Bol. Soc. Paran. Mat.
©SPM – E-ISSN-2175-1188
SPM: www.spm.uem.br/bspm

(3s.) **v. 2026 (44)** : 1–24.
ISSN-0037-8712
doi:10.5269/bspm.79294

A Bootstrap Approach for Constructing Control Charts for Beta-Binomial Processes with Applications in Manufacturing Quality

Mohammed Faris Abed, Ali Akbar Heydari and Hossein Jabbari Khamnei

ABSTRACT: Statistical process control (SPC) for attribute data traditionally relies on p-charts or c-charts based on Binomial or Poisson distributions, which are often inadequate for processes exhibiting over-dispersion. This paper proposes a novel bootstrapbased framework for constructing Shewhart-type control charts specifically designed for processes following a Beta-Binomial distribution (BBD). The methodology utilizes parametric bootstrap sampling to estimate empirical sampling distributions of relevant statistics, enabling accurate control limit derivation without relying on restrictive normality assumptions. Comprehensive Monte Carlo simulations demonstrate that the proposed bootstrap M and S charts effectively maintain desired in-control average run lengths (ARL) while exhibiting high sensitivity in detecting various parameter shifts. The practical utility of this approach is further validated through a real-world case study involving defective transformer counts in manufacturing, where the Beta-Binomial model provided superior fit compared to standard distributions and established reliable control limits for quality monitoring.

Key Words: Statistical process control, beta-binomial distribution, Bootstrap methods, control charts, average run length.

Contents

Submitted October 01, 2025. Published February 04, 2026
2020 *Mathematics Subject Classification:* 62P30, 62G09.

1. Introduction and Literature Review

1.1. Background and Motivation

Achieving success in modern manufacturing and business environments depends heavily on the quality of production. Various approaches can be taken to ensure high quality, with Statistical Process Control (SPC) being a fundamental methodological framework for monitoring, controlling, and improving processes [24]. SPC entails the use of statistical techniques to oversee production processes and detect any potential issues early on, thereby enhancing quality control practices and boosting organizational success.

At the heart of SPC lie control charts, the primary tools used to distinguish between common-cause variation (inherent to the process) and assignable-cause variation (indicative of a process shift). These charts employ random sampling to evaluate and oversee quality parameters during manufacturing processes, pinpointing any deviations from anticipated control levels. The efficacy of any control chart is fundamentally tied to the accurate estimation of its control limits, which are traditionally derived from an assumed underlying probability distribution of the quality characteristic, most commonly the normal distribution.

However, real-world processes often violate the assumption of normality. As manufacturing processes and components grow in complexity, the quality characteristics may follow skewed, heavy-tailed, or complex distributions such as the Weibull [26], Gamma [11], Birnbaum-Saunders [19], Inverse Gaussian [18], log-symmetric [17], or generalized Pareto [4]. Furthermore, processes in reliability engineering [17] and those monitoring system availability [22] often present unique distributional challenges. In such cases, using traditional parametric control charts can lead to inflated false-alarm rates or reduced sensitivity to detect actual process shifts.

1.2. The Bootstrap Methodology in SPC

To address these challenges, the bootstrap methodology, introduced by Efron [12], has emerged as a powerful non-parametric and semi-parametric alternative for constructing control limits. The bootstrap does not rely on strong distributional assumptions but instead uses resampling from the empirical data to approximate the sampling distribution of a statistic and estimate its quantiles for control limits.

The application of bootstrap methods in SPC has a rich and varied history. Seminal work by [32] introduced the subgroup bootstrap technique for constructing control charts. This was followed by [20] and [40], who expanded the concept for both dependent and independent data. [15] provided a critical early performance analysis of bootstrap control charts, establishing their credibility and exploring their properties.

Since these foundational studies, the bootstrap paradigm has been successfully adapted to a vast array of scenarios: for percentiles of specific non-normal distributions [26,19,18,4,21]; for multivariate quality control using Hotelling's T^2 [29,25] and maximum multivariate CUSUM [16]; for profile monitoring [27]; for process capability indices [41]; for monitoring proportion data [7], availability indices [22], and reliability data [17]; and for robust control chart design using model selection approaches [6]. Recent advancements continue to focus on robustness and handling non-normal data [30], including applications for geometric percentiles of the Lindley distribution [1] and monitoring Poisson-Lindley distributed counts [2].

1.3. Monitoring Proportion Data and the Beta-Binomial Distribution

A pervasive and critical scenario in SPC involves monitoring *proportion* or *count* data, which are inherently non-normal. The p-chart is commonly used to monitor the proportion of defective items or parts in a production process, helping to detect any shifts or trends in the proportion of defects over time and allowing for timely intervention to maintain quality standards. Similarly, the c-chart serves as the standard tool for count data. However, the performance of these charts relies on large sample sizes to justify the normal approximation. For processes with low defect counts, highly variable sample sizes, or inherent over-dispersion, these assumptions often fail spectacularly.

The binomial distribution is the natural model for defect counts, but its parameter p (the probability of a defect) may itself be variable from subgroup to subgroup due to unobserved factors or latent heterogeneity. This extra-binomial variation is elegantly captured by the *Beta-Binomial* distribution, where the success probability p follows a Beta distribution, leading to a more dispersed count distribution than the standard binomial [10,14]. The Beta-Binomial distribution, its generalizations [9,28], and its statistical inference [39,13,23,31,5] have been extensively studied and applied in various fields, including biometric identification [38] and analysis of binary measurement methods [34]. Its approximations have also been a subject of research [35,36,37].

While the Beta-Binomial model is well-established for modeling over-dispersed counts, the focus of this article is on the related but distinct *Beta-Binomial* distribution. This distribution, which can be conceptualized as a compound distribution where a binomial random variable serves as the parameter for a beta distribution, offers a flexible framework for modeling random variables that represent proportions or rates which are themselves the result of a prior binomial process.

The Beta-Binomial distribution finds applications in various fields such as:

- **Reliability Engineering**: Analyzing the reliability of systems or components, especially when the number of trials is fixed and the probability of success is known.
- **Bayesian Inference**: Updating beliefs or probabilities based on observed data, making it a valuable tool for decision-making under uncertainty.
- **Marketing and Market Research**: Modeling the success rate of marketing campaigns, customer responses, or survey results.
- **Health Sciences**: Clinical trials, epidemiological studies, and medical research to model the probability of success or failure in treatment outcomes.

Although less commonly discussed in the SPC literature than its Beta-Binomial counterpart, the theoretical properties and potential of the Beta-Binomial distribution for modeling complex random phenomena have been explored in more theoretical statistical contexts [3,33]. It provides a different approach to modeling variance in proportional data.

1.4. Contributions and Structure of This Paper

The primary objective of this paper is to bridge the gap between theoretical developments of the Beta-Binomial distribution and its practical application in SPC. We formally derive and characterize the Beta-Binomial distribution within the context of quality control and demonstrate its practical utility by developing novel parametric bootstrap control charts for monitoring processes where the quality characteristic follows, or can be approximated by, this distribution.

By leveraging the bootstrap's flexibility, as demonstrated in the extensive literature, we can accurately estimate control limits without relying on asymptotic approximations or restrictive assumptions. The performance of these new control charts will be evaluated by calculating the Average Run Length (ARL) index, followed by demonstrating their implementation through a real-world example. This work builds upon the established foundation of bootstrap SPC [8] and contributes a new model-specific tool to its growing arsenal, offering quality practitioners a robust method for handling complex, non-normal proportion data that the Beta-Binomial distribution characterizes effectively.

The remainder of this paper is organized as follows: Section 2 provides the theoretical foundation of the Beta-Binomial distribution. Section 3 details the construction of control charts for Beta-Binomial processes. Section 4 presents a performance evaluation of bootstrap control charts for Beta-Binomial processes. Section 5 demonstrates illustrative examples, including a simulated dataset and a real-world application. Finally, Section 6 concludes the paper.

2. Beta-Binomial Distribution

The Beta-Binomial distribution is a *compound probability distribution* that arises when the probability parameter p of a Binomial distribution follows a Beta distribution. This hierarchical structure makes it particularly useful for modeling over-dispersed count data.

2.1. Definition and Motivation

In Statistical Process Control (SPC), particularly when monitoring proportion data using tools like the p-chart, the fundamental assumption is that the number of nonconforming items X in a sample of size n follows a Binomial distribution: $X \sim \text{Binomial}(n, p)$. This model assumes the probability of a nonconforming item p is constant across all samples [24].

However, this assumption is often violated in real-world processes due to unobserved factors, latent heterogeneity, or external influences that cause p to vary from sample to sample. This phenomenon, known as *over-dispersion*, results in a variance of the observed counts that is larger than the Binomial variance $np(1-p)$. To model this extra variability, the Beta-Binomial distribution provides a powerful and elegant solution by treating the parameter p not as a fixed constant, but as a random variable following a Beta distribution [10].

2.2. Formal Definition

The hierarchical model is defined as:

$$\begin{aligned} P &\sim \text{Beta}(\alpha, \beta), \\ X \mid P = p &\sim \text{Binomial}(n, p), \end{aligned}$$

where $\alpha > 0$ and $\beta > 0$ are the shape parameters of the Beta distribution, and n is the number of trials in the Binomial distribution.

2.3. Probability Mass Function and Derivation

The Beta-Binomial distribution arises as a marginal distribution when integrating over the random probability parameter p. The Binomial distribution describes the conditional distribution of counts:

$$P(X = x \mid p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

while the Beta distribution describes the variation of the probability parameter:

$$f(p; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} p^{\alpha-1}(1-p)^{\beta-1}, \quad 0 \le p \le 1$$

where $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ is the Beta function, $\alpha > 0$ and $\beta > 0$.

