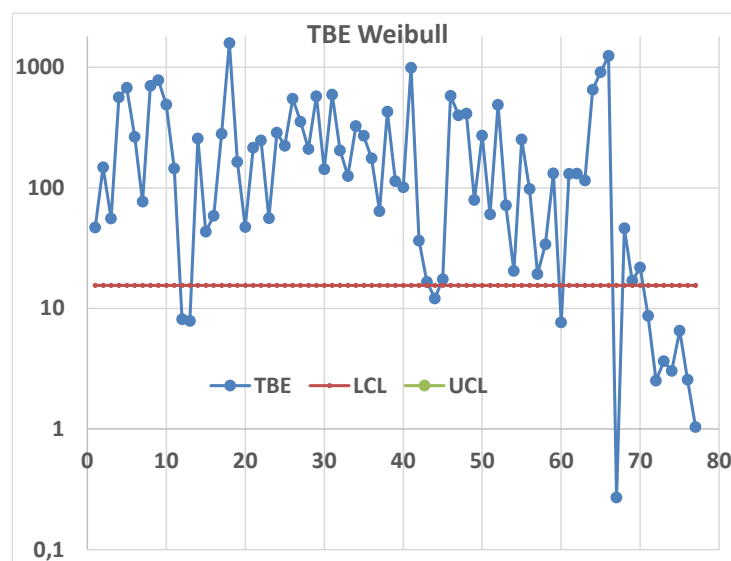


Figure 12. F. Galetto Control Chart of the “time intervals between consecutive earthquakes ...”, using the Weibull distribution; vertical axe logarithmic; UCL is >10000.(RIT used). Process OOC.

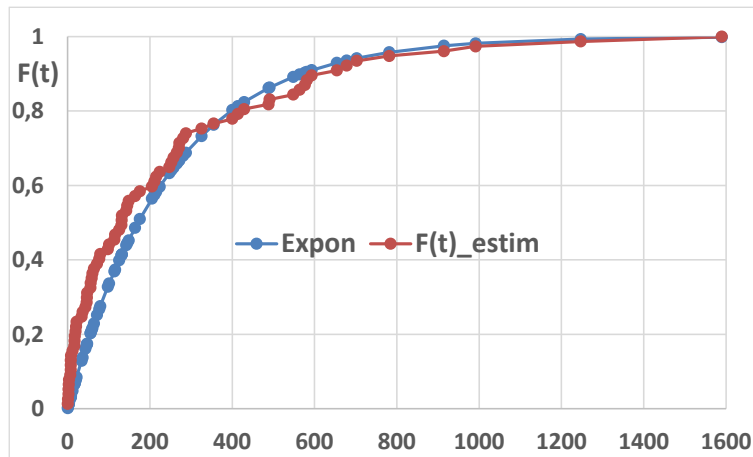


Figure 13. Distribution (red) of the “time intervals between consecutive earthquakes”, compared to the Exponential (blue).

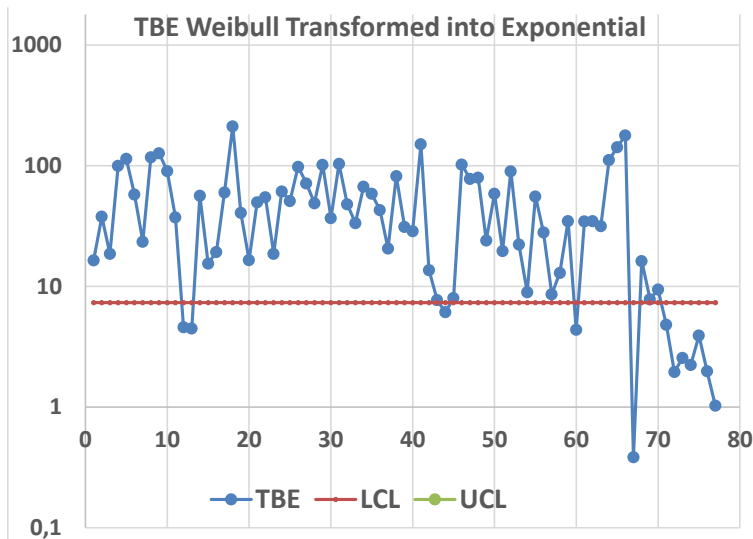


Figure 14. F. Galetto Control Chart of the “transformed (into Exponential) time intervals between consecutive earthquakes”; vertical axe logarithmic; UCL is >1000.(RIT used). Process OOC.

