The Korean Journal of Applied Statistics (2014) DOI: http://dx.doi.org/10.5351/KJAS.2014.27.5.787
27(5), 787–796

Poisson GLR Control Charts

Jaeheon Lee[a,1] · Jongtae Park[b]

[a]Department of Applied Statistics, Chung-Ang University
[b]Department of Digital Information and Statistics, Pyeongtaek University

(Received August 13, 2014; Revised September 11, 2014; Accepted September 12, 2014)

Abstract

Situations where sample size is not constant are common when monitoring a process with Poisson count data. In this paper, we propose a generalized likelihood ratio(GLR) control chart to detect shifts in the Poisson rate when the sample size varies. The performance of the proposed GLR chart is compared with the performance of several cumulative sum(CUSUM) type charts. It is shown that the overall performance of the GLR chart is comparable with CUSUM type charts and is significantly better in cases where the actual value of the shift is different from the pre-specified value in CUSUM type charts.

Keywords: Attribute data, CUSUM chart, GLR chart, Poisson process.

1. 서론

통계적 공정관리에서 공정변수 X가 결점수를 나타내는 계수형 자료(attribute data)일 경우, 일반적으로 X는 모수가 λ인 Poisson 분포를 따른다고 가정한다. 이와 같은 공정에서 표본추출 시점마다 표본 크기를 동일하게 유지하는 것은 어려운 경우가 많으며, 따라서 매 시점마다 표본 크기를 다르게 하는 경우가 흔하게 발생하게 된다.

Poisson 분포를 따르는 결점수를 관측하여 공정을 관리할 때, 매 시점마다 표본 크기가 다른 경우 가장 먼저 생각할 수 있는 것은 Shewhart의 u 관리도이다. 이 관리도 절차를 간단하게 설명하기 위해 다음과 같은 공정을 고려해 보자. $X_1, X_2, \ldots, X_k$는 각각 표본 크기 $n_1, n_2, \ldots, n_k$에서 관측한 결점수이고, 이 결점수는 서로 독립인 Poisson 분포를 따른다고 가정하자. 여기서 X_i의 관리상태에서의 평균은 $n_i\lambda_0$이고, 이상원인으로 인하여 평균이 $n_i\lambda_1$으로 변화한다고 가정한다($i = 1, 2, \ldots, k$). 이때 관리상태일 때 공정모수 값인 λ_0는 알려져 있거나, 제1국면(Phase I)에서 정확하게 추정할 수 있음을 가정한다. 또한 일반적으로 생산공정의 상태가 나빠지는 경우에 관심이 있기 때문에, 이 논문에서는 이상원인으로 인하여 λ_0가 $\lambda_1(> \lambda_0)$으로 커지는 경우만 고려하기로 한다.

Shewhart의 u 관리도는 관리통계량으로 X_k/n_k, 그리고 관리한계(control limit)로는

$$\lambda_0 \pm L_u\sqrt{\frac{\lambda_0}{n_i}}$$

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(2010-0009571).
[1]Corresponding author: Department of Applied Statistics, Chung-Ang University, 84 Heukseok-Ro, Dongjak-Gu, Seoul 156-756, Korea. E-mail: jaeheon@cau.ac.kr

를 사용한다. L_u는 관리상태에서의 특성을 만족하도록 설정할 수 있으며, 3을 사용하는 경우가 많다.

Shewhart의 u 관리도는 간편하게 적용할 수 있다는 장점은 있지만, 공정모수의 작은 변화 탐지에 효율이 많이 떨어진다고 알려져 있다. 공정모수의 작은 변화 탐지에 효율적인 관리도가 CUSUM(cumulative sum) 관리도와 EWMA(exponentially weighted moving average) 관리도이다. 표본 크기가 다른 Poisson 공정을 관리하는 CUSUM 관리도에 대한 연구를 살펴보면, Yashchin (1989), Hawkins와 Olwell (1998), 그리고 Mei 등 (2011)은 로그우도비를 사용하는 CUSUM 관리도를 연구하였고, Rossi 등 (1999)은 결점수 자료의 표준화 변환을 사용하는 CUSUM 관리도를 제안하였다. EWMA 관리도에 대한 연구로는 Dong 등 (2008)과 Ryan과 Woodall (2010) 등을 들 수 있다.

