

타겟 샘플링 검사를 통한 출하품질 향상에 관한 사례 연구

김준세 · 이창기[†] · 김경남 · 김창우 · 송혜미 · 안성수 · 오재원 · 조현상 · 한상섭

삼성전자 Global CS센터 품질혁신팀

A Case Study on the Target Sampling Inspection for Improving Outgoing Quality

Junse Kim · Changki Lee[†] · Kyungnam Kim · Changwoo Kim · Hyemi Song · Seoungsu Ahn ·
Jaewon Oh · Hyunsang Jo · Sangseop Han

Quality Innovation Team, Global CS Center, SAMSUNG ELECTRONICS

ABSTRACT

Purpose: For improving outgoing quality, this study presents a novel sampling framework based on predictive analytics.

Methods: The proposed framework is composed of three steps. The first step is the variable selection. The knowledge-based and data-driven approaches are employed to select important variables. The second step is the model learning. In this step, we consider the supervised classification methods, the anomaly detection methods, and the rule-based methods. The applying model is the third step. This step includes the all processes to be enabled on real-time prediction. Each prediction model classifies a product as a target sample or random sample. Thereafter intensive quality inspections are executed on the specified target samples.

Results: The inspection data of three Samsung products (mobile, TV, refrigerator) are used to check functional defects in the product by utilizing the proposed method. The results demonstrate that using target sampling is more effective and efficient than random sampling.

Conclusion: The results of this paper show that the proposed method can efficiently detect products that have the possibilities of user's defect in the lot. Additionally our study can guide practitioners on how to easily detect defective products using stratified sampling

Key Words: Predictive Analytics, Stratified Sampling, Machine Learning, Random Sampling, Acceptable Quality Level

● Received 5 August 2021, 1st revised 15 August 2021, accepted 20 August 2021

† Corresponding Author(yjck1220.lee@samsung.com)

© 2021, Korean Society for Quality Management

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-Commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

검사(Inspection)란 제품의 품질을 보증하기 위한 목적으로 개별 제품 혹은 로트(Lot)를 정의된 방법으로 측정하고 적합여부(혹은 합격여부)를 판정하는 것을 의미한다. 검사 방식에는 전수검사와 샘플링검사가 있다. 전수검사는 로트 안에 속한 모든 개별 제품에 대하여 적합여부를 판정하는 것이며, 샘플링검사는 로트 내 일부 제품을 검사하여 해당 로트의 적합여부를 판정하는 것이다. 샘플링검사에서는 로트 크기, 검사 수준, 표본수 등이 고려되어야 한다. 이에 국제표준화기구(International standard organization, ISO)는 샘플링검사 시 고려해야 할 표준을 제정하여 관리하고 있다.

샘플링검사는 품질특성치(Quality characteristic)에 따라 계수형(Attribute) 샘플링검사와 계량형(Quantitative) 샘플링검사로 분류된다. 계수형 샘플링검사는 품질특성치를 적합, 부적합으로 분류하여 로트 적합여부를 판정한다. 계량형 샘플링검사는 품질특성치를 연속형(Continuous)으로 측정하고 이 측정값을 기준으로 로트 적합여부를 판정한다. 실무에서는 품질특성치의 분포 가정이 없고 적용이 용이하여 AQL(Acceptance quality limit, 합격품질수준)기반 계수형 샘플링검사가 일반적으로 활용된다(Yang and Hwang, 2018). AQL기반 계수형 샘플링검사는 로트에 속한 불량률이 사전에 정의한 AQL에 부합하도록 엄격도(수월, 보통, 까다로운 검사)를 조정하는 검사방식이다(KS Q 2859-1, 2019).

스마트팩토리 단계에 접어든 현 시점에서 제품품질은 무결점을 추구하여(Chong et al., 2020) 품질수준은 PPM(Parts per million)으로 산출해야 할 정도로 개선되었으나 퍼센트단위 품질수준을 보증하는 기존 AQL 방식은 대부분의 로트를 합격시키는 한계를 나타내고 있다(Choe, 2007; Choe, 2011). 그럼에도 불구하고 대부분의 기업들은 관성적으로 AQL기반 샘플링 검사를 입고 및 출하검사에 적용하고 있다. 입고검사(Incoming inspection)는 납입된 부품의 적합여부를 판정하며, 출하검사(Outgoing inspection)는 완성품의 적합여부를 판정한다.

출하검사에서 불량을 인지하지 못할 경우 고객은 불량을 경험하게 되므로 높은 수준의 불량검출력(Defect detection) 확보가 필요하다. 그러나 한정된 검사자원으로 인해 제조 현장에서 상시 전수검사를 운영하는 것은 어렵기에 가용 자원 내에서 불량검출력을 최대한 높이는 방안이 요구되어 왔다. Ferrell and Chhoker(2002)는 다구찌(Taguchi) 손실함수를 이용하여 샘플링검사에서 비용이 최소가 되는 경제적 모형을 제안하였고 Hsu and Hsu(2012)는 내부실패(Internal failure)와 판매후실패(Post-sale failure) 비용 함수를 최소화 하는 경제적 모형을 개발하였다. 그러나 언급된 선행연구는 제품특성을 고려하는 것이 아닌 비용에만 초점을 둔 샘플링 방식이다.