The joint distribution of X and p is obtained by combining these components:

$$\begin{aligned} P(X = x, p) &= f(p; \alpha, \beta) \cdot P(X = x \mid p) \\ &= \frac{1}{B(\alpha, \beta)} p^{\alpha-1}(1-p)^{\beta-1} \cdot \binom{n}{x} p^x (1-p)^{n-x} \\ &= \frac{\binom{n}{x}}{B(\alpha, \beta)} p^{\alpha+x-1}(1-p)^{\beta+n-x-1} \end{aligned}$$

The marginal probability mass function of X is found by integrating out the random variable p:

$$\begin{aligned} P(X = x) &= \int_0^1 P(X = x, p) dp \\ &= \frac{\binom{n}{x}}{B(\alpha, \beta)} \int_0^1 p^{\alpha+x-1}(1-p)^{\beta+n-x-1} dp \\ &= \frac{\binom{n}{x}}{B(\alpha, \beta)} B(\alpha + x, \beta + n - x) \end{aligned}$$

This yields the Beta-Binomial probability mass function:

$$P(X = x \mid n, \alpha, \beta) = \binom{n}{x} \frac{B(x+\alpha, n-x+\beta)}{B(\alpha, \beta)}, \quad x = 0, 1, \ldots, n \tag{2.1}$$

Figure 1 depicts the probability mass function for the Beta-Binomial distribution, demonstrating how it varies at various p-values.

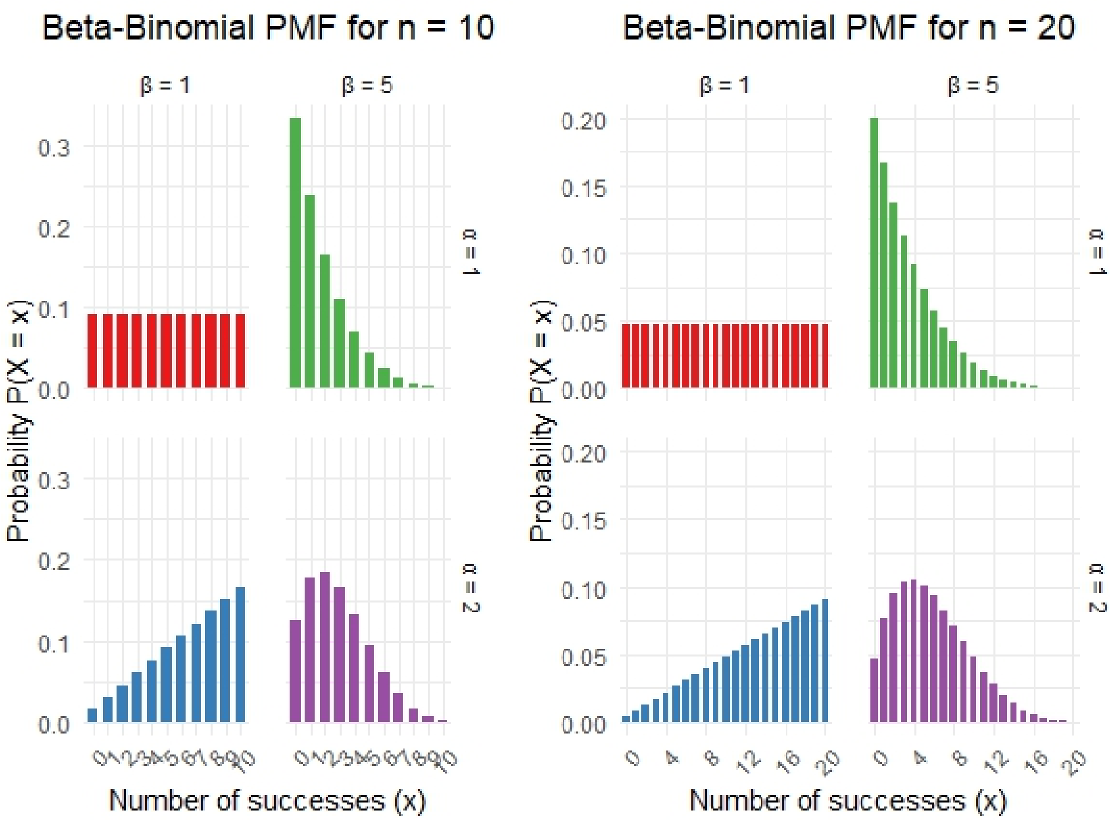

Figure 1: Density of Beta-Binomial distribution for different parameters

2.4. Bayesian Interpretation

The posterior distribution of p given $X = x$ is:

$$f(p \mid X = x) \propto p^{\alpha+x-1}(1-p)^{\beta+n-x-1}$$

which shows that:

$$p \mid X = x \sim \text{Beta}(\alpha + x, \beta + n - x)$$

This demonstrates the conjugate relationship between the Beta prior and Binomial likelihood.

2.5. Moments

The mean and variance of the Beta-Binomial distribution are:

$$E(X) = \mu = \frac{n\alpha}{\alpha + \beta} \tag{2.2}$$

$$\text{Var}(X) = \sigma^2 = \frac{n\alpha\beta(\alpha + \beta + n)}{(\alpha + \beta)^2(\alpha + \beta + 1)} \tag{2.3}$$

2.6. Properties

- **Over-dispersion**: The variance exceeds that of a Binomial distribution with the same mean
- **Conjugacy**: The Beta prior is conjugate to the Binomial likelihood
- **Flexibility**: Can model a wide range of dispersion patterns through parameters α and β

2.7. Random Number Generation

To generate random numbers from the Beta-Binomial distribution:

1. Generate $p \sim \text{Beta}(\alpha, \beta)$
2. Generate $X \sim \text{Binomial}(n, p)$

The resulting X follows the Beta-Binomial distribution with parameters n, α, and β.

2.8. Parameter Estimation

2.8.1. Method of Moments. Define the factorial moment ratio:

$$R_{r+1} = \frac{\mu'_{[r+1]}}{\mu'_{[r]}}$$

For the Beta-Binomial distribution:

$$R_{r+1} = \frac{(n-r)(r+\alpha)}{r+\alpha+\beta}$$

For $r = 0$ and $r = 1$:

$$R_1 = \frac{n\alpha}{\alpha+\beta}$$
$$R_2 = \frac{(n-1)(1+\alpha)}{1+\alpha+\beta}$$

Solving these equations yields the moment estimators:

$$\tilde{\alpha} = \frac{m'_1(m'_2 - nm'_1)}{(n-1)m'^2_1 - n(m'_2 - m'_1)} \tag{2.4}$$
$$\tilde{\beta} = \left(\frac{n}{m'_1} - 1\right)\tilde{\alpha} \tag{2.5}$$

where m'_1 and m'_2 are the first and second sample moments.

2.8.2. Maximum Likelihood Estimation. The likelihood function is:

$$L(\alpha, \beta) = \prod_{i=1}^{m} \frac{\binom{n}{x_i} B(x_i + \alpha, n - x_i + \beta)}{B(\alpha, \beta)}$$

The log-likelihood function is:

$$l(\alpha, \beta) = \sum_{i=1}^{m} \log \binom{n}{x_i} + m[\log B(\alpha + x_i, \beta + n - x_i) - \log B(\alpha, \beta)]$$

The maximum likelihood estimates, $\hat{\alpha}$ and $\hat{\beta}$ are obtained by solving:

$$\frac{\partial l}{\partial \alpha} = 0 \tag{2.6}$$

and

$$\frac{\partial l}{\partial \beta} = 0 \tag{2.7}$$

These equations require numerical optimization methods such as Newton-Raphson [23].

2.9. Applications in SPC and Beyond

The primary application of the Beta-Binomial distribution in SPC is for monitoring over-dispersed proportion data [7]. Traditional p-charts, based on the Binomial assumption, will have improperly set control limits in the presence of over-dispersion, leading to an inflated false-alarm rate. Control charts designed using the Beta-Binomial model are more robust and accurate for such processes.

Beyond SPC, the distribution has wide-ranging applications:

- **Biometrics and Epidemiology**: Modeling disease incidence in clusters with varying risk factors [14]
- **Reliability Engineering**: Modeling failure counts with varying failure probabilities [17]
- **Bayesian Analysis**: Canonical example of conjugate analysis
- **Market Research and Genetics**: Analyzing consumer behavior and allele frequencies

3. Control Charts for Beta-Binomial Processes

In statistical quality control, monitoring the number of nonconforming items in an inspection unit is a common task. While the inspection unit can be a single item, it often comprises a fixed-size group of products for practical reasons such as easier data collection and management. This section introduces control charts designed for scenarios where the number of nonconformities per inspection unit follows a Beta-Binomial distribution. This distribution is particularly useful when the probability of a nonconformity, p, is not constant but varies from unit to unit according to a Beta distribution.

3.1. Control Charts When the Parameters Are Known

Let the number of nonconformities, X, in an inspection unit be modeled by a Beta-Binomial distribution with parameters $\alpha > 0$, $\beta > 0$, and a fixed size parameter n (the number of trials per unit). Its probability mass function (pmf) is given by (2.1).

Given the mean μ and variance σ^2 of this distribution (see Eqs. (2.2) and (2.3)), a Shewhart-type control chart for individual observations (X chart) with three-sigma limits is constructed as follows:

$$\mathrm{UCL} = \mu + 3\sigma = \frac{n\alpha}{\alpha+\beta} + 3\sqrt{\frac{n\alpha\beta(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)}} \tag{3.1}$$

$$\mathrm{CL} = \mu = \frac{n\alpha}{\alpha+\beta} \tag{3.2}$$

$$\mathrm{LCL} = \max(0, \mu - 3\sigma) = \max\left(0, \frac{n\alpha}{\alpha+\beta} - 3\sqrt{\frac{n\alpha\beta(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)}}\right) \tag{3.3}$$

The Lower Control Limit (LCL) is set to zero if the calculation yields a negative value, as a negative count of nonconformities is not possible.

