김일성종합대학학보 주체107(2018)년
수 학 제64권 제3호

령－수정된 기하분포를 가지는 우연환경 옹근수값 1차자기회귀과정

유강혁, 김철광

론문에서는 현실에서 많이 제기되는 옹근수값시계렬자료들을 모형화하는데 리용할수 있는 부이항감소연산자에 의한 령－수정된 기하분포를 가지는 우연환경 옹근수값1차자기회귀과정의 성질에 대하여 연구하였다. 선행연구에서는 부이항감소연산자에 의한 령－수정된 기하분포에 관한 옹근수값1차자기회귀과정[2]과 우연환경을 가지는 비정상옹근수값1차자기회귀과정[1]에 대하여 론의되였다.

1. RrZMGINAR(1)과정

정의 1 부아닌 옹근수값을 취하는 우연량 X의 확률함수가

$$P_k = P(X=k) = \begin{cases} \pi + (1-\pi)\dfrac{1}{1+\mu} & (k=0) \\ (1-\pi)\dfrac{\mu^k}{(1+\mu)^{k+1}} & (k=1,\ 2\cdots) \end{cases} \tag{1}$$

일 때 X는 령－수정된 기하분포에 따른다고 말하고 $\mathrm{X} \sim \mathrm{ZMG}(\pi,\ \mu)$으로 표시한다.[4] 여기서 $\mu>0$이고 $\pi \in (-1/\mu,\ 1)$이며 $\pi \in (-1/\mu,\ 0)$일 때에는 $P(X=0)$가 보통의 기하분포 $\mathrm{Geo}(\mu/(1+\mu))$에 비하여 작고 $\pi \in (0,\ 1)$일 때에는 크다. 따라서 π를 팽창파라메터라고 부른다. $\mathrm{ZMG}(0,\ \mu) = \mathrm{Geo}(\mu/(1+\mu))$가 성립한다.

정의 2 $\{Z_t,\ t \in \mathbf{I}\}$가 $E = \{1,\ 2,\ \cdots,\ r\}$에서 값을 취하는 마르꼬브사슬이라면 $\{Z_t,\ t \in \mathbf{I}\}$를 r개의 상태를 가지는 우연환경과정이라고 부른다. 여기서 r는 정의옹근수이다.[1] 여기서 E는 서로 다른 환경전부의 모임을 나타낸다.

정의 3 $\mathrm{X_t} \sim \mathrm{ZMG}(\pi,\ \mu)$인 부아닌 옹근수값우연과정 $\{X_t\}_{t=0}^{\infty}$가 다음의 회귀방정식

$$X_t = \alpha(Z_t) * X_{t-1} + \varepsilon_t \tag{2}$$

를 만족시킨다고 하자. 여기서 《*》은 부이항감소연산자로서

$$\alpha * X = \sum_{i=1}^{X} G_i$$

로 정의하며 $\alpha>0$이고 계수렬 $\{G_i\}_{i\ge1}$는 X와 독립인 $\mathrm{Geo}(\alpha/(1+\alpha))$에 따르는 독립동일분포하는 우연량렬이다. 또한 $\{\varepsilon_t\}_{t=0}^{\infty}$는 계수렬 $\{G_i(Z_t)\}_{i\ge1}$과 독립이며 $X_{t-l},\ l \ge 1$과도 독립인 우연량렬이다. 이때 $\{X_t\}_{t=0}^{\infty}$를 부이항감소연산자에 의한 령－수정된 기하분포를 가지는 우연환경 옹근수값1차자기회귀과정이라고 말하고 간단히 RrZMGINAR(1)로 표시한다.

보조정리 1 $\mathrm{X} \sim \mathrm{ZMG}(\pi,\ \mu)$라면 X의 확률생성함수는

$$\varphi_X(s)=\frac{1+\pi\mu(1-s)}{1+\mu(1-s)} \tag{3}$$

이다.

증명 X의 확률생성함수 $\varphi_X(s)$는 무한등비수렬의 합공식을 리용하면

$$\varphi_X(s)=\mathrm{E}(s^X)=\sum_{k\ge0}P_k s^k=\left[\pi+(1-\pi)\frac{1}{1+\mu}\right]s^0+\sum_{k\ge1}(1-\pi)\frac{\mu^k}{(1+\mu)^{k+1}}s^k=$$

$$=\pi+(1-\pi)\frac{1}{1+\mu}+(1-\pi)\sum_{k\ge1}\frac{(\mu s)^k}{(1+\mu)^{k+1}}=\frac{1+\pi\mu(1-s)}{1+\mu(1-s)}$$

라는것이 곧 나온다.(증명끝)

따름 $X\sim\mathrm{ZMG}(\pi,\ \mu)$이라면 X의 기대값과 분산은

$$\mathrm{E}(X)=\mu(1-\pi),\ \ \mathrm{Var}(X)=\mu(1-\pi)[1+\mu(1+\pi)]$$

이다.

