Journal of the Korean
Data & Information Science Society
1999, Vol. 10, No. 2, pp. 319 ~ 327

이변량 영과잉-포아송모형에서 변화시점에 관한 추론 [1]

김 경 무 [2]

요약

영과잉-포아송분포는 여러 형태의 불량률을 줄이는 생산공정과정에서 유용하게 이용되어 왔다. 또한 생산공정과정 중 미지의 변화시점 이후 불량률의 변화가 있는 지를 알아보는 것은 흥미있는 일이고 연구된바 있다. 만약 불량품들이 서로 두가지 다른 형태의 규격에 의해 발생되었다면, 이는 일변량이 아닌 이변량 영과잉-포아송 분포를 이용해야 할 것이다. 본 논문은 이변량 영과잉-포아송모형에서 어느 미지의 시점 이후 분포의 변화가 있는지를 우도비 검정을 통해 알아본다. 또한 변화가 있다면 변화시점과 그리고 여러 형태의 모수들에 대한 점추정량을 알아보려 한다.

주제어: 영과잉-포아송분포

1. 서 론

영과잉-포아송(Zero-Inflated Poisson: ZIP)분포라 함은 이산형 확률분포에 있어서 정상적인 포아송 확률분포보다 영의 값이 과잉관측되는 분포를 의미한다. 최근 제품을 만들어 내는 기술의 고급화로 인하여 ZIP분포가 불량률을 줄이는데 유용하게 이용되어 왔다. 이러한 경우, 생산공정과정에서 나타나는 단위당 불량품수가 이러한 ZIP분포를 따른다고 할 때, 불량품 수가 어떤 시점 이후 변화가 있다고 하자. 이때 미지의 시점을 변화시점(changepoint)이라 한다. 이러한 ZIP분포는 Cohen(1963)과 Singh(1963) 그리고 Johnscn-Kotz(1969)에 소개되었다. 그러나 ZIP 분포는 수학적분포로만 단순히 이용되어오고 사회현상에는 응용되지 않았다. 그 이후 처음으로 Yip(1988)는 나무에서 잎사귀 당 나타나는 벌레의 수를 ZIP를 이용해 모형화 시켰다. Lambert(1992)는 ZIP 회귀모형을 제시하고 공변량들의 효과를 실제 예를 들어 연구하였다. 최근 Li외 5인(1999)는 일변량 ZIP에서 다변량 $ZIP(MZIP)$ 로 확장하고 분포의 성질과 이의 적용방법을 생각하였다. 또한 김경무(1998)는 ZIP 모형에 변화시점문제를 적용하였고 우도비검정과 추정법을 소개하였다.

[1]이 논문은 1999학년도 대구대학교 학술연구비 지원에 의한 논문임.
[2](712-714) 경북 경산시 진량읍 내리리 15, 대구대학교 자연과학대학 통계학과 교수

본 논문은 일변량 ZIP가 아닌 서로 종속관계가 있는 두 변수(예, 불량규격)가 이변량 ZIP ($BZIP$)를 따르는 모형을 소개한다. 그리고 $BZIP$ 모형에서 어떤 미지의 변화시점 이후 불량률이 증가하는 분포의 변화가 있는 경우, 변화시점의 유 무에 대한 우도비검정을 소개하려 한다. 여기에서 변화시점 이전과 이후의 두 분포는 서로 다른 두 형태의 $BZIP$를 이용한다. 변화시점 이전의 $BZIP$ 보다는 이후의 $BZIP$가 불량률이 커지는 대립가설을 설정하였다. 이는 미지의 변화시점 이후 생산공정의 여러가지 원인들로 인하여 불량품 수가 증가하는 지를 알아보기 위함이다. 이러한 일변량 ZIP에서 $BZIP$로 확장된 모형에서 변화시점문제는 그 응용분야가 광범위하리라 생각된다. 이변량에서 다변량으로 확장은 가능하지만 모형안에 많은 모수들이 있기 때문에 적용하기가 쉽지 않다. 대립가설이 참인 경우, 즉 분포의 변화가 있는 경우, 변화시점에 대한 점추정량, 변화된 $BZIP$에서의 모수 추정량을 최소제곱법 및 영에 대한 빈도수를 이용하여 구해보았다.

