LVI SBPO SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL
DE 04 ATÉ 07 NOVEMBRO 2024
FORTALEZA | CEARÁ

$\overline{X}$ Control Chart to Monitor Weibull Process by Markov Chain Approach

Roberto. C. Ferreira
Production Engineering Dept, USP
05508-220, São Paulo, Brazil
roberto.ferreira@usp.br

Linda Lee Ho
Production Engineering Dept, USP
05508-220, São Paulo, Brazil
lindalee@usp.br

Roberto C. Quinino
Statistics Dept, UFMG
31270-901, Belo Horizonte, Brazil
robertoquinino@gmail.com

ABSTRACT

Shewhart control charts are widely used to monitor the mean of a process. However, when the process variable follows a Continuous Weibull distribution, the Control Limits cannot be directly calculated due to the lack of a closed-form expression for the sum of Weibull random variables for $n > 1$. This study proposes a Markov chain approach to determine the control limits for a Shewhart control chart monitoring the mean of the Weibull distribution. The Markov chain approach discretizes the Continuous Weibull distribution, allowing for the calculation of the sum's distribution and subsequent determination of the control limits. The proposed method's performance is evaluated using the Average Run Length metric and compared to existing approaches. The results demonstrate the effectiveness of the proposed Markov chain approach in monitoring the mean of a Weibull distribution.

KEYWORDS. Discrete Approximation. Continuous Distribution. Control Limits.
PAPER TOPICS. Control Charts for Weibull Processes, Monitoring Weibull Distribution Mean, Markov Chain Approach to Control Limits

1. Introduction

The Weibull distribution is a versatile statistical model widely used in engineering, reliability, economics, and survival analysis. Although the distribution offers versatility in modeling real-world phenomena, its application in statistical process control (SPC), particularly with Shewhart control charts, is restricted by the lack of a closed-form expression for the sum of Weibull random variables for sample sizes larger than one. This limitation inhibits the calculation of control limits (CLs), which are essential for effective process monitoring.

In reliability engineering, the sum of Weibull distributions models system lifetimes (e.g., [Fernandes et al., 2019], while in manufacturing and survival analysis, it characterizes time-to-failure intervals (e.g., [Vasconcelos et al., 2022] and [Silva et al., 2024]). Additionally, in logistics and inventory management, it helps optimize inventory levels (e.g., [Saaty and Siskin, 1980]). Despite its broad applicability, determining CLs for continuous Weibull distributions (CWD) remains problematic, with existing methods either yielding inconsistent results or presenting challenges in usability.

To address these challenges, this study proposes a novel Markov chain (MC) approach to discretize the CWD, enabling the calculation of the sum's distribution and subsequent determination of CLs. By discretizing the continuous variable into small ranges, cumulative probabilities are calculated, facilitating the determination of CLs to achieve a predefined average run length (ARL_0). Moreover, this approach allows for the computation of the resulting distribution and ARL_1 when both CWD parameters vary, without the need of Monte Carlo simulations.

The efficacy of the proposed method is evaluated using the ARL metric, offering a comprehensive comparison with existing approaches. Overall, the study aims to enhance process control and facilitate the detection of out-of-control conditions in systems following CWD, thereby contributing to advancements in statistical process monitoring.

2. Weibull Distribution - a brief review

The Weibull distribution, denoted as $W(\lambda, k)$, is a continuous probability distribution characterized by two parameters: the scale parameter λ and the shape parameter k. Its probability density function (PDF) is given by: $f(x;\lambda,k) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}$ where x is a continuous random

https://proceedings.science/p/193714?lang=pt-br

variable, $\lambda > 0$ the scale parameter, $k > 0$ the shape parameter. The Weibull distribution exhibits diverse shapes depending on the value of the shape parameter k. When $k < 1$, the distribution indicates a decreasing failure rate, implying that the probability of failure decreases over time. For $k = 1$, the distribution represents a constant failure rate, also known as the exponential distribution. When $k > 1$, the distribution shows an increasing failure rate, indicating that the probability of failure increases over time. The mean (E[X]) and variance (Var[X]) of the Weibull distribution are given, respectively, by: $E(X) = \lambda\Gamma\left(1+\frac{1}{k}\right)$ and $\text{Var}(X) = \lambda^2\left[\Gamma\left(1+\frac{2}{k}\right) - \left(\Gamma\left(1+\frac{1}{k}\right)\right)^2\right]$ where $\Gamma(.)$ is the Gamma function. When monitoring a process using SPC, it is important to consider how changes in the CWD parameters can affect the distribution of observations. When changing one or two parameters of the Weibull distribution, the characteristic of the curve may obscure the detection of an out-of-control process, resulting in high ARL_1 values (average run length to the first signal).