이 논문에서는 매 시점마다 표본 크기가 다른 경우 Poisson 공정을 관리하는 GLR(generalized likelihood ratio) 관리도 절차를 제안하고자 한다. GLR 관리도는 우도비검정을 기반으로 제안된 관리도로서, 최근 그 연구가 활발하게 진행되고 있다. CUSUM 관리도에서는 공정모수의 변화값을 사전에 지정해야 하고 EWMA 관리도에서는 가중치(weight)를 사전에 지정해야 하는데, GLR 관리도에서는 관리한계 외에는 사전에 지정할 관리모수가 없다는데 큰 장점이 있다 (Reynolds와 Lou, 2010; Reynolds 등, 2013; Choi와 Lee, 2014a, 2004b). 일반적으로 GLR 관리도는 사전에 공정모수를 지정하는 CUSUM 관리도와 EWMA 관리도에 비해 효율이 좋은 경우는 많지 않지만, 공정모수의 변화값을 매 시점마다 추정하여 사용하기 때문에 어떠한 공정모수의 변화에도 효율이 크게 나빠지지 않는다는 특징을 가지고 있다. Shewhart의 u 관리도는 전반적으로 효율이 많이 떨어지고 CUSUM 관리도와 EWMA 관리도는 그 효율이 크게 다르지 않기 때문에, 이 논문에서는 제안된 GLR 관리도를 여러 가지 CUSUM 관리도와 비교하여 그 효율을 평가할 것이다.

2장에서는 기존에 제안된 CUSUM 관리도들을 소개하고, 3장에서는 Poisson 공정을 관리하는 GLR 관리도 절차를 제안한다. 4장에서는 제안된 GLR 관리도의 효율을 살펴보기 위해 모의실험을 통하여 CUSUM 관리도들과 비교하고, 마지막 5장에서는 본 논문에 대한 결론를 제시하고 있다.

2. Poisson 공정을 관리하는 CUSUM 관리도

Rossi 등 (1999)은 Poisson 분포를 따르는 결점수 자료를 근사적인 정규확률변수로 변환하는 3가지 방법을 제안하였다. 첫번째 변환은

$$Z_{1,k} = \frac{X_k - n_k\lambda_0}{\sqrt{n_k\lambda_0}}$$

이며, 이 변환은 X_k의 관리상태에서의 평균과 표준편차가 각각 $n_k\lambda_0$와 $\sqrt{n_k\lambda_0}$임을 이용한 것이다. 두번째 변환은

$$Z_{2,k} = 2\left(\sqrt{X_k} - \sqrt{n_k\lambda_0}\right)$$

이며, 이 변환은 $\sqrt{X_k}$의 관리상태에서의 평균과 표준편차가 근사적으로 각각 $\sqrt{n_k\lambda_0}$와 1/2가 됨을 이용한 것이다. 세번째 변환은 위의 두 확률변수의 평균을 취한 것으로

$$Z_{3,k} = \frac{1}{2}Z_{1,k} + \frac{1}{2}Z_{2,k} = \frac{X_k - 3n_k\lambda_0 + 2\sqrt{X_k n_k\lambda_0}}{2\sqrt{n_k\lambda_0}}$$

로 나타낼 수 있다. 관리상태, 즉 $\lambda = \lambda_0$에서 $Z_{3,k}$의 평균과 표준편차는 근사적으로 각각 0과 1이 되고, 이상상태, 즉 $\lambda = \lambda_1$에서 평균과 표준편차는 근사적으로 각각

$$E(Z_{3,k} \,|\, \lambda_1) = \frac{n_k\lambda_1 - 3n_k\lambda_0 + 2n_k\sqrt{\lambda_0\lambda_1}}{2\sqrt{n_k\lambda_0}}$$

과 1이 됨을 알 수 있다. Rossi 등 (1999)은 표준화된 3가지 통계량 중 $Z_{3,k}$를 사용한 CUSUM 관리도를 제안하고, $Z_{1,k}$과 $Z_{2,k}$를 사용한 CUSUM 관리도보다 효율이 좋음을 보였다. $Z_{3,k}$를 사용한 CUSUM 관리도를 표준화 CUSUM(standardized CUSUM) 관리도라고 부르기로 하며, 다음과 같이 관리통계량 C_k^S가 정의된다.

$$C_k^S = \max\left(0\,,\, C_{k-1}^S + Z_{3,k} - \frac{E(Z_{3,k}\,|\,\lambda_1)}{2}\right). \tag{2.1}$$

표준화 CUSUM 관리도 절차는 시점 k에서 $C_k^S \geq h_S$인 경우 이상상태의 신호를 주는 것이다. 여기서 관리한계 h_S는 주어진 관리상태에서의 특성을 만족하도록 설정할 수 있다. 일반적으로 CUSUM 관리도의 특성은 Markov 연쇄(chain) 또는 모의실험을 통하여 계산할 수 있다. 따라서 특정한 값이 주어질 경우 관리한계는 여러 가지 값을 대입시키는 시행착오방법(trial and error method)을 사용하여 설정할 수 있다.