2. 이변량 영과잉-포아송 모형

확률변수 X는 생산공정에서 일정 단위당 불량품이 나타나는 수로서 일변량 ZIP 분포를 따른다고 하자. ZIP분포는 포아송분포와 0에서 퇴화된(degenerated) 분포와의 혼합분포로 볼 수 있다. 즉,

$$\begin{aligned} X \quad &\sim \quad 0, && p \text{ 의 확률로} \\ &\sim \quad Poisson(\lambda), && 1-p \text{ 의 확률,} \end{aligned}$$

여기에서 $0 \le p \le 1$는 불량품이 전혀 나타나지 않는 상태 (perfect state)의 확률이며, $\lambda \ge 0$는 포아송분포의 평균이다. 이때 확률질량함수 (probability mass function : pmf)는 아래와 같이 된다.

$$\begin{aligned} P(X=k) &= p+(1-p)e^{-\lambda} \ , && k=0 \\ &= (1-p)\lambda^k e^{-\lambda}/k! \ , && k=1,2,\cdots . \end{aligned}$$

앞으로 위 분포를 $ZIP(p,\lambda)$로 표기하기로 한다. 다음으로 일변량 ZIP를 확장시켜서 변화시점 이전과 이후가 서로 다른 분포가 되는 두 형태의 $BZIP$를 소개하려 한다. $BZIP$ 혹은 $MZIP$는 그 분포가 유일하게 나타나지 않기 때문에 모형을 어떻게 설계하느냐에 따라 서로 다른 분포를 유도할 수 있다.

2.1 모형 I

서로 종속인 두 확률변수 X, Y가 각각 일변량 ZIP분포를 따르고, 이들 이변량 확률변수 (X, Y)가 다음과 같은 $BZIP$분포를 따른다 하자. 또한 이들 주변분포(marginal distribution)가 역시 일변량 ZIP분포가 되도록 모형을 설계하려 한다. 즉,

$$(X,Y) \sim \begin{cases} (0,0) & , p_{00} \text{ 의 } \quad \text{확률로,} \\ (0,V) & , p_{01} \text{ 의 } \quad \text{확률로,} \\ (U,0) & , p_{10} \text{ 의 } \quad \text{확률로,} \\ (U,V) & , p_{11} \text{ 의 } \quad \text{확률로} \end{cases} \tag{1}$$

여기에서 $\sum_{i=0}^{1}\sum_{j=0}^{1} p_{ij}$ 이다. 또한 U와 V는 각각 평균이 λ_1 및 λ_2의 일변량 포아송분포를 따르고 (U,V)는 이변량 포아송분포 $Poisson(\lambda_{10},\lambda_{01},\lambda_{11})$를 따른다. 즉, 그 pmf는 (2)식과 같다.

$$\begin{aligned} P(U=u, V=v) &= \sum_{a=0}^{min(u,v)} \frac{\lambda_{11}^{a}\lambda_{10}^{u-a}\lambda_{01}^{v-a}}{a!(u-a)!(v-a)!e^{-(\lambda_{10}+\lambda_{01}+\lambda_{11})}}\ , \\ & \qquad u=0,1,2,\cdots, \quad v=0,1,2,\cdots. \end{aligned} \tag{2}$$

이러한 이변량 ZIP분포는 6개의 모수를 포함하고 있고 $BZIP_I(p_{00},p_{10},p_{01},\lambda_{10},\lambda_{01},\lambda_{11})$로 표기하기로 한다. 이 분포의 pmf는 다음과 같게 된다.

$$\begin{aligned} P(Y_1=0,Y_2=0) &= p_{00}+p_{10}e^{-\lambda_1}+p_{01}e^{-\lambda_2}+p_{11}e^{-\lambda}, \\ P(Y_1=k_1,Y_2=0) &= p_{10}\frac{\lambda_1^{k_1}}{k_1!}e^{-\lambda_1}+p_{11}\frac{\lambda_{10}^{k_1}}{k_1!}e^{-\lambda}, \\ P(Y_1=0,Y_2=k_2) &= p_{01}\frac{\lambda_2^{k_2}}{k_2!}e^{-\lambda_2}+p_{11}\frac{\lambda_{01}^{k_2}}{k_2!}e^{-\lambda}, \\ P(Y_1=k_1,Y_2=k_2) &= p_{11}\sum_{a=0}^{min(k_1,k_2)}\frac{\lambda_{11}^{a}\lambda_{10}^{k_1-a}\lambda_{01}^{k_2-a}}{a!(k_1-a)(k_2-a)!}e^{-\lambda}, \end{aligned}$$

