조건부 종속성에 대한 Markov 연쇄의 최적차수 추정

李 在 浚*

1. 서 론

일 강수계열과 같은 수문사상은 습윤계속기간과 건조계속기간이 교대로 계속되면서 습윤일의 강수량으로 구성되어 있다. 일 강수계열은 추계학적 종속성을 나타내는 경향 즉, 습윤일과 건조일에는 각각의 계열을 이루고 군집하려는 경향을 보이고 있으므로 일 강수의 발생은 Bernoulli 시행에 의해 기술할 수 있는 독립 확률사상으로는 적절하게 나타낼 수가 없다. 이러한 기상학적 지속성의 존재는 물리현상의 조건부 종속성의 차수에 대응되는 적절한 차수를 갖는 Markov 연쇄에 의해 기술될 수 있으며, 간단한 1차 Markov 연쇄가 매우 유용 하게 사용되고 있다.

그러나 이 모형은 임의 일의 강수확률은 전날에 강수가 도래하였느냐에만 종속된다는 것을 가정하고 있으며, 많은 경우에서 일 강수의 발생과정에 널리 이용되어 왔으나 어떤 경우에는 적합도가 좋지 않게 되는 것으로도 알려져 있다.

Markov 연쇄의 적합도 검정에 Bartlett(1951)은 최초로 우도법(likelihood approach)을 고려하였으며, 관측자료에 대한 Markov 연쇄 모형의 적합도 검정을 위하여 우도비 검정법(likelihood ratio test)을 제안하였다. 이 방법은 관측계열이 최소한 r차 종속인지 아닌지를 검정하기 위한 것이었으며, 천이확률은 알려져 있거나 적어도 추정될 유한개의 매개변수에 종속된다고 가정하였다. 그리고 Hoel(1954)은 천이확률을 전혀 알 수 없을 경우에는 약간의 수정이 필요하다고 하였으며, Bartlett의 방법과 Hoel의 방법은 둘 다 연쇄는 "r-1차 종속이다"라는 귀무가설과 "r차 종속이다"라는 대립가설하에서 유의성이 얻어질 때까지 연쇄의 차수를 연속적으로 낮추어 가면서 검정을 한다. 또한 Good(1955)은 귀무가설 H_N과 대립가설 H_A($N<A$)에 대한 것을 검정하기 위하여 Hoel의 방법을 확장하였다.

Lowry and Guthrie(1968)는 Markov 연쇄 모형의 차수 추정치와 적합도 검정을 동시에 할 수 있는 방법을 제시하였으며, Dickey and Lientz (1970)는 Markov 연쇄의 차수를 검정하기 위해 Bayes 방법의 이용을 제안하였다.

그리고 Boza(1971)는 초기상태가 고정되고 정상 천이거동을 하는 이산형 유한 Markov 연쇄에 대한 검정방법을 제안하였다.

그러나 가설검정의 고전적인 Neyman-Pearson 이론에서는 결정에 의해 발생되는 손실(loss)을 정의하기 위해 단지 올바른 가설을 기각하거나 틀린 가설을 승인하는 확률만을 고려하였다. 실상황에서 가정된 귀무가설은 단지 미지수에 대한 근사값으로서 대부분 사실과는 다르다(Billingsley, 1961). 그러한 조건하에서 정의된 손실은 모형의 등정(identification)을 위한 적절한 기준으로 보기 어려우므로 Akaike(1974)는 Kullback-Leibler의 평균정보량에 기초하여 최우원리의 확장에 바탕을 둔 모형의 등정문제에 관한 객관적인 방법(Aka-

* 정회원, 금오공과대학교 토목공학과 부교수, 공학박사

ike information criteria; AIC)을 제안하였다. Gates and Tong(1976)은 AIC를 이용하여 유한 ergodic Markov 연쇄의 차수를 결정하는 방법을 제안하였고, Eidsvik(1980)은 풍속과 파고 및 일 강수량에 대한 Markov 연쇄 모형을 등정하기 위해 AIC를 사용하였다.