Yashchin (1989), Hawkins와 Olwell (1998), 그리고 Mei 등 (2011)은 로그우도비를 사용하는 CUSUM 관리도를 제안하고, 관리통계량 C_k^G를

$$C_k^G = \max\left(0\,,\, C_{k-1}^G + X_k - \frac{n_k(\lambda_1 - \lambda_0)}{\ln(\lambda_1) - \ln(\lambda_0)}\right) \tag{2.2}$$

로 정의하였다. 이 관리도를 일반우도비 CUSUM(generalized-likelihood ratio CUSUM) 관리도라고 부르며, 시점 k에서 $C_k^G \geq h_G$인 경우 이상상태의 신호를 주게 된다.

또한 Mei 등 (2011)은 같은 논문에서 일반우도비 CUSUM 관리도를 조금 변형하여 가중우도비 CUSUM(weighted-likelihood ratio CUSUM) 관리도를 제안하였는데, 관리통계량 C_k^W를

$$C_k^W = \max\left(0\,,\, C_{k-1}^W + \frac{X_k}{n_k} - \frac{\lambda_1 - \lambda_0}{\ln(\lambda_1) - \ln(\lambda_0)}\right) \tag{2.3}$$

로 정의하였다. 이 관리도 절차는 시점 k에서 $C_k^W \geq h_W$인 경우 이상상태의 신호를 주게 된다.

만일 표본의 크기가 동일한 경우 일반우도비 CUSUM 관리도와 가중우도비 CUSUM 관리도는 서로 같은 절차가 됨을 알 수 있다.

3. Poisson 공정을 관리하는 GLR 관리도

2장에서 소개한 CUSUM 관리도들은 이상상태에서의 공정모수 $\lambda_1(>\lambda_0)$를 사전에 지정해야 실제 공정에 적용할 수 있다. 그러나 공정에 따라서는 이 값을 예상하기 어려운 경우가 있으며, CUSUM 관리도는 공정모수가 사전에 지정한 모수의 값과 유사하게 변화했을 경우 효율이 매우 좋지만 지정한 값과 차이가 많이 나게 변화했을 경우 그 효율이 크게 떨어질 수 있다고 알려져 있기 때문에 다른 대안이 필요할 수 있다. 이런 상황에서 유용하게 사용할 수 있는 관리도가 GLR 관리도이다.

Poisson 공정을 관리하는 GLR 관리도 절차는 로그우도비를 사용하여 관리통계량을 구성한다는 점에서 Mei 등 (2011)이 제안한 일반우도비 CUSUM 관리도와 유사한 면이 있지만, 이상상태에서의 공정모수 λ_1을 사전에 지정하지 않고 매 시점마다 추정하여 사용한다는 점에서 가장 큰 차이가 있다. 제안하는 GLR 관리도를 Poisson GLR 관리도라고 부르기로 하며, 다음과 같은 과정을 거쳐 관리도 절차를 유도할 수 있다.

먼저 시점 τ와 $\tau+1$에서 이상원인이 발생했다고 가정하자. 즉, 시점 τ까지는 관리상태이고 시점 $\tau+1$부터 이상상태가 되는 것이다. 이때 τ를 공정의 변화시점(process change point)이라고 한다.

이상원인으로 인하여 공정모수 λ가 λ_0에서 λ_1으로 변화했음을 나타내는 대립가설 하에서의 우도함수(likelihood function)는 다음과 같다.

$$L(\tau, \lambda_1 | X_1, X_2, \ldots, X_k) = \frac{e^{-\lambda_0 \sum_{i=1}^{\tau} n_i} \prod_{i=1}^{\tau} (n_i \lambda_0)^{X_i} e^{-\lambda_1 \sum_{i=\tau+1}^{k} n_i} \prod_{i=\tau+1}^{k} (n_i \lambda_1)^{X_i}}{\prod_{i=1}^{k} X_i!}. \tag{3.1}$$

대립가설 하에서 λ_1의 최대우도추정량(maximum likelihood estimator; MLE)은 식 (3.1)의 우도함수로부터 유도할 수 있는데, $\lambda_1 > \lambda_0$이라는 제약조건 때문에

$$\widehat{\lambda}_{1,\tau,k} = \max\left(\lambda_0, \frac{\sum_{i=\tau+1}^{k} X_i}{\sum_{i=\tau+1}^{k} n_i}\right)$$

와 같이 설정할 수 있다.

다음으로 공정이 관리상태라고 가정하는 귀무가설 하에서의 우도함수는

$$L(\infty, \lambda_0 | X_1, X_2, \ldots, X_k) = \frac{e^{-\lambda_0 \sum_{i=1}^{k} n_i} \prod_{i=1}^{k} (n_i \lambda_0)^{X_i}}{\prod_{i=1}^{k} X_i!} \tag{3.2}$$

가 된다.