여기에서 $\lambda_1=\lambda_{10}+\lambda_{11},\quad \lambda_2=\lambda_{01}+\lambda_{11}, \lambda=\lambda_{10}+\lambda_{01}+\lambda_{11}$ 이다.

2.2 모형 II

이변량 확률변수 (X,Y) 가 모형 (1)을 따른다고 하자. 이때 U와 V는 각각 모형 I에서와 같이 포아송분포를 따르지 않고 $ZIP(1-e^{-(\lambda_{01}+\lambda_{11})},\lambda_{10})$ 그리고 $ZIP(1-e^{-(\lambda_{10}+\lambda_{11})},\lambda_{01})$ 를 따른다고 하자. 그리고 (U,V)는 이변량 포아송분포 $Poisson(\lambda_{10},\lambda_{01},\lambda_{11})$ 를 따르도록 설계한다. 이러한 이변량 ZIP분포를 $BZIP_{II}(p_{00},p_{10},p_{01},\lambda_{10},\lambda_{01},\lambda_{11})$ 라 하고 그 pmf는 다음과 같다.

$$\begin{aligned} P(Y_1=0,Y_2=0) &= p_{00}+p_{10}(1-e^{-\lambda_2}+e^{-\lambda})+p_{01}(1-e^{-\lambda_1}+e^{-\lambda})+p_{11}e^{-\lambda}, \\ P(Y_1=k_1,Y_2=0) &= (p_{10}+p_{11})\frac{{\lambda_{10}}^{k_1}e^{-\lambda}}{k_1!}, \end{aligned}$$

$$P(Y_1=0,Y_2=k_2) = (p_{01}+p_{11})\frac{\lambda_{01}{}^{k_2}e^{-\lambda}}{k_2!},$$
$$P(Y_1=k_1,Y_2=k_2) = p_{11}\sum_{a=0}^{min(k_1,k_2)}\frac{\lambda_{11}^{a}\lambda_{10}^{k_1-a}\lambda_{01}^{k_2-a}}{a!(k_1-a)(k_2-a)!}e^{-\lambda}.$$

3. 우도비 검정

서로 독립인 이변량 확률변수 $(X_1,Y_1),(X_2,Y_2),\cdots,(X_n,Y_n)$들이 시간의 흐름에 따라 연속적으로 얻을 수 있는 관측자료라 하자. 이때 생산공정과정 중 여러가지 원인들로 인하여 미지의 변화시점 c 이후 분포의 변화가 있는 다음과 같은 귀무, 대립가설을 생각할 수 있다.

$$H_0: \quad (X_1,Y_1),(X_2,Y_2),...,(X_n,Y_n) \sim BZIP_{II}(p_{00},p_{10},p_{01},\lambda_{10},\lambda_{01},\lambda_{11}),$$
$$H_1: \quad (X_1,Y_1),(X_2,Y_2),...,(X_c,Y_c) \sim BZIP_{II}(p_{00},p_{10},p_{01},\lambda_{10},\lambda_{01},\lambda_{11}),$$
$$(X_{c+1},Y_{c+1}),(X_{c+2},Y_{c+2}),...,(X_n,Y_n) \sim BZIP_{I}(p_{00},p_{10},p_{01},\lambda_{10},\lambda_{01},\lambda_{11}).$$

