응용통계연구 제 9권 1호
1996년 3월 pp. 111-124.

주변분포가 음이항 분포를 따르는 INAR(1)모형에서 추정량의 점근분포

김 회 영[1)], 박 유 성[2)]

요 약

본 논문은 비음의 정수값을 가지는 시계열 모형중 시계열의 상관관계가 연속형 AR(1) 모형과 비슷한 행태를 가지는 INAR(1)(Integer Valued Autogressive of order 1) 모형을 고려하고 있다. 주변분포가 음이항분포를 따르는 INAR(1)모형에 포함된 모수의 다양한 추정량을 도출하고, 이 추정량들의 점근분포를 유도하였다. 또한, 추정량들의 비교를 위하여 모의실험을 실시한 결과 본 논문에서 제시한 통계량이 Klimko and Nelson(1978)이 제시한 통계량보다 우수하다는 것을 볼 수 있다. 응용으로써 M/M/1 대기행렬과정에서의 모수를 추정하였다.

1. 서론

관공서등에서 기다리고 있는 민원인의 숫자, 교통신호등에서 대기하고 있는 차량의 숫자, 자동전화기의 전화통화량, 인구추이등은 시간의 흐름에 따라 관찰되는 정수값들을 취하여, 그들은 서로간에 독립이라고 할 수 없다. 이와 같이 상관관계를 가지는 정수값계열의 모형화는 최근 수년동안 많은 연구자들의 관심분야로 떠오르고 있다.

Mckenzie(1986)는 주변분포가 음이항분포와 기하분포를 따를때의ARMA 모형을 제시하였고, Grave and Lewis(1980), Jacobs and Lewis(1977), Lawrence(1982) 등은 주변분포로 감마와 지수분포를 가지는 정상성 과정을 제안하였다. 또한 Alzaid and Al-Osh(1990), Mckenzie(1988)는 Poisson ARMA 모형에 대한 성질들을 제시하였고, 최근에 Aly and Bouzar(1994)가 Poisson geometric, Negative Binomial, Poission Logarithmic ARMA 모형을 제안하였으며, Park and Kim(1994)은 주변분포가 음이항분포를 따르는 INAR(1) 모형에서의 표본자기공분산 함수의 점근분포를 구하였다.

INAR(1) 모형을 정의하기 전에 Stueutal and Van Harn(1979)에 의해 제시된 연산자를 소개한다. 비음의 정수값을 갖는 확률변수 X와 임의의 $\alpha \in [0,1]$ 에 대하여 $\circ$ 연산자는 다음과 같이 정의된다.

$$\alpha \circ X = \sum_{i=1}^{X} Y_i$$

---

1) (136-701) 서울 특별시 성북구 안암동 5-1, 고려대학교 정경대학 통계학과 이학석사.
2) (136-701) 서울 특별시 성북구 안암동 5-1, 고려대학교 정경대학 통계학과 조교수.

단, Y_i 는 성공의 확률이 α 인 i.i.d 베르누이 시행이고, Y_i 는 X와 독립이다.

이때 Al-Osh and Alzaid(1987)에 의해 정의된 INAR(1) 모형은 다음과 같다.

$$X_t = \alpha \circ X_{t-1} + \varepsilon_t\ ,\quad t = 0, \pm 1, \pm 2, \cdots \tag{1.1}$$

여기에서 $\alpha \in [0, 1]$이고, ε_t는 비음의 정수값을 갖는 i.i.d한 확률변수로서 이의 2차 적률이 존재한다고 가정한다.

Klimko and Nelson(1978)는 적절한 조건을 만족하는 확률과정의 조건부 최소제곱추정량(conditional least squares estimatior, C.L.S)를 제시하였고, Al-Osh and A.Alzaid는 이 결과의 단순한 응용으로써, 모형 (1.1)에서 ε_t 의 분포가 Poisson일 때의 모수에 대한 추정치들과 점근분포를 구하였다. 또한 Park and Kim(1995)은 X_t의 분포가 모수 r, p를 갖는 음이항분포(NB(r,p))일 때의 표본평균과 표본자기공분산함수들의 점근분포를 제시하였다.