Markov 연쇄에 대한 AIC 추정량은 표본의 크기가 큰 경우에도 진의차수(true order)를 과도 추정하는 확률이 상당히 커서 불일치성이 있는 추정량임이 밝혀지고 있어 Schwartz(1978)는 새로운 BIC 추정량을 제안하였다. Katz(1977, 1981)는 AIC에 의한 방법은 Markov 연쇄의 차수를 과도 추정하는 경향이 있음을 일부 확인하고 새로운 Schwartz의 기준(BIC)을 사용하여 Markov 연쇄의 차수를 결정하는 방법을 설명하였다.

또한 Chin(1977)은 Markov 연쇄 모형의 최적 차수를 AIC로 구하여 등차수도를 작성하고 사전 검정 없이 어떤 차수의 Markov 연쇄 모형을 사용하는 것은 오류를 범할 수 있다고 하였으며, Kavvas et al.(1977)은 터어키 앙카라 지방의 일 강수 발생에 관해 1차와 2차 Markov 연쇄 모형을 적용하고 최적 차수를 Anderson and Goodman(1957)의 χ^2 적합도 검정과 우도비 검정에 의해서 검정하였다.

2. Markov 연쇄의 적합도 검정

시간에 따라 어떠한 방법으로 변화되는 상황에 대한 수학적 모형을 확률과정(stochastic processes)이라고 하며, 이러한 모형중에서 주어진 현재와 과거의 상태가 미래의 상태에 영향을 주지않고 근접하는 두 상황의 바로 앞의 상태가 바로 뒤의 상황에만 영향을 주고 그 이전과 이후의 모든 상황에는 전혀 영향을 주지 않는 성질을 갖는 확률과정을 Markov 과정이라고 하고, 특히 이산적인 시간 매개변수를 갖는 Markov 과정을 Markov 연쇄라고 한다.

Markov 연쇄는 이와 같이 이산형 확률변수의 계열로서, 만일 k가 조건부 확률에 관련된 식 (2-1)을 만족하는 가장 작은 양의 정수이면 차수가 k인 Markov 연쇄라고 한다.

$$\Pr\{X_n \mid X_{n-1}, X_{n-2}, \cdots, X_{n-k}, \cdots\} = \Pr\{X_n \mid X_{n-1}, X_{n-2}, \cdots, X_{n-k}\} \qquad (2\text{-}1)$$

이 관계를 이용하면 독립 확률변수계열은 차수가 0인 연쇄로 정의할 수가 있다.

강수발생과정에 대한 Markov 연쇄는 일반적으로 다음과 같이 가정한다.

① Markov 연쇄는 homogeneous(또는 stationary)인 것으로 가정한다.
② Markov 연쇄는 ergodic이라고 가정한다.
③ Markov 연쇄의 상태는 강수발생 유무인 경우의 두가지(wet 또는 dry) 상태만을 고려 한다.

2.1 AIC에 의한 방법

Akaike(1970)는 자기회귀(AR) 모형에서 예측자(predictor)의 평균제곱 예측오차(mean square prediction error)로 정의되는 최종예측오차(final prediction error; FPE)의 개념을 유도하였으며, 그 후의 논문에서는(Akaike, 1974) 다음 식과 같이 정의한 정보기준을 도입하여 일반적인 통계학적 모형화에 FPE 방법을 적용할 수 있도록 확장하였다.

$$\mathrm{AIC} = -2\ln(\text{최우도}) + 2(\text{모형에서 독립적인 매개변수의 수}) \qquad (2\text{-}2)$$

이 통계량은 Kullback-Leibler 정보의 의미에서 보면 진의구조(true structure)와 적합모형 (fitting model)간의 편차의 척도로서 설명되며, 주어진 몇 개의 모형중에서 AIC를 최소화 하는(최소 AIC 추정치 ; MAICE) 모형을 최적모형으로서 채택한다. 이 MAICE를 사용하면 통계적 등정문제는 추정문제로 수식화되며, 가설검정에서 요구되는 주관적인 판단은 필요없게 된다.