위 대립가설에서 변화시점 c 이후 분포의 변화있다. 이때 분포 자체에서, 변화시점 이전과 이후 불량품이 전혀 나타나지 않는 확률, $P(X=0,Y=0)$는 $BZIP_{II}$가 $BZIP_I$보다 크게 된다. 다시 말하면 위 모형은 변화시점 이후 불량품수가 더 많이 나타나는 모형을 의미한다. 여기에서 유의할 점은 두 분포에서 모수들은 변화가 없다는 사실이다. 단지 분포형태에만 변화가 있다. 이러한 경우 관심의 대상이 되는 대립가설은 변화시점 이후 불량률이 늘어나는 모형이 된다. 모형I, II에 대한 불량품이 전혀 나타나지 않는 확률, $P(X=0,Y=0)$를 각각 P_I, P_{II}라 하고, 만약 대립가설이 참이라면 변화시점 이후와 이전의 차이는

$$P_I - P_{II} = (p_{10}+p_{01})(e^{-\lambda_1}+e^{-\lambda_2}-e^{-\lambda}-1) \tag{3}$$

이다. 위 (3)식 우변 두 번째 항에서, $\lambda_1=\lambda_{10}+\lambda_{11},\quad \lambda_2=\lambda_{01}+\lambda_{11}, \lambda=\lambda_{10}+\lambda_{01}+\lambda_{11}$ 는 모두 양의 값을 갖는다. 또한 $1=e^{-0}$이고 나머지 지수 제곱항을 살펴보면, 위 (3)식은 항상 음이 됨을 알 수 있다. 즉, 변화시점 이후 불량품이 전혀 나타나지 않을 확률이 작게 나타된다. 이는 대립가설이 참이라면 변화시점 이후 불량률이 증가하는 모형이라 볼 수 있다. 위 경우 귀무가설에 대한 로그-우도함수 (log-likelihood function)는

$$\begin{aligned} &l(x,y;BZIP_{II}(p_{00},p_{10},p_{01},\lambda_{10},\lambda_{01},\lambda_{11})) \\ &= \sum_{i=1}^{n} ln[(p_{00}+p_{10}e^{-\lambda_1}+p_{01}e^{-\lambda_2}+p_{11}e^{-\lambda})I(x_i=0,y_i=0) \end{aligned}$$

$$+(p_{10}\frac{\lambda_1^{k_1}}{k_1!}e^{-\lambda_1}+p_{11}\frac{\lambda_{10}^{k_1}}{k_1!}e^{-\lambda})I(x_i>0,y_i=0)$$
$$+(p_{01}\frac{\lambda_2^{k_2}}{k_2!}e^{-\lambda_2}+p_{11}\frac{\lambda_{01}^{k_2}}{k_2!}e^{-\lambda})I(x_i=0,y_i>0)$$
$$+(p_{11}\sum_{a=0}^{min(k_1,k_2)}\frac{\lambda_{11}^{a}\lambda_{10}^{k_1-a}\lambda_{01}^{k_2-a}}{a!(k_1-a)(k_2-a)!}e^{-\lambda})I(x_i,y_i>0)]$$

이다. 여기에서 $I(\cdot)$는 지시함수이다. 그리고 대립가설에 대한 로그-우도함수는 다음과 같다.

$$l(x,y;BZIP_{II}(p_{00},p_{10},p_{01},\lambda_{10},\lambda_{01},\lambda_{11}),c,BZIP_I(p_{00},p_{10},p_{01},\lambda_{10},\lambda_{01},\lambda_{11}))$$
$$=\sum_{i=1}^{c}ln[(p_{00}+p_{10}e^{-\lambda_1}+p_{01}e^{-\lambda_2}+p_{11}e^{-\lambda})I(x_i=0,y_i=0)$$
$$+(p_{10}\frac{\lambda_1^{k_1}}{k_1!}e^{-\lambda_1}+p_{11}\frac{\lambda_{10}^{k_1}}{k_1!}e^{-\lambda})I(x_i>0,y_i=0)$$
$$+(p_{01}\frac{\lambda_2^{k_2}}{k_2!}e^{-\lambda_2}+p_{11}\frac{\lambda_{01}^{k_2}}{k_2!}e^{-\lambda})I(x_i=0,y_i>0)$$
$$+(p_{11}\sum_{a=0}^{min(k_1,k_2)}\frac{\lambda_{11}^{a}\lambda_{10}^{k_1-a}\lambda_{01}^{k_2-a}}{a!(k_1-a)(k_2-a)!}e^{-\lambda})I(x_i,y_i>0)]$$
$$+\sum_{i=c+1}^{n}ln[(p_{00}+p_{10}(1-e^{-\lambda_2}+e^{-\lambda})+p_{01}(1-e^{-\lambda_1}+e^{-\lambda})+p_{11}e^{-\lambda})I(x_i=0,y_i=0)$$
$$+(p_{10}+p_{11})\frac{\lambda_{10}{}^{k_1}e^{-\lambda}}{k_1!}I(x_i>0,y_i=0)$$
$$+(p_{01}+p_{11})\frac{\lambda_{01}{}^{k_2}e^{-\lambda}}{k_2!}I(x_i=0,y_i>0)$$
$$+p_{11}\sum_{a=0}^{min(k_1,k_2)}\frac{\lambda_{11}^{a}\lambda_{10}^{k_1-a}\lambda_{01}^{k_2-a}}{a!(k_1-a)(k_2-a)!}e^{-\lambda}I(x_i>0,y_i>0)].$$