본 논문에서는 X_t의 주변분포가 NB(r,p)를 따를 때의 모형(1.1)에서 모수벡터 (α, r, p)의 추정량들과 그 점근분포를 알아보고자 한다. 제 2절에서는 Klimko and Nelson이 제시한 방법에 의해 C.L.S 추정량, Park and Kim의 결과를 이용한 Yule-Walker 추정량을 구하고 각 추정량의 극한분포를 구한다. 특히 Yule-Walker 추정량의 부산물로써 정상 M/M/1 대기행렬에로의 응용을 고려하고자 한다. 제 3절에서는 모의실험을 통하여 제 2절에서 제시한 추정량들을 비교 검토한다.

2. 추정량들과 점근분포의 유도

2. 1 조건부 최소제곱 추정량

모형 (1.1)에서 X_t의 분포가 NB(r,p)를 따를 때 ε_t의 p.g.f (probablity generating function)와 기대값, 분산은 아래와 같다. (Park and Kim (1995))

$$P(t) = \left(\frac{1-q(1-\alpha+\alpha t)}{1-qt} \right)^r$$

$$E(\varepsilon_t) = \frac{rq(1-\alpha)}{p}$$

$$Var(\varepsilon_t) = \frac{rq}{p^2}(1-\alpha)(1+\alpha q)$$

한편 Klimko and Nelson(1978)에 의해 전개된 C.L.S 추정량은 다음 식

$$Q_n(\theta) = \sum_{t=1}^{n}(X_t - E(X_t|X_{t-1}))^2 , \quad \theta = (\alpha, r, p)$$

을 최소화시키는 추정량이 된다.

모형(1.1)로 부터 $E(X_t| X_{t-1}) = \alpha X_{t-1} + \frac{rq}{p}(1-\alpha)$ 이므로,

$$Q_n(\theta) = \sum_{t=1}^{n} [\ X_t - (\alpha X_{t-1} + \frac{rq}{p}(1-\alpha))]^2 \tag{2.1.1}$$

이 된다

그러나, 식 (2.1.1) 을 최소화 시키는 $\hat{\alpha}, \hat{p}, \hat{r}$ 는 유일하지 않게 됨을 쉽게 알 수 있다. 따라서, 우리는 모수벡터 (α, r, p)를 조정하여 r이 알려져 있다고 가정하고 다음과 같이 전개한다. 여기서 $\alpha \equiv \lambda_1$, $rq(1-\alpha)/p \equiv \lambda_2$ 라고 하면 식 (2.1.1)로부터

$$H_n(\lambda) = \sum_{t=1}^{n}(X_t - \lambda_1 X_{t-1} - \lambda_2) \tag{2.1.2}$$

이 된다.

식 (2.1.2) 을 최소화시키는 $\widehat{\lambda_1}$, $\widehat{\lambda_2}$는

$$\widehat{\lambda_1} = \frac{n\sum_{t=1}^{n} X_{t-1}X_t - (\sum_{t=1}^{n} X_{t-1})(\sum_{t=1}^{n} X_t)}{n\sum_{t=1}^{n} X_{t-1}^2 - (\sum_{t=1}^{n} X_{t-1})^2}$$

$$\widehat{\lambda_2} = \frac{1}{n}(\sum_{t=1}^{n} X_t - \widehat{\lambda_1}\sum_{t=1}^{n} X_{t-1})$$

이 됨을 알 수 있다.

그런데, $g(\lambda_1, \lambda_2, X_{t-1}) = E(X_t| X_{t-1})$로 정의할 때 함수 g , $\partial g/\partial \lambda_1$, $\partial g/\partial \lambda_2$, $\partial^2 g/\partial \lambda_1 \ \ \partial \lambda_2$는 모두 Klimko and Nelson(1978)의 Theorem (3.1) 조건을 만족시킨다. 그러므로, $\widehat{\lambda_1}$, $\widehat{\lambda_2}$ 은 모두 강일치(strongly consistent) 추정량이 된다. 더우기 $E(X_t^3) < \infty$, $E(\varepsilon_t^3) < \infty$ (Park and Kim (1995))이므로, Klimko and Nelson의 Theorem (3.2) 에 의하여

$$\sqrt{n}\left(\begin{pmatrix}\widehat{\lambda_1}\\ \widehat{\lambda_2}\end{pmatrix} - \begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}\right) \longrightarrow AN(0, W)$$

이 된다. 여기서, $W=(w_{ij}) \quad i,j=1,2$

$$w_{11}=\alpha(1-\alpha)p(1+q)\frac{1}{rq}+(1-\alpha^2)$$

$$w_{12}=w_{21}=-(1-\alpha)q\left[\ \alpha+\frac{r}{p}(1+\alpha)\right] \tag{2.1.3}$$

$$w_{22}=\frac{rq}{p}(1-\alpha)\left[\ \alpha q+\frac{rq}{p}(1+\alpha)+\frac{1}{p}(1+\alpha q)\right]$$

