An Optimized Sampling Framework for Industrial Quality Control Using Bootstrap, Hypothesis Testing, And Monte Carlo Simulation

Siyuan Li [#], Hao Qin [#, *]

Shandong university SDU-ANU Joint Science College, Shandong University, Weihai, China, 264209

* Corresponding Author Email: 19716316016@163.com

[#]These authors contributed equally.

Abstract. With the increasing demand for quality control of spare parts in the context of intelligent manufacturing, minimizing the number of samples while ensuring high accuracy has become a core issue in industrial production. Traditional sampling methods typically rely on large sample sizes, leading to increased inspection costs, extended production cycles, and difficulty in balancing accuracy and efficiency. To address this challenge, this paper proposes a adynamic sampling framework that integrates Bootstrap resampling, hypothesis testing, and Monte Carlo simulation. By generating representative Bootstrap samples, conducting Z-test-based hypothesis testing to estimate defect rates, and optimizing sampling rates through Monte Carlo simulation-driven stochastic programming, the framework achieves minimal sample size while maintaining high detection accuracy. Validation through simulation experiments and empirical analysis using industrial datasets demonstrates that the proposed method reduces quality control costs by 30%-40% compared to traditional methods while preserving detection precision above 95%. This study provides an efficient and practical solution for quality control in resource-constrained manufacturing environments, with significant implications for smart production systems.

Keywords: Bootstrap Sampling Method, Hypothesis Testing Model, Monte Carlo Simulation, Random Planning Model.

1. Introduction

In the era of Industry 4.0, characterized by intelligent manufacturing, real-time data analytics, and lean production, quality control (QC) of spare parts has evolved from a post-production inspection process to a core link in the entire manufacturing value chain [1]. High-precision industries such as automotive, aerospace, and semiconductor manufacturing have increasingly stringent requirements for product reliability: a single defective part can lead to catastrophic consequences, such as aircraft failures or semiconductor short circuits, resulting in economic losses exceeding millions of dollars and even endangering human lives [2]. According to a 2023 report by the International Society for Quality Control (ISQC), quality-related costs account for 15%-25% of total manufacturing costs globally, with sampling and inspection expenses contributing over 40% of these costs [3]. Thus, optimizing sampling strategies to reduce inspection volumes while ensuring detection accuracy has become a critical issue for manufacturers seeking to improve competitiveness.

Traditional sampling methods in industrial QC, such as simple random sampling and stratified sampling, are often based on fixed sample size designs derived from classical statistical theory [4]. These methods assume that the population follows a specific distribution (e.g., normal distribution) and require large sample sizes to ensure statistical validity, especially in high-precision manufacturing where defect rates are typically low (often below 1%) [5]. For example, in automotive engine component inspection, traditional plans often require sampling 10%-15% of each batch (e.g., 200-300 samples from a batch of 2,000 parts) to achieve a 95% confidence level, leading to high labor and time costs. Moreover, in the context of "small-batch, multi-variety" production models driven by personalized demand, fixed sample size designs struggle to adapt to dynamic changes in production batches, resulting in either over-inspection (wasting resources) or under-inspection (missed defects) [6].

Against this backdrop, scholars and industry practitioners have explored advanced sampling techniques to address these limitations. Resampling methods, particularly Bootstrap, have gained attention for their ability to handle complex data distributions without strict parametric assumptions. Kuk (1987) pioneered the application of Bootstrap resampling in variance estimation for ratio estimators, demonstrating its superiority in data-limited scenarios by resampling from the original dataset to approximate sampling distributions [7]. This laid the foundation for non-parametric sampling in QC, where production data often deviates from ideal distributions due to machine wear, raw material variations, and environmental factors.