To monitor the process using subgroups, let each subgroup i consist of k inspection units. Let $\overline{X}_i = (X_{i1} + \ldots + X_{ik})/k$ be the average number of nonconformities per inspection unit in the ith subgroup. The center line and control limits for this $\overline{X}$ chart are derived from the properties of the individual (X) chart. Since the variance of the sample mean is $\sigma^2_{\overline{X}} = \sigma^2/k$, the limits are:

$$\mathrm{UCL} = \mu + 3\frac{\sigma}{\sqrt{k}} = \frac{n\alpha}{\alpha+\beta} + 3\sqrt{\frac{n\alpha\beta(\alpha+\beta+n)}{k(\alpha+\beta)^2(\alpha+\beta+1)}} \tag{3.4}$$

$$\mathrm{CL} = \frac{n\alpha}{\alpha+\beta} \tag{3.5}$$

$$\mathrm{LCL} = \max\left(0, \mu - 3\frac{\sigma}{\sqrt{k}}\right) = \max\left(0, \frac{n\alpha}{\alpha+\beta} - 3\sqrt{\frac{n\alpha\beta(\alpha+\beta+n)}{k(\alpha+\beta)^2(\alpha+\beta+1)}}\right) \tag{3.6}$$

3.2. Control Charts When the Parameters Are Unknown

In practical applications, the parameters α and β are often unknown and must be estimated from historical data.

3.2.1. Control Chart for Individual Observations. Let $x_1, x_2, \ldots, x_m$ be a reference sample of m observations from the Beta-Binomial distribution with known inspection unit size n. The parameters α and β can be estimated using the Method of Moments (MoM). By equating the sample mean ($\bar{x}$) and sample variance (s^2) to their theoretical counterparts and solving, the estimators are:

$$\hat{\alpha} = \bar{x}\left(\frac{n(\bar{x}/n)(1-\bar{x}/n) - s^2}{s^2 - (n\bar{x}/(m+n))(1-\bar{x}/n)(1-1/n)}\right) \tag{3.7}$$

(An approximate or iterative solution is often used)

$$\hat{\beta} = (n-\bar{x})\left(\frac{n(\bar{x}/n)(1-\bar{x}/n) - s^2}{s^2 - (n\bar{x}/(m+n))(1-\bar{x}/n)(1-1/n)}\right) \tag{3.8}$$

Alternatively, Maximum Likelihood Estimation (MLE) can be employed for more efficient estimates, typically requiring numerical optimization methods. Once the estimates $\hat{\alpha}$ and $\hat{\beta}$ are obtained, they are substituted into the formulas for the control limits in Section 3.1.

3.2.2. Control Chart for Subgroups. Suppose we have m subgroups, each of size k. Let x_{ij} represent the jth observation in the ith subgroup, and $\overline{x}_i$ and S_i^2 be the mean and variance of the ith subgroup sample. The grand mean is $\overline{\overline{x}} = \frac{1}{m}\sum_{i=1}^{m}\overline{x}_i$, and the total sample variance is $\hat{S}^2 = (n-1)\sum_{i=1}^{m} S_i^2/(m*(n-1))$.

To implement the $\overline{X}$ chart with unknown parameters, we use the Equations (3.4), (3.5), and (3.6), such that instead of α and β we use $\hat{\alpha}$, and $\hat{\beta}$ of Equations (3.7), and (3.8) respectively. Also in computing the $\hat{\alpha}$, and $\hat{\beta}$, instead of $\bar{x}$ and s^2 we use $\overline{\overline{x}}$ and $\hat{S}^2$ respectively.

3.3. Performance Evaluation and Implementation Note

The primary metric for evaluating the performance of a control chart is the Average Run Length (ARL), which measures the average number of samples collected before a chart signals a change. Calculating the exact ARL for the Beta-Binomial control chart requires knowledge of the distribution of the charting statistic (either X or $\overline{X}$) in both in-control and out-of-control states.

Deriving the exact distribution of the subgroup average $\overline{X}$ for the Beta-Binomial distribution is complex. Therefore, simulation-based methods, such as the bootstrap technique, are highly recommended for estimating the ARL and determining the practical implementation of these control charts. This approach involves resampling the historical data to establish empirical control limits and evaluate the chart's sensitivity to process shifts.

3.4. Bootstrap Control Charts for Beta-Binomial Processes

To construct a bootstrap control chart, we only use the sample data to estimate the sampling distribution of the parameter estimator, and then, to obtain appropriate control limits. Thus, only the usual assumptions of Phase II of SPC are required: stable process and independent and identically distributed subgroup observations. The following Algorithm, similar to the ones proposed in [26] and [19,18], can be used to implement bootstrap control charts for subgroup samples of size n, to monitor the process mean value and the process standard deviation of a Beta-Binomial distribution, respectively. This algorithm can be easily modified in order to implement bootstrap control charts for other parameters of interest.

In order to get information about the robustness of the bootstrap control limits, we must repeat Steps 1–6 of Algorithm 1 a large number of times, say $r = 100$, and then compute the average of the obtained control limits, UCL and LCL, and their associated variances.

Algorithm 1 Bootstrap Control Chart Construction

Phase I: Estimation and computation of the control limits

1. From in-control and stable process, observe k, say 25 or 30, random samples of size n, assuming the observations are independent and come from a Beta-Binomial distribution, $BB(n, \alpha, \beta)$.

2. Compute the MLE estimates $\hat{\alpha}$ and $\hat{\beta}$ of Equations (2.6) and (2.7), using the pooled sample of size $k \times n$.

3. Generate a parametric bootstrap sample of size n, $(x_1^*, \ldots, x_n^*)$, from a Beta-Binomial distribution using the MLEs obtained in Step 2 as the distribution parameters.

4. For the bootstrap sample of size n generated at step 3, compute the MLE estimates of $\hat{\alpha^*}$ and $\hat{\beta}^*$. Then select the step associated to the chart you want to implement:

   - Two-sided bootstrap M-chart to monitor the process mean value $BBM = \frac{n\alpha}{\alpha+\beta}$: from the bootstrap subgroup sample obtained in Step 3, compute the estimated sample mean, $BB\hat{M}^* = \frac{n\hat{\alpha^*}}{\hat{\alpha^*}+\hat{\beta^*}}$
   - Upper one-sided bootstrap S-chart to monitor the process standard deviation $BBSd = \sqrt{\frac{n\alpha\beta(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)}}$: from the bootstrap subgroup sample obtained in Step 3, compute the estimated sample standard deviation, $BB\hat{S}d^* = \sqrt{\frac{n\hat{\alpha^*}\hat{\beta^*}(\hat{\alpha^*}+\hat{\beta^*}+n)}{(\hat{\alpha^*}+\hat{\beta^*})^2(\hat{\alpha^*}+\hat{\beta^*}+1)}}$

5. Repeat Steps 3–4, a large number of times, say $B = 10000$ times, obtaining B bootstrap estimates of the parameter of interest (the process mean value or the standard deviation).

6. Let γ be the desired false alarm rate (FAR) of the chart. Using the B bootstrap estimates obtained in Step 5:

   - Find the $100(\gamma/2)^{\text{th}}$ and $100(1-\gamma/2)^{\text{th}}$ quantiles of the distribution of $BB\hat{M}^*$, i.e., the lower control limit LCL and the upper control limit UCL for the bootstrap M-chart of FAR $= \gamma$, respectively.
   - Find the $100(1-\gamma)^{\text{th}}$ quantile of the distribution of $BB\hat{S}d^*$, i.e., the upper control limit UCL for the bootstrap S-chart of FAR $= \gamma$.

Phase II: Process monitoring

7. Take subgroup samples of size n from the process at regular time intervals. For each subgroup, estimate α and β, and compute $B\hat{B}M$ and $B\hat{B}Sd$.

8. **Decision:**

   - If $B\hat{B}M$ falls between LCL and UCL of the M-chart, the process is assumed to be in-control (targeting the nominal mean value); otherwise, if the estimate falls below the LCL or above the UCL, the chart signals that the process may be out-of-control.
   - If $B\hat{B}Sd$ falls below the UCL of the S-chart, the process is assumed to be in-control (targeting the nominal standard deviation); otherwise, the chart signals that the process may be out-of-control.

4. Performance Evaluation of Bootstrap Control Charts for Beta-Binomial Processes

This section presents a comprehensive simulation study evaluating the performance of bootstrap-based M and S control charts for monitoring Beta-Binomial processes. The investigation focuses on two key aspects: (1) the accuracy of bootstrap control limit estimation, and (2) the detection capability measured through average run length (ARL) performance under both in-control and out-of-control conditions.

4.1. Simulation Design and Methodology

The simulation framework was designed to assess chart performance across diverse scenarios encompassing various sample sizes ($n = 5, 6, 10, 30, 50$) and different parameter configurations of the Beta-Binomial distribution (α, β combinations). The false alarm rate was fixed at $\gamma = 0.0027$ throughout all experiments, corresponding to the traditional 3-sigma control limits in normal-based charts.

For control limit estimation, $k = 30$ subgroups of size n were generated from the specified Beta-Binomial distribution following the procedure outlined in Algorithm 1. Maximum likelihood estimates of α and β were obtained from the pooled data, and $B = 10,000$ bootstrap samples were generated to construct the empirical distribution of the charting statistics. This entire process was repeated $r = 100$ times using Monte Carlo simulation to obtain stable estimates of the control limits.

ARL performance was evaluated by introducing deliberate parameter shifts from baseline values (α_0, β_0) to altered values (α_1, β_1). For each combination, the mean ARL (MARL) and standard deviation of ARL (SDARL) were computed based on 100 independent replications. All simulations were implemented using R statistical software.

4.2. Bootstrap Control Limit Estimation

Table 1 presents the estimated control limits for both M and S charts across different parameter configurations. The results demonstrate several important patterns regarding the behavior of bootstrap control limits.