이제, (2.1.3)을 이용하여 (α, p)의 추정량 $\hat{\alpha}^L$, $\hat{p}^L$ 들과 그의 점근분포를 구하도록 하자. $k_1(\lambda_1,\lambda_2)=\lambda_1$, $k_2(\lambda_1,\lambda_2)=(r\lambda_1-r)/(r\lambda_1-\lambda_2-r)$ 로 두자. $\lambda_1=\alpha$, $\lambda_2=rq(1-\alpha)/p$ 이므로 $k_1(\lambda_1,\lambda_2)=\alpha$, $k_2(\lambda_1,\lambda_2)=p$ 가 된다. 따라서,

$$\hat{\alpha}^L = k_1(\ \hat{\lambda}_1\ ,\ \hat{\lambda}_2\ ) = \frac{n\sum_{t=1}^{n}X_tX_{t-1}-(\sum_{t=1}^{n}X_{t-1})(\sum_{t=1}^{n}X_t)}{n\sum_{t=1}^{n}X_{t-1}^{\ 2}-(\sum_{t=1}^{n}X_{t-1})^2}$$

$$\hat{p}^L = k_2(\hat{\lambda}_1,\hat{\lambda}_2) = \frac{r(\hat{\lambda}_1-1)}{r(\hat{\lambda}_1-1)-\hat{\lambda}_2} \tag{2.1.4}$$

을 얻게 된다.

식 (2.1.4)와 함수 k_1, k_2가 모두 연속인 1차 미분함수가 존재하므로,

$$\sqrt{n}\begin{pmatrix}\hat{k}_1-k_1\\ \hat{k}_2-k_2\end{pmatrix} \longrightarrow AN\left(\begin{pmatrix}0\\0\end{pmatrix}\ HWH'\right)$$

이 된다. 여기서, $H=\begin{pmatrix}\partial k_1/\partial\lambda_1 & \partial k_1/\partial\lambda_2\\ \partial k_2/\partial\lambda_1 & \partial k_2/\partial\lambda_2\end{pmatrix}$ 이다. 그러므로,

$$\sqrt{n}\left(\begin{pmatrix}\hat{\alpha}^L\\ \hat{p}^L\end{pmatrix}-\begin{pmatrix}\alpha\\ p\end{pmatrix}\right) \longrightarrow AN(0\ ,L\ ) \tag{2.1.5}$$

이고 L 는 다음과 같이 구성되어 있다.

$$l_{11}=\frac{\alpha}{q}(1-\alpha)p(1+q)+(1-\alpha^2)$$

$$l_{12} = l_{21} = pq\left(-\alpha - \frac{\alpha p}{rq} + \frac{\alpha p}{r} + \frac{p}{r}\right)$$

$$l_{22} = \frac{p^2 q(1+\alpha)}{r(1-\alpha)}$$

2. 2 율-워커(Yule-Walker) 추정량

2. 2. 1 r 이 알려져 있지 않을 경우

Park and Kim (1995)은 모형식 (1.1)에 대하여 $\overline{X}_n = \frac{1}{n}\sum_{t=1}^{n} X_t$, $\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(X_t - \overline{X})(X_{t+h} - \overline{X})$, $\hat{\rho}(h) = \hat{\gamma}(h)/\hat{\gamma}(0)$ 라고 정의할 때, 다음의 점근분포를 도출하였다.

$$\sqrt{n}\left(\begin{pmatrix} \overline{X}_n \\ \hat{\gamma}(0) \\ \hat{\rho}(1) \end{pmatrix} - \begin{pmatrix} \mu \\ \gamma(0) \\ \rho(1) \end{pmatrix}\right) \longrightarrow AN(0, P) \qquad (2.2.1)$$

