

정부 품질보증활동 데이터 활용을 위한 Zero-Inflated 포아송 분포 적용

김지훈[†] · 이창우

국방기술품질원

Application of Zero-Inflated Poisson Distribution to Utilize Government Quality Assurance Activity Data

Kim, JH[†] · Lee, CW

Defense Agency for Technology and Quality

ABSTRACT

Purpose: The purpose of this study was to propose more accurate mathematical model which can represent result of government quality assurance activity, especially corrective action and flaw.

Methods: The collected data during government quality assurance activity was represented through histogram. To find out which distributions (Poisson distribution, Zero-Inflated Poisson distribution) could represent the histogram better, this study applied Pearson's correlation coefficient.

Results: The result of this study is as follows; Histogram of corrective action during past 3 years and Zero-Inflated Poisson distribution had strong relationship that their correlation coefficients was over 0.94. Flaw data could not re-parameterize to Zero-Inflated Poisson distribution because its frequency of flaw occurrence was too small. However, histogram of flaw data during past 3 years and Poisson distribution showed strong relationship that their correlation coefficients was 0.99.

Conclusion: Zero-Inflated Poisson distribution represented better than Poisson distribution to demonstrate corrective action histogram. However, in the case of flaw data histogram, Poisson distribution was more accurate than Zero-Inflated Poisson distribution.

Key Words : Quality Performance Index(QPI), Poisson Distribution, Zero-Inflated Poisson Distribution

● Received 30 May 2018, 1st revised 10 July, accepted 30 July 2018

† Corresponding Author(nobless@dtq.re.kr)

© 2018, The 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

급변하는 국내외 안보환경 속에서 우리나라의 자주국방을 위한 국방 예산은 꾸준히 증가하고 있다. 국방기술품질원에서 품질보증활동 하는 양산 및 운용유지 단계의 군수품에 대한 계약금액은 2011년 약 5조 9,000억에서 2017년 약 8조 7,600억으로 150% 정도 증가하였다. 또한 군수품에 대한 고객들의 눈높이가 높아지면서, 단순히 요구 조건을 만족하는 제품을 납품하는 것보다 계약 시 고려되지 않은 무형의 필요조건까지 만족하는 제품을 납품하도록 주변 환경이 요구하고 있다.

변화하는 환경에 맞춰 정부 품질보증 담당원은 계약업체의 정보 및 제품품질 이력 데이터를 활용하여 품질보증활동을 수행해야 하는 필요성이 대두되고 있다. 이를 위해 품질보증 담당원의 정성적인 업체 평가보다 과거에 이루어진 품질보증활동 데이터를 기반으로 하는 체계적이며 정량적인 업체 성과 평가 기준이 필요하게 되었다.

국방기술품질원에서는 이러한 상황을 지원하기 위하여 제품에 대한 단순 평가로 끝나는 것이 아닌 업체에 대한 종합적인 평가를 통해 제품의 수준과 고객의 만족도를 향상시킬 수 있는 기준을 수립할 수 있도록 업체 품질수준 평가 체계(QPI; Quality Performance Index)를 제안하였다. 더불어 품질보증 담당원은 QPI를 활용하여 납품 후에도 업체의 군수품 양산 단계의 프로세스 효율화 및 효과성에 대한 지속적 점검 및 개선을 위한 지표로써 활용할 수 있게 된다.

QPI는 업체의 품질경영시스템, 제품품질 수준, 고객만족도, 계약이행 충실도를 평가지표로 정의하고 4개 평가지표를 종합하여 업체에 대한 최종 점수를 산정한다. 품질경영시스템은 국방품질경영체계 이행 정도를 평가하며, 제품품질 수준은 업체의 시정조치현황 및 하자발생현황, 적기납품률에 대한 적합한 수확 모델을 활용하여 평가한다. 고객만족도는 운용성 향상 노력도와 사용자불만 적기조치율을 평가하며, 계약이행 충실도는 업체 품질보증계획의 적절성, 문서 품질 및 업무 협조성, 지속적 품질개선 노력을 평가한다.

QPI의 기존 연구에서 제품품질 수준의 시정조치현황 및 하자발생현황은 시정조치와 하자발생을 단위 시간 당 발생하는 사건으로 정의하고 포아송 분포를 따르는 것으로 가정하였다. 하지만 국방기술품질원 품질정보서비스에 입력된 시정조치현황과 하자발생현황 데이터를 정리한 결과 사건이 0회 발생할 확률이 포아송 분포의 $P(X=0)$ 값에 비해 훨씬 큰 값을 가지는 것을 확인하였다.

일반적으로 신뢰성 있는 제조 공정의 불량품 발생률은 포아송 분포를 따르는 것으로 알려져 있지만 제조 공정의 실제 불량품 발생률이 포아송 분포로 예측한 불량품 발생률보다 훨씬 낮을 때, 불량품 발생률은 포아송 분포를 따른다고 할 수 없어, 이를 보완하기 위하여 Zero-Inflated Poisson (ZIP) 분포가 제안되었다. 이 외에도 다양한 산업 분야에서 포아송 분포를 따른다고 알려진 모델들에 대해 ZIP 분포를 활용하여 데이터의 분포를 분석한 연구 결과들이 있다.

따라서 본 논문은 제안된 QPI의 평가지표 중 제품품질 수준의 시정조치현황과 하자발생현황의 수확 모델을 포아송 분포를 따르지만 데이터의 $P(X=0)$ 값이 포아송 분포보다 클 때 활용할 수 있는 ZIP 분포로 제안한다. 데이터의 히스토그램과 두 확률분포의 확률질량함수(PMF; Probability Mass Function)에 대해 Pearson 상관계수(correlation coefficient)를 이용하여 상관관계를 비교하였으며, 분석을 위한 데이터는 국방기술품질원 유도전자4팀에서 담당하는 31개 업체의 데이터를 이용하였다. 본 연구를 통해 업체 시정조치현황과 하자발생현황 분석 시 기존보다 정확한 분석이 가능할 것으로 기대된다.