샘플링 방법의 변화를 통해 불량검출력을 개선한 연구도 진행되어 왔다(Boetticher, 2005; Sunter et al., 2014). Boetticher(2005)는 랜덤 샘플링과 같은 전통적인 샘플링 방법이 아닌 최근접 이웃 샘플링(Nearest neighbor sampling)을 통해 검출력을 개선할 수 있음을 주장하였고, Sunter et al.(2014)는 아날로그 불량 시뮬레이션(Analog fault simulation)에서 랜덤 샘플링보다 2배 높은 성능을 보이는 결합가능도 가중 샘플링(Defect-likelihood weighted sampling)을 제안하였다.

예측모형(Prediction model)을 이용하여 불량검출력을 개선하기 위한 연구도 진행되어 왔다. Kim and Baek(2012)은 DRAM 출하검사에 앙상블 기법을 적용하여 불량 로트를 예측하였다. 이 연구에서는 1) Stability selection (Meinshausen and Buhlmann, 2010)을 이용하여 총 61개 변수 가운데 41개 변수를 선택하고 2) 선택된 변수로 앙상블 모형을 학습하여 불량 로트를 예측하였다. 연구 결과 머신러닝 기법을 통해 불량검출력을 높일 수 있음을 시사하였다. 그러나 동일 로트라 할지라도 공정변동을 포함한 다양한 원인으로 인해 로트 내 개별 제품품질에 차이가 존재한다(Kang et al., 2012). 따라서 로트 단위 예측이 아닌 개별 제품품질차이를 고려한 예측모형이 제안

되어야 한다.

이에 본 연구는 선행 연구 결과를 참고하여 출하검사 불량검출력 향상을 위한 예측기반 층화표본추출 방법(Prediction-based stratified sampling)을 제안하고 그 효과에 대해 다룬다. 예측기반 층화표본추출 방법(이후, 타겟 샘플링 검사)은 제품생산과정에서 발생하는 데이터를 분석하여 로트 내 개별 제품을 불량가능성에 따라 분류한다. 여기서 데이터 분석은 불량을 예측하기 위한 변수선택(Variable selection)과 예측모형을 개발하는 단계를 의미한다. 제안 방법은 검사 자원을 효율적으로 활용할 수 있도록 하고 출하검사 불량검출력을 상승시킨다. 본 연구에서는 삼성전자의 다양한 제품군(모바일 기기, TV, 냉장고)에 제안 방법을 적용하여 그 효과를 검증한다.

본 연구의 이후 구성은 다음과 같다. 2장에서는 제안된 방법론에 대하여 자세히 서술하였으며 이를 적용한 품질개선 사례는 3장에 서술되어있다. 마지막으로 4장에서는 본 연구의 결론과 시사점을 서술하였다.

2. 제안 방법

정보기술의 발달로 제조 기업들은 제품생산과정에서 발생하는 방대한 양의 데이터를 기록 및 관리하고 있다(Song et al., 2021). 타겟 샘플링 검사(Target sampling inspection, 이후 TSI)는 제조 빅데이터를 분석하여 불량발생 확률이 높은 샘플을 실시간으로 알려주고 이를 출하검사에서 집중적으로 검사하는 방식이다. TSI의 궁극적인 목적은 출하검사에서 더 많은 불량을 정확하게 검출하여 고객 사용 중 발생하는 불량(Market defects, 이후 시장불량)을 효과적으로 개선하는 것이다. TSI는 층화표본추출(Stratified sampling)에 이론적 기반을 두고 있지만 모집단의 정보를 모르는 상태에서 예측 분석(Predictive analytics)을 통해 로트 내 개별 제품을 타겟 층(Target stratum)과 랜덤 층(Random stratum)으로 분할한다. 타겟 층은 불량발생 확률이 높은 제품(이후, TS)으로 구성되며 랜덤 층은 그렇지 않은 제품(이후, RS)으로 구성된다.

TSI는 크게 변수선택(Variable selection), 예측모형학습(Learning models), 모형적용(Applying models) 단계로 구분된다. 변수선택 단계는 TSI 성공여부를 결정짓는 가장 중요한 단계로 제품생산과정의 검사항목을 분석하여 시장불량과 연관성이 높은 중요변수(Important predictor variables, 이후 IPV)를 선별하는 과정이다. IPV를 선정하는 과정은 지식 기반(Knowledge-based) 방식과 데이터 기반(Data-driven) 방식을 조합하여 적용한다. 지식 기반 방식은 현장에서 관리하고 있는 CTF(Critical to function), CTQ(Critical to quality) 등 실무자 전문지식을 활용하여 변수를 선택하는 것이다. 데이터 기반 방식은 시장불량과의 연관성을 고려한 필터 방식(Filter method) 변수선택을 의미한다. IPV는 제품특성과 생산환경을 고려하여 선정하였고 제품별 IPV 선정 과정은 3장에서 자세히 다루도록 하겠다.