위 대립가설에 대한 로그-우도함수는 7개의 모수로 이루어진 복잡한 함수이다. 각 모수들의 최우추정량은 우도함수 자체가 복잡하기 때문에 간단한 형태의 추정량으로 구할 수 없기 때문에 수치해석적인 방법으로 구해야 할 것이다. 저자는 수치해석적인 방법으로 Powell(1964)방법을 추천하고 싶다.

우도함수에 대한 각 $c,p_{00},p_{10},p_{01},\lambda_{10},\lambda_{01},\lambda_{11}$ 모수들의 최우추정량 (MLE)을 $\tilde{c},\tilde{p_{00}},\tilde{p_{10}},\tilde{p_{01}}$, $\tilde{\lambda_{10}}$, $\tilde{\lambda_{01}}$, 그리고 $\tilde{\lambda_{11}}$ 라 한다면 우도비 검정통계량의 형태는

$$T_n=-2((l(x,y;BZIP_{II}(\tilde{p_{00}},\tilde{p_{10}},\tilde{p_{01}},\tilde{\lambda_{10}},\tilde{\lambda_{01}},\tilde{\lambda_{11}}))$$

$$-l(x, y; BZIP_{II}(\tilde{p_{00}}, \tilde{p_{10}}, \tilde{p_{01}}, \tilde{\lambda_{10}}, \tilde{\lambda_{01}}, \tilde{\lambda_{11}}), \tilde{c}, BZIP_I(\tilde{p_{00}}, \tilde{p_{10}}, \tilde{p_{01}}, \tilde{\lambda_{10}}, \tilde{\lambda_{01}}, \tilde{\lambda_{11}})))$$

이 되고 이는 대표본에서 점근적으로 자유도 2인 카이제곱분포를 따르게 된다. 그러므로 위 검정통계량의 값이 크면 귀무가설이 기각되고 이는 변화시점이 존재함을 의미한다.

4. 점추정

대립가설 H_1이 참이라면, 변화시점과 6개의 모수 그리고 변화시점 이후와 이전의 불량품이 전혀나타나지 않을 확률의 차이, 즉 식 (3)에 관한 모수들을 추정해 볼 필요가 있다. 먼저 변화시점에 대한 추정량은 변화시점 이전과 이후의 표본들이 평균과의 제곱합이 최소가 되는 최소제곱추정량을 제시하려 한다. 즉, 식 (4)를 최소로 하는 양의 정수 $j(1 \le j < n)$ 값을 변화시점에 대한 추정량 $\hat{c}$ 로 설정하였다.

$$\begin{aligned} \textstyle\sum_{i=1}^{i} \quad & [(x_i - \sum_{i=1}^{j} x_i/j)^2 + (y_i - \sum_{i=1}^{j} y_i/j)^2] + \\ \textstyle\sum_{i=j+1}^{n} \quad & [(x_i - \sum_{i=j+1}^{n} x_i/(n-j))^2 + (y_i - \sum_{i=j+1}^{n} y_i/(n-j))^2]. \end{aligned} \tag{4}$$