여기서, $\gamma(h) = E(X_t - \mu)(X_{t+h} - \mu)$, $E(X_t) = \mu$, $\rho(h) = \gamma(h)/\gamma(0)$.
그리고, $P = (p_{ij})\ i, j = 1, 2, 3$

$$p_{11} = \frac{rq}{p^2}\frac{1+\alpha}{1-\alpha}$$

$$p_{12} = p_{21} = \frac{rq}{p^3(1-\alpha^2)}[\ (1+\alpha)^2 + (1+\alpha^2)q]$$

$$p_{13} = p_{31} = \frac{1+\alpha}{p} - \frac{q}{p}\frac{\alpha(1+\alpha+\alpha^2)}{1-\alpha^2}$$

$$p_{22} = \frac{rq}{p^4(1-\alpha^2)}[\ (1+\alpha)^2 + q\{4(\alpha^2+\alpha+1) + 2r(1+\alpha^2)\} + (1-\alpha)^2 q^2]$$

$$p_{23} = p_{32} = \frac{\alpha}{(1-\alpha^2)q^2}[\ (1-\alpha^2) - 2q\{(1-\alpha)^2 - r(1-\alpha^2)\}$$
$$+ q^2(3\alpha^2 - 4\alpha + 1)]$$

$$p_{33} = \frac{1}{rq}\alpha(1-\alpha)(1-q^2) + (1-\alpha^2)$$

이다.

이제 모수 (α, r, p)의 추정량 $(\hat{\alpha}^{Y_1}, \hat{r}^{Y_1}, \hat{p}^{Y_1})$를 구해보자. 아래와 같이 정의된 함수 h_1, h_2, h_3 를 고려하기로 하자.

$$h_1(\mu, r(0), \rho(1)) = \rho(1) = \alpha$$

$$h_2(\mu, r(0), \rho(1)) = \frac{\mu^2}{r(0)-\mu}$$

$$h_3(\mu, r(0), \rho(1)) = \frac{\mu}{r(0)}$$

그러면, $E(X_t) = \mu = rq/p$, $Var(X_t) = r(0) = rq/p^2$ 이므로, $h_1(\cdot) = \alpha$, $h_2(\cdot) = r$, $h_3(\cdot) = p$ 이 된다

그러므로, 추정량은 다음과 같이 얻어진다.

$$\hat{\alpha}^{Y} = h_1(\overline{X}_n, \hat{r}(0), \hat{\rho}(1)) = \hat{\rho}(1)$$

$$= \frac{\sum_{t=1}^{n}(X_t - \overline{X}_n)(X_{t+1} - \overline{X}_n)}{\sum_{t=1}^{n}(X_t - \overline{X}_n)^2}$$

$$\hat{r}^{Y_1} = h_2(\overline{X}_n, \hat{r}(0), \hat{\rho}(1))$$

$$= \frac{n\,\overline{X}_n^{\,2}}{\sum_{t=1}^{n}(X_t - \overline{X})^2 - n\,\overline{X}_n}$$

$$\hat{p}^{Y_1} = h_3(\overline{X}_n, \hat{\rho}(0), \hat{\rho}(1))$$

$$= \frac{n\,\overline{X}_n}{\sum_{t=1}^{n}(X_t - \overline{X}_n)^2}$$

실제응용에서는, r이정수이므로 $\hat{r}^{Y_1}$대신 [$\hat{r}^{Y_1}$]을 사용하는 것이 바람직하다. 즉 $\hat{r}^{Y_1}$을 초과하지 않는 최대의 정수값을 r의 추정치로 사용하는것이 타당하다.

식 (2.2.1)과 함수 h_1, h_2, h_3가 연속인 1차 미분함수가 있으므로 다음과 같은 결과를 도출하게 된다.

$$\sqrt{n}\left(\begin{pmatrix}\hat{\alpha}^{Y}\\ \hat{p}^{Y_1}\\ \hat{r}^{Y_1}\end{pmatrix}-\begin{pmatrix}\alpha\\ p\\ r\end{pmatrix}\right)\longrightarrow AN(0,A) \qquad (2.2.2)$$

$$A = (a_{ij}) \qquad i,j = 1,2,3.$$

여기에서, $a_{11} = \frac{1}{rq}\alpha(1-\alpha)(1-q^2) + (1-\alpha^2)$

$$a_{12} = \frac{p}{rq} - \frac{p}{r}\frac{\alpha(1+\alpha+\alpha^2)}{1-\alpha^2} + \frac{\alpha p}{rq(1-\alpha^2)}[\ 2q((1-\alpha)^2 - r(1-\alpha^2)) - (3\alpha^2+4\alpha+1)q^2]$$

$$a_{22} = \frac{2p^2(1+\alpha)}{r(1-\alpha)} + \frac{2p^2(1+\alpha^2)}{1-\alpha^2} + \frac{p^2q(1-\alpha)}{r(1+\alpha)}$$

$$a_{23} = \frac{p}{q^2(1-\alpha^2)}\left[\ q((\alpha^2+4\alpha+1) + (1+\alpha^2)(p+2r)) + q^2(1-\alpha)^2\right]$$

$$a_{33} = \frac{r(2-p)^2}{q^3}\frac{1+\alpha}{1-\alpha}[\ (1+\alpha)^2(2p-3) + q\{4(\alpha^2+\alpha+1) + 2(1+\alpha^2)(r+p-2)\} + q^2(1-\alpha)^2\,]$$

이다.