논문의 구성은 2장에서 업체 품질 수준평가 체계 중 제품품질 수준에 대해 간략히 정리하고, 3장에서 분석 데이터에 대한 특성과 ZIP 분포에 대해 설명한다. 4장에서 분석 데이터와 포아송 분포, ZIP 분포 간의 Pearson 상관계수를

이용한 상관관계 비교를 통해 결과에 대해 확인하고 마지막 5장에서 결론을 맺는다.

2. 업체 품질수준 평가 체계

2.1 업체 품질수준 평가 체계 제안

국방 분야에서는 품질수준 향상을 위해 80년대 초부터 품질경영시스템의 제도 정착을 유도하여 왔으며, 1993년 ISO 9000의 모태가 되는 MIL-Q-9858A, MIL-I-45208A, MIL-STD-45662를 한국화한 품질시스템 규격인 국방 0050-0015(연), 국방 0050-0029(연), 국방 0050-0030(연)을 제정하여 품보형태 4, 5형 품목에 대해 절차평가(Procedure Evaluation)를 실시하여 왔다.

1996년 MIL-A-9858A와 MIL-I-45208A가 폐지되고 ISO 9000을 적용하도록 하는 등 변화에 따라 기존의 군수품 품질보증체계를 우리의 실정에 적합한 품질보증체제로 개선해야 했으며, 이를 위해서 국방기술품질원(당시 국방품질연구소)에서는 군수업체의 품질보증능력을 평가할 수 있는 모델을 1993년도부터 개발했다. 이 모델에 의해 1996년 시범적으로 업체 평가를 실시하고, 그 결과를 분석하여 향후 군수품 품질보증체계 및 업체 평가 방법의 개선을 모색하였다.

이후 국방기술품질원은 군수품 품질향상과 국방비용절감을 통한 고객만족 달성을 위하여 총 수명 주기 균형적 품질경영(BQM; Balanced Quality Management)을 수행하며, 이를 위한 세부 실행 계획의 일환으로 2011년 업체 품질수준 평가 체계(QPI; Quality Performance Index)를 구축 및 제안하였다.

QPI의 목적은 업체 품질수준을 평가 및 진단하여 업체의 자율적 품질보증 기반 확립을 위한 솔루션을 제공하고 평가서를 업체와 관련기관에 연간 단위로 제시하여 군수품 양산 단계의 프로세스 효율화와 프로세스 효과성에 대한 지속적인 점검 및 개선을 돕는 데에 있다. 또한 정부 품질보증 담당원이 품질보증활동을 위해 위험 식별 및 품질보증 활동 계획 수립 시 QPI를 활용하면 정부 품질보증 담당원이 변경되더라도 동일한 업체 평가 결과를 얻을 수 있어 일관성있는 품질보증활동이 가능하게 된다.

QPI는 업체의 품질경영시스템, 제품품질 수준, 고객만족도, 계약이행 충실도를 평가지표로 정의하고 있으며, 4개의 평가지표를 종합하여 업체에 대한 최종 점수를 표 1과 같이 산정한다. 본 논문의 2장에서는 QPI의 4가지 평가지표 중 연구 내용과 직접 연관되는 제품품질 수준에 대해 간략히 정리한다.

Table 1. Example of total score of QPI

구분	등급	평점비율
평가 종합 결과	Very Good	85.9
1. 품질경영시스템 이행능력	B	84
2. 제품품질 수준	A	80.6
3. 고객 만족도	B	88
4. 계약이행 충실도	A	91

2.2 QPI 제품품질 수준

QPI의 평가지표 중 하나인 제품품질 수준은 업체가 생산한 제품의 품질 수준을 나타내는 지표로 시정조치현황, 하자발생현황, 적기납품률로 분류한다. 제품품질 수준의 평점은 세부 3항목에 대한 가중평균값이며, 3항목의 평점은 평가 당해 연도의 시정조치 수, 하자발생 수, 적기납품률을 각각의 특성에 맞게 처리한 값이 과거 3년 동안의 평균값과 비교하고 백분율 점수로 환산하여 부여한다.

2.2.1 시정조치현황

시정조치(CA; Corrective Action)는 정부 품질보증 과정 중 발견되는 불합리한 사항 및 규정위반 사항 또는 계약 조건에 위배되는 사항 등에 대하여 계약업체에 시정 및 재발방지대책을 요구하는 내용이다. QPI 제품품질 수준에서의 시정조치현황은 정부 품질보증 담당원이 평가 당해 연도 업체에 발행한 시정조치 횟수를 의미한다.

시정조치는 일반적으로 하나의 시정조치가 다른 시정조치에 영향을 미치지 않기 때문에 두 시정조치는 서로 독립이라고 가정할 수 있다. 시정조치현황을 수학적으로 표현하기에 적합한 모델은 이산확률분포의 포아송 분포이다. 포아송 분포는 사건의 평균 발생률(λ)이 알려져 있고, 각각의 사건(X)이 독립적일 때, 단위 시간당 발생하는 사건의 수를 식 (1)의 확률로 표현하는 이산확률분포이다.

$$P(X=x) = \frac{\lambda^x e^{-\lambda}}{x!}, (\lambda > 0) \quad (1)$$

QPI 제품품질 수준에서 시정조치현황을 포아송 분포로 표현하기 위해 필요한 파라미터는 사건의 평균 발생률이다. QPI 평가대상이 되는 업체들의 과거 3년 동안 시정조치 수의 평균값을 사건의 평균 발생률(λ)로 하였으며, 식으로 표현하면 식 (2)와 같다.

$$\begin{aligned} \lambda_{CA} &= \frac{\text{평가 대상 업체들의 3년 동안 시정조치 수}}{\text{평가 대상 업체 총 수}} \\ &= \frac{1}{n} \sum_{i=1}^n \lambda_i \end{aligned} \quad (2)$$