3. The $\overline{X}$ Control Chart for Continuous Weibull Distribution

This section details the development of a $\overline{X}$ control chart aimed at monitoring the stability of $\mathbb{E}(X)$ when $X \sim \text{CWD}(\lambda, k)$. The methodology applies to scenarios where events occur between periods, with a sample size $n \geq 1$. Consider a random sample $X_1, X_2, ..., X_n$ drawn from a CWD. Monitoring $\overline{X}$ is essentially equivalent to monitoring $Y = \sum_{i=1}^{n} X_i$. Although both variables are significant, we opt for $\overline{X}$ due to its straightforward interpretation as a single summary measure of the sample results, while Y represents the total of the sampled units, influenced by sample size n. To construct the $\overline{X}$ control chart, it is crucial to determine the distribution function of $Y = \sum_{i=1}^{n} X_i$. The possible values of Y belong to the set $\mathbb{R}^+$, and the probability of values being less than or equal to a specific value y is given by: $P(Y \leq y) = P\left(\sum_{i=1}^{n} X_i \leq y\right)$. For CWD X_i, this expression does not yield a closed-form solution, but can be computed numerically. A recursive function can be implemented to evaluate this for possible values of Y up to a threshold T, ensuring $P(Y \leq T) \approx 1$, thus approximating the distribution function of Y. However, this computational approach demands substantial processing power, especially for larger sample sizes n. To mitigate this, we propose a method that involves a CWD discretization and a MC approach.

Consider the transition probability matrix $\mathbf{M}$ of an MC characterized by states $S_1, S_2, \ldots, S_f$ with values $a_1, a_2, \ldots, a_f$. Here, $a_1 = 0$ (always) and $a_f = T$, where $T = t \times n$ and $t \in \mathbb{R}^+$ such that $P(X \leq t) \approx 1$, assuming $X \sim \text{CDW}(\lambda, k)$. Thus, $P(Y \leq T) \approx 1$, where $Y_i = \sum_{j=1}^{i} X_j$. The range between consecutive values of the states (e.g., $a_5 - a_4$) is referred to as R. The total number of states f is given by $f = \frac{T}{R} + 1$. In this context, the transition probabilities are defined as follows: from state S_1 to S_1 is the cumulative probability from 0 to the midpoint between a_1 (which is 0) and a_2; from state S_1 to S_2 is the cumulative probability from the midpoint between a_1 and a_2 to the midpoint between a_2 and a_3; this pattern continues for subsequent transitions until the transition between S_1 and state S_{f-1}; and from state S_1 to S_f is the difference between 1 and the sum of the probabilities of the other transitions from S_1, ensuring that the sum of each row equals 1. The elements of the transition matrix $M(i, i+1)$, for $i = 1, 2, \ldots, f-1$, represent the conditional transition probability $P(Y_{i+1} = a_{i+1} \mid Y_i = a_i)$, where $i \geq 1$ and $a_{i+1} > a_i$. The transition matrix $\mathbf{M}$ is:

	S_1	S_2	S_3	$\cdots$	S_f
	$a_1 = 0$	$a_2 = \text{R}$	$a_3 = 2 \times \text{R}$	$\cdots$	$a_f = (f-1) \times \text{R}$
S_1 - a_1	$P(0 \leq X < \frac{\text{R}}{2})$	$P(\frac{\text{R}}{2} \leq X < \frac{3\times\text{R}}{2})$	$P(\frac{3\times\text{R}}{2} \leq X < \frac{5\times\text{R}}{2})$	$\cdots$	$1 - \sum_{x=1}^{f-1} P(X = S_x)$
S_2 - a_2	0	$P(0 \leq X < \frac{\text{R}}{2})$	$P(\frac{\text{R}}{2} \leq X < \frac{3\times\text{R}}{2})$	$\cdots$	$1 - \sum_{x=1}^{f-1} P(X = S_x)$
S_3 - a_3	0	0	$P(0 \leq X < \frac{\text{R}}{2})$	$\cdots$	$1 - \sum_{x=1}^{f-1} P(X = S_x)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
S_f - a_f	0	0	0	$\cdots$	1

Define $\mathbf{b} = (1, 0, 0, ..., 0)$ as a vector of dimension $f+1$. Then, $\mathbf{v} = \mathbf{bM}^n$ is the probability vector of $Y_n (= \sum_{i=1}^{n} X_i)$. For illustration, consider $X \sim \text{CWD}(\lambda = 4, k = 3)$, with $n = 3$. Through the CWD with MC approach, we find $t = 13.30$ that ensures $P(Y \leq t \times n) \approx 1$. And we obtain a CDF of Y.