OM을 상태가 S=m인 Markov 연쇄 $\{X_t, t=1, 2, \cdots\}$로 부터의 관측계열이라고 하자.

$$\mathrm{OM} = \{X_1, X_2, \cdots, X_{n-1}, X_n\} \qquad (2\text{-}3)$$

일반적으로 주어진 계열 OM을 얻는 확률은 식

(2-4)와 같으며

$$Pr(OM)=Pr(X_1)\ Pr(X_2 \mid X_1)\ Pr(X_3 \mid X_2, X_1)\cdots Pr(X_n \mid X_{n-1},\cdots,X_1) \quad (2\text{-}4)$$

우도(likelihood)는 식 (2-5)와 같이 된다.

$$L=Pr(X_1)\prod_{i=2}^{n} Pr(X_i \mid X_{i-1},\cdots,X_1) \quad (2\text{-}5)$$

연쇄는 최소한 k차 종속 즉, 차수가 k라는 가정하에서 우도를 다음과 같이 쓸 수 있다.

$$L=Pr(X_1)Pr(X_2 \mid X_1)Pr(X_3 \mid X_2,X_1)\cdots Pr(X_n \mid X_{n-1},\cdots,X_1) \prod_{i=1}^{n-k} Pr(X_{k+i} \mid X_{k+i-i},\cdots,X_i) \quad (2\text{-}6)$$

$$\ln(L)=\sum_{j=1}^{k}\ln Pr(X_j \mid X_{j-1},\cdots,X_1) + \sum_{i=1}^{n-k}\ln Pr(X_{k+i} \mid X_{k+i-1},\cdots,X_i) \quad (2\text{-}7)$$

윗 식에서 우변의 첫 항은 표본의 크기 n이 증가함에 따라 증가하지 않으므로 표본의 크기 n이 증가하면 할 수록 두번째 항이 지배적이 된다.

r차 연쇄의 천이확률은 $p_{ij,\cdots,kl}$이라 하면 귀무가설 H_{r-1}은 식 (2-8)과 같이 쓸 수 있다.

$$H_{r-1}: P_{ij,\cdots,kl}=P_{j\cdots,kl},\ i=1,2,\cdots,m \quad (2\text{-}8)$$

또한 관측계열에서 $i \rightarrow j \rightarrow \cdots \rightarrow k \rightarrow l$로 천이되는 돗수를 $n_{ij,\cdots,kl}$로 놓으면 우도함수는 식 (2-9)와 같이 된다.

$$L=\prod_{ij,\cdots,kl} p_{ij,\cdots,kl}^{\,n_{ij,\cdots,kl}} \quad (2\text{-}9)$$

여기서 첨자 i, j,... , k, l 은 m 개의 가능한 상태에 대응하여 1부터 m 까지이다. 따라서 구성된 검정방법은 근본적으로는 혼합가설에 대한 우도비 검정의 점근적인 형태로서, 임의의 합리적인 모형이 포함되도록 r을 충분히 크게 택하고 차수가 불합리하게 될 때까지 낮은 차수에 대해 연속적으로 검정할 수가 있다.

지금 $p_{ij,\cdots,kl}$은 미지수이므로 $p_{ij,\cdots,kl}$의 최우추정치(maximum likelihood estimate)를 사용하기로 하며, 이는 식(2-10)에 의해서 추정된다.

$$\hat{p}_{ij,\cdots,kl}=\frac{n_{ij,\cdots,kl}}{n_{ij,\cdots,k}} \quad (2\text{-}10)$$

여기서, $n_{ij,\cdots,k}=\sum_{l=1}^{m} n_{ij,\cdots,kl}$

귀무가설 H_{r-1}하의 추정치를 '(prime)으로 표시하면 식 (2-11)이 되고,

$$\hat{p}_{ij,\cdots,kl}=p_{j,\cdots,kl}=\frac{n_{j,\cdots,kl}}{n_{j,\cdots,k}} \quad (2\text{-}11)$$

H_{r-1}과 H_r 가설을 검정하기 위한 우도비는 식 (2-12)의 형태가 된다.

$$\lambda_{r-1,r}=\frac{L(\hat{p}'_{ij,\cdots,kl})}{L(\hat{p}_{ij,\cdots,kl})}=\frac{L_{r-1}(\max)}{L_r(\max)} \quad (2\text{-}12)$$