2. 2. 2 r 이 알려져 있을 경우

Klimko and Nelson의 결과를 이용한 C.L.S와 비교하기 위하여 r의 값이 알려져 있을 경우의 p와 α의 추정량과 그 점근분포를 고려하여 보자. 앞 절에서의 식 (2.2.1)을 이용하여 r의 값을 알 경우, 우리는 아래의 결과를 쉽게 구할 수 있다. 즉,

$$\sqrt{n}\left(\begin{pmatrix}\overline{X}_n\\ \hat{\rho}(1)\end{pmatrix}-\begin{pmatrix}\mu\\ \alpha\end{pmatrix}\right)\xrightarrow{d} AN(0,V) \qquad (2.2.3)$$

$$V = (v_{ij}) \quad i,j=1,2$$

여기에서 $v_{11} = p_{11}$, $v_{12} = p_{13}$, $v_{22} = p_{33}$ 이다. (식(2.2.1)참조)

2 .2. 1절에서와 같이 함수 f_1, f_2를 $f_1(\mu, \alpha) = r/(\mu + r)$, $f_2(\mu, \alpha) = \alpha$로 정의하면 $E(\overline{X}_n) = \mu = rq/p$ 이므로 $f_1(\cdot) = p$, $f_2(\cdot) = \alpha$ 이 된다. 따라서, p와 α의 추정량은 각각

$$\hat{p}^{Y_2} = f_1(\overline{X}_n, \hat{\rho}(1)) = \frac{r}{\overline{X}_n + r}$$

$$\hat{\alpha}^{Y} = f_2(\overline{X}_n, \hat{\rho}(1)) = \hat{\rho}(1)$$

이 된다.

이제 2. 2. 1절에서와 동일한 방법으로 $\hat{p}^{Y_2}$, $\hat{\alpha}^{Y_2}$ 의 점근분포를 구하면 다음과 같다.

$$\sqrt{n}\left(\begin{pmatrix} \hat{\alpha}^{Y} \\ \hat{p}^{Y_2} \end{pmatrix} - \begin{pmatrix} \alpha \\ p \end{pmatrix}\right) \longrightarrow AN(0, B) \qquad (2.2.4)$$

$$B = (b_{ij}) \quad , \; i, j = 1, 2$$

여기서, $b_{11} = p_{33}$, $b_{12} = b_{21} = -\frac{p}{r}[\,1 + \alpha - q\alpha(1 + \alpha + \alpha^2)\,]$, $b_{22} = \frac{p^2 q}{r}\frac{1+\alpha}{1-\alpha}$이다.

이상에서 살펴 본 바와 같이 조건부 최소제곱추정량과 r이 알려져 있을 경우 율-워커 추정량의 점근분포는 식(2.1.5)와 식(2.2.4)에서 볼 수 있듯이 같다는 것을 알 수 있다. 그러나, r이 알려져 있지 않을 경우에는 CLS추정량과 $\hat{p}^{Y_2}$를 사용할 수 없고, 단지 $\hat{p}^{Y_1}$만을 사용할 수 있다. 따라서, $\hat{p}^{Y_1}$의 MSE가 r 이 알려져 있을 때의 추정량들의 MSE와 크게 차이가 나타나지 않는다면 실제적인 문제에서는 $\hat{p}^{Y_1}$, $\hat{\alpha}^{Y}$를 사용해도 큰 무리가 없을 듯하다.

3. 응용

M/M/1 대기행렬과정에서 X_t가 t 시점에서의 size이면, X_t의 분포는 모수 $\eta(=\lambda_1/\lambda_2)$, λ_1는 도착율, λ_2는 서비스 공급율)을 갖는 기하분포라는 사실이알려져 있다. 또한, M/M/1 대기행렬과정에서 자기상관함수 ρ(h)가 점근적으로

$$\exp\left(-\lambda_2 h \frac{(1-\eta)^2}{\eta}\right)$$

를 따르게 됨이 밝혀져 있다. (Morse(1955), Steudel and Wu(1977))

그러므로, 모형 (1.1)에서 r = 1, $\alpha = \exp(-\lambda_2(1-\eta)^2/\eta)$, $p = \eta$ 로 두면 X_t의 분포는 M/M/1 대기행렬과정에서의 t 시점에서의 size가 된다. 그러므로, 이제 우리는 $\widehat{\lambda_1}$, $\widehat{\lambda_2}$ 를 구할 수 있다.