Poisson GLR 관리도의 관리통계량 R_k는 두 우도함수, 식 (3.1)과 식 (3.2)의 로그비의 최댓값으로 정의된다. 즉,

$$\begin{aligned} R_k &= \ln \frac{\max_{0 \le \tau < k, \lambda_1 > \lambda_0} L(\tau, \lambda_1 | X_1, X_2, \ldots, X_k)}{L(\infty, \lambda_0 | X_1, X_2, \ldots, X_k)} \\ &= \max_{0 \le \tau < k} \left\{ \left(\ln \widehat{\lambda}_{1,\tau,k} - \ln \lambda_0\right) \left(\sum_{i=\tau+1}^{k} X_i\right) - \left(\widehat{\lambda}_{1,\tau,k} - \lambda_0\right) \left(\sum_{i=\tau+1}^{k} n_i\right) \right\} \end{aligned}$$

가 된다. Poisson GLR 관리도 절차는 시점 k에서 $R_k \ge h_{\text{GLR}}$인 경우 이상상태의 신호를 주게 된다. 관리통계량 R_k는 CUSUM 관리도와 달리 Markov 연쇄의 조건을 만족하지 못하기 때문에, 모의실험을 통해서만 관리도의 특성을 계산할 수 있다. 따라서 관리한계 h_{GLR}은 모의실험에서 시행착오방법을 사용하여 주어진 관리상태에서의 특성을 만족하도록 설정할 수 있다.

GLR 관리도에서 관리통계량을 계산하기 위해서는 매 시점마다 그 이전 시점의 모든 데이터에 대해 최댓값을 찾아야 하는데 k가 커짐에 따라 이는 매우 번거로운 일이 된다. 이에 대한 해결책으로 윈도우 크기(window size)를 사전에 설정하고, 윈도우 크기의 과거 표본에 대해서만 최댓값을 계산하는 방법을 고려할 수도 있다. 그러나 이 논문에서는 윈도우 크기를 설정하지 않고 모든 데이터에 대해 최댓값을 찾는 방법을 사용하였다.

4. Poisson GLR 관리도의 효율

4.1. 관리도 효율의 측도 및 모의실험의 기본 설정

관리도의 효율을 나타내는 측도로 일반적으로 평균런길이인 ARL(average run length)을 사용한다. 관리도에서 공정의 시작부터 이상상태의 신호를 줄 때까지 추출한 표본의 수를 런길이(run length)라고 하는데, 관리상태의 런길이는 공정이 관리상태에서 시작하여 오경보(false alarm)를 줄 때까지 추출한 표본의 수가 되고, 이상상태의 런길이는 이상원인이 발생한 이후부터 이상상태의 신호까지 추출한 표본의 수가 된다. 런길이는 확률분포를 가지며, 관리도의 효율을 비교할 때 일반적으로 평균인 ARL을 사

Table 4.1. ARL values for the Poisson GLR, the Standardized CUSUM, and the GLR(WLR) CUSUM chart when $\lambda_0 = 10$ and the sample sizes are fixed.

	Poisson GLR	Standardized CUSUM			GLR(WLR) CUSUM		
$\lambda_1 =$		12	14	18	12	14	18
λ	[1]	[2]	[3]	[4]	[5]	[6]	[7]
10.25	124.90	117.26	131.86	143.07	**114.65**	131.22	143.23
10.50	79.94	73.31	88.39	108.30	**71.47**	86.94	108.24
10.75	53.69	49.27	61.97	80.28	**47.37**	60.88	79.92
11.00	38.32	34.65	44.73	62.15	**33.36**	43.15	61.96
11.50	22.31	19.93	25.62	37.94	**19.38**	24.88	37.72
12.00	14.81	13.43	15.88	24.45	**13.10**	15.41	24.14
13.00	8.21	**7.66**	8.30	11.62	7.67	8.10	11.50
14.00	5.44	5.40	5.35	6.67	5.44	**5.30**	6.63
15.00	3.99	4.19	3.91	4.47	4.27	**3.89**	4.44
16.00	3.14	3.49	**3.12**	3.28	3.53	3.13	3.26
18.00	2.19	2.63	2.26	**2.13**	2.64	2.28	**2.13**
20.00	1.70	2.17	1.84	**1.62**	2.16	1.83	**1.62**
h	4.043	4.68	2.95	1.56	16.33	10.50	6.39

용하고 있다. 이상상태의 ARL을 계산할 때 다음 2가지 방법을 사용할 수 있다. 하나는 공정이 처음부터 이상상태라고 가정하고 ARL을 계산하는 것이고, 다른 하나는 공정이 관리상태로 유지되다가 어느 시점부터 이상상태로 변화했다고 가정하고 ARL을 계산하는 것이다. 전자를 초기상태(zero-state) ARL이라 하고, 후자를 안정상태(steady-state) ARL이라고 한다. 후자의 경우가 좀 더 현실적이라 할 수 있기 때문에 이 논문에서도 안정상태 ARL을 이용하여 관리도의 효율을 비교하고자 한다.