Hoel(1954)은 ergodic Markov 연쇄에서 $-2\ln\lambda_{r-1,r}$은 점근적으로 $\nabla^2 S^{r+1}$의 자유도를 갖는 χ^2변량임을 보였다. 여기서 ∇은 superscript에 관한 차분 연산자(difference operator)로서 식 (2-13)과 같다.

$$\nabla S^r=S^r-S^{r-1} \quad (2\text{-}13a)$$

$$\nabla^2 S^r=\nabla(\nabla S^r) \quad (2\text{-}13b)$$

Hoel의 검정방법을 적용하기 위하여 식 (2-14)를 계산한다.

$$-2\ln\lambda_{r-1,r}=2\sum_{i,\cdots,l} n_{ij,\cdots,kl}\left(\ln\frac{n_{ij,\cdots,kl}}{n_{ij,\cdots,k}}-\ln\frac{n_{j,\cdots,kl}}{n_{j,\cdots,k}}\right) \quad (2\text{-}14)$$

이 식은 H_{r-1} 가설하에서 자유도 $\nabla^2 S^{s+1}$인 χ^2 확률변수가 된다. $-2\ln\lambda_{r-1,r}$을 다시 ${}_{r-1}\eta_r$로 놓고, 연쇄는 k차 종속이라는 가설 H_k를 고찰해보면 이 가설은 $k<r$ 일 때 마다 연쇄는 r차 종속이라는 가설 H_r을 암암리에 포함하고 있으므로 가설 H_k는 가설 H_r의 부분집합이 된다. 따라서 일단 그러한 가설검정이 성립되면 사실상 H_r 내에서 H_k를 검정하는 것으로 볼 수가 있다.

이와 같은 경우에 대하여 $\lambda_{r-1,r}$을 확장하여 H_r 하의 우도에 대한 H_k 하의 우도비를 $\lambda_{k,r}$로 쓰면 식 (2-15)와 같이 된다.

$$\lambda_{k,r}=\lambda_{k,k+1}\lambda_{k+1,k+2}\lambda_{r-1,r} \qquad (2\text{-}15)$$

또는

$$ {}_k\eta_r=-2\ln\lambda_{k,r}=-2\ln\lambda_{k,k+1}-2\ln\lambda_{k+1,\ k+2}-\cdots -2\ln\lambda_{r-1,r} \qquad (2\text{-}16)$$

Good(1955)은 이 식이 H_k 가설하에서 다음과 같은 자유도를 갖는 χ^2 분포를 이루고 있음을 증명하였다.

$$\nabla S^{r+1}-\nabla S^{k+1},\ k\geq 0$$
$$\nabla S^{r+1},\ k=-1 \qquad (2\text{-}17)$$

그런데 통계학적 동정절차를 결정절차로 고려할 경우에는 위험함수(risk function; expected loss function)의 적절한 선택이 가장 기본적인 문제가 된다. Markov 연쇄 모형화에서 Tong(1975)은 AIC 접근방법에 바탕을 두고 Good(1955)의 결과를 이용하여 식 (2-18)과 같은 손실함수(loss function)를 제안하였다.

$$RAIC(k)={}_k\eta_M-2(\nabla S^{M+1}-\nabla S^{k+1}) \qquad (2\text{-}18)$$

여기서, M; 가설검정에서 고려한 가장 큰 차수의 모형
k; 가설검정에서 고려한 적합한 모형의 차수

따라서 Markov 연쇄의 차수에 관한 최소 AIC 추정치(MAICE)는 RAIC(k) 값을 최소화 시키는 k에 대한 RAIC(k) 값이 된다.

2.2 BIC에 의한 방법

Markov 연쇄에 대한 AIC 추정량은 표본의 크기가 큰 경우에도 진의차수를 과도 추정하는 확률이 상당히 커서 불일치성(inconsistency)이 있는 추정량임이 밝혀지고 있다. Schwarz(1978)는 Bayesian 논리를 활용하여 분포형이 지수분포이고 독립적으로 동일하게 분포된(i.i.d.) 관측치에 대한 모형의 차원을 구하는 수정된 추정량(BIC)을 유도하였으며, 이 BIC 추정량은 penalized 우도함수 또는 penalized 우도비 통계량에 바탕을 두고 있다는 점 에서 AIC 추정량과 아주 비슷하다.