다음 단계는 IPV를 이용하여 TS와 RS를 분류하는 예측모형을 학습하는 것이다. 예측모형은 제품과 데이터 특성을 고려하여 머신러닝을 포함한 다양한 예측모형 가운데 가장 우수한 모형을 채택한다. 이 과정은 예측 정확도 이외에도 실시간 예측 및 검사를 가능하게 하는 연산 소요시간 또한 고려 대상이다. 따라서 연산시간이 짧은 규칙기반모형(Rule-based classifier)과 머신러닝모형을 병행 활용하는 최종 예측모형을 생성한다. 마지막 모형적용 단계는 출하검사에 학습된 모형을 적용하는 단계이다. 이 단계는 예측 시작 및 완료 시점 선정, 예측 소요시간 등 예측된 제품을 출하검사로 연계하는 과정을 검사현장과 조율한다. 이외에도 공정, 제품, 예측모형의 변경에도 검사중단이 발생하지 않도록 하는 Fool-Proof 운영사항도 포함된다. 최종적으로 TSI를 통해 출하검사원은 제품의 TS여부를 확인할 수 있다. 아래의 Figure 1은 TSI 프레임워크를 도식화 한 것이다.

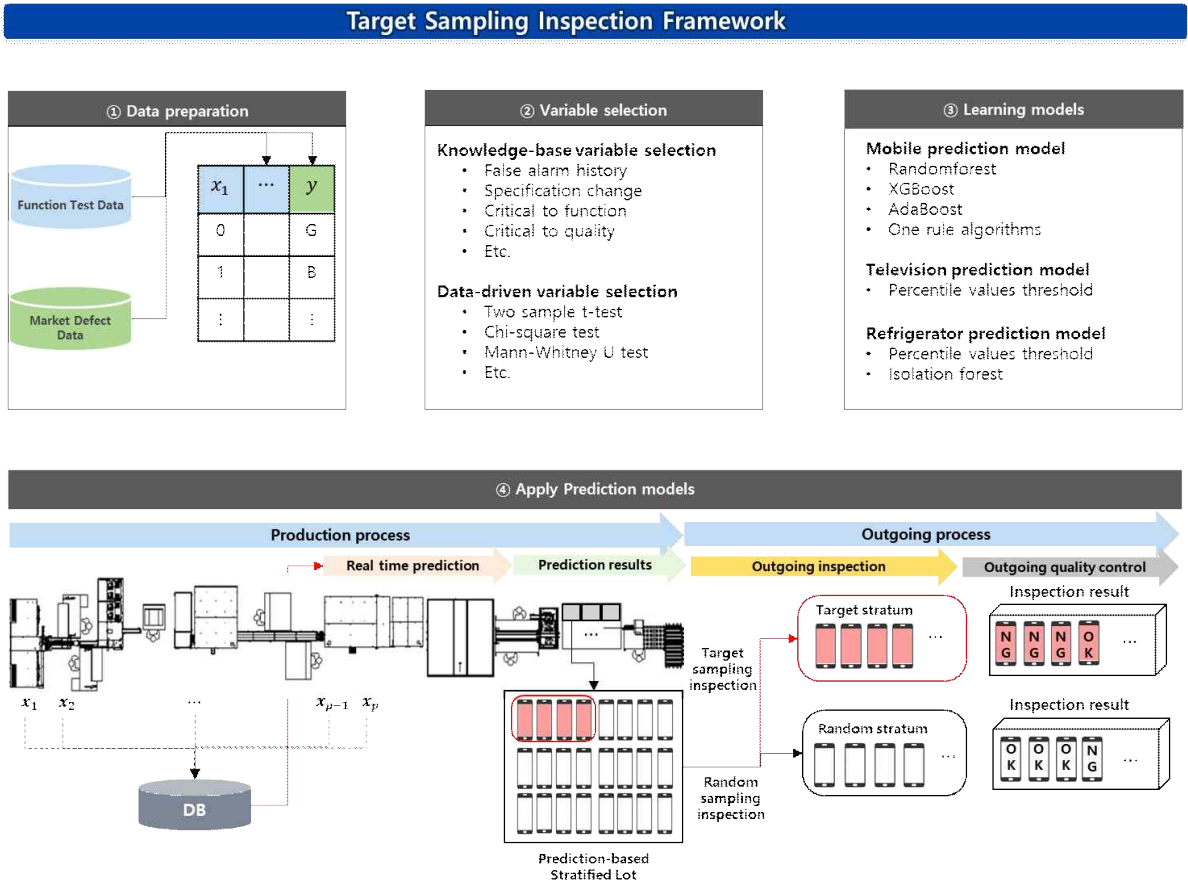


Figure 1. Target sampling inspection framework

3. 사례 연구

성공적인 TSI 적용을 위해서는 데이터 통합 수집 인프라, 데이터 관련 전문지식, 적합한 분석 방법론, 조직간 원활한 협력체계 등이 종합적으로 고려되어야 한다. 이번 장에서는 삼성전자 SET 부문 불량검출력 개선 성과를 살펴보고자 한다. 사례 연구 대상이 되는 제품군은 모바일 기기, TV, 냉장고이며 각 제품의 특성과 생산환경을 고려한 변수 선택과 예측모형학습 과정을 상세하게 서술하였다.

3.1 모바일 기기 사례 연구

모바일 기기는 고객 실사용기간이 짧고 다수의 기능(송수화, 인터넷, 카메라 등)이 결합되어 있는 특징을 가지고 있다. 이로 인해 다양한 기능을 확인하는 검사항목이 3만여개를 넘고 신제품이 출시될 경우 신기능으로 인해 이전 모델에서 진행하지 않았던 검사가 추가되어 검사항목은 더욱 증가한다. 따라서 모바일 기기의 IPV 선정에서는 공통적인 검사항목 선별과 차원의 저주(Curse of dimensionality)에 대한 고려가 필수적이다. 차원의 저주란 검사항목이 늘어남에 따라 예측모형의 성능이 저하되는 현상을 의미한다.