보조정리 2 $X\sim\mathrm{ZMG}(\pi,\ \mu)$이고 $\alpha*X$의 계수렬 $G_i\sim\mathrm{Geo}\left(\frac{\alpha}{1+\alpha}\right),\ i\ge1$이라면 $\alpha*X$의 확률생성함수는

$$\varphi_{\alpha*X}(s)=\frac{1+\alpha(1+\pi\mu)(1-s)}{1+\alpha(1+\mu)(1-s)} \tag{4}$$

이다. 즉 $\alpha*X\sim\mathrm{ZMG}\left(\frac{1+\pi\mu}{1+\mu},\ \alpha(1+\mu)\right)$이다.

증명 $\{G_i\}_{i\ge1}$들이 독립동일분포에 따른다는것과 보조정리 1을 리용하면

$$\varphi_{\alpha*X}(s)=\mathrm{E}(s^{\alpha*X})=\mathrm{E}(\mathrm{E}(s^{\alpha*X}\mid X=k))=\sum_{k\ge0}\mathrm{E}(s^{\alpha*X}\mid X=k)P_k=$$

$$=\sum_{k\ge0}\mathrm{E}\left(s^{G_1+\cdots+G^k}\middle|X=k\right)P_k=\sum_{k\ge0}\varphi_{G_1}^kP_k=\left[\pi+(1-\pi)\frac{1}{1+\mu}\right]+$$

$$+\sum_{k\ge1}\left(\frac{1}{1+\alpha(1-s)}\right)^k(1-\pi)\frac{\mu^k}{(1+\mu)^{k+1}}=\frac{1+\alpha(1+\pi\mu)(1-s)}{1+\alpha(1+\mu)(1-s)}$$

가 성립하고 여기로부터 $\alpha*X\sim\mathrm{ZMG}\left(\frac{1+\pi\mu}{1+\mu},\ \alpha(1+\mu)\right)$라는것을 곧 알수 있다.(증명끝)

현실에서는 $X_{t-1}=x_{t-1}$이 주어지면 우연환경과정의 한걸음앞 실현값이 주어진다. 즉 $Z_t=z_t$라고 가정할수 있다.[2] 따라서 $\alpha(Z_t)=\alpha(z_t)=\alpha_t$로 표시하면 식 (2)는

$$X_t=\alpha_t*X_{t-1}+\varepsilon_t \tag{5}$$

로 쓸수 있다. 따라서 이제부터 식 (5)로 표시되는 $\{X_t\}_{t=0}^{\infty}$를 RrZMGINAR(1)과정이라고 한다.

보조정리 3 RrZMGINAR(1)과정의 ε_t의 확률생성함수는 $\mathrm{ZMG}\left(\frac{\alpha_t(1+\mu)}{\mu},\ \mu\right)$에 따르는 우연량의 확률생성함수와 $\mathrm{ZMG}\left(\frac{\pi\mu}{\alpha_t(1+\pi\mu)},\ \alpha_t(1+\pi\mu)\right)$에 따르는 우연량의 확률생성함수의 적으로 표시된다.

증명 ε_t 의 확률생성함수를 $\varphi_{\varepsilon_t}(s)$ 라고 하면 RrZMGINAR(1)과정에 대한 가정으로부터

$$\varphi_X(s)=\varphi_{\alpha_t * X}(s)\varphi_{\varepsilon_t}(s)$$

이다. 따라서 보조정리 1, 2를 리용하면

$$\varphi_{\varepsilon_t}(s)=\varphi_X(s)\frac{1}{\varphi_{\alpha_t * X}(s)}=\frac{1+\pi\mu(1-s)}{1+\mu(1-s)}\times\frac{1+\alpha_t(1+\mu)(1-s)}{1+\alpha_t(1+\pi\mu)(1-s)}=$$

$$=\frac{1+\alpha_t(1+\mu)(1-s)}{1+\mu(1-s)}\times\frac{1+\pi\mu(1-s)}{1+\alpha_t(1+\pi\mu)(1-s)}=\varphi_1(s)\varphi_2(s)$$

가 성립하고 여기서 $\varphi_1(s)$ 는 $\mathrm{ZMG}\left(\frac{\alpha_t(1+\mu)}{\mu},\ \mu\right)$, $\varphi_2(s)$ 는 $\mathrm{ZMG}\left(\frac{\pi\mu}{\alpha_t(1+\pi\mu)},\ \alpha_t(1+\pi\mu)\right)$ 에 따르는 우연량의 확률생성함수라는것을 알수 있다.(증명끝)