BIC는 식 (2-19)와 같이 정의된다.

$$BIC=-2\ln(\text{최우도})+(\text{모형에서 독립적인 매개변수의 수})\cdot\ln(n) \qquad (2\text{-}19)$$

식 (2-19)를 Markov 연쇄의 차수추정 문제에 적용하기 위하여 손실함수로 표시하면 식 (2-20)과 같고,

$$RBIC(k)={}_k\eta_M-(\nabla S^{M+1}-\nabla S^{k+1})\cdot\ln(n) \qquad (2\text{-}20)$$

여기서, n; 표본의 크기

모형의 최적차수를 결정하는 BIC 추정량 k는 다음 기준에 의해 정해진다.

$$RBIC(\hat{k})=\min_{0\leq k\leq M} Rbic(\hat{k}) \qquad (2\text{-}21)$$

2.3 Markov 연쇄의 최적차수의 결정

강수발생과정의 지속성의 길이를 결정하는 연쇄의 차수는 모형의 동정에 상당히 중요한 영향을 미친다. 즉, 차수가 크면 클수록 모형과 실측치간의 잔차는 줄어들게 되나 매개변수의 수가 증가하게 되어 모형화가 복잡하게 되며, 차수가 작아지면 잔차의 분산은 증가하게 되나 간단한 모형이 되어 사용이 간편한 잇점이 있게 된다(Billingsley, 1961; Katz, 1979). 따라서 모형의 동정은 가능한 한 매

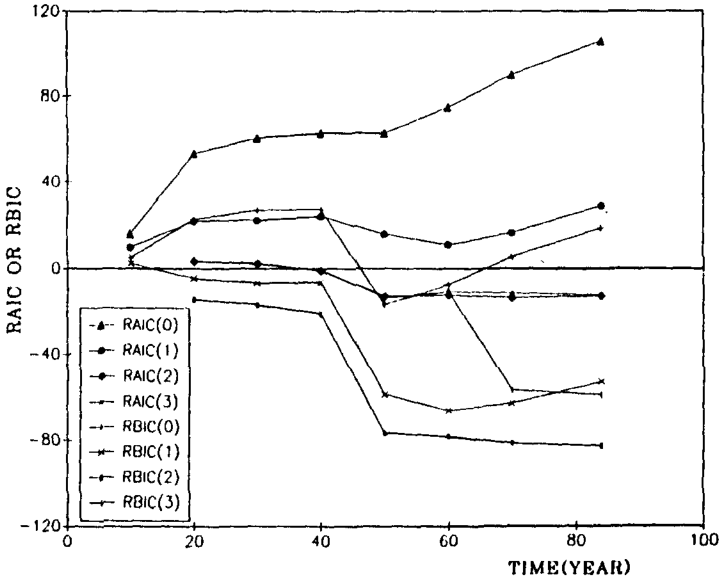

Fig. 1 RAIC and RBIC Values versus Sample Size(Tae Ku-Apr.)

개변수가 적으면서도 잔차의 분산을 줄일 수 있는 모형을 선정하는데 주안을 두고 있다.

그러므로 조건부 종속성이 있는 수문자료의 모형화에는 2.1절과 2.2절에 기술한 AIC와 BIC를 사용하여 Markov 연쇄의 적정차수를 구하는 것이 바람직할 것으로 판단된다.

3. 적용 및 고찰

여기에서는 앞에 설명한 이론적 접근방법을 대상지점에 적용하여 그 결과를 고찰하기로 한다.

3.1 표본의 크기에 따른 AIC, BIC의 변화

Markov 연쇄의 적정차수를 추정하기 위해 RAIC와 RBIC의 안정성을 검토하기 위하여 표본의 크기를 변화시켜 가면서 조사하였으며, 표본의 크기는 자료의 끝단에서 부터 연속적으로 관측치의 수를 줄여 나가면서 RAIC(k)와 RBIC(k) 값을 계산하였고 그 결과의 일부를 Fig. 1에 나타내었다.