모의실험에서 관리한계는 관리상태에서의 ARL, 즉 ARL_0가 200이 되도록 설정하였으며, 이를 비롯한 기본적인 설정은 Ryan과 Woodall (2010)의 경우와 동일하게 하였다. 그 이유는 Poisson GLR 관리도와 효율을 비교하는 식 (2.1)의 표준화 CUSUM 관리도, 식 (2.2)의 일반우도비 CUSUM 관리도, 그리고 식 (2.3)의 가중우도비 CUSUM 관리도의 경우 Ryan과 Woodall (2010)에 제시된 모의실험 결과를 그대로 이용하였기 때문이다. Poisson GLR 관리도에 대한 모의실험에서 안정상태 ARL을 계산하기 위하여 50번째 시점까지는 관리상태에서, 51번째 시점부터 이상상태에서 표본을 추출하였고, 관리상태에서 오경보가 난 경우는 반복에서 제외시켰다. 즉, 공정의 변화시점으로 $\tau = 50$을 사용한 것이다. 이 방법은 CUSUM 관리도들에 대한 모의실험에서 Ryan과 Woodall (2010)이 사용하였기 때문에, Poisson GLR 관리도에 대한 모의실험에 동일하게 적용하였다. 또한 Poisson GLR 관리도에서 안정상태 ARL 계산을 위해 100,000번 반복을 수행하였다.

4.2. 표본 크기가 동일한 경우 모의실험 결과

먼저 표본 크기가 모두 동일한 경우를 고려해 보자. 이때 표본 크기는 일반성을 잃지 않고 $n_i = 1$을 가정하였다. 관리상태에서의 공정모수 λ_0는 10과 6, CUSUM 관리도에서 사전에 지정해야 하는 공정모수 λ_1은 $\lambda_0 = 10$일 때 12, 14, 18인 경우를 고려했으며, $\lambda_0 = 6$일 때 7, 8, 9인 경우를 고려하였다. 이에 대한 모의실험 결과를 Table 4.1과 Table 4.2에 제시하였다.

Table 4.1과 Table 4.2에서 열 [1]은 Poisson GLR 관리도, 열 [2], [3], [4]는 각각 λ_1을 12, 14, 18 (Table 4.2에서는 7, 8, 9)로 지정했을 때 표준화 CUSUM(Standardized CUSUM) 관리도, 그리고 열 [5], [6],

Table 4.2. ARL values for the Poisson GLR, the Standardized CUSUM, and the GLR(WLR) CUSUM chart when $\lambda_0 = 6$ and the sample sizes are fixed.

	Poisson GLR	Standardized CUSUM			GLR(WLR) CUSUM		
$\lambda_1 =$		7	8	9	7	8	9
λ	[1]	[2]	[3]	[4]	[5]	[6]	[7]
6.25	111.56	92.69	106.15	<u>111.60</u>	**89.68**	103.50	109.80
6.50	63.48	51.83	61.62	<u>67.72</u>	**49.78**	59.54	66.21
6.75	40.05	33.03	38.93	<u>44.33</u>	**32.18**	37.54	43.16
7.00	27.52	23.05	26.76	<u>30.33</u>	**22.29**	25.50	29.42
7.25	20.20	17.07	19.40	<u>21.67</u>	**17.03**	18.46	21.20
7.50	15.67	13.61	14.92	<u>16.85</u>	**13.42**	14.53	16.15
7.75	12.50	**11.27**	11.62	<u>12.87</u>	11.29	11.40	12.46
8.00	10.33	**9.44**	9.67	<u>10.38</u>	9.58	9.49	10.20
9.00	5.82	5.97	5.61	5.63	<u>6.09</u>	5.63	**5.59**
10.00	3.96	4.36	3.93	3.83	<u>4.50</u>	4.00	**3.80**
11.00	2.97	3.56	3.10	**2.93**	<u>3.63</u>	3.21	2.94
12.00	2.39	2.99	2.60	**2.38**	<u>3.02</u>	2.64	2.41
13.00	**1.99**	2.61	2.27	2.04	<u>2.65</u>	2.31	2.08
14.00	**1.73**	2.35	1.99	1.78	<u>2.38</u>	2.03	1.84
h	3.964	5.71	4.00	3.01	16.14	11.20	8.61

[7]은 일반우도비(가중우도비) CUSUM(GLR(WLR) CUSUM) 관리도의 결과이다. 앞에서 언급한 바와 같이 표본 크기가 동일한 경우 일반우도비 CUSUM 관리도와 가중우도비 CUSUM 관리도는 동일한 절차이다. 모의실험 결과를 좀 더 명확하게 살펴보기 위해, 공정모수 λ의 변화에 대해 이를 가장 빨리 탐지한 결과는 볼드체로 나타내었고 가장 늦게 탐지한 결과에는 밑줄을 그어 나타내었다. 또한 마지막 행의 h는 각 관리도에서 $\mathrm{ARL}_0 = 200$을 근사적으로 만족하는 관리한계의 값을 나타낸다.