정리 1 RrZMGINAR(1)과정의 ε_t 의 확률함수는

$$P_{\varepsilon_t}(0)=P(\varepsilon_t=0)=\left(\alpha_t+\frac{1}{1+\mu}\right)\frac{1+\pi\mu}{1+\alpha_t(1+\pi\mu)} \tag{6}$$

$$P_{\varepsilon_t}(k)=\frac{(1+\pi\mu)[\mu-\alpha_t(1+\mu)]\mu^{k-1}}{[1+\alpha_t(1+\pi\mu)](1+\mu)^{k+1}}+\frac{[1+\alpha_t(1+\mu)][\alpha_t(1+\pi\mu)-\pi\mu][\alpha_t(1+\pi\mu)]^{k-1}}{(1+\mu)[1+\alpha_t(1+\pi\mu)]^{k+1}}+$$

$$+\frac{(1+\mu)[\mu+\alpha_t(1+\mu)][\alpha_t(1+\pi\mu)-\pi\mu]}{[\alpha_t(1+\pi\mu)-\mu][\alpha_t(1+\pi\mu)]^k}\times$$

$$\times\{\alpha_t^k(1+\pi\mu)]^k+\alpha_t(1+\pi\mu)[\alpha_t^k(1+\pi\mu)^k+(1+\alpha_t(1+\pi\mu))^k]\} \tag{7}$$

이다.

증명 보조정리 3과 확률생성함수의 성질[5]로부터 ε_t 는 $\xi_1\sim\mathrm{ZMG}\left(\frac{\alpha_t(1+\mu)}{\mu},\ \mu\right)$, $\xi_2\sim\mathrm{ZMG}\left(\frac{\pi\mu}{\alpha_t(1+\pi\mu)},\ \alpha_t(1+\pi\mu)\right)$ 인 독립인 두 우연량 ξ_1, ξ_2의 합으로 표시되므로

$$P_{\varepsilon_t}(0)=P(\xi_1=0)P(\xi_2=0)$$

$$P_{\varepsilon_t}(k)=P(\xi_1=0)P(\xi_2=k)+P(\xi_1=k)P(\xi_2=0)+\sum_{l=1}^{k-1}P(\xi_1=l)P(\xi_2=k-l)$$

이고 정의 1을 리용하면 정리의 결과가 쉽게 나온다.(증명끝)

2. RrZMGINAR(1)과정의 성질

보조정리 4 RrZMGINAR(1)과정의 조건부확률함수와 조건부확률생성함수는

$$p_t(m,\ n)=\sum_{l=0}^{\min(m,\ n)}\binom{l+m-1}{l}\frac{\alpha_t^l}{(1+\alpha_t)^{l+m}}P_{\varepsilon_t}(n-l) \tag{8}$$

$$\varphi_{X_t|X_{t-1}}(s)=\mathrm{E}\left(s^{(X_t|X_{t-1})}\right)=\left(\frac{1}{1+\alpha_t(1-s)}\right)^{X_{t-1}}\varphi_{\varepsilon_t}(s) \tag{9}$$

이다.

증명 $G_i \sim \mathrm{Geo}\left(\dfrac{\alpha}{1+\alpha}\right)$, $i \ge 1$ 일 때 $\sum_{i=1}^{k} G_i \sim \mathrm{NegBi}\left(\dfrac{\alpha}{1+\alpha},\ k\right)$이고 ε_t 는 $\{G_i(t)\}_{i\ge 1}$과 독립이라는것을 리용하면 조건부확률함수와 조건부확률생성함수는

$$p_t(m,n) = P(X_t = n \mid X_{t-1} = m) = P(\alpha_t * X_{t-1} + \varepsilon_t = n \mid X_{t-1} = m) =$$
$$= P(G_1(t) + G_2(t) + \cdots + G_m(t) + \varepsilon_t = n) =$$
$$= \sum_{l=0}^{\min(m,\,n)} \binom{l+m-1}{l} \left(\frac{\alpha_t}{1+\alpha_t}\right)^{l} \left(\frac{1}{1+\alpha_t}\right)^{m} P_{\varepsilon_t}(n-l) =$$
$$= \sum_{l=0}^{\min(m,\,n)} \binom{l+m-1}{l} \frac{\alpha_t^l}{(1+\alpha_t)^{l+m}} P_{\varepsilon_t}(n-l)$$
$$\varphi_{X_t|X_{t-1}}(s) = \mathrm{E}\left(s^{(X_t|X_{t-1})}\right) = \mathrm{E}\left(s^{(X_t|X_{t-1})}\right) = \mathrm{E}\left(s^{(\alpha_t * X_{t-1} + \varepsilon_t|X_{t-1})}\right) =$$
$$= \mathrm{E}\left(s^{(\alpha_t * X_{t-1}|X_{t-1})}\right)\mathrm{E}\left(s^{(\varepsilon_t|X_{t-1})}\right) = \left(\frac{1}{1+\alpha_t(1-s)}\right)^{X_{t-1}} \varphi_{\varepsilon_t}(s)$$

이다.(증명끝)

보조정리 5 RrZMGINAR(1)과정의 X_{t-1} 과 X_t 의 동시적확률생성함수는

$$\varphi_{X_{t-1},\,X_t}(s,\ t) = \varphi_{\varepsilon_t}(t) \frac{(1+\pi\mu)[1+\alpha_t(1-t)] - \pi\,\mu s}{(1+\mu)[1+\alpha_t(1-t)] - \mu s} \tag{10}$$

이다.