Fig. 1은 대구의 4월 자료에 대한 것을 도시한 것이며, 이 그림으로부터 표본의 크기가 증가함에 따라서 독립성(0차 연쇄)이 채택될 가능성은 점점 희박해지고, 2차 연쇄가 채택될 가능성에는 큰 변화없이 일정한 수준으로 안정감을 보이고 있음을 알 수 있다.

3.2 Markov 연쇄의 최적차수의 결정

Table 1은 낙동강 유역의 4개 관측소의 일 강수계열 자료에 대한 Markov 연쇄의 최대차수를 M=3 및 4로 하였을 때 그 이하 차수에 대한 RAIC 및 RBIC 값을 수록한 것이다. Table 1에서 보면, 낙동강 유역의 대구 관측소의 경우는 RAIC 하에서 7, 8, 11, 12월은 1차 연쇄가 나머지 월은 2차 연쇄가, RBIC 하에서는 4월에 2차 연쇄가 적합한 것으로 나타났고, 그 외에서는 모두 1차 연쇄가 적합함을 보였다. 고령 관측소의 경우는 RAIC 기준하 에서 모든 월에서 1차 연쇄의 적합을 보였고, RBIC 기준하에서는 4월의 경우만 0차 연쇄의 적합을 나타냈으며 그 외에는 모두 1차 연쇄가 가장 적합한 것으로 나타났다.

밀양 관측소의 경우는 RAIC 기준하에서 역시 모든 월에서 1차 연쇄의 적합을 보였고, RBIC 기준하에서도 모든 월에서 1차 연쇄의 적합을 보였

Table 1. Values of RAIC and RBIC for the Model Selection(The Nakdong River Basin)

Station		Tae Ku		Ko Ryung		Mil Yang		Young Ju	
Month	Order	RAIC	RBIC	RAIC	RBIC	RAIC	RBIC	RAIC	RBIC
Jan.	0	58.24	22.04	25.17	11.22	17.44	3.48	10.58	-21.99
	1	4.07	-26.96*	-3.51*	-12.82*	-3.82*	-10.12*	-8.86*	-36.78*
	2	-5.31*	-25.99					-5.43	-24.04
Feb.	0	60.88	25.39	43.37	29.71	33.14	19.48	24.44	-7.42
	1	3.42	-27.00*	3.02*	-6.08*	-1.53*	10.63*	-4.37	-31.68*
	2	-3.06*	-23.34					-7.65*	-25.85
Mar.	0	55.74	19.54	28.21	14.25	12.66	-1.30	11.50	-21.07
	1	2.17	-28.86*	3.96*	-5.35*	4.29*	-5.02*	-1.81	-29.73*
	2	-1.82*	-22.51					-5.71*	-24.32
Apr.	0	72.42	36.43	10.54	-3.32*	30.63	16.76	11.69	20.65
	1	30.41	-0.43	6.27*	-2.97	20.19*	10.95*	0.52	-27.20*
	2	0.11*	-20.44*					-6.23*	-24.71
May	0	53.99	17.79	15.63	1.67	18.96	5.01	19.04	-13.53
	1	8.05	-22.98*	1.40*	-7.91*	3.75*	-5.56*	3.08	-24.84*
	2	-1.13*	-21.82					-4.47*	-23.08
Jun.	0	68.90	32.93	31.52	17.66	22.20	8.34	26.23	-6.11
	1	4.46	-26.37*	8.54*	-0.70*	-1.96*	-11.20*	-3.47*	-31.19*
	2	-2.34*	-22.90					0.35	-18.13
Jul.	0	152.34	116.14	91.51	77.55	86.28	72.32	43.99	11.42
	1	-6.17*	-37.20*	-3.33*	-12.64*	10.87*	1.57*	-4.90*	-32.82*
	2	-3.46	-24.14					-2.70	-21.31
Aug.	0	119.59	83.39	77.80	63.84	58.86	44.90	63.60	31.03
	1	-6.48*	-37.51*	-1.51*	-10.82*	-0.81*	-10.11*	-5.36*	-33.28*
	2	-6.06	-26.74					-4.78	-23.39
Sep.	0	129.18	93.21	64.73	50.87	65.06	51.20	43.69	11.35
	1	5.54	-25.29*	-3.74*	-12.98*	-0.41*	-9.65*	-4.66*	-32.38*
	2	0.71*	-19.85					-3.44	-21.92
Oct.	0	72.92	36.72	39.60	25.64	32.23	18.27	14.56	-18.01
	1	4.08	-26.95*	-0.06*	-9.36*	-1.28*	-10.59*	-7.76*	-35.68*
	2	-2.91*	-23.60					-5.30	-23.92
Nov.	0	53.68	17.70	15.29	1.43	9.54	-4.32	6.81	-25.53
	1	2.24*	-28.59*	5.92*	-3.32*	-1.38*	-10.62*	-9.15*	-36.87*
	2	5.56	-14.99					-6.90	-25.38
Dec.	0	51.62	15.42	10.15	-3.81	10.64	-3.32	4.74	-27.83
	1	-7.59*	-38.62*	-1.63*	-10.93*	-3.51*	-12.82*	-11.46*	-39.38*
	2	-4.10	-24.78					-7.78	-26.39