모바일 기기 IPV 선정은 시장불량을 활용하여 다음과 같은 절차로 진행한다. 1) TSI 적용 모델 및 참조 모델 선정

2) 적용 모델과 참조 모델의 공통 검사항목 선별 3) 정보성 변수, CTQ, CTF 선별 4) 시장불량 비율과 연관성이 높은 변수선택이다. 적용 모델이 시장불량 사례가 없는 출시전 신모델인 경우 유사 모델을 참조 모델로 활용한다. 정보성 변수는 검사 재시도 횟수, 가성 불량 등의 검사 정보를 의미한다. 본 연구에서는 시장불량이 없는 최신 플래그십 모바일 기기를 TSI 적용 모델로, 이전 플래그십 모델을 참조 모델로 선정하였다. 분석에 사용된 적용 모델 데이터는 2020년 12월부터 2021년 1월까지 생산된 제품 데이터이며 출시전에 분석을 진행하여 시장불량은 존재하지 않았다. 참조 모델 데이터는 2020년 1월부터 5월까지 생산된 플래그십 모델의 검사 데이터와 시장불량 데이터이다. 적용/참조 모델에서 공통으로 검사하는 2만 여개 항목중 실무자와 조율을 통해 586개를 선별하고 해당 변수와 시장불량과의 연관성을 검토하였다. 586개의 변수중 시장불량과 연관성을 보이는 145개의 변수를 IPV로 선정하였다. 시장불량 연관성은 불량 여부에 따라 수치형 변수는 평균과 분산, 정보성 변수는 비율의 차이가 존재하는지 살펴보았다.

다음으로 참조 모델 IPV와 시장불량을 이용하여 예측모형을 학습하였다. 예측모형으로는 지도 학습(Supervised model)을 사용하였다. 후보 모형으로는 XGBoost, Randomforest, Adaboost를 선정하였고 성능 지표로는 AUC(Area under the curve)를 사용했다. AUC는 수신자작특성곡선(Receiver operating characteristic curve, ROC curve) 아래 면적을 의미하고 1에 가까울수록 좋다. 모형 훈련 및 검증을 위해 데이터를 8:2로 분할하였고 훈련데이터 수는 725,703이며 시험데이터 수는 181,426이다. 훈련데이터는 제조 데이터의 특성상 불균형비율(Imbalance ratio=소수 클래스/다수 클래스) 값이 매우 작다. 이에 Random undersampling을 사용하여 훈련데이터 불균형 비율을 0.05 수준으로 조정하였다.

다음으로 각 모형의 하이퍼파라미터는 5-교차타당법(5-fold cross-validation)을 이용하여 결정하였고 평가기준으로 AUC를 사용하였다. XGBoost에서 고려한 하이퍼파라미터는 학습률(0.01, 0.05, 0.1)과 트리 깊이(3, 9, 15)이며 Randomforest는 트리 수(50, 100, 200)와 트리 깊이(5, 10, 최대), Adaboost는 학습률(0.01, 0.05, 0.1)과 트리 수(50, 100, 200)이다. 시험데이터에서 AUC를 계산한 결과 XGBoost는 0.7271, Randomforest는 0.7490, Adaboost는 0.7417를 나타내어 Randomforest를 예측모형으로 채택하였다.

Randomforest와 함께 단순한 구조로 불량을 예측하는 OneR(One rule algorithm, 이후 OneR)을 예측모형에 추가하였다. Holte(1993)에 따르면 매우 간단한 단 하나의 룰(OneR)이 때로는 데이터 기반 머신러닝 모형 이상의 정확도를 나타낼 수 있다. OneR 규칙은 IPV에 포함된 정보성 변수가 값을 가지는 경우에 TS로 예측한다. 예를 들면, IPV 재검사 횟수가 1회 이상이면 TS로 분류하는 유형의 규칙이다.

3.2 TV 사례 연구

TV 검사항목은 모바일 기기에 비하여 상대적으로 적은 500여 개이지만 여전히 차원의 저주가 존재한다. 따라서 TV도 모바일 기기와 유사한 분석 절차를 거쳐 IPV를 선정하였다. 먼저, TSI 적용 모델은 TV 전체모델로 선정하였고 전체 검사항목 중에서 실무자 전문지식을 기반으로 27개 변수를 1차 선별하였다. 1차 선별된 27개 검사항목에 대한 통계 검증을 수행하여 양품과 불량품 간 유의미한 차이를 나타내는 9개 변수를 IPV로 선정하였다.

TV 예측모형은 데이터 특성을 고려하여 거리기반 이상탐지(Anomaly detection) 기법을 적용하였다. TV의 경우 생산 공장의 환경적 요인으로 인해 정상 제품임에도 이를 측정하는 측정값 분포와 범위 변동이 빈번하게 발생한다. 이런 특수성을 고려하여 볼 때, 변동 사항을 매번 모형에 반영하여 적용하기에는 많은 인력과 비용을 수반한다. 따라서 본 연구에서는 지도학습 모형보다는 간단한 방식이라도 유연하게 현장에 적용될 수 있는 방법이 더 적합하다고 판단하였다.