다음과 같은 함수를 정의하여 보자.

$$k_1(\gamma(0), \rho(1)) = -\frac{\log \rho(1)}{\gamma(0)^2}$$

$$k_2(\gamma(0), \rho(1)) = -\frac{(1+\gamma(0))}{\gamma(0)^2} \log \rho(1) \quad .$$

그러면, $\gamma(0) = p/q^2$, $\rho(1) = \alpha$ (Park and Kim(1995)) 이므로, $k_1(\gamma(0)\ , \rho(1)) = \lambda_1$, $k_2(\gamma(0), \rho(1)) = \lambda_2$이 된다.

그러므로, 추정량은 다음과 같다.

$$\widehat{\lambda_1} = k_1(\hat{\gamma}(0),\ \hat{\rho}(1)) \quad = -\frac{\log \hat{\rho}(1)}{\hat{\gamma}(0)^2}$$

$$\widehat{\lambda_2} = k_2(\hat{\gamma}(0),\ \hat{\rho}(1)) \quad = -\frac{(1+\hat{\gamma}(0))}{\hat{\gamma}(0)^2} \log \hat{\rho}(1)$$

여기서, $\hat{\gamma}(0) = \frac{1}{n} \sum_{t=1}^{n} (X_t - \overline{X})^2$, $\hat{\rho}(1) = \dfrac{\sum_{t=1}^{n-1} (X_t - \overline{X})(X_{t+1} - \overline{X})}{\sum_{t=1}^{n} (X_t - \overline{X})^2}$

또한 제 2절에서와 같은 방법으로 $\widehat{\lambda_1}$, $\widehat{\lambda_2}$ 의 점근분포를 아주 쉽게 구할 수 있다. 그러나, 공분산행렬이 상당히 복잡하므로 여기서는 생략하기로 한다.

지금까지 INAR(1)모형을 소개하고, 주변분포가 음이항분포를 따를 때의 모형에 포함된 모수들의 다양한 추정량들과 그의 점근분포를 구하여 보았다.

제 4절의 모의실험의 결과에서 나타난 점은 Klimko and Nelson(1978)이 제시한 C.L.S 추정량이 갖는 제한(r을 모를때)을 Yule-Walker 추정량이 극복하였다는 점이다.

4. 모의실험

제 2절에서 구한 모수 (α, p)의 2가지 추정치들의 장단점을 비교 검토하기 위하여 모의실험을 실시하였다. 실험의 편의상 r=3 으로 설정하였다. X_t의 기대값이 $rp/q\,(=3p/q\,, q=1-p)$이므로, 초기치 X_0를 [$3q/p$]로 두고 실험하였다. 그리고 Park and Kim(1994)에 의해

$$\varepsilon_t = \begin{cases} 0 & \text{with} \quad \text{probability} \quad \alpha^r \\ N(j, p) & \text{with} \quad \text{probability} \quad \binom{r}{r-j}\alpha^{r-j}(1-\alpha)^j \ , \ 1 \le j \le r \end{cases}$$

이므로 모의실험은 α, p의 값을 0.1부터 0.9까지 0.2씩 증가시켜 각각 25가지의 다양한 조합에 대하여, 표본의 크기를 $n=50, 100, 150, 200$으로 증가시켜가면서 500번 반복하였다. 다음의 <표1> ($n=50$)과 <표2> ($n=200$)는 이중 일부의 결과를 보여준다.