cf) * : minimum, Tae Ku : 1907-1948 Data

다. 영주 관측소의 경우는 RAIC 기준하에서 2, 3, 4, 5월의 경우가 2차 연쇄의 적합을, 나머지 월은 1차 연쇄의 적합을 나타냈으며, RBIC 기준하에서는 모두 1차 연쇄의 적합이 얻어졌다. 또한 대구 관측소의 84년간의 전 자료에 대한 경우는 RAIC 기준하에서는 10개월의 경우가 2차 연쇄의 적합성을 보였고, RBIC 기준하에서는 9개월의 경우가 1차 연쇄가 적합한 것으로 나타났으나 RAIC와 RBIC 간에 큰 차이를 보여주고 있다. 그러나 RAIC에 의한 방법은 Markov 연쇄의 차수를 과도 추정하는 경향이 있음이 있다고 하였으므로(Katz, 1977, 1981), RBIC의 기준을 채택하여 최적

Markov 연쇄 차수를 선정하기로 한다.

이상의 결과로 부터 낙동강 유역의 고령, 밀양, 영주 관측소의 강수상태는 1차 연쇄의 특성을 보이고 있으나, 대구 관측소의 3, 4, 5월(봄철)의 강수상태는 2차 연쇄의 특성을 보이고 있는 점이 특이하다.

4. 결 론

본 고에서는 시간 종속성을 가지고 있는 건습일 계열에 대한 ergodic Markov 연쇄의 차수를 결정하기 위하여 본질적으로 최우원리의 확장에 바탕을 두고 있는 Akaike's information criterion(AIC)와 Bayesian information criterion(BIC)를 이용하는 것을 기술하였다.

낙동강 유역의 4개 관측소의 건습일 계열을 대상으로 얻어진 결과를 고찰해 보면, 표본 크기의 변화에 따른 RAIC의 변화는 해석에 사용된 25~84년간 자료기간이 안정함을 보이고 있었다. 종속성의 판단 기준으로 사용한 RAIC와 RBIC의 결과를 비교해 보면, RAIC에 의한 방법은 Markov 연쇄의 차수를 과도추정하는 경향이 있음이 있다고 지적한 Katz(1977, 1981)의 결론과 마찬가지로 RAIC에 의한 최적차수가 RBIC에 의한 최적차수보다 다소 크게 되는 것을 확인할 수 있었다.

낙동강 유역의 강수상태는 대체로 1차 종속이 강한 특성이 있으나, 대구 관측소의 봄철(3 ~5월)의 강수 상태는 2차 종속성이 강한 것으로 판단된다.

따라서 우리나라 지역에서의 강수발생 과정에 Markov 연쇄 모형을 적용할 때는 단순히 1차 Markov 연쇄를 적용하기 이전에 보다 고차의 Markov 연쇄를 본 고에서 기술한 방법 등으로 검토를 한 후에 적용하는 것이 바람직할 것으로 사료된다.