Table 4.1의 결과를 살펴보면 다음의 내용을 알 수 있다. 먼저 CUSUM 관리도들의 효율을 비교하면, 전반적으로 일반우도비(가중우도비) CUSUM 관리도가 표준화 CUSUM 관리도에 비해 공정모수의 변화 탐지에 효율적임을 알 수 있다. (CUSUM 관리도들에 대한 상세한 비교는 Ryan과 Woodall (2010)을 참고할 수 있다.) 당연한 결과일 수 있지만, 공정모수가 미리 지정한 값과 유사하게 변화한 경우 CUSUM 관리도들의 효율이 좋지만, 매우 상이하게 변화한 경우에는 효율이 나빠질 수 있다는 것을 알 수 있다. 그러나 Poisson GLR 관리도는 공정모수의 변화에 대해 최적인 경우는 없지만, 어떠한 변화에도 효율이 크게 나빠지는 경우는 없으며 대부분 최적인 결과와 큰 차이가 없게 나타났다.

Table 4.2의 경우에도 유사한 경향이 나타난다. Poisson GLR 관리도는 공정모수의 변화가 작은 경우 전반적으로 CUSUM 관리도들에 비해 효율이 좋지 않지만 최악인 경우는 없으며, 변화가 커질수록 효율이 좋아짐을 알 수 있다. 특히 공정모수가 CUSUM 관리도에서 지정한 λ_1의 최댓값인 9보다 더 크게 변화한 경우($\lambda \geq 13$) 효율이 가장 좋은 것으로 나타났다.

4.3. 표본 크기가 동일하지 않은 경우 모의실험 결과

이제 표본 크기가 서로 다른 경우를 고려해 보자. 표본 크기는 Ryan과 Woodall (2010)의 경우와 동일하게 이산형 균등분포(discrete uniform distribution)에서 랜덤하게 생성하였다. 최솟값이 a이고 최댓값이 b인 균등분포를 $U(a, b)$라고 표시할 때, $U(10, 15)$와 $U(10, 50)$인 경우를 고려하였다. 또한 관리상

Table 4.3. ARL values for the Poisson GLR, the Standardized CUSUM, the GLR CUSUM, and the WLR CUSUM chart when $\lambda_0 = 1$ and $n_i \sim U(10, 15)$.

	Poisson GLR	Standardized CUSUM			GLR CUSUM			WLR CUSUM		
$\lambda_1 =$		1.2	1.5	2	1.2	1.5	2	1.2	1.5	2
λ	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
1.025	117.70	110.59	131.61	148.81	**108.71**	130.95	147.06	110.01	131.43	148.44
1.05	72.83	67.04	91.73	114.58	**65.24**	90.10	111.54	65.82	91.51	113.44
1.075	47.68	44.65	63.92	86.41	**43.13**	63.02	83.41	43.72	63.49	85.36
1.1	33.41	30.04	46.23	66.83	**29.27**	45.02	64.44	29.57	45.90	65.58
1.15	19.10	17.23	25.80	41.54	**16.92**	24.90	39.43	16.94	25.31	40.92
1.2	12.61	11.49	16.03	26.69	**11.18**	15.51	24.88	11.46	15.76	25.79
1.3	6.96	**6.57**	7.77	12.57	6.58	7.51	11.69	6.69	7.72	12.05
1.4	**4.63**	4.67	4.76	7.07	4.68	4.69	6.52	4.74	4.74	6.91
1.5	3.42	3.61	3.37	4.39	3.63	**3.34**	4.13	3.67	3.41	4.28
1.6	2.69	3.01	**2.65**	3.08	2.99	**2.65**	2.97	3.04	2.66	2.99
1.8	1.89	2.29	**1.86**	1.91	2.28	1.87	1.87	2.31	1.90	1.89
2.0	1.49	1.90	1.51	1.43	1.89	1.52	**1.41**	1.90	1.50	1.43
h	4.112	4.37	2.25	1.06	16.97	9.04	4.68	1.38	0.74	0.39

Table 4.4. ARL values for the Poisson GLR, the Standardized CUSUM, the GLR CUSUM, and the WLR CUSUM chart when $\lambda_0 = 1$ and $n_i \sim U(10, 50)$.