Table 1: Bootstrap M and S charts' Control limits for BBD with different values of n, α, and β

n	α	β	LCL (M-chart)	UCL (M-chart)	UCL (S-chart)	True Mean	True SD
5	3	2	1.200 00	4.800 00	2.408 32	3.000 00	1.414 21
5	3	4	0.400 00	3.800 00	2.302 17	2.142 86	1.355 26
5	3	6	0.200 00	3.200 00	2.167 95	1.666 67	1.247 22
5	5	2	1.800 00	4.800 00	2.190 89	3.571 43	1.237 18
5	5	4	1.000 00	4.400 00	2.302 17	2.777 78	1.314 68
5	5	6	0.600 00	3.800 00	2.280 35	2.272 73	1.285 65
5	7	2	2.200 00	5.000 00	2.190 89	3.888 89	1.099 94
5	7	4	1.400 00	4.600 00	2.302 17	3.181 82	1.242 05
5	7	6	1.200 00	4.400 00	2.280 35	2.692 31	1.263 98
6	3	2	2.000 00	5.500 00	2.658 32	3.600 00	1.624 81
6	3	4	0.833 33	4.666 67	2.658 32	2.571 43	1.545 24
6	3	6	0.666 67	4.000 00	2.338 09	2.000 00	1.414 21
6	5	2	2.500 00	5.666 67	2.449 49	4.285 71	1.410 60
6	5	4	1.500 00	5.083 56	2.483 28	3.333 33	1.490 71
6	5	6	1.000 00	4.666 67	2.562 55	2.727 27	1.451 70
6	7	2	3.000 00	5.833 33	2.345 21	4.666 67	1.247 22
6	7	4	2.000 00	5.333 33	2.366 43	3.818 18	1.402 48
6	7	6	1.666 67	5.000 00	2.422 12	3.230 77	1.422 56
10	3	2	3.500 00	8.100 00	3.683 30	6.000 00	2.449 49
10	3	4	2.000 00	6.300 00	3.565 26	4.285 71	2.281 25
10	3	6	1.400 00	5.450 14	3.496 03	3.333 33	2.054 80
10	5	2	5.249 87	9.000 00	3.335 00	7.142 86	2.082 48
10	5	4	3.300 00	7.300 00	3.306 56	5.555 56	2.165 95
10	5	6	2.600 00	6.400 00	3.212 85	4.545 45	2.082 99
10	7	2	5.749 87	9.300 00	3.224 90	7.777 78	1.812 17
10	7	4	4.500 00	8.050 14	2.981 42	6.363 64	2.012 36
10	7	6	3.700 00	7.100 00	2.951 48	5.384 62	2.020 60
30	3	2	14.616 62	21.483 38	8.434 26	18.000 00	6.480 74
30	3	4	9.833 24	16.316 71	7.819 01	12.857 14	5.829 20
30	3	6	7.216 62	12.800 00	6.928 21	10.000 00	5.099 02
30	5	2	18.466 67	24.116 71	7.101 99	21.428 57	5.321 31
30	5	4	13.800 00	19.650 05	7.098 74	16.666 67	5.374 84
30	5	6	10.600 00	16.133 33	6.526 35	13.636 36	5.041 15
30	7	2	20.866 67	25.666 67	6.319 13	23.333 33	4.496 91
30	7	4	16.533 33	21.516 71	6.279 05	19.090 91	4.870 22
30	7	6	13.449 96	18.466 67	6.213 22	16.153 85	4.785 33
50	3	2	25.269 97	34.330 03	12.933 68	30.000 00	10.488 09
50	3	4	17.149 87	25.260 05	11.510 87	21.428 57	9.340 50
50	3	6	13.469 97	20.530 03	10.523 12	16.666 67	8.096 64
50	5	2	31.860 00	39.160 05	10.791 63	35.714 29	8.526 67
50	5	4	24.069 97	31.410 03	10.654 68	27.777 78	8.534 61
50	5	6	19.720 00	26.370 03	9.964 15	22.727 27	7.938 30
50	7	2	35.679 95	41.770 03	9.222 88	38.888 89	7.140 56
50	7	4	28.720 00	35.190 03	9.655 81	31.818 18	7.669 12
50	7	6	23.789 92	30.000 00	9.227 87	26.923 08	7.477 78

Note: Control limits calculated with $\gamma = 0.0027$ (probability of false alarm).

4.2.1. Impact of Sample Size. The control limit range (UCL - LCL) for the M-chart expands substantially with increasing sample size. For instance, with $\alpha = 3$, $\beta = 2$, the range increases from 3.6 ($n = 5$) to 9.06 ($n = 50$). Similarly, the S-chart UCL values scale approximately with $\sqrt{n}$, consistent with the variance properties of the Beta-Binomial distribution. Larger sample sizes provide enhanced discrimination between in-control and out-of-control states, as evidenced by the wider control bands.

4.2.2. Effect of Shape Parameters. Increasing α while maintaining β constant elevates both LCL and UCL for the M-chart while generally reducing the control limit range. For example, with $n = 10$, $\beta = 2$, the range decreases from 4.6 ($\alpha = 3$) to 3.55 ($\alpha = 7$). Concurrently, the S-chart UCL values diminish, indicating reduced process dispersion.

Conversely, increasing β while holding α constant lowers both M-chart control limits and substantially reduces the S-chart UCL values. The β parameter exerts a more pronounced effect on variance reduction than the α parameter.

4.2.3. Joint Parameter Effects. The ratio $\alpha/(\alpha + \beta)$ primarily determines the centerline position, as reflected in the true mean values. The most constrained control limits occur with high α and high β combinations. For fixed α/β ratios, the control limit range expands with n but at a diminishing rate. The bootstrap methodology effectively accommodates the asymmetric nature of the Beta-Binomial distribution, particularly for small sample sizes where normal approximations prove inadequate.

4.3. Average Run Length Performance

Table 2 summarizes the MARL and SDARL results for various parameter shift scenarios. The findings reveal crucial insights into the detection capability of the proposed control charts.

Table 2: Mean and Standard Deviation of ARL (100 replications) for BBD Parameter Shifts

α_0	β_0	α_1	β_1	$n = 5$		$n = 6$		$n = 10$		$n = 30$		$n = 50$	
				MARL	SDARL	MARL	SDARL	MARL	SDARL	MARL	SDARL	MARL	SDARL
3	2	3	2	356.76	317.27	420.72	405.17	455.53	428.97	333.87	307.74	441.56	446.64
3	2	3	4	18.99	19.06	7.65	6.77	5.21	4.85	1.06	0.24	1.00	0.00
3	2	3	6	4.34	3.80	1.73	1.04	1.72	1.30	1.00	0.00	1.00	0.00
3	2	5	2	121.14	132.53	92.39	91.09	14.02	13.54	2.15	1.53	1.16	0.44
3	2	5	4	174.68	180.53	86.47	86.30	76.86	74.30	68.38	68.54	55.91	56.97
3	2	5	6	28.51	24.58	8.81	7.79	21.81	20.35	1.24	0.51	1.01	0.10
3	2	7	2	189.34	174.56	24.89	23.04	3.20	2.61	1.03	0.17	1.00	0.00
3	2	7	4	207.06	161.06	101.99	101.45	94.54	89.52	86.00	80.09	83.39	76.16
3	2	7	6	248.44	223.78	59.43	67.25	77.01	72.52	36.40	29.59	22.74	23.00
3	4	3	2	13.17	12.89	21.51	16.00	3.01	2.12	1.11	0.37	1.00	0.00
3	4	3	4	249.97	246.60	459.51	495.59	340.20	332.34	457.29	544.08	285.31	258.58
3	4	3	6	89.81	86.71	75.68	76.13	72.54	77.95	3.02	2.38	1.59	0.93
3	4	5	2	3.20	2.55	4.48	4.15	1.07	0.26	1.00	0.00	1.00	0.00
3	4	5	4	110.45	109.99	31.01	29.43	7.23	8.86	1.34	0.67	1.00	0.00
3	4	5	6	243.81	206.39	151.72	133.48	42.76	40.81	43.87	46.12	49.84	46.13
3	4	7	2	2.02	1.47	1.87	1.50	1.00	0.00	1.00	0.00	1.00	0.00
3	4	7	4	20.98	22.41	12.38	14.42	1.89	1.37	1.01	0.10	1.00	0.00
3	4	7	6	54.80	45.75	219.57	215.61	11.92	11.92	2.19	1.50	1.08	0.27
3	6	3	2	2.61	2.17	3.79	3.13	1.25	0.54	1.00	0.00	1.00	0.00
3	6	3	4	21.48	25.73	74.33	68.96	21.03	21.34	1.61	1.02	1.20	0.47
3	6	3	6	210.60	188.80	234.50	251.14	737.66	699.52	329.64	350.02	550.46	603.83
3	6	5	2	1.29	0.57	1.66	1.16	1.02	0.14	1.00	0.00	1.00	0.00
3	6	5	4	4.64	4.72	7.12	6.69	1.63	1.02	1.00	0.00	1.00	0.00
3	6	5	6	16.67	15.19	58.25	59.67	12.35	13.03	1.15	0.46	1.03	0.17
3	6	7	2	1.10	0.30	1.20	0.43	1.00	0.00	1.00	0.00	1.00	0.00
3	6	7	4	2.01	1.21	2.42	1.84	1.12	0.33	1.00	0.00	1.00	0.00
3	6	7	6	5.09	4.56	13.18	13.43	2.19	1.59	1.00	0.00	1.00	0.00
5	2	3	2	34.49	34.93	17.79	15.69	5.87	5.58	1.72	1.05	1.13	0.34
5	2	3	4	4.05	3.24	2.21	1.30	1.06	0.24	1.00	0.00	1.00	0.00
5	2	3	6	1.58	0.99	1.29	0.62	1.00	0.00	1.00	0.00	1.00	0.00
5	2	5	2	661.97	676.45	271.12	327.54	259.09	216.30	286.21	295.99	471.12	462.56
5	2	5	4	24.85	23.98	11.39	11.13	2.29	1.76	1.02	0.14	1.00	0.00
5	2	5	6	5.94	6.08	2.91	2.42	1.18	0.44	1.00	0.00	1.00	0.00
5	2	7	2	180.02	207.83	199.11	172.15	27.41	24.82	5.48	5.30	2.56	2.40
5	2	7	4	143.09	140.24	58.15	65.93	19.75	21.39	5.24	5.31	2.10	1.44
5	2	7	6	21.26	20.15	8.76	9.02	2.43	1.64	1.00	0.00	1.00	0.00
5	4	3	2	175.78	165.97	100.78	81.19	19.40	15.80	13.13	13.35	4.93	4.30
5	4	3	4	43.55	44.64	25.90	25.98	12.66	12.86	1.21	0.48	1.02	0.14
5	4	3	6	9.31	7.59	5.19	4.00	2.30	1.71	1.00	0.00	1.00	0.00
5	4	5	2	20.22	17.10	17.83	16.59	2.33	2.28	1.02	0.14	1.00	0.00
5	4	5	4	357.63	289.12	686.69	661.26	209.32	221.59	355.09	344.66	471.35	476.21
5	4	5	6	114.03	100.98	63.02	69.08	44.70	42.89	1.90	1.51	1.19	0.44
5	4	7	2	5.50	5.40	5.76	4.97	1.25	0.63	1.00	0.00	1.00	0.00
5	4	7	4	170.55	180.97	85.57	81.53	13.28	11.47	3.46	2.71	1.56	1.04
5	4	7	6	217.77	260.59	179.54	163.90	83.09	82.85	85.52	81.36	36.10	37.18
5	6	3	2	21.30	20.06	12.20	11.70	2.91	2.16	1.07	0.26	1.01	0.10
5	6	3	4	198.16	187.94	124.50	147.13	84.95	72.85	55.93	54.55	14.86	14.77
5	6	3	6	37.00	38.00	24.47	20.66	8.18	8.35	1.51	0.97	1.00	0.00
5	6	5	2	4.14	3.16	3.90	3.35	1.13	0.34	1.00	0.00	1.00	0.00
5	6	5	4	104.53	116.57	32.40	34.88	9.13	9.86	1.25	0.46	1.05	0.26
5	6	5	6	211.50	201.55	583.05	590.16	305.67	256.20	205.91	193.69	312.88	319.17
5	6	7	2	2.14	1.40	2.21	1.40	1.00	0.00	1.00	0.00	1.00	0.00
5	6	7	4	19.73	17.28	9.11	8.27	2.02	1.60	1.00	0.00	1.00	0.00
5	6	7	6	163.46	113.89	71.20	65.84	14.12	12.46	1.66	1.00	1.37	0.63
7	2	3	2	9.33	9.17	5.89	5.12	2.93	2.54	1.00	0.00	1.00	0.00
7	2	3	4	1.88	1.24	1.37	0.73	1.03	0.17	1.00	0.00	1.00	0.00
7	2	3	6	1.34	0.61	1.02	0.14	1.00	0.00	1.00	0.00	1.00	0.00
7	2	5	2	89.61	85.59	57.98	51.26	42.34	44.96	3.84	3.27	2.16	1.56