참 고 문 헌

이재준, 이정식, 김홍태, "AIC를 이용한 건습일 계열에 대한 Markov 연쇄의 차수추정", 금오공과대학교 논문집, 제14집, 1993, pp.129~139.

Akaike, H., "Statistical Predictor Identification", Ann. Inst. Statist. Math., Vol.22, 1970, pp. 203-217.

Akaike, H., "A New Look at the Statistical Model Identification", Proc. IEEE Trans. Aut. Cont., Vol.AC-19, No.6, 1974, pp.716-723.

Anderson, T. W. and L. A. Goodman, "Statistical Inference about Markov Chains", Ann. of Math. Statist., Vol.28, No.1, 1957, pp.89-110.

Bartlett, M. S., "The Frequency Goodness of Fit Test for Probability Chains", Proc. Camb. Phil. Soc., Vol.47, 1951, pp.86-95.

Billingsley, P., "Statistical Methods in Markov Chains", Ann. of Math. Statist., Vol.32, No.1, 1961, pp.12-40.

Boza, L. B., "Asymptotically Optimal Tests for Finite Markov Chains", Ann. of Math. Statist., Vol.42, No.6, 1971, pp.1992-2007.

Chin, E. H., "Modeling Daily Precipitation Occurrence Process with a Markov Chain", Water Resour. Res., Vol.13, No.6, 1977, pp.949-956.

Dickey, J. M. and B. P. Lientz, "The Weighted Likelihood Ratio, Sharp Hypothesis about Changes, the Order of a Markov Chain", Ann. of Math. Statist., Vol.41, No.1, 1970, pp.214-226.

Eidsvik, K. J., "Identification of Models for Some Time Series of Atmospheric Origin with Akaike's Information Criteria", Jour. Appl. Meteor., Vol.19, No.4, 1980, pp.357-369.

Gates, P. and H. Tong, "On Markov Chain Modeling to Some Weather Data", Jour. Appl. Meteor., Vol.15, 1976, pp.1145-1151.

Good, I. J., "The Likelihood Ratio Test for Markov Chains", Biometrika, Vol.42, 1955, pp.531-533.

Hoel, P. G., "A Test for Markov Chains", Biometrika, Vol.41, 1954, pp.430-433.

Katz, R. W., "Computing Probabilities Associated with the Markov Chain Model for Precipitation", Jour. Appl. Meteor., Vol.13, 1974, pp. 953-954.

Katz, R. W., "Precipitation as a Chain-Dependent

Process", Jour. Appl. Meteor., Vol.16, 1977, pp.671–676.

Katz, R. W., "Estimating the Order of a Markov Chain: Another Look at the Tel Aviv Rainfall Data", Sixth Conf. on Prob. and Statist. in Atmosp. Sci., Oct. 9–12, 1979, Banff, Alta., Canada, pp.217–221.

Katz, R. W., "Parsimony in Modeling Daily Precipitation", Water Resour. Res., Vol.15, No.6, 1979, pp.1628–1630.

Katz, R. W., "On Some Criteria for Estimating the Order of a Markov Chain", Techno-metrics, Vol.23, No.3, 1981, pp.243–249.

Kavvas, M. L., A. A. Aksit, and Y. K. Tulunay, "A First Order Nonhomogeneous Markov Chain for the Daily Rainfall Occurrences in Ankara", Proc. Third Inter. Hydrol. Symp. on Theoretical and Applied Hydrology, Jul. 27–29, 1977, Fort Collins, pp.44–59.

Lowry, W. P. and D. Guthrie, "Markov Chains of Order Greater Than One", Mon. Weath. Rev., Vol.96, No.11, 1968, pp.798–801.

Schwartz, G., "Estimating the Dimension of a Model", Ann. of Statist., Vol.6, No.2, 1978, pp.461–464.

Stone, M., "Comments on Model Selection Criteria of Akaike and Schwartz", Jour. Royal Statist. Soc. B, Vol.41, No.2, 1979, pp.276–278.

Tong, H., "Determination of the Order of a Markov Chain by Akaike's Information Criterion", Jour. Appl. Prob., Vol.12, 1975, pp.488–497.