예측모형은 다음과 같은 절차로 개발되었다. 1) IPV별 품질수준을 LB(Larger-the-better), NB(Nominal-the

-best), SB(Smaller-the-better) 특성으로 분류 2) 각 특성에 대하여 평균 기반 이상치(Outlier) 정의 3) 실시간으로 IPV별 측정값 분포와 범위 변동을 고려하여 주별 데이터에서 이상치 조건에 맞는 제품을 TS로 그렇지 않은 경우 RS로 분류한다. 각 특성에 대하여 이상치는 다음과 같이 정의하였다. 주별 생산데이터의 평균과 표준편차를 이용하여 LB는 평균보다 5표준편차 작은 값을, SB는 평균보다 5표준편차 큰 값을 이상치로 정의하였다. NB 특성은 평균보다 ± 2.5 표준편차 크거나 작은 값을 이상치로 정의한다. 해당 임계값은 실무자 협의 및 출하검사 가용 자원을 고려하여 정의하였다.

3.3 냉장고 사례 연구

이번 절에서는 냉장고 IPV 선정 과정과 예측모형에 대하여 다룬다. 냉장고에는 냉장 기능 이외에 다양한 부가 기능이 추가되고 있어 모델별로 상세 검사항목에 차이가 있으나 공통적으로는 301개의 검사를 진행한다. 시장에서 접수되는 주요 불량은 외관과 냉장 기능으로 압축된다. 이 중에서 냉장 기능 불량은 외관 불량보다 불량 처리 비용이 상대적으로 크기 때문에 본 연구는 냉장 기능 불량에 한정한다.

냉장고 IPV 선정은 다음과 같은 과정으로 진행되었다. 적용 모델은 삼성전자에서 생산하는 냉장고 전체 모델이다. 다음으로 IPV 는 1) 검사 규격 상한(Upper specification limit)과 하한(Lower specification limit)이 없는 변수 제거 2) 변동이 없는 상수 변수 제거 3) 시장불량 연관성이 높은 변수를 실무자 전문지식을 활용하여 선별하였다. 그 결과 전체 301개 검사항목 중에 23개의 검사항목이 채택되었다. 이 중 시장불량과 높은 연관성을 가지고 있는 변수가 시계열 특성을 가지고 있고, 이 변수는 요약 통계량(Summary statistic)으로 관리되고 있음을 확인하였다. 그리하여 해당 변수의 전체 시계열 자료를 수집하는 체계를 구축하고 이를 예측모형 학습에 활용하였다.

냉장고 데이터 역시 TV 데이터와 유사하게 제조회사 요인으로 인해 데이터 변동이 발생하는 특징을 지니고 있다. 따라서 전체 데이터 대신 최근 생산 제품 데이터만을 고려한다. 연속적인 제품 생산 과정에서 i 번째 생산된 제품의 시계열 데이터를 x_i 라고 할 때, i 번째 생산된 제품의 TS여부는 다음과 같은 절차로 진행된다.

- 1) 최근 k 대 제품($x_{i-k}, x_{i-k+1}, \dots, x_{i-1}$)의 시점 별 중앙값(Median)을 이용하여 기준 시계열(y_i)을 생성
- 2) 대상 시계열 x_i 와 기준 시계열 y_i 의 유클리드 거리 $D_i = D(x_i, y_i)$ 를 계산

- 3) 직전 k 대 제품의 유클리드 거리 평균 $\mu_i = \frac{1}{k} \sum_{j=1}^k D_{i-j}$ 과 표준편차를 계산 $\sigma_i = \sqrt{\frac{1}{k-1} \sum_{j=1}^k (D_{i-j} - \mu_i)^2}$

- 4) $D_i \geq \mu_i + 9\sigma_i$ 또는 $D_i \leq \mu_i - 9\sigma_i$ 이면 TS로 분류

추가적으로 거리 기반 방식이 모든 시점을 고려하는 단점을 보완하고자 특정 시점 혹은 구간에서의 이상치를 탐지하기 위한 이상탐지모형 Isolation forest도 함께 적용하였다. Isolation forest에서 이상치는 상대적으로 적은 분할 규칙으로도 쉽게 분리되는 관측값으로 정의한다(Liu et al., 2008). 따라서 각 시점을 하나의 변수로 간주하여 Isolation forest를 적용하면 시점 또는 구간에서 발생하는 이상치를 탐지할 수 있으며 그 경우 TS로 분류된다.

3.4 사례 연구 결과

이번 절에서는 실제 삼성전자 출하 TSI 적용 결과를 랜덤 샘플링 검사(Random sampling inspection, 이후 RSI)와 비교하여 제품별로 제시하고자 한다. 모바일 기기는 2달간 541,903대를 검사하였으며 TS 대상은 43,765대(약 8.1%), RS 대상은 498,138대(약 91.9%)이다. TV는 11개월 간 총 323,802대를 검사하였으며 TS 대상은 16,733대

(약 5.2%), RS 대상은 307,069대(약 94.8%)이다. 마지막으로 냉장고는 10개월간 36,696대를 검사하였으며 TS 대상은 1,402대(약 3.8%), RS 대상은 35,294대(약 96.2%)이다. Table 1은 제품별 데이터 수집 기간, 총 검사 제품 수, TS 대상 검사 제품 수, RS 검사 제품 수를 제시하였다.