Continued on next page

Table 2 (continued)

α_0	β_0	α_1	β_1	$n = 5$		$n = 6$		$n = 10$		$n = 30$		$n = 50$	
				MARL	SDARL	MARL	SDARL	MARL	SDARL	MARL	SDARL	MARL	SDARL
7	2	5	4	6.29	5.19	3.67	3.21	1.66	1.04	1.00	0.00	1.00	0.00
7	2	5	6	2.17	1.51	1.37	0.69	1.02	0.14	1.00	0.00	1.00	0.00
7	2	7	2	582.11	636.35	445.00	372.68	802.67	771.53	301.90	297.99	406.99	412.77
7	2	7	4	21.05	20.96	13.06	11.75	6.03	5.94	1.02	0.14	1.00	0.00
7	2	7	6	5.72	5.05	3.14	2.51	1.40	0.75	1.00	0.00	1.00	0.00
7	4	3	2	152.70	151.91	96.34	92.05	39.85	32.56	9.36	8.74	6.25	5.76
7	4	3	4	9.22	8.45	5.19	5.24	1.70	1.18	1.00	0.00	1.00	0.00
7	4	3	6	3.09	2.59	2.01	1.49	1.03	0.17	1.00	0.00	1.00	0.00
7	4	5	2	98.23	97.00	61.30	65.06	12.75	11.42	2.09	1.76	1.49	0.90
7	4	5	4	136.22	147.61	63.76	60.19	17.44	17.32	2.66	1.72	1.22	0.48
7	4	5	6	15.92	14.77	8.09	6.30	2.31	1.54	1.00	0.00	1.00	0.00
7	4	7	2	33.04	32.97	15.90	14.03	3.34	3.25	1.00	0.00	1.00	0.00
7	4	7	4	466.11	519.39	420.50	415.70	260.73	304.07	221.90	208.46	423.41	386.26
7	4	7	6	102.45	87.94	48.93	50.28	10.36	9.09	1.54	0.89	1.10	0.33
7	6	3	2	83.17	81.25	80.72	78.77	13.26	12.28	2.77	2.08	1.93	1.20
7	6	3	4	20.70	18.30	13.36	14.00	4.56	4.66	1.47	0.69	1.05	0.22
7	6	3	6	5.01	3.79	2.95	2.37	1.38	0.66	1.00	0.00	1.00	0.00
7	6	5	2	14.99	16.31	12.33	9.65	2.00	1.81	1.01	0.10	1.00	0.00
7	6	5	4	206.73	191.39	210.47	192.64	69.49	72.00	27.96	24.28	25.96	22.92
7	6	5	6	32.22	30.81	28.80	30.18	12.27	10.10	2.41	1.56	1.22	0.50
7	6	7	2	5.85	5.07	4.14	3.76	1.08	0.27	1.00	0.00	1.00	0.00
7	6	7	4	128.81	113.64	86.49	92.15	7.93	7.94	1.24	0.57	1.03	0.17
7	6	7	6	249.20	236.55	269.02	336.45	239.56	235.63	284.28	280.60	307.75	339.76

4.3.1. In-Control Performance. When no parameter shift occurs ($(\alpha_0, \beta_0) = (\alpha_1, \beta_1)$), the MARL values approximate the theoretical expectation of 370 (corresponding to $\gamma = 0.0027$), with SDARL values roughly equivalent to MARL. This indicates proper calibration of the bootstrap-based control limits across different sample sizes, maintaining the desired in-control performance.

4.3.2. Out-of-Control Detection Capability. Under parameter shifts, MARL decreases substantially as sample size increases. For small n ($n = 5$ or 6), MARL remains elevated, indicating delayed shift detection. In contrast, for larger n ($n = 30$ or 50), MARL often drops to near 1, demonstrating rapid detection of out-of-control conditions.

The detection sensitivity varies with the nature of the parameter shift. Changes in β are generally detected more rapidly than equivalent changes in α for the same sample size. Some joint shifts in α and β (particularly those that preserve the mean while altering variance) exhibit slower detection, even with large sample sizes.

4.3.3. Detection Consistency. SDARL values decrease with increasing sample size, reflecting more consistent detection times. High SDARL values for small n indicate variable detection performance, while low SDARL values for large n demonstrate consistent and reliable shift detection.

4.4. Conclusions from Simulation Study

The simulation results lead to several important conclusions:

1. **Sample size significantly impacts detection performance**: Larger sample sizes ($n \geq 30$) provide substantially better shift detection capability across all parameter configurations.

2. **Bootstrap limits maintain nominal false alarm rates**: The proposed methodology consistently achieves the target in-control ARL of approximately 370, validating the bootstrap approach for control limit estimation.

3. **Detection sensitivity depends on shift direction**: Changes affecting process dispersion (β shifts) are detected more rapidly than changes primarily affecting location (α shifts).

4. **Joint parameter changes may challenge detection**: Simultaneous changes in α and β that offset each other's effects on the mean may require larger sample sizes for prompt detection.

5. **Consistency improves with sample size**: Larger samples not only enhance detection speed but also improve detection consistency, as evidenced by reduced SDARL values.

These findings underscore the importance of selecting appropriate sample sizes and understanding the nature of potential process changes when implementing control charts for Beta-Binomial processes. The bootstrap methodology provides a robust framework for constructing effective control charts that accommodate the distributional characteristics of Beta-Binomial data.

5. Illustrative Examples

This section demonstrates the practical application of bootstrap control charts for Beta-Binomial processes through both simulated data and real-world case studies. The examples showcase the methodology's effectiveness in detecting process changes and monitoring quality characteristics.

5.1. Simulation Study

A comprehensive simulation was conducted to evaluate the performance of the proposed bootstrap control charts under controlled conditions with known parameter shifts.

5.1.1. Simulation Design. The simulation followed a two-phase approach:

- **Phase I (In-control process):** Generated $m = 25$ subgroups of size $n = 5$ from a Beta-Binomial distribution $BB(\alpha = 3, \beta = 6)$ to establish baseline performance.

- **Phase II (Out-of-control process):** Generated 10 additional subgroups of size $n = 5$ from a shifted Beta-Binomial distribution $BB(\alpha = 5, \beta = 2)$ to simulate a process change.

- The statistics $BB\hat{M}$ (mean estimator) and $BB\hat{S}_d$ (dispersion estimator) were computed for all subgroups using Algorithm 2.

5.1.2. Simulation Results. Table 3 presents the calculated values for $BB\hat{M}$ and $BB\hat{S}_d$ across all subgroups.