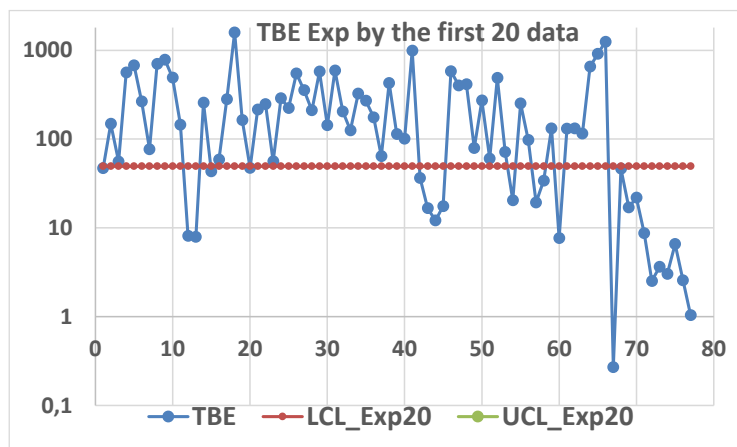


Figure 15. F. Galetto Control Chart of the “time intervals between consecutive earthquakes”, with Control Limits found by the Phase I 20 data; vertical axe logarithmic; UCL is >100000. (RIT used). Process OOC.

Since $1 \in CI_{\beta}$ we can assume, with Confidence Level 95%, that the data are exponentially distributed and the Control Limits can be derived with the exponential distribution. Our analysis is in the Figure 12.

Compare Figure 12 with the figures in the Excerpt 7: Control Chart (on the right) from the paper published by QREI 2025 (time intervals between consecutive earthquakes, ln vertical scale).

You see that the K&C analysis (based on the wrong formulae, in the excerpts 8 and 9) is not coherent with the Theory [19–74].

All the results are found via RIT (Reliability Integral Theory).

4. Discussion

We decided to use the data from the papers [16–18] and the analysis by the authors.

We got different results from those authors: the cause is that they use the Probability Limits of the PI (Probability Interval) as they were the Confidence Limits (Control Limits 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), $q=1,2, \dots, n$, and any other mean $\mu_{\bar{x}(t_r)}$ (parameter), $r=1,2, \dots, n$, are different, with risk α , if their estimates are not both included in their common Confidence Interval as the CI of the grand mean $\mu_{\bar{x}} = \mu$ (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{X} (vertical interval) and the Random Interval comprising the unknown mean μ (horizontal interval) (Figure 16)

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

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

This is particularly important for the Individual Control Charts for Exponential, Weibull, Inverted Weibull, General Inverted Weibull, Maxwell and Gamma distributed data: this is what Deming calls “*Profound Knowledge* (understanding variation)” [29,30]. In this case, you see that the Confidence Interval is the realisation of the horizontal Random Interval. The same happens for any distribution.

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 [19].

Since a lot of papers (related to Exponential, Weibull, Inverted Weibull, General Inverted Weibull, Maxwell and Gamma distributions), with the same problem as that of “*The garden of flowers*” [19], are published in reputed Journals we think that an “*alternative*” title “*History is written by the winners. Reflections on Control Charts for Process Control*” should be suitable for this paper: *the authors of the wrong papers [19] are the winners.*