<표 1> n=50 일때의 α, p의 다양한 조합에 따른 각 추정치들의 M.S.E

p	α	M. S. E. (p) CLS	M. S. E. (p) Y-W (r을 모름)	M. S. E. (p) Y-W (r을 앎)	M. S. E (α) CLS	M. S. E (α) Y-W
0.1	0.1	.00008	.00090	.00007	.01935	.01861
	0.3	.00054	.00701	.00051	.01888	.01820
	0.5	.00099	.01933	.00094	.01951	.01865
	0.7	.00128	.04063	.00123	.02186	.02104
	0.9	.00068	.04419	.00065	.02319	.02246
0.3	0.1	.00011	.00100	.00011	.02006	.01988
	0.3	.00084	.00896	.00079	.01997	.01964
	0.5	.00151	.02524	.00142	.02296	.02254
	0.7	.00173	.04464	.00166	.02074	.02044
	0.9	.00103	.05024	.00099	.03128	.03105
0.5	0.1	.00022	.00176	.00020	.01757	.01846
	0.3	.00132	.01429	.00122	.02015	.02089
	0.5	.00241	.03955	.00210	.02206	.02250
	0.7	.00308	.06207	.00273	.02053	.02065
	0.9	.00175	.06629	.00155	.04013	.04012
0.7	0.1	.00046	.00409	.00040	.01321	.01535
	0.3	.00285	.03502	.00251	.01753	.01967
	0.5	.00499	.09221	.00442	.01749	.01959
	0.7	.00557	.16843	.00499	.02095	.02314
	0.9	.00378	.27820	.00290	.04494	.04597
0.9	0.1	.00307	.03556	.00109	.01362	.01778
	0.3	.02527	.30372	.00666	.01601	.02230
	0.5	.02794	.82650	.01251	.01747	.02323
	0.7	.02525	1.44780	.01241	.01702	.02199
	0.9	.05920	7.49031	.00583	.06342	.06887

<표 1>로부터 다음과 같은 사실들을 알 수 있다.

모수 p를 추정하는 문제에 있어서는, r을 알고 있을때 Yule-Walker 추정량인 $\hat{p}^{Y_2}$의 M.S.E

가 전체 20경우중에서 모두 CLS추정량인 $\hat{p}^L$의 M.S.E보다 작다.

또한, r을 모르는 경우일때도 $\hat{p}^{Y_1}$의 M.S.E는 대부분 $\hat{p}^L$의 M.S.E보다 약 10배의 크기로 크다. 그 절대치는 $10^{-4} \sim 10^{-1}$의 범위안에 있다. 그러나, α값이 0.7이상이고, p값이 0.7이상인 경우는 나머지 두 추정량에 비하여 약 100배의 크기로 커짐을 알 수 있다. 따라서 α값이 0.7이상, p값이 0.7이상인 때를 제외하면 r을 모르는 경우에도 Yule-Walker 추정량을 사용해도 큰 무리는 없을 듯하다.

모수 α를 추정하는 경우에는 α값이 0.3이하인 경우는 Yule-Walker 추정량인 $\hat{\alpha}^Y$의 M.S.E가 CLS추정량인 $\hat{\alpha}^L$의 M.S.E보다 작다. 그러나, α값이 0.5이상인 경우 p값의 변화를 고려한 15경우중 14경우에서 $\hat{\alpha}^L$의 M.S.E가 $\hat{\alpha}^Y$의 M.S.E보다 작다.

<표 2> n=200 일때의 α, p의 다양한 조합에 따른 각 추정치들의 M.S.E

p	α	M. S. E. (p) CLS	M. S. E. (p) Y-W (r을 모름)	M. S. E. (p) Y-W (r을 앎)	M. S. E (α) CLS	M. S. E (α) Y-W
0.1	0.1	.00002	.00018	.00002	.00480	.00476
	0.3	.00013	.00144	.00013	.00503	.00499
	0.5	.00024	.00365	.00024	.00445	.00442
	0.7	.00029	.00779	.00029	.00499	.00495
	0.9	.00016	.01114	.00016	.00624	.00620
0.3	0.1	.00003	.00019	.00003	.00449	.00450
	0.3	.00020	.00152	.00020	.00578	.00574
	0.5	.00038	.00466	.00037	.00537	.00538
	0.7	.00051	.00857	.00050	.00540	.00536
	0.9	.00024	.01342	.00023	.00743	.00740
0.5	0.1	.00005	.00029	.00005	.00334	.00336
	0.3	,00033	.00210	.00032	.00421	.00424
	0.5	.00064	.00676	.00063	.00481	.00491
	0.7	.00073	.01396	.00072	.00530	.00533
	0.9	.00040	.01604	.00039	.00761	.00772
0.7	0.1	.00010	.00051	.00009	.00277	.00289
	0.3	.00053	.00400	.00052	.00278	.00292
	0.5	.00126	.01290	.00121	.00326	.00342
	0.7	.00150	.02753	.00146	.00432	.00445
	0.9	.00086	.03099	.00084	.00643	.00661
0.9	0.1	.00037	.00258	.00034	.00171	.00197
	0.3	.00229	.02584	.00207	.00187	.00216
	0.5	.00400	.05954	.00373	.00194	.00230
	0.7	.00415	.10932	.00383	.00214	.00236
	0.9	.00284	.08458	.00216	.00885	.00941

<표 2>로 부터 알 수 있는 점들은 다음과 같다.