To set CLs with a Type I error of $\alpha = 0.27$, for bilateral shifts, interpolate between the two states whose cumulative probabilities encompass the desired value ($\alpha/2$ or $1-\alpha/2$) to calculate the CLs. In our example, the bilateral shifts are LCL = 1.490 and UCL = 5.863.

4. $\bar{X}$ Performance Evaluation

To assess the efficacy of the proposed MC approach in monitoring the mean of a CWD, we conducted a thorough performance evaluation. The primary metric used in this evaluation was the ARL, representing the average number of samples taken before an out-of-control signal is detected. Under the null hypothesis (H_0), assuming that the process is in control with the base parameters λ_0 = 4.0 and k_0 = 3.0 for the CWD, we calculated ARL_0. In contrast, under the alternative hypothesis (H_1), where the process is considered out of control with altered parameters λ_1 and/or k_1, we computed ARL_1. Using the proposed MC approach, we calculated the ARL values (see DMC in Table 1) and compared them with those obtained through Monte Carlo simulations (see SIMin Table 1).

Table 1 provides a summary of the ARL values obtained for sample sizes (n = 3, 5, 10, 30, 100), scale (λ = 2.4, 3.2, 4.8, 5.6) and shape parameters (k =0.5, 5, 10), indicating the performance under both H_0 and H_1 scenarios. For example, under H_0 with $n = 10$, the simulated ARL_0 was 370.39, closely matching the MC-calculated value of 369.72. ARL values greater than 10,000 were indicated as ∞. The results demonstrated that the MC approach provided ARL values consistent to simulation results. This highlighted the accuracy and reliability of the MC method in calculating CLs for Weibull-distributed processes. Analysis of ARL_0 showed that the MC approach maintained values close to the expected value under the null hypothesis, ensuring robust performance. Additionally, under the alternative hypothesis, the MC approach effectively detected shifts in process parameters while maintaining a high degree of sensitivity, as evidenced by comparable ARL_1 values to simulation results. Importantly, a key distinction between the methods lies in the fact that the simulation-based approach generates CLs that vary with each iteration, whereas our method consistently produces the same values. This difference ensures that our approach provides stable and predictable CLs, simplifying interpretation and implementation.

Table 1: Performance Evaluation

Sample size	Shape	Scale									
		λ = 2.40		λ = 3.20		λ_0 = 4.00		λ = 4.80		λ = 5.60	
n	k	SIM	DMC	SIM	DMC	SIM	DMC	SIM	DMC	SIM	DMC
3	3	14.01	13.95	118.20	118.06	369.70	370.37	23.50	23.48	4.86	4.86
	0.5	1.74	1.70	1.68	1.65	1.61	1.58	1.55	1.52	1.49	1.46
	5	106.76	107.32	5661.87	5657.22	∞	∞	273.69	273.88	6.95	6.95
	10	∞	∞	∞	∞	∞	∞	∞	∞	16.39	16.38
5	3	3.78	3.79	51.84	51.85	370.58	370.38	15.05	15.07	3.01	3.01
	0.5	1.69	1.66	1.60	1.58	1.50	1.49	1.41	1.41	1.35	1.34
	5	9.37	9.41	1794.11	1766.19	∞	∞	65.13	64.83	2.82	2.81
	10	266.72	268.96	∞	∞	∞	∞	∞	∞	2.04	2.03
10	3	1.20	1.20	13.91	13.94	369.72	370.39	7.22	7.22	1.62	1.62
	0.5	1.62	1.61	1.46	1.46	1.33	1.33	1.23	1.24	1.17	1.18
	5	1.15	1.16	188.74	189.96	∞	∞	10.60	10.58	1.24	1.24
	10	1.15	1.15	∞	∞	∞	∞	15.75	15.77	1.01	1.01
30	3	1.00	1.00	1.88	1.88	369.45	370.39	2.03	2.04	1.02	1.02
	0.5	1.47	1.48	1.20	1.22	1.08	1.10	1.04	1.04	1.02	1.02
	5	1.00	1.00	3.37	3.40	∞	∞	1.35	1.36	1.00	1.00
	10	1.00	1.00	146.28	149.31	∞	∞	1.00	1.00	1.00	1.00
100	3	1.00	1.00	1.00	1.00	371.39	370.35	1.02	1.02	1.00	1.00
	0.5	1.22	1.27	1.02	1.04	1.00	1.00	1.00	1.00	1.00	1.00
	5	1.00	1.00	1.00	1.00	4357.30	4512.36	1.00	1.00	1.00	1.00
	10	1.00	1.00	1.00	1.00	6013.23	5949.17	1.00	1.00	1.00	1.00