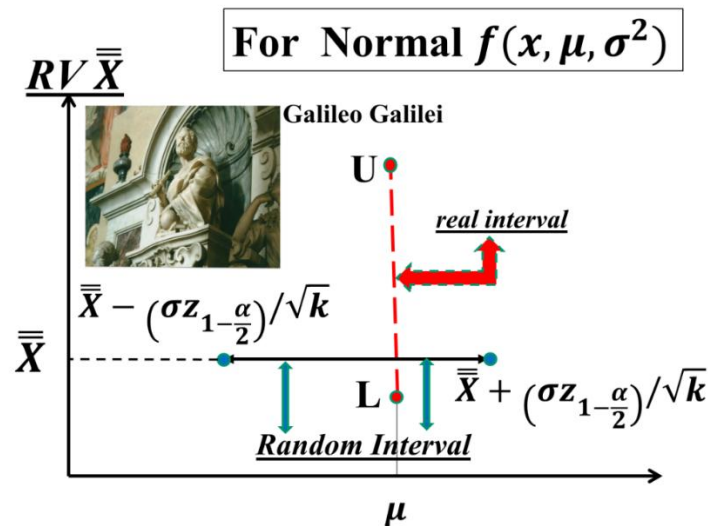


Figure 14. Probability Interval L–U (vertical line) versus Random Intervals comprising the “mean” μ (horizontal random variable lines), for Normally distributed RVs $\bar{X} \sim N(\mu, \sigma^2)$.

Further studies should consider other distributions which cannot be transformed into the above distributions considered before.

5. Conclusions

With our figures (and the Appendix, that is a short extract from the “Garden ... [19]”) we humbly ask the readers to look at the references [19–74] 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 ... [19]”, 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). 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. The basics of RIT have been given.

We could have shown many other cases (from papers not mentioned here, that you can find in [19–75]) 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 (Galletto 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. [24]

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 [19–74]. 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.

We informed the authors and the Journals who published wrong papers by writing various letters to the Editors [1–15] ...: 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 [19–74].

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.

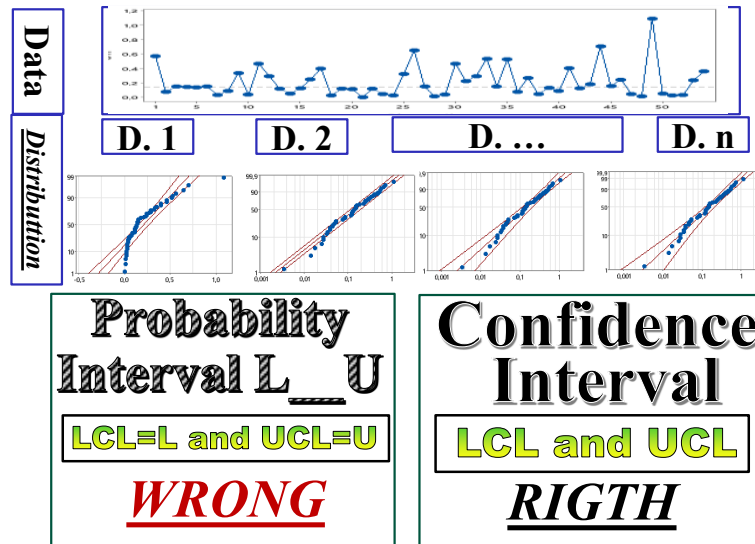


Figure 15. Probability Intervals L—U versus Confidence Intervals LCL—UCL in Control Charts.

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 [29,30].

Unfortunately, many authors do not know Scientifically the role (concept) of Confidence Intervals for Hypothesis Testing. The same happens for AIg...