식 (2)의 λ_{CA} 는 평가년도 이전 3년 동안의 시정조치 평균 발생률이며, n 은 평가 대상이 되는 업체의 총 수, λ_i 는 i 번째 업체의 3년 동안 발생한 시정조치 수를 의미한다. 식 (2)를 이용하여 3년간 시정조치 평균발생률을 구한 후, 업체 시정조치현황 점수는 평가 당해 연도에 발생한 시정조치 수를 파라미터로 하여 식 (1)로 확률 값을 계산한다. 시정조치는 수가 적을수록 업체의 품질관리 수준이 좋다는 것을 의미하기 때문에 1에서 평가연도의 시정조치 수에 해당하는 누적확률분포(CDF; Cumulative Distribution Function)를 뺀 값에 100을 곱하여 식 (3)과 같이 산정한다.

$$S_{CA} = 100 \times \left(1 - \sum_{j=0}^k \frac{\lambda_{CA}^j e^{-\lambda_{CA}}}{j!} \right) \quad (3)$$

S_{CA} 는 평가 대상 업체의 시정조치현황 점수를 나타내며, k 는 평가 당해 연도에 받은 시정조치 수가 된다. 예를 들어, 평가대상 업체들이 3년 동안 평균 5개의 시정조치를 받고, A 업체가 평가 당해 연도에 2개의 시정조치를 받았다면, A 업체의 시정조치현황 점수는 식 (3)에 의해 87.53점을 얻게 된다.

2.2.2 하자발생현황

하자는 계약품질 요구조건과 일치하지 않는 계약업체의 잘못에 의해 발생한 제품의 결함 사항, 수량 부족, 손상,

이종품, 기능/성능 미 충족 등을 가리킨다. QPI 제품품질 수준에서의 하자발생현황은 정부 완성품 감사 후 납품된 제품에 대해 고객이 만족스럽지 않다고 판단하면 사용자불만을 보내게 되고, 원인 검토 결과 하자로 판단된 사항을 의미한다.

하자 역시 시정조치와 마찬가지로 하나의 하자와 다른 하자는 서로 영향을 미치지 않기 때문에 서로 다른 하자는 서로 독립이라고 가정할 수 있다. 따라서 하자현황의 수학적 모델로 이산확률분포의 포아송 분포를 선택하였다.

포아송 분포를 이용하기 위해 필요한 파라미터인 사건의 평균 발생률은 QPI 평가 대상 업체들의 3년 동안 발생한 하자 수의 평균값을 사건의 평균 발생률(λ)로 하였으며, 식으로 표현하면 식 (4)와 같다.

$$\begin{aligned}\lambda_f &= \frac{\text{평가대상업체들의 3년 동안 하자수}}{\text{평가대상업체총수}} \\ &= \frac{1}{n} \sum_{i=1}^n \lambda_i\end{aligned}\quad (4)$$

식 (4)의 λ_f 는 평가년도 이전 3년 동안 하자의 평균 발생률이며, n 은 평가대상이 되는 업체의 총 수, λ_i 는 i 번째 업체의 3년 동안 발생한 하자 수를 의미한다. 식 (4)를 이용하여 3년 동안의 평균 하자발생 수를 구한 후, 업체 하자발생현황에 대한 점수는 평가 당해 연도에 발생한 하자 수를 파라미터로 하여 계산한다. 하자 역시 발생한 수가 적을수록 업체의 품질관리 수준이 좋다는 것을 의미하기 때문에, 1에서 평가연도의 하자발생 수에 해당하는 CDF를 뺀 값에 100을 곱하여 식 (5)와 같이 산정한다.

$$S_f = 100 \times \left(1 - \sum_{j=0}^k \frac{\lambda_f^j e^{-\lambda_f}}{j!} \right) \quad (5)$$

S_f 는 평가 대상 업체의 하자발생현황 점수를 나타내며, k 는 평가 당해 연도에 받은 하자발생 수가 된다. 예를 들어, 평가대상 업체들의 3년 동안 하자발생 이력이 3개이고, 평가 당해 연도에 A 업체가 1개의 하자가 발생했다면, A업체의 하자발생현황 점수는 식 (5)에 의해 80.09점을 받게 된다.

2.2.3 적기납품률

적기납품률이란 평가 당해 연도에 납품하기로 계약된 제품들이 적기에 납품되었는가를 나타내는 지표이다. 전체 계약품목 수와 납품품목 수를 기준으로 점수를 부여하며, 이를 위해 적기납품률의 수학적 모델로 지수분포를 선택하였다. 본 논문에서는 시정조치현황과 하자발생현황의 데이터를 분석하였기 때문에, 적기납품률에 대해서는 간략히 기술하였다.

지수분포는 일정한 비율로 발생하는 독립적인 두 사건 간의 시간 간격을 나타내는 확률분포이다. 3년 동안 전체 계약 품목 수와 적기 납품 품목 수의 비율을 계산하여 지수분포의 확률값으로 삼고, 평가 당해 연도의 적기 납품 품목 수를 이용하여 적기납품률의 평점을 계산하게 된다.

QPI 제품품질 수준의 평점은 위에서 기술한 시정조치현황 및 하자발생현황, 적기납품률의 평점들에 대한 가중평균값으로 산출할 수 있다.

3. 분석대상 및 방법

3.1 분석대상

본 논문은 2012년부터 2017년의 기간 동안 국방기술품질원의 품질보증체계에 기록된 유도전자4팀 담당의 31개 업체에 대해 시정조치 발생 횟수, 하자 발생 횟수를 토대로 분석 대상을 설정하였다.

3.1.1 시정조치현황 데이터

표 2는 2012년부터 2017년의 기간 동안 31개 업체의 시정조치 요구 횟수를 나타낸다. 2016년도에 59회의 시정조치를 받은 업체가 있으나, 이는 예외적인 상황이며, 대체로 시정조치 발생 현황은 0회에서 17회 정도임을 알 수 있다.