Table 1. Description of the data

Products	Periods	Number of total inspection	Number of TS inspection	Number of RS inspection
Mobile	2020.12 ~ 2021.01	541,903	43,765	498,138
TV	2020.02 ~ 2020.12	323,802	16,733	307,069
Refrigerator	2020.03 ~ 2021.12	36,696	1,402	35,294

수집 기간 동안 TSI 불량검출력은 RSI 불량검출력에 비해 모바일 기기, TV, 냉장고에서 각각 4.17, 6.55, 8.39배가 높게 나타났으며 통계적인 검증을 위해 부트스트래핑(Bootstrapping)을 이용하여 결과를 10,000회에 걸쳐 재생산하였다. 부트스트래핑은 단일 데이터셋에서 복원추출(Sampling with replacement)을 통해 모수(Parameter)를 추정하는 방법이다(Efron, 1997). 부트스트래핑 결과 모바일 기기, TV, 냉장고 TSI 평균불량검출력(Average defect detection)은 RSI 평균불량검출력에 비해 4.26, 6.67, 8.48배로 높게 나타났다. 기존의 결과와 비교하여 모든 제품에 대하여 불량검출력이 소폭 증가(모바일 기기: 4.17 → 4.26, TV: 6.55 → 6.67, 냉장고: 8.39 → 8.48)했다. Figure 2는 제품별 평균불량검출력을 도식화한 것이다. 여기서 x 축은 제품군, y 축은 부트스트래핑을 이용하여 얻은 평균불량검출력을 의미한다.

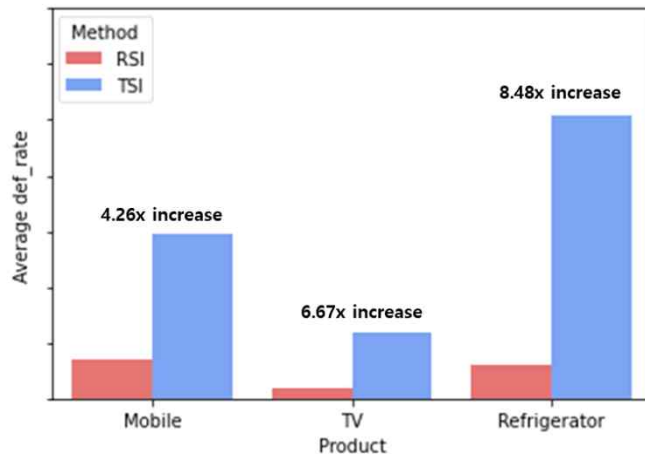


Figure 2. Result of the bootstrapping

각 제품별 불량검출력 히스토그램은 Figure 3와 같다. 여기서 x 축은 불량검출력을 의미하며 그 값이 클수록 동일한 검사 수량에 대하여 많은 불량을 찾아냄을 의미하고 y 축은 해당 불량검출력이 나타난 횟수이다. Figure 3를 참고하면, RSI 분포는 TSI 분포와 비교하여 상대적으로 치우친 분포를 나타내었다. 분포 형태를 정량적으로 살펴보기 위해 왜도(Skewness)를 계산하였고 RSI에서 모바일 기기의 왜도는 1.10, TV는 2.15, 냉장고는 1.18으로 나타났다. TSI에서 모바일 기기 왜도는 0.50, TV는 0.80, 냉장고는 0.41으로 나타났다.

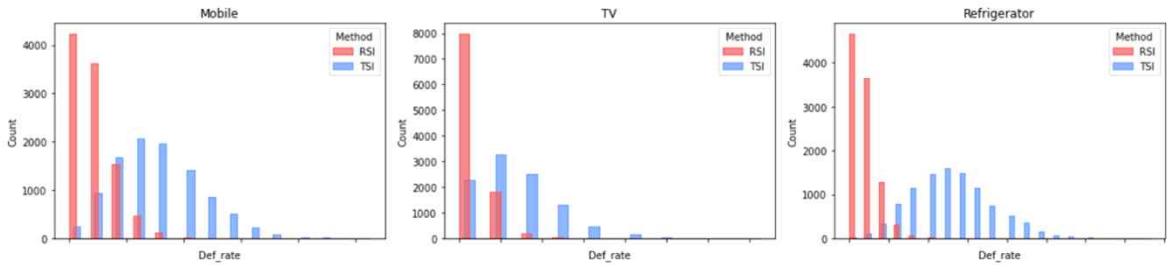


Figure 3 Distribution of defect detection by the products

Figure 2에서는 모든 제품에서 TSI 평균불량검출력이 RSI보다 높게 나타났으며 Figure 3에서는 분포 차이를 확인할 수 있다. 위 사실을 토대로 다음과 같은 가정을 유도하였다.

H_1 : TSI 불량검출력 분포는 RSI 불량검출력 분포와 차이가 있다.

H_{1-1} : 모바일 기기 TSI 불량검출력 분포는 RSI 불량검출력 분포와 차이가 있다.

H_{1-2} : TV TSI 불량검출력 분포는 RSI 불량검출력 분포와 차이가 있다.

H_{1-3} : 냉장고 TSI 불량검출력 분포는 RSI 불량검출력 분포와 차이가 있다.

H_2 : TSI 불량검출력은 RSI 불량검출력보다 우수하다.

H_{2-1} : 모바일 기기 TSI 불량검출력은 RSI 불량검출력보다 우수하다.

H_{2-2} : TV TSI 불량검출력은 RSI 불량검출력보다 우수하다.