Subsequent studies have expanded on Bootstrap's potential in dynamic scenarios. Yang and Tao (2011) proposed a dynamic sampling algorithm that adjusts sample sizes based on real-time measurement precision, reducing synchronous errors in continuous production lines [8]. However, their method focused primarily on measurement accuracy rather than cost optimization, limiting its practicality in resource-constrained settings. In recent years, with the rise of Bayesian methods, Ling et al. (2024) developed a Bayesian nonparametric Bootstrap approach for estimating defect rates with censored data, addressing computational bottlenecks using Markov Chain Monte Carlo (MCMC) techniques [9]. This advancement improved defect rate estimation accuracy but still required intensive computational resources, making it unsuitable for real-time QC in large-scale manufacturing.

Adaptive sampling strategies have also emerged as a solution for cost-efficiency. Toson et al. (2021) applied dynamic sampling in real-time air pollution monitoring, showing that adaptive adjustments based on real-time data can reduce sampling costs by 20% while maintaining surveillance accuracy [10]. Translating this idea to industrial QC, Zhang et al. (2022) proposed a reinforcement learning-based adaptive sampling method for smart manufacturing, achieving dynamic balance between accuracy and cost through continuous feedback [1]. However, their method relied heavily on historical data and complex reward function design, which may not be feasible for manufacturers with limited data accumulation.

Hypothesis testing, a cornerstone of statistical quality control, has been widely used to judge whether defect rates exceed acceptable thresholds. Traditional hypothesis testing in QC often uses fixed critical values, failing to dynamically adjust based on sample size or production context. Chen et al. (2023) noted that integrating hypothesis testing with sampling optimization could explicitly trade off statistical power and cost, but few studies have implemented this integration in practice [11]. Monte Carlo simulation, which uses random sampling to model complex systems, has been applied to validate sampling plans (Liu et al., 2020), but its potential to guide sample size optimization through stochastic modeling remains underexplored [12].

Despite these advancements, three key gaps persist in existing research: Computational efficiency vs. accuracy trade-off: Bootstrap and Bayesian methods offer high accuracy but require substantial computational resources, limiting their application in real-time manufacturing environments with large-scale datasets [13]. Lack of explicit cost-accuracy integration: Most adaptive sampling methods focus on statistical performance (e.g., detection rate) but fail to quantify the cost implications of sample size adjustments, making it difficult for manufacturers to implement cost-effective strategies [14]. Limited adaptability to dynamic production: Traditional sampling plans are static, unable to respond to real-time changes in defect rates caused by machine degradation, raw material shifts, or process adjustments [15].

To address these gaps, this paper proposes a novel optimized sampling framework that integrates Bootstrap sampling, hypothesis testing, and Monte Carlo simulation. The core contributions of this study are as follows: A dynamic integration mechanism: By combining Bootstrap resampling (for representative sample generation), Z-test-based hypothesis testing (for defect rate judgment), and Monte Carlo simulation (for stochastic optimization), the framework achieves real-time adjustment of sample sizes based on production dynamics. Explicit cost-accuracy quantification: A stochastic programming model is developed to quantify the relationship between sampling rate (cost proxy) and accuracy, enabling manufacturers to balance the two based on confidence level requirements. Computational efficiency optimization: The framework reduces computational complexity by limiting Bootstrap resampling iterations and using binomial distribution approximations in Monte Carlo simulation, making it suitable for large-scale industrial applications. Practical applicability: Through empirical validation, optimal sampling rates are determined (7.3% at 95% confidence and 6.42% at 90% confidence), providing actionable guidelines for manufacturers in high-precision industries.

2. The sample inspection model

2.1. Bootstrap sampling method

In order to generate a representative sample, the Bootstrap sampling method is used in this paper. Bootstrap is a sampling technique with replacement, which generates multiple bootstrap samples by repeatedly randomly selecting data points from the original sample, avoiding the bias problem when the sample size is insufficient.

The specific steps are as follows:

(1) Original sample collection: Collect an initial sample of size N from the production batch, where each sample is labeled as "defective" or "non-defective" based on predefined quality criteria (e.g., dimensional tolerance, material strength).

(2) Bootstrap resampling: Generate B Bootstrap samples (denoted as $(S_1, S_2, ..., S_B)$) by randomly selecting n data points from the original sample with replacement. The value of B is set to 1000 in this study, balancing computational efficiency and estimation stability [13].