Table 3: Randomly generated subgruops of BBD with calculated values of $BB\hat{M}$ and $BB\hat{S}_d$ for each subgroup

Sample ID	Sample Values	$BB\hat{M}$	$BB\hat{S}_d$
1	2,1,2,0,0	0.99998	0.89585
2	0,4,0,3,1	1.53023	1.59974
3	2,1,2,1,2	1.60090	1.04464
4	1,0,2,2,5	2.10037	1.70112
5	2,3,4,0,1	1.98312	1.41768
6	1,5,2,3,1	2.45615	1.49241
7	4,0,1,3,4	2.32872	1.59606
8	2,4,0,1,4	2.16318	1.57209
9	3,0,3,0,1	1.35763	1.38635
10	2,1,0,1,3	1.39998	1.02092
11	3,2,0,1,3	1.79673	1.18082
12	3,1,3,2,1	2.00030	1.09679
13	1,2,3,3,2	2.20043	1.11122
14	1,3,1,1,4	2.00405	1.24721
15	1,2,1,0,2	1.20056	0.95659
16	2,1,1,4,0	1.60879	1.33007
17	0,3,1,2,1	1.39998	1.02092
18	0,0,1,1,0	0.40074	0.60825
19	3,2,2,3,1	2.20043	1.11122
20	1,1,3,2,2	1.80056	1.07476
21	0,2,2,3,1	1.60011	1.04451
22	3,1,0,2,0	1.18693	1.19936
23	1,4,1,3,0	1.79385	1.44488
24	2,1,1,0,1	1.00080	0.89612
25	0,0,1,3,0	0.77908	1.17478
26	5,1,4,5,5	4.03245	1.47970
27	1,3,0,4,4	2.32872	1.59606
28	5,4,4,4,5	4.39895	0.72846
29	2,3,5,4,3	3.39989	1.04451
30	4,4,3,4,5	3.99920	0.89612
31	4,4,5,1,4	3.57679	1.29381
32	5,2,1,5,4	3.46977	1.59974
33	5,2,4,5,2	3.64237	1.38635
34	3,4,4,3,5	3.79944	0.95659
35	4,5,4,5,5	4.59926	0.60825

Figure 2 shows the bootstrap M-chart for simulated data, demonstrating effective monitoring of process mean with established control limits. Figure 3 presents the bootstrap S-chart for simulated data, illustrating the monitoring of process variability with its upper control limit.

5.1.3. Key Observations. The simulation results reveal several important patterns:

- **Phase I stability:** Samples 1–25, generated from $BB(3,6)$, show consistent behavior in both $BB\hat{M}$ and $BB\hat{S}_d$ values, indicating process stability.
- **Phase II shift detection:** Samples 26–35, generated from $BB(5,2)$, exhibit a pronounced increase in $BB\hat{M}$ values, clearly signaling the parameter shift.
- **Dispersion changes:** The $BB\hat{S}_d$ values show noticeable variations between phases, reflecting changes in process variability.

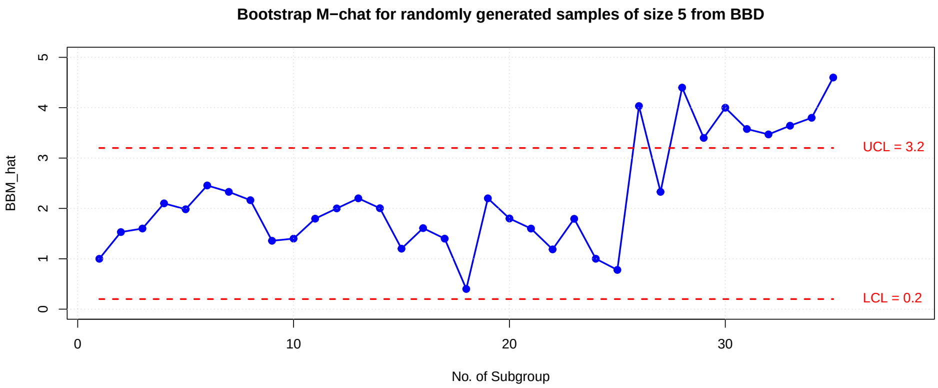

Figure 2: Bootstrap M-chart for simulated data from Beta-Binomial distribution with sample size 5.

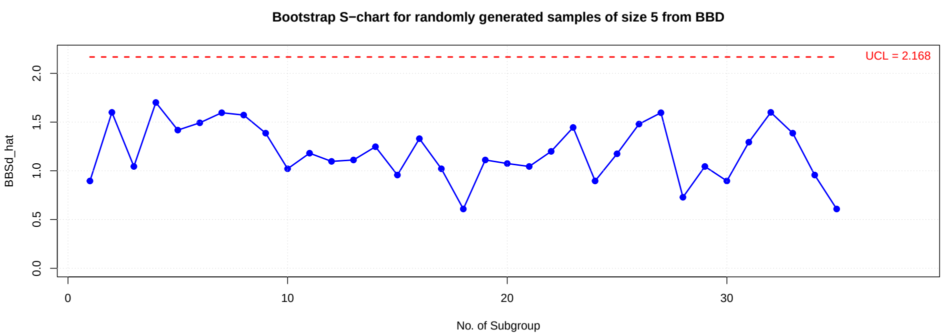

Figure 3: Bootstrap S-chart for simulated data from Beta-Binomial distribution with sample size 5.

- **Chart sensitivity:** The bootstrap control charts effectively detected the introduced shift, demonstrating their sensitivity to process changes.

These findings confirm that the bootstrap methodology provides an effective framework for monitoring Beta-Binomial processes using both location (M-chart) and dispersion (S-chart) control charts.

5.2. Real-World Application: Transformer Manufacturing Quality

To demonstrate practical utility, we applied the methodology to quality control data from a transformer manufacturing company in Diyala, Iraq.

5.2.1. Data Description. The dataset comprises monthly defective transformer counts from 2017 to 2021, totaling 60 monthly observations. The average production volume was $n = 274$ transformers per month, with substantial variation in monthly production levels. Table 4 summarizes the annual manufacturing and defect data.

Table 4: Monthly Transformer Production and Defect Counts (2017–2021)

Year	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**	**Total**
2017													
Manufactured	432	87	36	32	161	101	256	220	223	362	294	352	2524
Defective	1	0	1	0	3	4	0	2	0	0	0	0	11
2018													
Manufactured	280	182	259	235	175	210	277	260	287	260	263	389	2902
Defective	0	0	0	0	0	0	0	0	0	0	0	0	0
2019													
Manufactured	375	367	420	479	468	338	281	505	457	498	497	457	5142
Defective	11	1	0	0	1	1	1	0	2	0	0	0	17
2020													
Manufactured	391	445	79	20	66	50	146	281	175	160	143	208	2164
Defective	0	0	0	0	0	0	0	0	0	0	0	0	0
2021													
Manufactured	180	127	251	220	262	530	272	334	308	194	370	454	3502
Defective	0	0	2	2	0	1	7	9	1	1	4	24	51

5.2.2. Parameter Estimation and Model Fitting. Maximum likelihood estimation was used to fit the Beta-Binomial distribution to the defect count data. The parameter estimates are presented in Table 5.

Table 5: Beta-Binomial Parameter Estimates for Transformer Defect Data

Parameter	Symbol	Estimate
Alpha	α	0.1744
Beta	β	7.5810
Defect probability	$p = \alpha/(\alpha+\beta)$	0.0225

5.2.3. Goodness-of-Fit Assessment. The Beta-Binomial distribution's fit was evaluated using multiple statistical measures, as shown in Table 6.

Table 6: Goodness-of-Fit Statistics for Beta-Binomial Distribution

Statistic	Value
Log-Likelihood	-80.30
AIC	164.60
BIC	168.79
Chi-square statistic	1.48
Degrees of freedom	2
p-value	0.477

Figure 4 displays the goodness-of-fit assessment, confirming the Beta-Binomial distribution adequately models the transformer defect data. The plot compares observed versus expected defect frequencies. Figure 5 shows the Q-Q plot, providing additional visual confirmation of the Beta-Binomial distribution's fit to the observed data.The alignment of points with the reference line indicates a good distributional fit.

5.2.4. Model Comparison. Four probability distributions were compared for their ability to model the defect data. Table 7 presents the comparison results, while Table 8 shows the Akaike weights indicating each model's relative support.

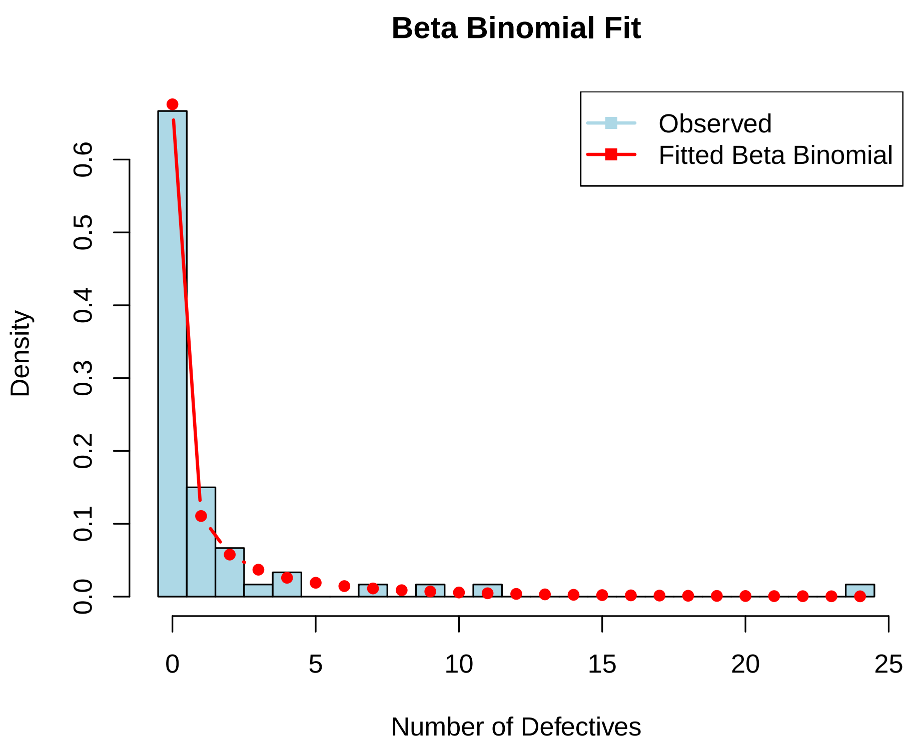

Figure 4: Goodness-of-fit assessment showing the Beta-Binomial distribution fitted to the transformer defect data.