표본의 크기가 n=200 으로 커짐에 따라 α, p각 추정치들의 M.S.E가 작아짐은 물론이다.

p의 추정에있어서는 n=50 일 때와 전체적인 양상은 같다. r을 모르는 경우일 때, Yule-Walker 추정량의 M.S.E의 절대치가 나머지 두 추정량의 M.S.E에 비해 여전히 10배 정도의 크기로 크지만, 그 절대치가 $10^{-4} \sim 10^{-2}$의 범위에 있으므로 극히 양호하다고 할 수 있다.

그리고, α의 추정에 있어서는 α=0.1 일 때는 Yule-Walker 추정량의 M.S.E가 보다 작으며, α값이 0.5 이상일 때는 CLS추정량의 M.S.E가 보다 작다.

참고 문헌

[1] Al-Osh, M. A. and Alzaid A. A. (1987). First-order integer-valued autorgressive (INAR(1)) process, *Journal of Times Series Analysis,* Vol. 8, No. 3, 261-275.

[2] Aly, A. A. and Bouzar, N. (1994). On some integer-valued autoregressive moving average models, *Journal of Multivariate Analysis,* Vol. 50, 132-151.

[3] Alzaid, A. A. and Al-Osh, M. (1990). An integer valued pth-order autoregressive struucture(INAR(P)) process, *Journal of Applied Probability,* Vol. *27,* 314-324.

[4] Gaver, D. P. and Lewis, P. A. W. (1980). First order autocoregressive gamma-sequences and point processes, *Advanced in Applied Probability,* Vol. *12,* 727-745.

[5] Jacobs, P. A. and Lewis, P. A. W. (1977). A mixed autoregressive moving average exponential sequences and point process(EARMA(1,1)), *Advanced in Probability,* Vol. 9, 87-104.

[6] Klimko. L. A. and Nelson. P. I (1978). On conditional least squares estimation for stochastic processes, *Annals of Statistics,* Vol. 6, 629-642.

[7] Lawrence, A. J. (1982). The innovation distribution of a gamma distributed autogressive process, *Scandinavian Journal of Statistics,* Vol. 9, 234-236.

[8] Lawrence, A. J. and Lewis, P. A. W. (1980). The exponential autoregressive moving average EARMA(p,q) process, *Journal of the Royal Statistical Society, B42,* 150-161.

[9] Mckenzie, E. (1986). Autoregressive moving-average processes with negative-binomial and geonetric marginal distribution, *Advances in Applied Probability,* Vol. 18, 679-705.

[10] Mckenzie, E. (1988). Some ARMA models for dependent sequences of Poission counts, *Advances in Applied Probability,*Vol. 20, 822-835.

[11] Morse, P. M. (1955). Stochastic properties of wating times, *Operations Research* Vol. 3, No. 3, 255-261.

[12] Park, Y. S. and Kim K. (1994). On autocovariance function in INAR(1) process with negative binomial marginals *unpublished.*

[13] Steutel. F. W. and Van Harn. K. (1979). Discrete analogues of self-decomposabilty and stability, *Annals of Probability,* Vol. 7, No. 5, 893-899.

[14] Steudel, H. J. and Wu, S. M. (1977). A times series approach to queueing systems with applications for modelling jobs-shop in process inventories, *Management Science,* Vol. *23,* 745-755.

Asymptotic Distribution of Estimatior in INAR(1) Process with Negative Binomial Marginal

Hee Young Kim[1] and You Sung Park[2]

Abstract

In this paper, we consider the first-order integer valued autoregressive(INAR(1)) model where correlation structure is similar to that of the continuous valued AR(1) process. Several methods for estimating the parameters of the INAR(1) process with negative binomial marginal are discussed. We derive asymptotic distributions of these estimators. The results of a simulation study for these estimators methods show that the estimator which we present in this paper is better than the estimator which Klimko and Nelson(1978) presented. As an application we considered the estimator of M/M/1 queue length.

1) Dept. of Statistics, Korea University. Anamdong 5-1, Sungbukku, Seoul, Korea. 136-701.
2) Dept. of Statistics, Korea University. Anamdong 5-1, Sungbukku, Seoul, Korea. 136-701.