Table 2. Observed Corrective Action Data for 6 past years

구분	2012년	2013년	2014년	2015년	2016년	2017년
0회	15	19	18	16	13	19
1회	4	4	3	5	6	4
2회	4	1	2	1	2	3
3회	2	1	2	3	0	1
4회	3	0	1	2	3	0
5회	0	0	1	1	1	0
6회	2	2	1	1	2	0
7회	0	1	0	1	1	1
8회	1	1	0	1	0	0
9회	0	1	1	0	0	1
10회	0	0	1	0	1	0
11회	0	0	0	0	0	1
12회	0	0	0	0	0	0
13회	0	0	0	0	0	0
14회	0	0	0	0	0	1
15회	0	0	0	0	0	0
16회	0	1	0	0	1	0
17회	0	0	1	0	0	0
18회	0	0	0	0	0	0
∴	∴	∴	∴	∴	∴	∴
59회	0	0	0	0	1	0

그림 1은 2015년부터 2017년 시정조치현황 히스토그램과 각 해의 시정조치 발생 평균값(λ_{CA})을 파라미터로 하는 포아송 분포를 나타내었다. 시정조치 발생 평균값은 평가 당해 연도 이전 3년 동안의 시정조치 발생 평균값을 의미하며, 2015년의 경우 2012년부터 2014년까지 시정조치 발생 횟수 평균값이 2015년 평가를 위한 λ_{CA} 가 된다. 2015년 시정조치 발생 평균값은 1.84, 2016년은 1.81, 2017년은 2.61이었으며, 3개년도 시정조치현황 히스토그램의 $P(X=0)$ 는 모두 QPI의 시정조치현황 수학 모델로 가정한 포아송 분포의 값에 비해 크다는 것을 확인할 수 있다.

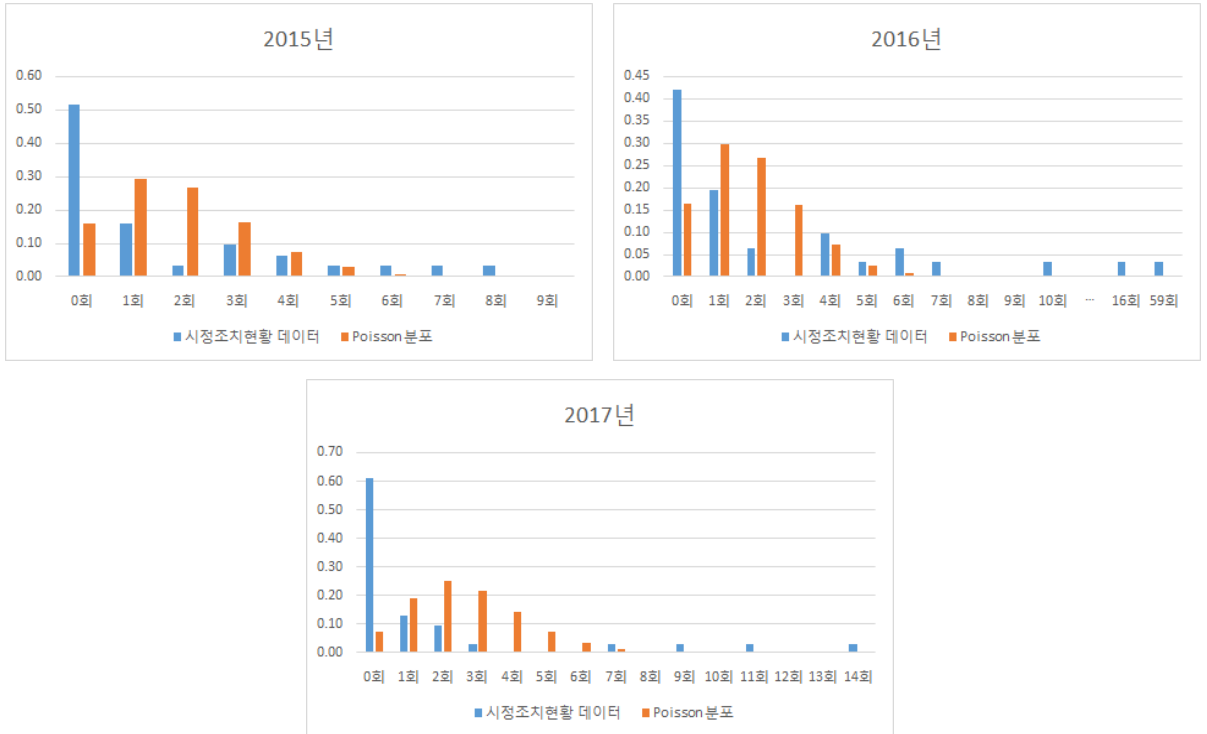


Figure 1. Histogram of Corrective Action and Poisson Distribution in 2015, 2016, 2017

3.1.2 하자발생현황 데이터

표 3은 2012년부터 2017년까지 31개 업체에 발생한 하자의 횟수를 나타낸다. 하자는 한 해에 발생 횟수가 3회 이하로 발생 빈도수가 낮았으며, 하자가 전혀 발생하지 않은 업체가 대부분임을 표 3으로부터 확인할 수 있다.

Table 3. Observed Flaw Data for 6 past years

구분	2012년	2013년	2014년	2015년	2016년	2017년
0회	28	28	29	27	28	26
1회	3	3	1	3	1	5
2회	0	0	1	0	2	0
3회	0	0	0	1	0	0

그림 2는 2015년부터 2017년의 하자발생현황 히스토그램과 포아송 분포를 나타내었다. 포아송 분포의 파라미터로 각 해의 하자발생 평균값(λ)은 평가 당해 연도 이전 3년 동안 하자발생 횟수의 평균값을 사용하였으며, 2015년의 경우 2012년부터 2014년까지의 하자발생 횟수의 평균값이 된다. 2015년 하자발생 평균값은 0.10, 2016년은 0.13, 2017년은 0.15였다. 하자발생현황의 $P(X=0)$ 값은 시정조치현황과는 다르게 포아송 분포의 $P(X=0)$ 값과 거의 차이가 나지 않는 것을 확인할 수 있다.