Table 7: Comparison of Distribution Fits for Defect Data

Distribution	Log-Likelihood	AIC	BIC	Parameters
Negative Binomial	-80.07	164.15	168.33	2
Beta Binomial	-80.30	164.60	168.79	2
Poisson	-161.80	325.60	327.70	1
Binomial	-163.13	328.26	330.36	1

Table 8: AIC Model Weights for Distribution Comparison

Distribution	AIC Weight
Negative Binomial	0.5097
Beta Binomial	0.4903
Poisson	0.0000
Binomial	0.0000

5.2.5. Interpretation and Practical Implications. The analysis yields several important insights:

- The estimated defect probability of 2.25% indicates generally high manufacturing quality.
- The substantial difference between α and β parameters (0.174 vs. 7.581) suggests significant over-dispersion in the defect data, making standard Binomial and Poisson models inappropriate.

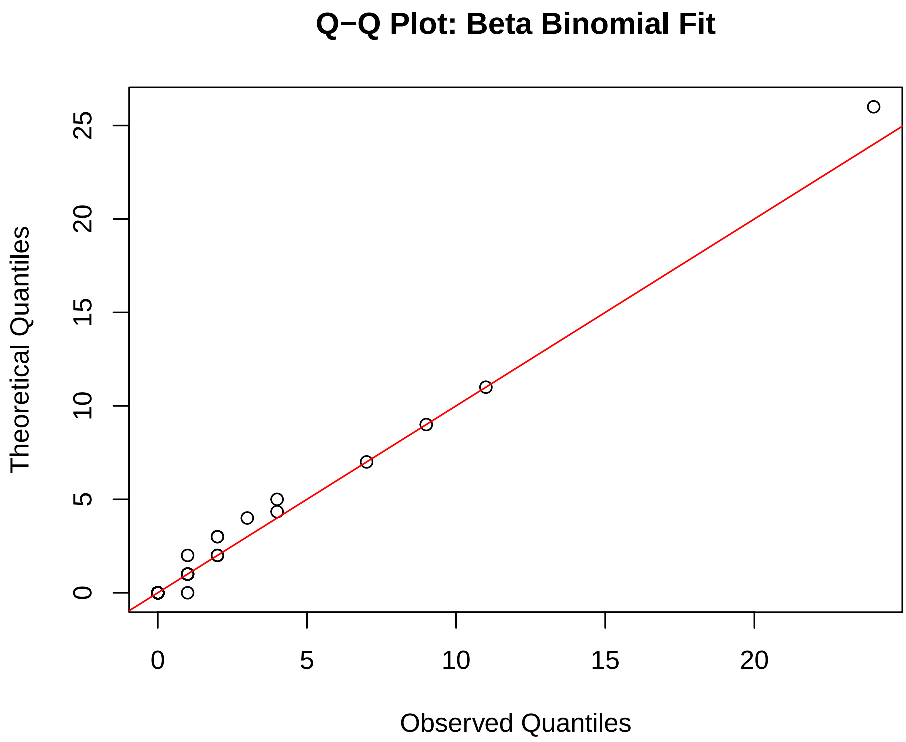

Figure 5: Q-Q plot assessing the fit of Beta-Binomial distribution to the transformer defect data.

- Both Beta-Binomial and Negative Binomial distributions provide excellent fits, with nearly equivalent statistical support (AIC weights of 0.49 and 0.51, respectively).

- The non-significant chi-square test result ($p = 0.477$) confirms that the Beta-Binomial distribution adequately represents the defect data.

- The complete lack of support for Poisson and Binomial models (AIC weights essentially zero) highlights the importance of accounting for over-dispersion in quality control applications.

5.2.6. Control Chart Implementation. To evaluate the bootstrap control charts for the real transformer data, the observations were organized into subgroups of six, corresponding to summer and winter periods for each year. The calculated statistics for each subgroup are presented in Table 9.

Table 9: Statistical summary of the transformer defect data organized in subgroups of six

Row	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	$BB\hat{M}$	$BB\hat{S}_d$
1	1	0	1	0	3	4	1.484	1.498
2	0	2	0	0	0	0	0.314	0.826
3	0	0	0	0	0	0	0.000	0.002
4	0	0	0	0	0	0	0.000	0.002
5	11	1	0	0	1	1	0.601	0.737
6	1	0	2	0	0	0	0.497	0.800
7	0	0	0	0	0	0	0.000	0.002
8	0	0	0	0	0	0	0.000	0.002
9	0	0	2	2	0	1	0.832	0.929
10	7	9	1	1	4	24	1.953	2.251

Subsequently, the bootstrap M-chart and S-chart for monitoring the process mean and standard deviation, respectively, were constructed based on this data (Figures 6 and 7). These charts provide the necessary control limits for ongoing quality surveillance of the transformer manufacturing process.

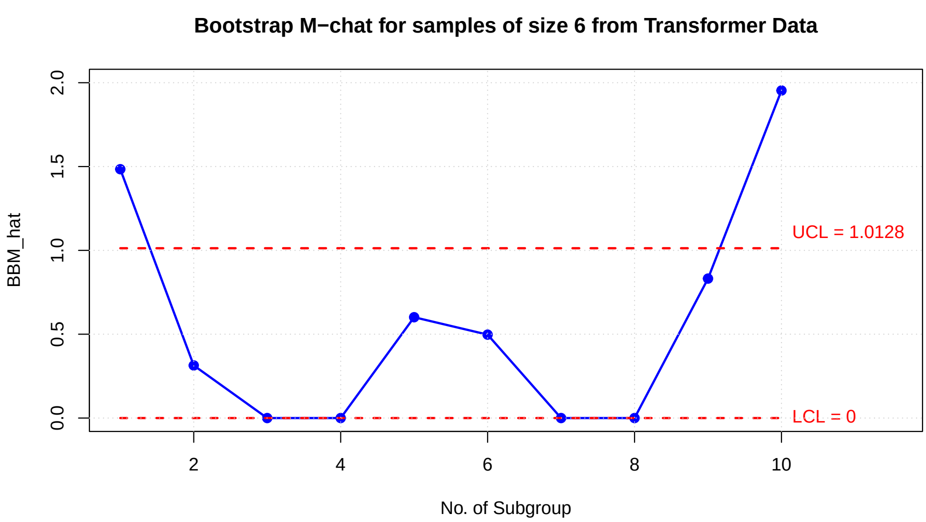

Figure 6: Bootstrap M-chart applied to the transformer manufacturing data with sample size 6. The chart monitors the process mean with control limits at 0 and 1.0128.

The charts effectively detected the increased defect rates observed in 2021, providing early warning of quality issues that warranted investigation and corrective action.

Both simulated and real-world examples demonstrate the practical utility of bootstrap control charts for Beta-Binomial processes. The methodology effectively detects process changes, accommodates over-dispersed count data, and provides a robust framework for quality monitoring in manufacturing environments characterized by variable production volumes and low defect rates.

6. Conclusion

This study successfully developed and validated a bootstrap-based methodology for constructing control charts for processes characterized by Beta-Binomial distributions. The primary challenge of deriving accurate control limits for such non-normal, over-dispersed data was effectively addressed by leveraging parametric bootstrapping, which empirically estimates the sampling distribution of the sample mean (M-chart) and standard deviation (S-chart).

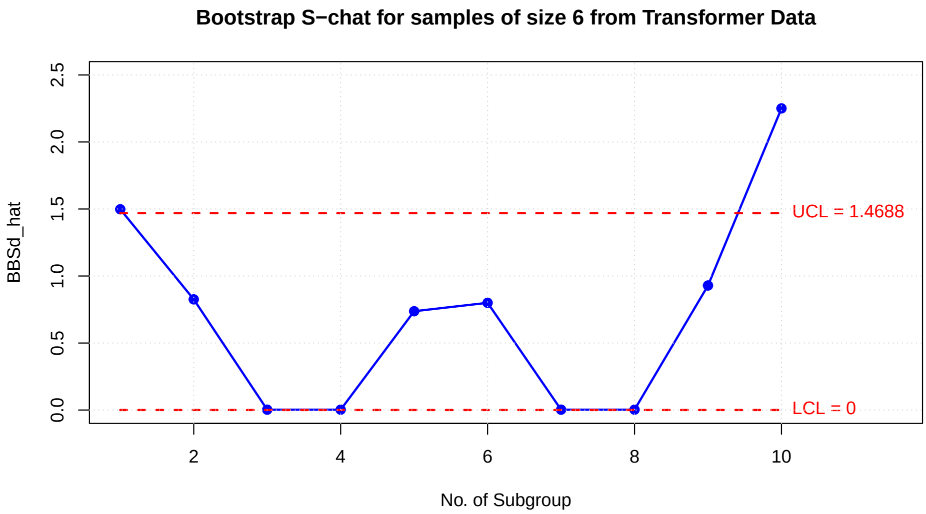

Figure 7: Bootstrap S-chart applied to the transformer manufacturing data. The chart monitors process variability and helps detect changes in dispersion.

The simulation study unequivocally demonstrated the effectiveness of the proposed charts. Under in-control conditions, the bootstrap limits consistently achieved the nominal false alarm rate ($ARL_0 \simeq 370$), confirming their statistical accuracy. Furthermore, the charts exhibited high sensitivity in detecting out-of-control conditions, with the average run length (ARL_1) decreasing significantly as the sample size increased or the magnitude of the parameter shift grew. The analysis also revealed that the charts are particularly adept at detecting increases in the defect probability and changes in process dispersion.