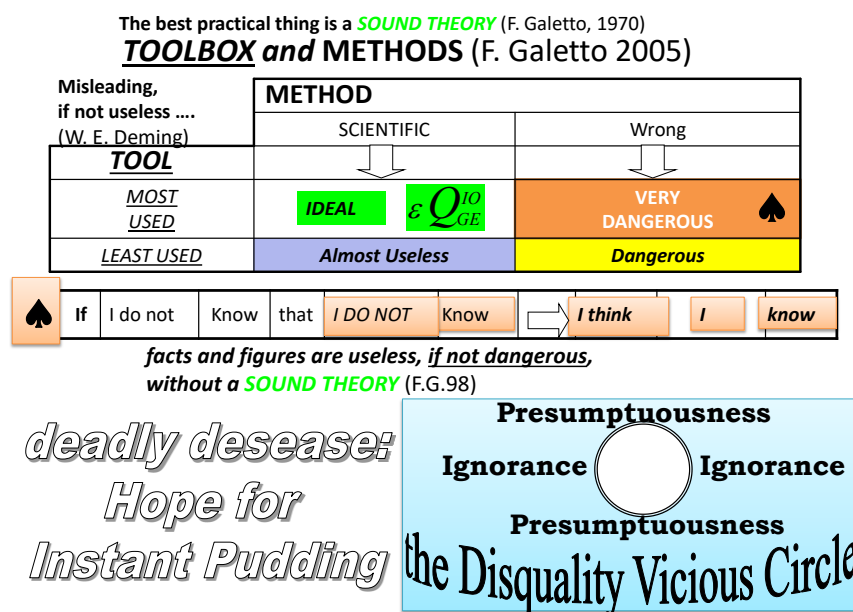


Figure 16. Knowledge versus Ignorance, in Tools and Methods. The same happens for AIg....

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, $H_0=\{\text{Process In Control}\}$ versus $H_1=\{\text{Process Out Of Control}\}$, 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. The same happens for AIg...

Humbly, given our commitment to Quality and our long-life love for it [19–74], 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.

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-December 2024 we *again* verified (through several *new* downloaded papers, not shown here) that the *Pandemic Disease about the (wrong) Control Limits, that are actually the Probability Limits of the PI* is still present ...

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* [29,30] *Profound Knowledge, Metanoia, Theory*. AIg is the opposite of Knowledge...

As mentioned before, when the parameter is unknown, it is estimated from a phase I sample of size m , say, Y_1, \dots, 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 $\hat{\lambda}_0$ into the control limits in Equation (1) as follows.

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

These control limits are known as the conditional control limits conditioned on a given phase I sample (or a given estimated value of λ_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 $\hat{\beta}(\delta) = P[X < \widehat{LCL} | \lambda_1] + P[X > \widehat{UCL} | \lambda_1]$.

Excerpt 10. From *“Conditional analysis of Phase II exponential chart... an event”*, *Q. Tech. & Quantitative Mgt*, '19.

We think that we provided the readers with several ideas and methods to be meditated in view of the applications, generating wealth for the companies using them.

The documents [19–74] are very important: ASSURE ...

There is no “free lunch”: *metanoia and study are needed and necessary.*

Funding: This research received no external funding.

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

Appendix A (Related to [19])

Typical statement by ALL ...

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 ($i=1, 2, \dots$) re independent and identically distributed random variables with pdf $f(x) = \theta^{-1} \exp(-x/\theta)$ for $x \geq 0$, 0 otherwise, where θ is the “mean time between events”.

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

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

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 $\log(2) \bar{t}$.”

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

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

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) α_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))

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

where $A_1 = \frac{\chi_{2r, \alpha_0/2}^2}{2}$, $A_2 = \frac{\chi_{2r, 1-\alpha_0/2}^2}{2}$ are the design constants and λ_0 is the known or specified IC rate parameter value. The $\chi_{2r, a}^2$ 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 $\text{CL} = \frac{\chi_{2r, 0.5}^2}{2\lambda_0}$.

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

ACTUALLY LCL=L and UCL=U

Chakraborti et. al. (with several papers...)

Excerpt A1. Typical statements in the “Garden ...[19]” where the authors name LCL and UCL what **actually** are the Probability Limits L and U. See the Figure 9 and the Excerpt 10.

Many other cases, with the same errors, can be found in the “Garden ...[19], and the Conclusions” where *the authors name LCL and UCL what actually are the Probability Limits L and U.*

There is no “free lunch”: metanoia and study are needed and necessary.