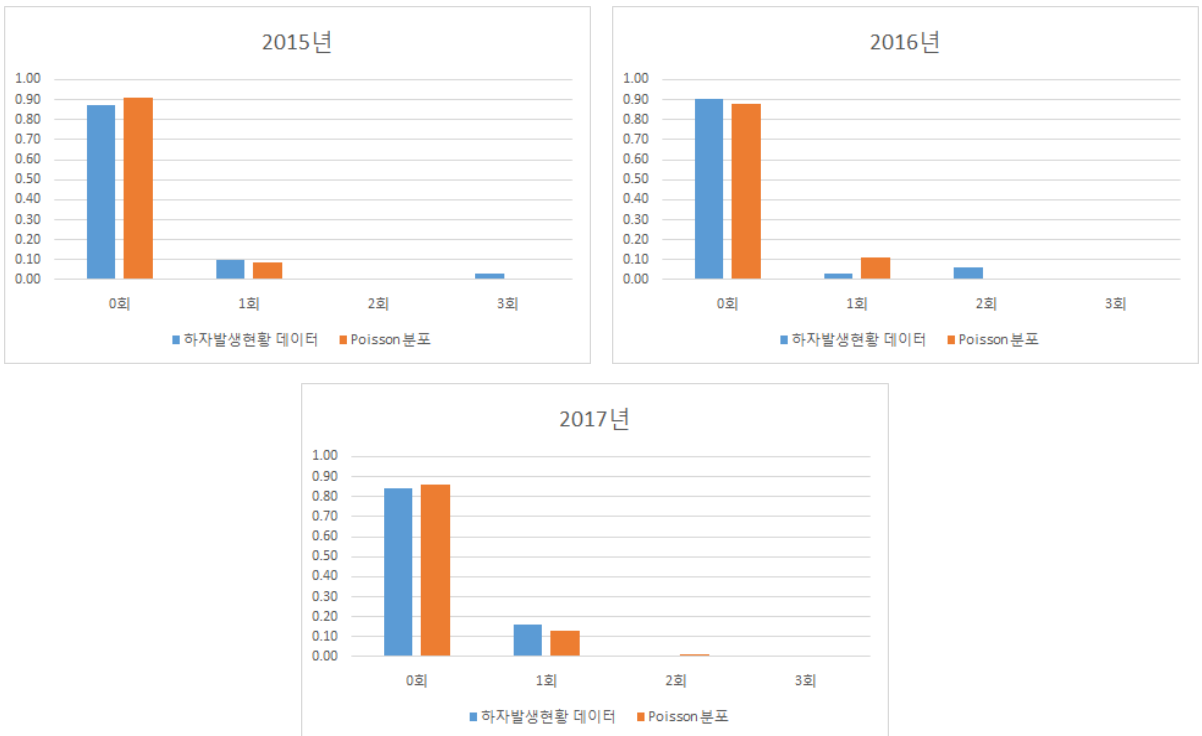


Figure 2. Histogram of Flaw and Poisson Distribution in 2015, 2016, 2017

3.2 Zero-Inflated 포아송 분포

Zero-Inflated Poisson (ZIP) 분포는 포아송 분포를 따르는 어느 한 표본의 $P(X=0)$ 값이 포아송 분포의 $P(X=0)$ 값보다 매우 클 때, 나타나는 포아송 분포의 특수한 경우이다. 일반적인 통계학에서 신뢰성 있는 제조 공정의 불량률은 포아송 분포를 따른다고 하지만 제조 공정이 포아송 분포로 예측한 불량률보다 실제로 더 낮은 불량률을 가지는 경우 그 공정은 명확히 포아송 분포를 따른다고 할 수 없기 때문에, 이러한 문제를 해결하기 위하여 1992년 Diane Lambert에 의해 ZIP 분포가 제안되었다. 이 외에도 치의학 분야에서 아이들의 씹거나, 빠지고 새로 난 이에 대한 역학 분석을 위해, 식품 산업 분야에서 샘플링 계획 수립 시 활용을 위해 ZIP 분포가 활용되기도 하였다.

ZIP 분포의 PMF는 포아송 분포의 PMF를 $j = 0$ 일 때와 $j = 1, 2, 3, \dots$ 일 때로 나누어 식 (6)과 같이 정리된다.

$$\begin{aligned}
 P(X=0) &= w + (1-w)e^{-\lambda} \\
 P(X=j) &= \frac{(1-w)e^{-\lambda}\lambda^j}{j!}, j = 1, 2, 3, \dots
 \end{aligned}
 \tag{6}$$

그림 3과 그림 4에서는 3.1 분석대상의 2015년부터 2017년까지 3개년도 시정조치현황 평균값과 하자발생 평균값을 파라미터 λ 로 하여 ZIP 분포와 시정조치현황 히스토그램 및 하자발생현황 히스토그램을 함께 나타내었다.