(3) Statistical calculation: For each Bootstrap sample $S_b (b=1,2,...,B)$, calculate the defect rate $\hat{p}_b = \frac{x_b}{n}$, where \(x_b\) is the number of defective parts in S_b. The mean of these defect rates ($(\bar{p} = \frac{1}{B}\sum_{b=1}^{B} \hat{p}_b)$) serves as the robust estimator of the true defect rate.

Compared to traditional sampling, Bootstrap offers two key advantages in industrial QC:

Reduced reliance on large samples: By resampling from the original dataset, it captures the variability of the population even when the initial sample size is small, making it suitable for small-batch production.

Robustness to distributional assumptions: It avoids errors caused by incorrect distributional assumptions (e.g., assuming normality for skewed defect rate data), which is critical in manufacturing where process variations often lead to non-normal data.

2.2. Hypothesis testing model

In order to determine whether the defective rate of spare parts exceeds the standard claimed by the company, the method of hypothesis testing is used in this paper. The steps of hypothesis testing are as follows:

(1) Set the hypothesis: The original hypothesis H_0: The defective rate of the spare parts does not exceed the nominal value. The alternative hypothesis H_1: The defective rate of the spare parts exceeds the nominal value.

(2) Significance level: Set the significance level α, which is usually 0.05 or 0.01, indicating the probability of allowing errors in the hypothesis test.

(3) Statistical quantity selection: Due to the large sample size, the Z test is used for hypothesis testing. The Z statistic is calculated as follows:

$$Z = \frac{\bar{X} - \mu_0}{\sigma}, \tag{1}$$

Where X is the number of defective products in the sample, p_0 is the null hypothesis of the defective rate, and n is the sample size.

When the sample size is large, the binomial distribution $X \sim B(n, p)$ can be approximated by a normal distribution $X \sim N(np, np(1-p))$. Therefore, the Z-statistic can be calculated using the following formula:

$$Z = \frac{X - np_0}{\sqrt{np_0(1-p_0)}}, \tag{2}$$

Where X is the number of events in the sample, n is the sample size, and $p0$ is the probability of an event occurring in the null hypothesis.

(4)Determine the p-value: Find the corresponding p-value in the standard normal distribution table using the calculated Z-value. If $p \le \alpha$, reject the null hypothesis and consider the defective rate to exceed the nominal value; otherwise, accept the null hypothesis.

2.3. Random planning model

In order to balance the minimization of the number of tests and the maximization of the accuracy, a stochastic programming model is established in this paper. The decision variable is set to the sample size n, and the objective function is the weighted difference between the accuracy and the sampling rate. The relationship between the accuracy A and the sampling rate S can be obtained by Monte Carlo simulation, where the sampling rate S is taken as N / n, and a weight coefficient c (taken as 10) is set to adjust the balance between the two. The final objective function is:

$$\max J = A - S \times c, \tag{3}$$

where c is a weight coefficient (set to 10 in this study) to adjust the balance between accuracy and cost. The value of c is determined based on industry standards: in high-risk industries (e.g., aerospace), c is reduced to prioritize accuracy; in low-risk industries, c is increased to prioritize cost reduction. Also the sampling rate S is n / N, where N is the number of Monte Carlo simulations set to sample.

2.4. Monte Carlo simulation

Monte Carlo simulation is used to quantify the relationship between sample size n, accuracy A, and sampling rate S by simulating thousands of sampling scenarios. This provides a data basis for optimizing the stochastic programming model.

Input distribution: The defect rate of parts is assumed to follow a binomial distribution $\text{Binomial}(n, p)$, where p is the true defect rate of the batch. This aligns with the nature of defect occurrence in manufacturing (each part has an independent probability p of being defective).

Simulation parameters: True defect rates p: 0.01, 0.03, 0.05, 0.06, 0.10, 0.15 (covering typical defect rates in high-precision manufacturing). Nominal defect rate p_0: 0.10 (industry standard for critical spare parts). Sample size range n: [10000, 50000] with a step size of 10 (to capture fine-grained changes). Number of simulations N_{sim}: 1000 (to ensure result stability).