	Poisson GLR	Standardized CUSUM	GLR CUSUM	WLR CUSUM
$\lambda_1 =$		2	2	2
λ				
1.025	**90.87**	149.96	145.13	152.00
1.05	**46.18**	113.26	106.33	117.65
1.075	**27.31**	84.92	77.37	89.52
1.1	**18.26**	64.90	56.69	69.70
1.15	**10.06**	37.49	31.45	40.55
1.2	**6.55**	23.07	18.60	25.29
1.3	**3.70**	9.52	7.60	10.39
1.4	**2.53**	4.61	3.84	4.91
1.5	**1.93**	2.71	2.33	2.85
1.6	**1.59**	1.82	1.68	1.94
1.8	1.26	1.24	**1.20**	1.27
2.0	1.12	1.07	**1.06**	1.08
h	4.142	0.46	2.82	0.16

태에서의 공정모수는 $\lambda_0 = 1$이고, CUSUM 관리도에서 사전에 지정해야 하는 공정모수는 $U(10, 15)$일 때 $\lambda_1 = 1.2$, 1.5, 2인 경우, $U(10, 50)$일 때 $\lambda_1 = 2$인 경우를 고려하였다. $\lambda_0 = 1$로 설정한 이유는 관리상태에서 Poisson 분포의 평균인 $n_i\lambda_0$가 표본 크기가 동일한 경우의 평균값과 유사하게 유지되기 때문이다.

$n_i \sim U(10, 15)$와 $n_i \sim U(10, 50)$에 대한 모의실험 결과를 각각 Table 4.3과 Table 4.4에 제시하였다. 또한 Table 4.3에서 열 [1]은 Poisson GLR 관리도, 열 [2], [3], [4]는 각각 λ_1을 1.2, 1.5, 2로 지정했

을 때 표준화 CUSUM(Standardized CUSUM) 관리도, 열 [5], [6], [7]은 일반우도비 CUSUM(GLR CUSUM) 관리도, 그리고 열 [8], [9], [10]은 가중우도비 CUSUM(WLR CUSUM) 관리도의 결과이다. Table 4.1과 Table 4.2에서와 마찬가지로 공정모수 λ의 변화를 가장 빨리 탐지한 결과는 볼드체로 나타내었고, 가장 늦게 탐지한 결과에는 밑줄을 그어 나타내었다. 또한 마지막 행의 h는 $\text{ARL}_0 = 200$을 근사적으로 만족하는 관리한계를 나타낸다.

Table 4.3의 결과를 살펴보면 전반적으로 일반우도비 CUSUM 관리도의 효율이 제일 좋게 나타났지만, 일반우도비 CUSUM 관리도의 경우 공정모수가 지정한 값과 상이하게 변화한 경우 효율이 나빠질 수 있다는 단점이 있다. 예를 들어, 공정모수가 작게 변화한 경우($\lambda \leq 1.2$) $\lambda_1 = 1.2$로 설정한 일반우도비 CUSUM 관리도의 효율이 Poisson GLR 관리도보다 좋게 나타나지만, 이 경우 $\lambda_1 = 2$로 설정한 일반우도비 CUSUM 관리도의 효율은 Poisson GLR 관리도에 비해 크게 떨어짐을 알 수 있다. 공정모수가 크게 변화한 경우에도 유사한 방법으로 효율을 설명할 수 있다.

Table 4.4에는 $\lambda_1 = 2$로 지정한 CUSUM 관리도의 결과만 제시하였다. Ryan과 Woodall (2010)에서 이 경우만 제시한 이유도 있지만, 다른 λ_1 값에 대해서는 앞에서와 유사한 경향이 있으며 이 경우에 Poisson GLR 관리도의 장점을 명확하게 설명할 수 있기 때문이다. 즉, 공정모수가 크게 변화할 것이라 예상하여 $\lambda_1 = 2$로 지정했지만, 실제 공정모수가 작게 변화한 경우($\lambda \leq 1.6$) Poisson GLR 관리도의 효율이 월등히 좋게 나타났으며, $\lambda \geq 1.8$인 경우 일반우도비 CUSUM 관리도를 비롯한 CUSUM 관리도의 효율이 좋게 나타났다. 그러나 $\lambda = 2.0$인 경우 큰 차이는 아니지만 Poisson GLR 관리도의 효율이 가장 나쁘게 나타났다.

표본 크기의 변동이 크지 않은 경우($n_i \sim U(10, 15)$) 표준화 CUSUM 관리도의 효율이 상대적으로 가장 나빴지만, 표본 크기의 변동이 심한 경우($n_i \sim U(10, 50)$) 가중우도비 CUSUM 관리도의 효율이 가장 나쁘게 나타난 것은 특이한 사항이라고 생각한다.