대립가설이 참이라면, 변화시점 이전의 분포는 $BZIP_{II}$이고 이후의 분포는 $BZIP_I$이다. 변화시점 이전의 분포에서, 모수들은 영-빈도법 (methods of zero-frequency)에 의해 추정가능하다. 이는 $BZIP_{II}$ 분포가 영이 과잉 관측되고 적률법으로 쉽게 유도되지 않기 때문이다. 영-빈도법은 관측치가 영인 빈도수를 이용하여 추정하는 방법이다. 그러나 변화시점 이후의 분포 $BZIP_I$에서의 모수들은 적률법으로 쉽게 추정되어 진다. 먼저 영-빈도법을 이용한 $BZIP_{II}$의 모수 추정은 다음과 같다. $f_{ij}(i, j = 0, 1, 2)$를 표본 $(X = i, Y = j)$의 총 표본에 대한 빈도율이라 하자. 예를 들면 f_{11}는 총 표본수에 대한 관측치 $(1, 1)$의 도수비율을 의미한다. 그리고

$$f_{.0} = \sum_{k_1=1}^{\infty} f_{k_1 0}, \quad f_{0.} = \sum_{k_2=1}^{\infty} f_{0 k_2}, \quad f_{..} = \sum_{k_1=1}^{\infty}\sum_{k_2=1}^{\infty} f_{k_1 k_2}$$

라 하자. 그리하면 $BZIP_{II}$의 pmf를 이용하면 다음과 같은 식을 얻을 수 있다.

$$f_{k_1 0} \doteq (p_{10} + p_{11})\lambda_{10}^{k_1} e^{-\lambda}/k_1!, \quad k_1 = 1, 2, \cdots \tag{5}$$

$$f_{0 k_2} \doteq (p_{01} + p_{11})\lambda_{01}^{k_2} e^{-\lambda}/k_2!, \quad k_2 = 1, 2, \cdots \tag{6}$$

$$f_{21} \doteq p_{11} e^{-\lambda}(\lambda_{10}^2 \lambda_{01}/2 + \lambda_{11} + \lambda_{10}) \tag{7}$$

$$f_{12} \doteq p_{11} e^{-\lambda}(\lambda_{10}\lambda_{01}^2/2 + \lambda_{11} + \lambda_{01}). \tag{8}$$

만약 영에 관한 빈도수가 거의 없다면 이 표본들은 $BZIP$분포에 적합하지 않을 것이다. 그러므로 $f_{k_1,0} \neq 0, f_{0,k_2} \neq 0$ 라 가정해도 무리가 없을 것이다. 식 (5)-(8)로부터 λ_{ij}들을 추정하면

$$\begin{aligned}
\hat{\lambda_{10}} &= Mean_{k_1 \geq 1}\{(k_1+1)f_{k_1+1,0}/f_{k_1,0}\} \\
\hat{\lambda_{01}} &= Mean_{k_2 \geq 1}\{(k_2+1)f_{0,k_2+1}/f_{0,k_2}\} \\
\hat{\lambda_{11}} &= \hat{\lambda_{10}}\hat{\lambda_{10}}/2
\end{aligned}$$

이 된다. $BZIP_{II}$분포에서

$$\begin{aligned}
f_{0.} &= (1-p_{00}-p_{10})(e^{-\lambda_1}-e^{-\lambda}) \\
f_{.0} &= (1-p_{00}-p_{01})(e^{-\lambda_2}-e^{-\lambda}) \\
f_{..} &= (1-p_{00}-p_{10}-p_{01})(1-e^{-\lambda_1}-e^{-\lambda_2}+e^{-\lambda})
\end{aligned}$$

를 얻을 수 있고 이를 이용하면 다음 추정량을 얻을 수 있다.

$$\begin{aligned}
\hat{p_{01}} &= \frac{f_{0.}}{E_1-E_3} - \frac{f_{..}}{1-E_1-E_2+E_3}, \\
\hat{p_{00}} &= 1-\hat{p_{01}} - \frac{f_{.0}}{E_2-E_3}, \\
\hat{p_{10}} &= 1-\hat{p_{00}} - \frac{f_{0.}}{E_1-E_3}, \\
\hat{p_{11}} &= 1-\hat{p_{10}}-\hat{p_{01}}-\hat{p_{00}},
\end{aligned}$$

여기에서, $E_1 = e^{-(\lambda_{10}+\lambda 11)}$, $E_2 = e^{-(\lambda 01+\lambda 11)}$, $E_3 = e^{-(\lambda_{10}+\lambda 01+\lambda 11)}$.