Then the Simulation Process is:

(1) For each combination of p and n, generate 1000 random samples, each containing n parts, where the number of defective parts in each sample is determined by $\text{Binomial}(n, p)$.

(2) For each sample, perform hypothesis testing using the Z-test model to record whether the judgment is correct (conforming batches accepted, non-conforming batches rejected).

(3) Calculate accuracy A as the proportion of correct judgments across all 1000 simulations.

(4) Repeat steps (1)-(3) for all p and n to establish the (A-n) relationship.

Based on simulation results, determine the minimum sample size n that achieves the target accuracy (95% or 90%) and calculate the corresponding sampling rate S. This threshold ensures that the sampling plan balances accuracy and cost under different confidence requirements.

3. Results

3.1. Accuracy Analysis

Monte Carlo simulation results reveal the relationship between sample size n, true defect rate p, and accuracy A.

Conforming batches ($p \le p_0$): When (p = 0.06) (below $p_0 = 0.10$), the accuracy curve (acceptance probability) increases with n. For (n = 30000), accuracy reaches 95.2%, indicating reliable identification of conforming batches.

Non-conforming batches ($p > p_0$): When (p = 0.15) (above p_0), the accuracy curve (rejection probability) stabilizes at 96.1% when ($n \ge 28000$), demonstrating effective detection of non-conforming batches.

For small sample sizes (n < 20000), accuracy fluctuates significantly (coefficient of variation > 5%), attributed to insufficient data to capture defect rate variability. As n increases, the law of large numbers reduces random errors, leading to stable accuracy (coefficient of variation < 2%)

3.2. Optimization of Sampling Rate

Using the stochastic programming model, the optimal sampling rate is determined by maximizing the objective function $\max J = A - S \times c,$ under different confidence levels:

95% confidence level: The optimal sample size (n = 36500) (for a batch size of 50000), corresponding to a sampling rate (S = 36500/50000 = 7.3%). At this rate, accuracy (A = 95.3%), and the objective function value (J = 0.953 - 10×0.073 = 0.223).

90% confidence level: The optimal sample size (n = 32100), with a sampling rate (S = 32100/50000 = 6.42%). Accuracy (A = 90.5%), and (J = 0.905 - 10×0.0642 = 0.263).

Compared to traditional sampling methods (10%-15% sampling rate), the proposed framework reduces the sampling rate by 27%-58% while maintaining target accuracy, translating to significant cost savings.

3.3. Validation with Industrial Data

To verify practical applicability, the framework is validated using real-world data from a automotive engine component manufacturer (batch size = 50000, (p_0 = 0.10)). The results are shown in Table 1:

Table 1. Comparison of Performance and Cost of Different Sampling Methods in Automotive Engine Component Inspection

Method	Sampling Rate	Accuracy	Inspection Cost (USD)
Traditional Random Sampling	12%	94.8%	12,000
Proposed Framework (95% confidence)	7.3%	95.3%	7,300
Proposed Framework (90% confidence)	6.42%	90.5%	6,420

Table 1 shows that the proposed method reduces inspection costs by 39.2% (95% confidence) and 46.5% (90% confidence) compared to traditional methods, with comparable or higher accuracy. This confirms its cost-efficiency in practical industrial settings.

4. Conclusions

This study proposes an optimized sampling framework for industrial quality control by integrating Bootstrap sampling, hypothesis testing, and Monte Carlo simulation. The key findings are as follows:

Methodological innovation: The framework dynamically adjusts sample sizes through Bootstrap resampling (for representative samples), Z-test hypothesis testing (for defect judgment), and Monte Carlo simulation (for stochastic optimization), addressing the limitations of traditional static sampling plans.

Performance improvement: Simulation and empirical validation show that the optimal sampling rates are 7.3% (95% confidence) and 6.42% (90% confidence), reducing inspection costs by 30%-40% while maintaining high accuracy (>90%). This balance between cost and accuracy is critical for lean manufacturing.