5. 결 론

Poisson 분포를 따르는 결점수를 관측하여 공정을 관리할 때 표본 크기를 동일하게 유지할 수 없는 경우가 종종 발생한다. 이 논문은 이러한 상황에서 사용할 수 있는 Poisson GLR 관리도 절차를 제안하였고, 그 효율을 표준화 CUSUM, 일반우도비 CUSUM, 그리고 가중우도비 CUSUM 관리도와 비교하였다. GLR 관리도는 CUSUM 관리도와는 달리 공정모수의 변화값을 미리 지정할 필요가 없다는 장점을 가지고 있다. 실제 공정에서 과거의 경험을 통하여 이상원인으로 인한 Poisson 공정모수 λ의 변화값을 예상할 수 있는 경우도 있지만, 예상할 수 없는 경우도 흔하게 발생할 수 있다. 즉, 이상원인으로 인하여 λ가 관리상태에서의 모수값인 λ_0에 비하여 크게 변화할지, 작게 변화할지에 대한 예상이 어려운 경우가 있을 수 있다. 이때 Poisson GLR 관리도를 유용하게 사용할 수 있을 것이다.

모의실험을 통하여 CUSUM 관리도들과 효율을 비교할 때, Poisson GLR 관리도는 공정모수의 어떠한 변화에도 효율이 크게 떨어지는 경우가 없으며 전반적으로 최적인 결과와 큰 차이가 발생하지 않음을 알 수 있었다. 결론적으로 Poisson 공정모수의 변화값을 어느 정도 예상할 수 있는 경우에는 일반우도비 CUSUM 관리도를 사용하고, 그렇지 않은 경우에는 Poisson GLR 관리도를 사용하는 것이 효율적인 공정관리를 위해 바람직하다고 생각한다.

References

Choi, M. L. and Lee, J. (2014a). GLR charts for simultaneously monitoring a sustained shift and a linear drift in the process mean, *Communications for Statistical Applications and Methods*, **21**, 69–80.

Choi, M. L. and Lee, J. (2014b). A GLR chart for monitoring a zero-inflated Poisson process, *The Korean Journal of Applied Statistics*, **27**, 345–355.

Dong, Y., Hedayat, A. S. and Sinha, B. K. (2008). Surveillance strategies for detecting changepoint in incidence rate based on exponentially weighted moving average methods, *Journal of the American Statistical Association*, **103**, 843–853.

Hawkins, D. M. and Olwell, D. H. (1998). *Cumulative Sum Charts and Charting for Quality Improvement*, New York, NY: Springer.

Mei, Y., Han, S. W. and Tsui, K. L. (2011). Early detection of a change in Poisson rate after accounting for population size effects, *Statistica Sinica*, **21**, 597–624.

Reynolds, M. R. J. and Lou, J. (2010). An evaluation of a GLR control chart for monitoring the process mean, *Journal of Quality Technology*, **42**, 287–310.

Reynolds, M. R. J., Lou, J., Lee, J. and Wang, S. (2013). The design of GLR control charts for monitoring the process mean and variance, *Journal of Quality Technology*, **45**, 34–60.

Rossi, G., Lampugnani, L. and Marchi, M. (1999). An approximate CUSUM procedure for surveillance of health events, *Statistics in Medicine*, **18**, 2111–2122.

Ryan, A. G. and Woodall, W. H. (2010). Control charts for Poisson count data with varying sample sizes, *Journal of Quality Technology*, **42**, 260–275.

Yashchin, E. (1989). Weighted cumulative sum technique, *Technometrics*, **31**, 321–338.

Poisson GLR 관리도

이재헌[a,1] · 박종태[b]

[a]중앙대학교 응용통계학과, [b]평택대학교 디지털응용정보학과

(2014년 8월 13일 접수, 2014년 9월 11일 수정, 2014년 9월 12일 채택)

요 약

Poisson 분포를 따르는 결점수를 관측하여 공정을 관리할 때 표본 크기를 동일하게 유지하기가 힘든 경우가 많다. 이 논문은 표본 크기가 동일하지 않은 경우 Poisson 공정모수의 변화를 탐지하는 GLR(generalized likelihood ratio) 관리도 절차를 제안하고 있다. 또한 제안된 GLR 관리도의 효율을 모의실험을 통하여 기존에 연구된 CUSUM 관리도들과 비교하였다. 모의실험 결과, 제안된 GLR 관리도는 공정모수의 다양한 변화에 대해 효율이 대체적으로 양호했으며, CUSUM 관리도에서 실제 공정모수의 변화값이 미리 지정한 값과 차이가 많이 날 경우 CUSUM 관리도에 비해 효율이 월등히 좋음을 알 수 있었다.

주요용어: 계수형 자료, CUSUM 관리도, GLR 관리도, Poisson 과정.

이 논문은 2010년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(2010-0009571).

[1]교신저자: (156-756) 서울시 동작구 흑석로 84, 중앙대학교 응용통계학과. Email: jaeheon@cau.ac.kr