증명 보조정리 4를 리용하고 정돈하면

$$\varphi_{X_{t-1},\,X_t}(s,t) = \mathrm{E}\left(s^{X_{t-1}} t^{X_t}\right) = \sum_i \sum_j s^i t^j P(X_{t-1} = i,\ X_t = j) =$$
$$= \sum_i \sum_j s^i t^j P(X_t = j \mid X_{t-1} = i) P(X_{t-1} = i) =$$
$$= \sum_i s^i P(X_{t-1} = i) \sum_j t^j P(X_t = j \mid X_{t-1} = i) = \sum_i s^i P(X_{t-1} = i) \varphi_{X_t|X_{t-1}}(t) =$$
$$= \varphi_{\varepsilon_t}(t) \frac{(1+\pi\mu)[1+\alpha_t(1-t)] - \pi\,\mu s}{(1+\mu)[1+\alpha_t(1-t)] - \mu s}$$

가 나온다.(증명끝)

정리 2 RrZMGINAR(1)과정의 조건부기대값과 조건부분산, 자기상관함수는

$$\mathrm{E}(X_t \mid X_{t-1}) = \mu(1-\pi)(1-\alpha_t) + \alpha_t X_{t-1} \tag{11}$$
$$\mathrm{Var}(X_t \mid X_{t-1}) = \mu(1-\pi)\{(1-\alpha_t)[1+\mu(1+\pi)(1+\alpha_t)] - 2\alpha_t^2\} + \alpha_t(1+\alpha_t)X_{t-1} \tag{12}$$
$$\gamma_t(k) = \frac{\rho_t(k)}{\rho_t(0)} = \alpha_t \alpha_{t-1} \cdots \alpha_{t-k} \tag{13}$$

이다.

증명 보조정리 1의 따름의 증명에서와 같이 확률생성함수의 성질을 리용하면 식 (11), (12)가 곧 나온다. 자기공분산함수는 RrZMGINAR(1)과정의 정의로부터

$$\rho_t(k) = \mathrm{Cov}(X_t,\ X_{t-k}) = \mathrm{Cov}(\alpha_t * X_{t-1} + \varepsilon_t,\ X_{t-k}) = \alpha_t \rho_{t-1}(k-1) =$$
$$= \alpha_t \alpha_{t-1} \cdots \alpha_{t-k} \rho_{t-k}(0)$$

이다. 그런데 임의의 t에 대하여 $X \sim \mathrm{ZMG}(\pi,\ \mu)$이므로 $\rho_t(0) = \rho_{t-k}(0)$이 성립한다.

따라서 RrZMGINAR(1)과정의 자기상관함수는

$$\gamma_t(k) = \frac{\rho_t(k)}{\rho_t(0)} = \alpha_t \alpha_{t-1} \cdots \alpha_{t-k}$$

이다.(증명끝)

따름 RrZMGINAR(1)과정의 X_t가 주어졌다는 조건밑에서 X_{t+k}의 조건부기대값은

$$\mathrm{E}(X_{t+k} \mid X_t) = \mu(1-\pi)(1 - \alpha_{t+k}\alpha_{t+k-1} \cdots \alpha_{t+1}) + \alpha_{t+k}\alpha_{t+k-1} \cdots \alpha_{t+1} X_t \tag{14}$$

이다.

식 (14)는 RrZMGINAR(1)과정의 예측에 리용된다.

참 고 문 헌

[1] A. S. Nastic et al.; J. Time Ser. Anal., 37, 267, 2016.

[2] W. B. Souza; J. Time Ser. Anal., 36, 839, 2015.

[3] Y Suhov et al.; Probability and Statistics by Example, Cambridge University Press, 54～59, 2005.

주체107(2018)년 3월 10일 원고접수

Random Enviroment Integer-Valued Autoregressive Process of Order 1 with Zero-Modified Geometric Distribution

Yu Kang Hyok, Kim Chol Gwang

In this paper, we propose a random enviroment integer-valued autoregressive process of order 1 based on negative binomial thinning operator, which can be used in dealing with count time series, and derive some properties of the process.

The proposed process has zero-modified geometric distribution.

Key words: integer-valued autoregressive process, zero-modified geometric distribution