The application to real-world data from transformer manufacturing underscored the practical relevance of the methodology. The analysis confirmed that the defect data was severely over-dispersed, rendering traditional Binomial and Poisson models entirely unsuitable. The Beta-Binomial distribution, in contrast, provided an excellent fit, as visually confirmed by the fitted distribution plot (Figure 4) and the quantile-quantile plot. The subsequent bootstrap M and S charts, derived from the fitted model, established statistically sound control limits (e.g., UCL $= 1.0128$ for the M-chart in Figure 6) that can be used for ongoing quality surveillance.

In conclusion, the bootstrap approach offers a powerful, flexible, and theoretically sound alternative for monitoring processes that exhibit extra-binomial variation. It eliminates the need for complex analytical derivations or unrealistic normality assumptions. For future research, this work can be extended by investigating multivariate scenarios, developing adaptive control schemes, and exploring non-parametric bootstrap methods for distributions beyond the Beta-Binomial family. The charts presented herein mark a significant contribution to the toolbox of modern quality control, especially for industries dealing with complex, low-defect, or highly variable production processes.

Acknowledgments

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data Availability Statement: The data was taken for the number of defective electrical transformers during manufacture for five years from the Diyala General Company for Electrical Industries in Diyala.

Code Availability: The authors composed the code for this article using R software, and you can acquire it by reaching out to the corresponding author via email.

Ethics Approval: We take deliberate steps to ensure that the following conditions are met for this manuscript: The content represents the authors' exclusive and original work, with no prior publication elsewhere. It is not under consideration for publication elsewhere, and it accurately and comprehensively presents the authors' research and analysis.

Conflicts of Interest: The authors assert that they have no conflicts of interest.

References

1. Al-Lami, M. A. H., Jabbari Khamnei, H., and Heydari, A. A., *A parametric bootstrap control chart for Lindley Geometric percentiles*, PLOS ONE, 20(2), e0316449, 2025.

2. Al-Nuaami, W. A. H., Heydari, A. A., and Jabbari Khamnei, H., *The Poisson-Lindley distribution: Some characteristics, with its application to SPC*, Mathematics, 11(11), 2428, 2023.

3. Bibby, B. M. and Væth, M., *The two-dimensional beta binomial distribution*, Statistics and Probability Letters, 81(7), 884-891, 2011.

4. Burkhalter, R. and Lio, Y., *Bootstrap control charts for the generalized Pareto distribution percentiles*, Chilean Journal of Statistics, 12(1), 1-22, 2021.

5. Cepeda-Cuervo, E. and Cifuentes-Amado, M. V., *Tilted beta binomial linear regression model: A Bayesian approach*, Journal of Mathematics and Statistics, 16(1), 1-8, 2020.

6. Chiang, J.-Y., Lio, Y. L., Ng, H. K. T., Tsai, T.-R., and Li, T., *Robust bootstrap control charts for percentiles based on model selection approaches*, Computers & Industrial Engineering, 123, 119-133, 2018.

7. Chowdhury, S., Kundu, A., and Modok, B., *Bootstrap beta control chart for monitoring proportion data*, International Journal of Quality & Reliability Management, 39(10), 2354-2377, 2022.

8. Coetzee, E. J., *The theory and application of bootstrap control charts for statistical process control*, Stellenbosch University, 2017.

9. Danaher, P. J., *A three-parameter generalisation of the beta-binomial distribution with applications*, Defense Technical Information Center, ADA185733, 1987.

10. Danaher, P. J., *Parameter estimation and applications for a generalisation of the beta-binomial distribution*, Australian Journal of Statistics, 30(3), 263-275, 1988.

11. Dopke, J., *Estimation of parameters in gamma distribution*, International Journal of Quality & Reliability Management, 11(8), 27-43, 1994.

12. Efron, B., *Bootstrap methods: Another look at the jackknife*, The Annals of Statistics, 7(1), 1-26, 1979.

13. Everson, P. J. and Bradlow, E. T., *Bayesian inference for the beta-binomial distribution via polynomial expansions*, Journal of Computational and Graphical Statistics, 11(1), 202-207, 2002.

14. Hogg, R. V., Tanis, E. A., and Zimmerman, D. L., *Probability and Statistical Inference*, Macmillan, 1977.

15. Jones, L. A. and Woodall, W. H., *The performance of bootstrap control charts*, Journal of Quality Technology, 30(4), 362-375, 1998.

16. Khusna, H., Mashuri, M., Ahsan, M., Suhartono, S., and Prastyo, D. D., *Bootstrap-based maximum multivariate CUSUM control chart*, Quality Technology & Quantitative Management, 17(1), 52-74, 2020.

17. Leiva, V., dos Santos, R. A., Saulo, H., Marchant, C., and Lio, Y., *Bootstrap control charts for quantiles based on log-symmetric distributions with applications to the monitoring of reliability data*, Quality and Reliability Engineering International, 39(1), 1-24, 2023.

18. Lio, Y. L. and Park, C., *A bootstrap control chart for inverse Gaussian percentiles*, Journal of Statistical Computation and Simulation, 80(3), 287-299, 2010.

19. Lio, Y. L. and Park, C., *A bootstrap control chart for Birnbaum-Saunders percentiles*, Quality and Reliability Engineering International, 24(5), 585-600, 2008.

20. Liu, R. Y. and Tang, J., *Control charts for dependent and independent measurements based on bootstrap methods*, Journal of the American Statistical Association, 91(436), 1694-1700, 1996.

21. Ma, Z., Park, C., and Wang, M., *A robust bootstrap control chart for the log-logistic percentiles*, Journal of Statistical Theory and Practice, 16(1), 3, 2022.

22. Mairinque, L. A., Pereira, R. B. D., Nascimento, K. M., Lauro, C. H., and Brandão, L. C., *A bootstrap control chart for the availability index*, The International Journal of Advanced Manufacturing Technology, 120(7), 5151-5161, 2022.

23. Minka, T., *Estimating a Dirichlet distribution*, Massachusetts Institute of Technology, 2000.

24. Montgomery, D. C., *Introduction to Statistical Quality Control*, John Wiley & Sons, 8th edition, 2019.

25. Mostajeran, A., Iranpanah, N., and Noorossana, R., *A new bootstrap based algorithm for Hotelling's T2 multivariate control chart*, Journal of Sciences, Islamic Republic of Iran, 27(3), 269-278, 2016.

26. Nichols, M. D. and Padgett, W. J., *A bootstrap control chart for Weibull percentiles*, Quality and Reliability Engineering International, 22(2), 141-151, 2006.

27. Noorossana, R. and Ayoubi, M., *Profile monitoring using nonparametric bootstrap T2 control chart*, Communications in Statistics - Simulation and Computation, 41(3), 302-315, 2012.

28. Olmo-Jiménez, M. J., Martínez-Rodríguez, A. M., Conde-Sánchez, A., and Rodríguez-Avi, J., *A generalization of the bivariate beta-binomial distribution*, Journal of Statistical Planning and Inference, 141(7), 2303-2311, 2011.

29. Phaladiganon, P., Kim, S. B., Chen, V. C. P., Baek, J.-G., and Park, S.-K., *Bootstrap-based T2 multivariate control charts*, Communications in Statistics - Simulation and Computation, 40(5), 645-662, 2011.

30. Saeed, N., Kamal, S., and Aslam, M., *Percentile bootstrap control chart for monitoring process variability under non-normal processes*, Scientia Iranica, 2021.

31. Scollnik, D. P. M., *Bayesian inference for three bivariate beta binomial models*, The Open Statistics & Probability Journal, 8(1), 27-38, 2017.

32. Seppala, T., Moskowitz, H., Plante, R., and Tang, J., *Statistical process control via the subgroup bootstrap*, Journal of Quality Technology, 27(2), 139-153, 1995.

33. Sonin, A. A., *A generalization of the*, Annals of Mathematical Statistics, 2004.

34. Takeshita, J.-I. and Suzuki, T., *Precision for binary measurement methods and results under beta-binomial distributions*, arXiv preprint, arXiv:2008.13619, 2020.

35. Teerapabolarn, K., *A bound on the binomial approximation to the beta binomial distribution*, International Mathematical Forum, 28, 1355-1358, 2008.

36. Teerapabolarn, K., *On pointwise binomial approximation for independent beta binomial random variables*, Applied Mathematical Sciences, 9(9-12), 459-463, 2015.

37. Teerapabolarn, K. and Sae-Jeng, K., *A non-uniform bound on binomial approximation to the beta binomial cumulative distribution function*, Songklanakarin Journal of Science and Technology, 41(1), 96-108, 2019.

38. Werner, G. and Hanka, L., *Using the beta-binomial distribution for the analysis of biometric identification*, IEEE 13th International Symposium on Intelligent Systems and Informatics (SISY), 209-215, 2015.

39. Wilcox, R. R., *Estimating the parameters of the beta-binomial distribution*, Educational and Psychological Measurement, 39(3), 527-535, 1979.

40. Wu, Z. and Wang, Q., *Bootstrap control charts*, Quality Engineering, 9(1), 143-150, 1996.

41. Xu, J. and Peng, C., *Parametric bootstrap process capability index control charts for both mean and dispersion*, Communications in Statistics - Simulation and Computation, 48(10), 2936-2954, 2019.

Mohammed Faris Abed ,

General Authority for Statistics and Geographic Information Systems,

Iraq.

E-mail address: fcb.prince1994@gmail.com

and

Ali Akbar Heydari,

Department of Statistics, Faculty of Mathematics, Statistics and Computer Science,

Iran.

E-mail address: heydari@tabrizu.ac.ir

and

Hossein Jabbari Khamnei,

Department of Statistics, Faculty of Mathematics, Statistics and Computer Science,

Iran.

E-mail address: h_jabbari@tabrizu.ac.ir