References

1. ResearchGate: M. Ananda Articles
2. ResearchGate: S. Chakraborti Articles
3. ResearchGate: S. Chowdhury Articles
4. ResearchGate: A. McCracken Articles
5. ResearchGate: O. Dag Articles
6. ResearchGate: W. Kossler Articles
7. ResearchGate: N. Kumar Articles
8. ResearchGate: S. Lian Articles
9. ResearchGate: B. Liu Articles
10. ResearchGate: H. Murakami Articles
11. ResearchGate: A. Mukherjee Articles
12. ResearchGate: A. Rakitzis Articles
13. ResearchGate: S. Weerahandi Articles
14. ResearchGate: J. Zhang Articles

15. Journals: Quality and Reliability Engineering International, Quality Engineering, Reliability Engineering and System Safety, IIE Transactions, Quality Technology and Quantitative Management, Quality & Quantity, J. of Applied Statistics, J. of Quality Technology, J. of Statistical Planning and Inference, J. of Statistical Computation and Simulation, International J. of Production Research, J. of the American Statistical Association, International J. for Quality Research, J. of Quality Technology, J. of Statistical Theory and Practice, J. of Statistical Theory and Applications, J. of Applied Probability and Statistics, International J. of Advanced Manufacturing Technology, J. of Systems Science and Information, European J. of Operational Research, J. of Nonparametric Statistics, J. of the Japanese Society of Computational Statistics, Communications in Statistics - Simulation and Computation, Statistics and Probability Letters, Communications in Statistics - Theory and Methods, Computational Statistics, Computational Statistics & Data Analysis,
16. Chakraborti S. et al., Distribution-free Control Chart for Joint Monitoring of Unknown Location and Scale Parameters, Quality and Reliability Engineering International (QREI), 2013 and ARL-Unbiased Exponential Control Charts With Estimated Parameters : Statistical Design and Implementation, Quality and Reliability Engineering International, May 2025
17. N. Kumar et al., Improved Shewhart-type Charts for Monitoring Times Between Events, Journal of Quality Technology, 2016, and Improved Phase I Control Charts for Monitoring Times Between Events, QREI, 2014,
18. N. Kumar, 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, 2021
19. Galetto, F., (2024), The garden of flowers (*including various of the above authors and Journals*), Academia.edu
20. Galetto F. (2020-25) Various Letters to the EDITORS of QREI (Quality Reliability Engineering International, Academia.edu
21. Galetto, F., (2024), JMP and Minitab betray Quality, Academia.edu
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., Quality of methods for quality is important. European Organisation for Quality Control Conference, Vienna. 1989
25. 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
26. Belz, M. Statistical Methods in the Process Industry: McMillan; 1973.
27. Casella, Berger, Statistical Inference, 2nd edition: Duxbury Advanced Series; 2002.
28. Cramer, H. Mathematical Methods of Statistics: Princeton University Press; 1961.
29. Deming W. E., Out of the Crisis, Cambridge University Press; 1986.
30. Deming W. E., The new economics for industry, government, education: Cambridge University Press; 1997.
31. Dore, P., Introduzione al Calcolo delle Probabilità e alle sue applicazioni ingegneristiche, Casa Editrice Patron, Bologna; 1962.
32. Juran, J., Quality Control Handbook, 4th, 5th ed.: McGraw-Hill, New York: 1988-98.
33. Kendall, Stuart, (1961) The advanced Theory of Statistics, Vol 2, Inference and Relationship:, Hafner Publishing Company; 1961.
34. Meeker, W., Hahn, G., Escobar, L. Statistical Intervals: A Guide for Practitioners and Researchers. John Wiley & Sons. 2017
35. Mood, Graybill, Introduction to the Theory of Statistics, 2nd ed.: McGraw Hill; 1963.
36. Rao, C. R., Linear Statistical Inference and its Applications: Wiley & Sons; 1965.