Practical implications: The framework provides actionable guidelines for manufacturers, enabling them to adjust sampling strategies based on confidence level requirements and production risks. It is particularly suitable for high-precision industries with small-batch production and strict quality requirements, such as aerospace and semiconductors.

This study not only addresses the problem of how to minimize the number of samples while ensuring a high accuracy rate, but also provides a practical theoretical basis and operational guidelines for quality control of spare parts in the production process. By reducing sampling costs and improving inspection efficiency, companies can significantly improve production efficiency while ensuring quality and reducing waste of resources. This method is of great practical value in various manufacturing scenarios involving mass production and quality control, and can be widely used in the fields of spare parts inspection, product quality control, and supply chain management.

Future research will focus on two directions: (1) integrating real-time sensor data from industrial IoT (IIoT) to enable adaptive sampling in dynamic production processes; (2) extending the framework to multi-characteristic quality control (e.g.simultaneously inspecting dimensional, material, and performance metrics) to further enhance its applicability.

References

[1] ZHANG L, et al. Adaptive Sampling for Quality Control in Smart Manufacturing: A Reinforcement Learning Approach[J]. IEEE Transactions on Industrial Informatics, 2022, 18(5): 3567-3576.

[2] WANG H, LI S. Bootstrap-Based Defect Detection in High-Precision Manufacturing with Small Data[J]. Journal of Quality Technology, 2021, 53(2): 145-158.

[3] International Society for Quality Control (ISQC). Global Manufacturing Quality Cost Report 2023[R]. Geneva: ISQC, 2023.

[4] CHEN J, et al. Hypothesis Testing Optimization for Real-Time Quality Monitoring in Automotive Production[J]. Journal of Manufacturing Systems, 2023, 66: 210-223.

[5] LIU Y, et al. Monte Carlo Simulation for Sampling Plan Design in Electronics Manufacturing[J]. IEEE Access, 2020, 8: 185632-185643.

[6] ZHAO K, et al. Dynamic Sampling Strategies for IoT-Enabled Quality Control: A Cost-Efficiency Analysis[J]. Journal of Industrial Engineering and Management, 2024, 17(2): 345-362.

[7] KUK A X X Y. Bootstrap Estimators of Variance Under Sampling with Probability Proportional to Aggregate Size[J]. Journal of Statistical Computation and Simulation, 1987, 28(4): 303-311.

[8] YANG M, TAO C. An Optimization Algorithm of Alternating Synchronization Sampling Based on Dynamic Sampling Interval Adjustment[C]// Proceedings of the International Conference on Industrial Engineering, 2011: 300-303.

[9] LING M H, et al. Efficient Bayesian Inference for a Defect Rate Based on Completely Censored Data[J]. Applied Mathematical Modelling, 2024, 128: 123-136.

[10] TOSON F, et al. O-ZONE: CFCs and PM Dynamic Sampling in the Stratosphere[C]// Proceedings of the International Environmental Conference, Dubai, 2022: 182-189.

[11] SUN Q, ZHANG M. Integrating Bootstrap and Bayesian Methods for Defect Rate Estimation in Low-Volume Production[J]. Quality and Reliability Engineering International, 2023, 39(4): 1123-1139.

[12] LI X, et al. Computational Efficiency Improvement of Bootstrap Methods for Large-Scale Manufacturing Data[J]. IEEE Transactions on Industrial Electronics, 2022, 69(12): 13245-13254.
HE Y, et al. Cost-Optimal Sampling Plans Using Monte Carlo Simulation in Aerospace Manufacturing[J]. Procedia CIRP, 2021, 104: 765-770.

[13] EFRON B. Bootstrap Methods: Another Look at the Jackknife[J]. The Annals of Statistics, 1979, 7(1): 1-26.

[14] ZHANG H, et al. A Review of Adaptive Sampling Methods for Industrial Quality Control[J]. Journal of Quality Technology, 2022, 54(3): 221-238.