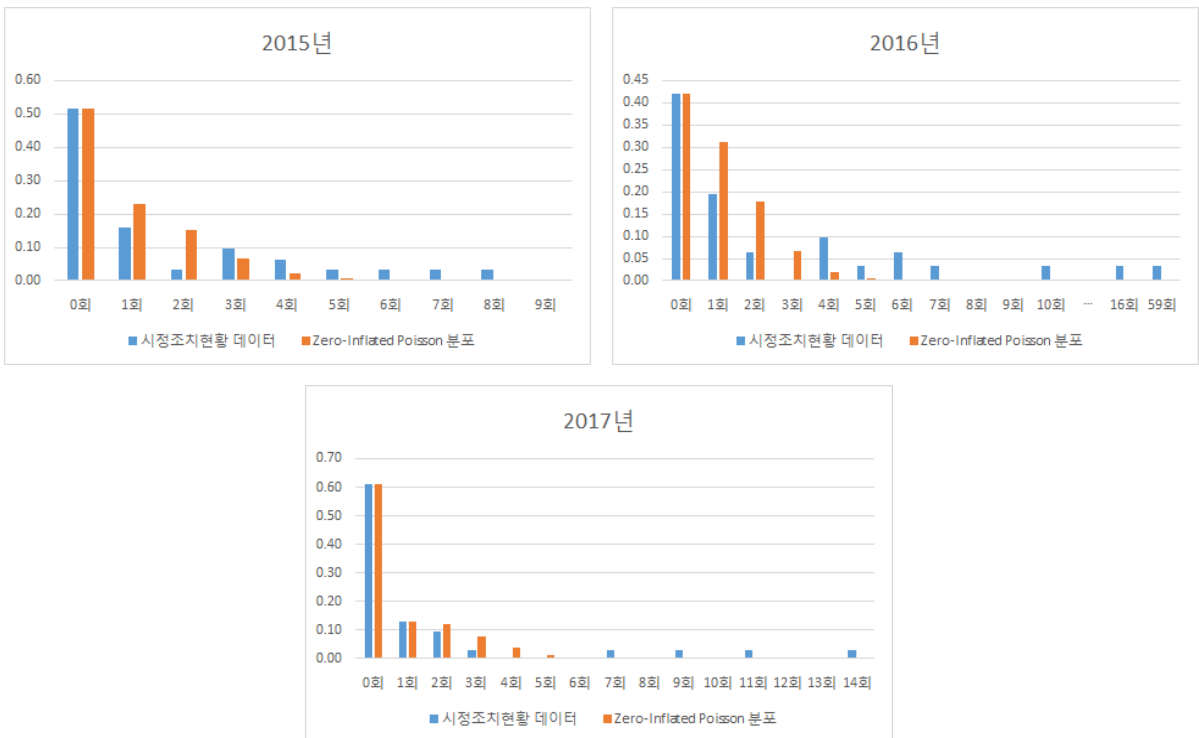


Figure 3. Histogram of Corrective Action and Zero-Inflated Poisson Distribution in 2015, 2016, 2017

그림 3으로부터 2015년과 2017년의 시정조치현황 히스토그램과 ZIP 분포의 히스토그램을 비교하면 두 그래프가 크게 다르지 않음을 확인할 수 있다.

하자발생현황은 그림 4에서 확인할 수 있듯이 2015년과 2017년에 ZIP 분포에 확률 값이 음수가 나옴을 확인할 수 있다. ZIP 분포 역시 포아송 분포를 따르기 때문에, 포아송 분포의 정의에 따라 ZIP 분포의 평균값은 0보다 큰

값을 가져야 한다. 데이터로부터 ZIP 분포의 평균값이 0보다 크려면 식 (7)을 만족하여야 하며, 식 (7)에서 λ 는 데이터의 평균값을 의미한다.

$$\lambda > 1 - P(X=0) \tag{7}$$

2015년의 $P(X=0)$ 은 0.8710이며, λ 는 0.0968이므로 식 (7)을 만족하지 않으며, 2016년은 $P(X=0)=0.9032$, $\lambda=0.1290$ 으로 만족, 2017년은 $P(X=0)=0.8387$ 은 $\lambda=0.1505$ 로 만족하지 않음을 확인할 수 있다. 따라서 식 (7)을 만족하지 않는 해인 2015년과 2017년은 그림 4에서 볼 수 있듯이 ZIP 분포에 음의 값을 가지는 확률 값이 존재한다.

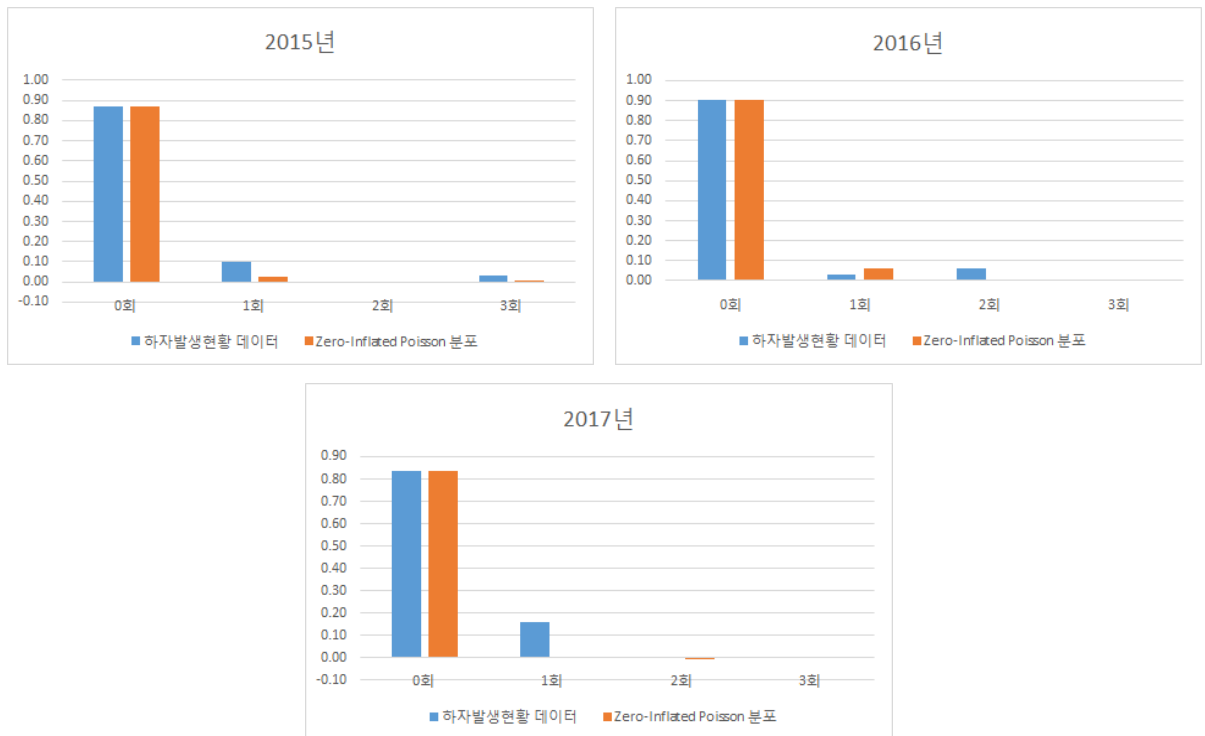


Figure 4. Histogram of Flaw and Zero-Inflated Poisson Distribution in 2015, 2016, 2017

4. 분석결과

4.1 시정조치현황 분석결과

그림 5는 2015년부터 2017년까지 3년 동안의 데이터를 scatterplot으로 나타낸 그림들로 X축은 시정조치현황의 확률 값을, Y축은 시정조치현황의 확률 값에 대응하는 포아송 분포(파란색 점)와 ZIP 분포(주황색 점)의 확률 값을 나타낸다. 그림 안의 검은 선은 데이터들의 분산 정도를 비교하기 위하여 추가하였다.