5. Application

In this section, we apply the proposed MC approach to a practical dataset previously examined in studies such as those by in [Fernandes et al., 2019]. The objective is to determine

whether there has been a shift in the mean of the breaking strengths of carbon fibers used in fibrous composite materials. For consistency and comparison, the parameters of the Weibull distribution the shape parameter k and scale parameter λ are assumed to be known and set at 4.8 and 3.2 yielding expected Mean and Variance 2.9312 and 0.4856 respectively. We analyzed the dataset to determine the CLs for the $\bar{X}$ control chart. The Lower and Upper CLs were 1.9658 and 3.8149 respectively. Samples that fall outside these CLs are indicative of a potential shift in the process mean. According to our findings, samples 14, 17, and 19 are identified as "Out of Control" (OC), suggesting that these observations deviate significantly from the expected process behavior. The summary of the results is included in Table 2, showing the calculated means for each sample and their respective control status. Overall, the application of the MC approach to real-world data reaffirms its effectiveness in statistical process control, offering a practical solution for monitoring Weibull-distributed processes in various industries.

Table 2: Summary of control chart analysis.

Sample	X_1	X_2	X_3	X_4	X_5	$\bar{X}$	Decision	Sample	X_1	X_2	X_3	X_4	X_5	$\bar{X}$	Decision
1	3.70	2.74	2.73	2.50	3.60	3.054	In Control	11	1.41	3.68	2.97	1.36	0.98	2.080	In Control
2	3.11	3.27	2.87	1.47	3.11	2.766	In Control	12	2.76	4.91	3.68	1.84	1.59	2.956	In Control
3	4.42	2.41	3.19	3.22	1.69	2.986	In Control	13	3.19	1.57	0.81	5.56	1.73	2.572	In Control
4	3.28	3.09	1.87	3.15	4.90	3.258	In Control	14	1.59	2.00	1.22	1.12	1.71	1.528	Out of Control
5	3.75	2.43	2.95	2.97	3.39	3.098	In Control	15	2.17	1.17	5.08	2.48	1.18	2.416	In Control
6	2.96	2.53	2.67	2.93	3.22	2.862	In Control	16	3.51	2.17	1.69	1.25	4.38	2.600	In Control
7	3.39	2.81	4.20	3.33	2.55	3.256	In Control	17	1.84	0.39	3.68	2.48	0.85	1.848	Out of Control
8	3.31	3.31	2.85	2.56	3.56	3.118	In Control	18	1.61	2.79	4.70	2.03	1.80	2.586	In Control
9	3.15	2.35	2.55	2.59	2.38	2.604	In Control	19	1.57	1.08	2.03	1.61	2.12	1.682	Out of Control
10	2.81	2.77	2.17	2.83	1.92	2.500	In Control	20	1.89	2.88	2.82	2.05	3.65	2.658	In Control

6. Conclusions

In this study, we proposed a novel approach to determine CLs to monitor the mean of a Weibull-distributed process using Shewhart control charts. By leveraging MC techniques, we discretized the Continuous Weibull Distribution (CWD), enabling the calculation of the sum's distribution and subsequent determination of CLs. The performance evaluation demonstrated the accuracy and reliability of the MC approach in calculating CLs, with ARL values consistent with simulation results. The method effectively detected shifts in process parameters while maintaining a high degree of sensitivity, making it a valuable tool for statistical process control across various applications.

Overall, the proposed MC approach offers a promising avenue for enhancing process monitoring and quality assurance practices in diverse industries, contributing to improved product quality, reliability, and customer satisfaction.

References

Fernandes, F. H., Lee Ho, L., and Bourguignon, M. (2019). About shewhart control charts to monitor the Weibull mean. *Quality and Reliability Engineering International*, 35(7):2343–2357.

Saaty, T. L. and Siskin, V. (1980). *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. McGraw-Hill.

Silva, L. A., Ho, L. L., and Quinino, R. C. (2024). Markov chain approach to get control limits for a shewhart control chart to monitor the mean of a Discrete Weibull distribution. *Journal of Process Control*, 134:103149.

Vasconcelos, R. M. R., Quinino, R. C., Ho, L. L., and Cruz, F. R. B. (2022). About shewhart control charts to monitor the Weibull mean based on a Gamma distribution. *Quality and Reliability Engineering Internacional*, 38:4210–22.

https://proceedings.science/p/193714?lang=pt-br