한편 변화시점 이후의 분포 $BZIP_I$에서의 모수추정은 적률법을 이용하는데 이는 참고문헌 [6, pp.32 − 33]에서 이변량으로 적용하면 다음과 같이 쉽게 유도된다.

$$\begin{aligned}
\check{\lambda_1} &= \frac{\sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i} - 1, \quad \check{\lambda_2} = \frac{\sum_{i=1}^n y_i^2}{\sum_{i=1}^n y_i} - 1, \\
\check{\lambda_{11}} &= \frac{(\sum_{i=1}^n x_i^2 y_i + \sum_{i=1}^n x_i y_i^2)\check{\lambda_1}\check{\lambda_2}/\sum_{i=1}^n x_i y_i - \check{\lambda_1}\check{\lambda_2}(2+\check{\lambda_1}+\check{\lambda_2})}{2(\check{\lambda_1}+\check{\lambda_2})+2-(\sum_{i=1}^n x_i^2 y_i + \sum_{i=1}^n x_i y_i^2)/\sum_{i=1}^n x_i y_i},
\end{aligned}$$

그리고

$$\check{p_{11}} = \frac{\sum_{i=1}^n x_i y_i/n}{\check{\lambda_1}\check{\lambda_2}+\check{\lambda_{11}}},$$

$$
\begin{aligned}
\check{p_{10}} &= \frac{\sum_{i=1}^{n} x_i/n}{\check{\lambda_1}} - \check{p_{11}}, \\
\check{p_{01}} &= \frac{\sum_{i=1}^{n} y_i/n}{\check{\lambda_2}} - \check{p_{11}}.
\end{aligned}
$$

물론 여기에서도 $\lambda_1 = \lambda_{10} + \lambda_{11}$, $\lambda_2 = \lambda_{01} + \lambda_{11}, \lambda = \lambda_{10} + \lambda_{01} + \lambda_{11}$ 이다.

참고 문헌

1. 김경무. (1998). 변화시점이 있는 영과잉-포아송모형, 통계이론방법연구, 9, 1, 1-9.

2. 김경무, 이성호, 김종태. (1998). 이변량 영과잉-포아송분포의 적률, 통계이론방법연구, 9, 1, 47-56.

3. Cohen, A. C. (1963). Estimation in Mixtures of Discrete distributions, *Proceedings of the International Symposium on discrete Distributions, Montreal*, 373-378.

4. Johnson, N. L., Kotz, S. (1969). *Distributions in Statistics: Discrete Distributions*, Boston: Houghton Mifflin.

5. Lambert, Diane. (1992). Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing, *Technometrics*, 34, 1-14.

6. Li, C. S., Lu, J. C., Park, J. H., Kim, K. M., Brinkly, P. A., Peterson, J. P. (1999). Multivariate Zero-Inflated Poisson Models And Their Applications, *Technometrics*, 41, 1, 29-38.

7. Powell, M. J. D. (1964). An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives, *Computer Journal*, 7, 155-162.

8. Singh, S. N. (1963). A Note on Inflated Poisson Distribution, *Journal of the Indian Statistical Association*, 1, 140-144.

9. Yip, P. (1988). Inference About the Mean of a Poisson Distribution in the Presence of a Nuisance Parameter, *Australian Journal of Statistics*, 30, 299-306.

Inferences for the Changepoint in Bivariate Zero-Inflated Poisson Model

Kyungmoo Kim [3]

Abstract

Zero-Inflated Poisson distributions have been widely used for defect-free products in manufacturing processes. It is very interesting to check the shift after the unknown changepoint. If the defectives are caused by the two different types of factor, we should use bivariate zero-inflated model. In this paper, likelihood ratio tests were used to detect the shift of changes after the changepoint. Some inferences for the parameters in this model were made.

Key Words and Phrases: Zero-inflated poisson model

[3]Professor, Department of Statistics, Taegu University, Kyungbuk 712-714, Korea