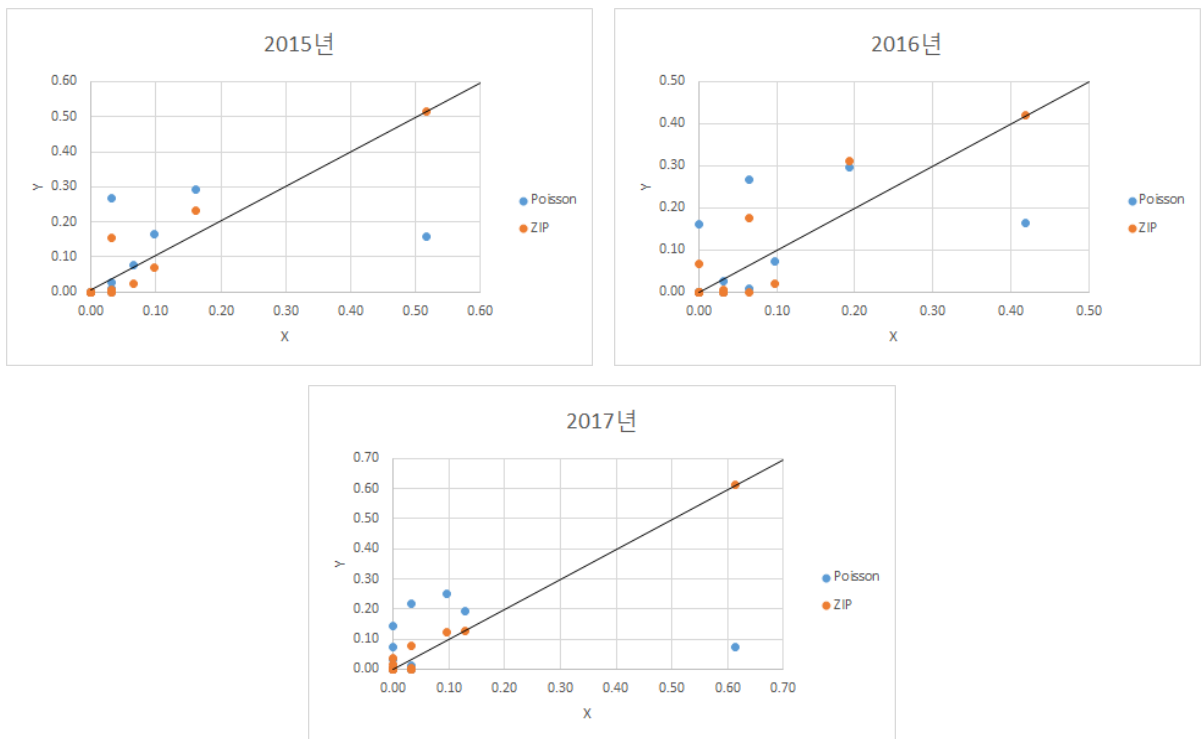


Figure 5. Scatterplot of CA data and Poisson, Zero-Inflated Poisson(ZIP) Distribution in 2015, 2016, 2017

먼저 시정조치현황 데이터와 포아송 분포의 Pearson 상관계수 값은 2015년부터 2017년까지 차례로 0.50, 0.56, 0.19이었으며, 시정조치현황 데이터와 ZIP 분포의 상관계수 값은 각각 0.96, 0.94, 0.99이었다. 상관계수 값이 1에 가까울수록 두 그룹 간의 상관관계가 크기 때문에, 시정조치현황 데이터와 ZIP 분포가 상관관계가 크다고 할 수 있다. 따라서 ZIP 분포가 포아송 분포보다 시정조치현황을 잘 나타내는 모델이라 할 수 있다.

4.2 하자발생현황 분석결과

하자발생현황 데이터와 포아송 분포의 Pearson 상관계수 값은 2015년부터 2017년까지 모두 0.99로 상관관계가 매우 크다는 것을 확인할 수 있었으며, 그림 6과 같이 scatterplot으로 나타내었을 때도 데이터의 분산이 적게 된 것을 확인할 수 있다.

반면 ZIP 분포는 3년 동안 하자발생 횟수 평균값이 충분히 크지 않아 하자발생현황 데이터에 적용할 수 없음이 확인되었다. 실제 그림 4를 확인하였을 때, 2015년과 2017년의 확률 값에 음수가 있음을 확인할 수 있으며, 이는 확률로서 활용할 수 없음을 의미한다.