37. Rozanov, Y., *Processus Aleatoire*, Editions MIR: Moscow, (traduit du russe); 1975.
38. Ryan, T. P., *Statistical Methods for Quality Improvement*: Wiley & Sons; 1989.
39. Galetto, F., (2025). *Sequential Tests and Reliability Integral Theory*. Eng OA, 3 (7), 01-37
40. Shewhart W. A., *Economic Control of Quality of Manufactured Products*: D. Van Nostrand Company; 1931.
41. Shewhart W.A., *Statistical Method from the Viewpoint of Quality Control*: Graduate School, Washington; 1936.
42. D. J. Wheeler, Various posts, Online available from Quality Digest
43. Galetto, F., *Affidabilità Teoria e Metodi di calcolo*: CLEUP editore, Padova (Italy); 1981-94.
44. Galetto, F., *Affidabilità Prove di affidabilità: distribuzione incognita, distribuzione esponenziale*: CLEUP editore, Padova (Italy); 1982, 85, 94.
45. Galetto, F., *Qualità. Alcuni metodi statistici da Manager*: CUSL, Torino (Italy); 1995-99).
46. Galetto, F., *Gestione Manageriale della Affidabilità*: CLUT, Torino (Italy); 2010.
47. Galetto, F., *Manutenzione e Affidabilità*: CLUT, Torino (Italy); 2015.
48. Galetto, F., *Reliability and Maintenance, Scientific Methods, Practical Approach, Vol-1*: www.morebooks.de.; 2016.
49. Galetto, F., *Reliability and Maintenance, Scientific Methods, Practical Approach, Vol-2*: www.morebooks.de.; 2016.
50. Galetto, F., *Statistical Process Management*, ELIVA press ISBN 9781636482897; 2019
51. 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
52. 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
53. 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
54. 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
55. Galetto F., (2021) *Minitab T charts and quality decisions*, Journal of Statistics and Management Systems, DOI: 10.1080/09720510.2021.1873257
56. Galetto, F., (2012) *Six Sigma: help or hoax for Quality?*, 11th Conference on TQM for HEI, Israel
57. Galetto, F., (2020) *Six Sigma_Hoax against Quality_Professionals Ignorance and MINITAB WRONG T Charts*, HAL Archives Ouvert, 2020.
58. Galetto, F., (2021) *Control Charts for TBE and Quality Decisions*, Academia.edu
59. Galetto F. (2023) *Control Charts, Scientific Derivation of Control Limits and Average Run Length*, International Journal of Latest Engineering Research and Applications (IJLERA) ISSN: 2455-7137 Volume – 08, Issue – 01, January 2023, PP – 11-45
60. Galetto, F., (2006) *Quality Education and quality papers*, IPSI, Marbella (Spain)
61. Galetto, F., (2006) *Quality Education versus Peer Review*, IPSI, Montenegro
62. Galetto, F., (2006) *Does Peer Review assure Quality of papers and Education?* 8th Conference on TQM for HEI, Paisley (Scotland)
63. Galetto, F., (1998), *Quality Education on Quality for Future Managers*, 1st Conference on TQM for HEI (Higher Education Institutions), Toulon (France)
64. Galetto, F., (2000), *Quality Education for Professors teaching Quality to Future Managers*, 3rd Conference on TQM for HEI, Derby (UK)

65. Galetto, F., (2001), Looking for Quality in "quality books", 4th Conference on TQM for HEI, Mons (Belgium)
66. Galetto, F., (2001), Quality and Control Charts: Managerial assessment during Product Development and Production Process, AT&T (Society of Automotive Engineers), Barcelona (Spain)
67. Galetto, F., (2001), Quality QFD and control charts, Conference ATA, Florence (Italy)
68. Galetto, F., (2002), Business excellence Quality and Control Charts", 7th TQM Conference, Verona (Italy)
69. Galetto, F., (2002), Fuzzy Logic and Control Charts, 3rd ICME Conference, Ischia (Italy)
70. Galetto, F., (2002), Analysis of "new" control charts for Quality assessment, 5th Conference on TQM for HEI, Lisbon (Portugal)
71. Galetto, F., (2009), The Pentalogy, VIPSI, Belgrade (Serbia)
72. Galetto, F., (2010), The Pentalogy Beyond, 9th Conference on TQM for HEI, Verona (Italy)
73. Galetto, F., (2021) ASSURE: Adopting Statistical Significance for Understanding Research and Engineering. Journal of Engineering and Applied Sciences Technology. SRC/JEAST-128. DOI: [https://doi.org/10.47363/JEAST/2021\(3\)118](https://doi.org/10.47363/JEAST/2021(3)118)
74. Galetto, F., (2024), News on Control Charts for JMP, Academia.edu

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.