H_{2-3} : 냉장고 TSI 불량검출력은 RSI 불량검출력보다 우수하다.

위의 가설을 통계적으로 입증하기 위해 H_1 에는 콜모고로프-스미르노프검정(Kolmogorov-Simrnov test) 실시하였다(Hodges, 1958). H_2 는 RSI 분포가 정규분포를 따르고 있지 않아 단측 맨-휘트니 U 검정(One-sided Mann-Whitney U test)을 수행하였다(Mann and Whitney, 1947). 그 결과 H_1 과 H_2 모두 통계적으로 유의미한 차이(p-value = 0.000)를 보여 총 6개의 가설은 모두 채택되었다. 각 가설에 대한 검정통계량과 p-value는 Table 2에 요약 되어있다.

Table 2. Results of hypothesis test

Statistical test	Hypothesis	Test statistic	p-value	Significant
Kolmogorov-Simrnov test	H_{1-1}	0.6663	0.000	Yes
	H_{1-2}	0.5730	0.000	Yes
	H_{1-3}	0.9136	0.000	Yes
Mann-Whitney U test	H_{2-1}	91070734.5	0.000	Yes
	H_{2-2}	82466664.5	0.000	Yes
	H_{2-3}	98850748.0	0.000	Yes

4. 결론 및 시사점

현대 품질관리 기법의 발달로 양산하는 제품의 각 단위공정별 불량률은 현저하게 감소하게 되었다. 그로 인해 RSI 방식 출하검사는 대부분의 로트를 합격시키고 있는 한계를 보이고 있으나 현장에서는 여전히 RSI를 고수하고 있다. 이에 본 논문은 출하검사 불량검출력을 극대화하기 위한 샘플링 방법론 TSI를 제안하였다. TSI는 제품생산과정에서 축적된 데이터를 예측분석하여 개별 제품의 불량가능성을 제공하는 예측기반 층화추출샘플링 방식이다. TSI가 구현되기 위해서는 제품생산과정에서 발생하는 품질데이터와 시장불량 데이터가 연계 분석되어야 한다. 본 논문에서는 TSI 프레임워크를 활용하여 삼성전자 모바일 기기, TV, 냉장고 제품의 실제 출하검사에 TSI를 적용하였다. 그 결과 RSI보다 더 우수한 불량검출력(모바일 기기: 4.17배 상승, TV: 6.55배 상승, 냉장고: 8.39배 상승)을 보였고 통계적으로도 유의미한 차이를 확인하였다.

본 연구의 시사점으로는 다음과 같다. 첫째 TSI는 제조 과정에서 축적된 데이터를 바탕으로 판매 후 제품의 불량가능성을 예측할 수 있으며, RSI 대비 더 많은 불량을 판매 전에 사전 검출함으로써 시장불량으로 인한 실패비용을 절감할 수 있다. 다른 한편으로는 더 적은 검사 자원으로 기존과 동등이상의 품질 수준을 확보할 수 있음을 시사한다. 둘째, 로트 단위의 불량예측 방법론이 아닌 각 개별 제품의 불량발생 위험도를 실시간으로 예측하는 TSI 프레임워크는 제조부문의 혁신적인 방법론이다. 본 연구결과는 실제 삼성전자 출하검사에 적용하여 품질을 개선한 사례로 TSI 도입을 고려하는 다른 제조업에게 좋은 지침이 될 것이다.

본 연구 시사점에도 불구하고 이번 연구에서는 품질관리 활동의 경제성을 품질비용(Quality cost) 중심으로 살펴 보지 않았다. Feigenbaum(1991)에 따르면 품질비용은 예방비용(Prevention cost), 평가비용(Appraisal cost), 내부 실패비용(Internal failure cost), 외부실패비용(External failure cost)으로 구성되며 각 품질비용 측면에서 TSI 유용성을 제시하는 것은 후속 연구로 남아있다. 다음으로 IPV 선정에 있어 필터방식 변수선택 방법만이 고려된 점을 보완하여 래퍼(Wrapper) 혹은 임베디드(Embedded) 변수선택 방법과 특징추출(Feature extraction) 등이 IPV 선정에 종합적으로 반영될 필요가 있다. 또한 범주형 변수에 대하여 Lee and Uk (2019)이 제안한 연관성 기반 범주형 변수선택 방법도 고려 대상이다. 마지막으로 제조 데이터의 특징인 고차원 불균형 데이터(High-dimensional imbalanced data)를 다루는 IPV 선정 기법을 개발하는 것 또한 흥미로운 연구주제이다.