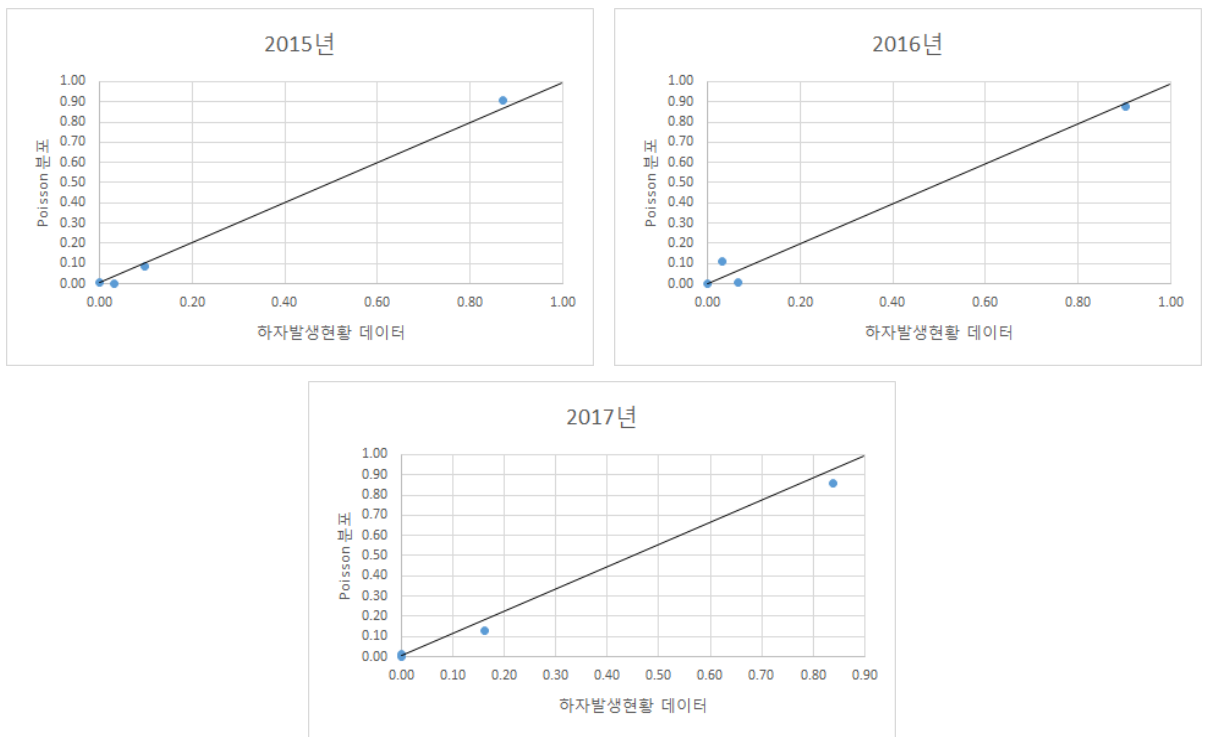


Figure 6. Scatterplot of Flaw data and Poisson Distribution in 2015, 2016, 2017

따라서 ZIP 분포보다 포아송 분포를 하자발생현황의 수학적 모델로 이용하는 것이 타당하다고 할 수 있다.

5. 결 론

본 연구는 국방기술품질원에서 정부 품질보증 업무를 수행하며 발생하는 시정조치현황과 하자발생현황이 어떠한 수학적 모델을 따르는지 상관계수를 통해 상관관계를 확인해보았다. 기존의 연구에서 시정조치현황과 하자발생현황을 포아송 분포를 따르는 것으로 가정하고 제품품질 수준 점수를 계산하여 QPI에 활용하고자 하였다.

국방기술품질원 품질정보서비스에 입력된 시정조치현황과 하자발생현황의 데이터들을 분석한 결과 시정조치현황 데이터의 경우 업체가 받은 시정조치 횟수가 0일 경우, 포아송 분포의 $P(X=0)$ 값에 비해 큰 값을 가지는 것을 확인하였다. 따라서 시정조치현황을 수학적 모델로 보다 정확히 표현하기 위하여 포아송 분포가 아닌 ZIP 분포를 제안하게 되었다.

2012년부터 2017년까지 6년 동안 국방기술품질원 유도전자4팀 담당 31개 업체의 시정조치현황 데이터를 정리하여 2015년부터 2017년까지의 데이터를 히스토그램으로 나타내고 포아송 분포 및 ZIP 분포의 히스토그램과 상관관계 확인을 위해 상관계수를 계산한 결과 포아송 분포는 최고 0.56 이었으며, ZIP 분포는 최하 0.94 였다. 따라서 시정조치현황 데이터는 포아송 분포가 아닌 ZIP 분포와 상관관계가 매우 높다는 것을 확인할 수 있다.

하자발생현황은 하자발생 횟수 평균값이 충분히 크지 않아, ZIP 분포를 하자발생현황에 적용할 수 없음을 확인하였으며, 히스토그램으로 나타낸 2015년부터 2017년까지의 하자발생현황 데이터는 기존 연구에서 제안한 포아송 분포의 히스토그램과 상관계수 값이 0.99로 상관관계가 매우 높음을 확인하였다.

실제 정부 품질보증활동하며 획득한 데이터들과 상관관계가 높은 수학적 모델을 사용하는 것이 QPI 제품품질 수준 평가 시 더 정확하게 평가할 수 있기 때문에, 시정조치현황은 본 연구에서 제안한 ZIP분포를, 하자발생현황은 기존 연구에서 제안한 포아송 분포를 적용하는 것이 타당하다고 할 수 있다.

본 연구의 결과는 정부 품질보증 담당원이 품질보증활동하며 확보한 시정조치현황과 하자발생현황 데이터를 활용하여 QPI의 평가지표 중 하나인 제품품질 수준을 보다 정확한 수학적 모델에 근거하여 평가할 수 있도록 도움을 준다. 또한 QPI 평가 결과를 업체에 지속적으로 제공하여 업체의 품질관리에 도움을 주며, 궁극적으로 업체의 자율적인 품질보증활동이 가능하도록 유도하고, 더 나아가 주 고객인 군의 제품에 대한 만족도를 높이는 데 유용한 자료로 활용할 수 있을 것으로 기대한다.

REFERENCES

- ChangWoo, Lee., and JungHo, Kim. 2011. "The Report on the Research of the Quality Performances Index System(QPI)." DtaQ-11-2828-R.
- Dankmar Bohning, and Ekkehart Dietz., 1999. "The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology." *Journal of the Royal Statistical Society Series A (Statistics in Society)*, 162(1999), 195-209.
- Diane Lambert., 1992, "Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing." *Technometrics* 34, 1-14.
- Fu-Kwun Wang., Shalemu Sharew Hailemariam. 2018. "Sampling plans for the zero-inflated Poisson distribution in the food industry." *Food Control* 85, 359-368.
- Ronald E. Walpole., Raymond H. Myers. 2012. "Probability & Statistics for Engineers & Scientists." Prentice Hall.
- Norman L. Johnson., and Adrienne W. Kemp. 2005. "Univariate Discrete Distributions." Wiley-Interscience.