REFERENCES

- Boetticher, G. D. 2005. Nearest neighbor sampling for better defect prediction. ACM SIGSOFT Software Engineering Notes 30(4):1-6.
- Cheo, Songwoon. 2007. Review of Acceptance Sampling Plans for Parts Per Million Fraction of Defectives. Journal of the Korea Safety Management and Science 9(4):137-142.
- Cheo, Songwoon. 2011. Practical Limitations of Theoretical Acceptance Sampling. Proceedings of the Safety Management and Science Conference. Korea Safety Management and Science, 2011.
- Chong, Hyeran, Bae, Kyoungghan, Lee, Minkoo, Kwon, Hyuckmoo, and Hong, Sunghoon. 2020. Quality strategy for building a smart factory in the fourth industrial revolution. J Korean Soc Qual Manag. 48(1):87-105.
- Efron, B. and Tibshirani, R. J. 1994. An introduction to the bootstrap. CRC press.
- Feigenbaum, A.V. 1991. Total quality control. New York.
- Ferrell Jr, W.G. and Chhoker A. 2002. Design of economically optimal acceptance sampling plans with inspection

- error. *Computers & Operations Research* 29(10): 1283-1300.
- Hodges, J. L. 1958. The significance probability of the Smirnov two-sample test. *Arkiv för Matematik* 3(5):469-486.
- Holte, R. C. 1993. Very simple classification rule perform well on most commonly used datasets. *Machine Learning* 11(1): 63-90.
- Hsu, L. F. and Hsu, J. T. 2012. Economic design of acceptance sampling plans in a two-stage supply chain. *Advances in Decision Sciences*.
- Kang, Pilsung, Kim, Dongil, Lee, Seungkyung, Doh, Seungyong, and Cho, Sungzoon. 2012. Estimating the reliability of virtual metrology predictions in semiconductor manufacturing: A novelty detection-based approach. *Journal of Korean Institute of Industrial Engineers* 38(1):46-56.
- Kim, Minseok. and Baek, Jungeol. 2012. Fail prediction of DRAM module outgoing quality assurance inspection using ensemble learning algorithm. *IE interfaces* 25(2):178-186.
- KS Q ISO 2859-1. 2019. Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection, Korean Agency for Technology and Standards.
- Lee, Changki and Jung, Uk. 2019. Association-based unsupervised feature selection for high-dimensional categorical data. *J Korean Soc Qual Manag.* 47(3):537-552.
- Liu, F. T., Ting, K.M., and Zhou, Z.H. 2008. Isolation forest. 2008 eighth IEEE international conference on data mining. IEEE.
- Mann, H. B. and Whitney, D. R. 1947. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 50-60.
- Meinshausen, Nicolai, and Peter B. 2010. Stability selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 72(4):417-473.
- Song, Hojun, Kim, Mingyu, and Shin, Wanseon. 2021. A Study on the relative importance of quality scorecard for open quality. *J Korean Soc Qual Manag.* 49(1):61-79.
- Sunter, S., Jurga, K., Dingene, P., and Vanhooren, R. 2014. Practical random sampling of potential defects for analog fault simulation. 2014 International Test Conference. IEEE.
- Yang, Inmo and Hwang, Inkeuk. 2018. *Statistical Quality Control Statistical Process Control*. Minyoungsa.

저자소개

- 김준세** 산업공학을 전공하고, 2009년부터 삼성전자 Global CS센터 재직 중이다. 전사 부품/공정/출하품질 업무를 9년간 담당해왔고 2018년부터 품질데이터 분석을 담당하고 있다. 주요 연구 관심분야는 머신러닝, OR, 시뮬레이션 등이다.
- 이창기** 현재 삼성전자 Global CS 센터에 재직 중이며 동국대학교에서 경영학 박사학위를, 동대학교에서 경영학과 회계학 학사 학위를 취득하였다. 주요 연구관심분야는 통계적 품질관리, 기계학습, 품질경영, 범주형 자료분석 등이다.
- 김경남** 공장자동화를 전공하였으며 삼성전자 무선사업부 출하품질업무를 거쳐 현재 Global CS센터에 재직 중이다. 주요 연구 관심분야는 통계적 공정관리, 이상탐지 등이다.
- 김창우** 전자공학을 전공하였으며 삼성전자 Global CS센터에서 제조 시스템 운영, 출하품질 개선업무를 담당하고 있다. 주요 연구 관심분야는 데이터 전처리, 머신러닝 등이다.

- 송혜미** 학부에서 수학을 전공하였으며 이후 삼성전자 Global CS센터에서 서비스팀을 거쳐 품질 팀에 재직 중이다. ISO9001과 CQE를 보유중이며 주요 관심분야는 통계적공정관리, 머신러닝, 딥러닝 등이다.
- 안성수** 학부에서 전기전자공학을 전공했으며 이후 삼성전자 H/W개발을 거쳐 현재 삼성전자 Global CS센터에 재직 중이다. 주요 관심분야는 품질경영, 머신러닝, 이상탐지 등이다.
- 오재원** 학부에서 기계공학을 전공하고 LG 디스플레이에서 TV 설계 업무를 거쳐 현재 삼성전자 Global CS 센터에서 TV 검출력혁신 업무를 담당하고 있다. 주요 관심분야는 통계적품질관리, 이상탐지이다.
- 조현상** 학부에서 산업공학을 전공하고, 미국 Indiana University (Kelly School of Business)에서 MBA학위를 취득하였다. 삼성전자 SET부문 품질보증체계 전반에 걸친 업무를 담당하는 Staff부서인 Global CS센터에서 근무중에 있으며 주요 연구 관심분야는 머신러닝과 AI 및 Simulation 등이다.
- 한상섭** 제어계측공학을 전공하고 경부고속전철 전자연동장치 등 국책연구과제를 수행하였으며 현재 삼성전자 Global CS센터 품질보증그룹장을 맡아 SET부문 품질혁신업무를 담당하고 있다. 고속신호제어 및 Fail Safe 로직 설계 전문가로 주요 연구 관심분야는 신뢰성, SW Test 플랫폼 개발 및 머신러닝 등이다.