The Korean Communications
in Statistics Vol. 5, No. 1, 1998
pp. 43-53

시계열모형에서 계절성에 대한 검정통계량들의 비교[1)]

이 성 덕[2)], 윤 여 창[3)], 김 인 규[4)], 이 철 영[5)]

요 약

본 연구에서는 승법 시계열모형에 있어서 계절성 존재에 대한 대표본 검정통계량으로서 Wald, 우도비, Rao 통계량을 제안했다. 그리고 세 통계량의 극한분포를 유도하였다. 모의실험을 통하여 세 검정통계량의 경험적 유의수준과 극한분포를 구하고, 세 검정통계량의 검정력을 비교하였다.

1. 서 론

시계열자료에 대한 모형은 여러 가지 모수를 통하여 만들어지고 이렇게 설정된 모형을 통하여 예측을 한다. 모형을 이루는 모수에 대한 추정은 모수축약의 원칙에 따라 구하게 된다. 계절 시계열자료를 인식하는 과정에서 계절성 요인을 찾지 못하여 비계절 시계열자료로 인식한다면 예측은 나쁜결과를 가져오게 될 것이고, 비계절성 시계열자료를 계절성이 있는 자료로 잘못 인식하면 매우 복잡한 구조를 가진 모형을 가지게 된다.

본 연구의 목적은 승법 모형에 있어서 계절성 요인의 존재에 대한 검정에 있다. 주어진 시계열자료를 그래프나 AIC, BIC등과 같은 모형식별과정에서 계절성의 존재를 주관적인 판단에서 찾지 않고, 객관적으로 자료로부터 이끌어낸 극한분포를 통하여 계절성의 존재를 확인하고, 여러 가지 모형식별과정을 가진다면 더욱 더 적절한 모형을 얻을 수가 있다.

Buse(1982)는 우도비, Wald 및 Rao 통계량에 대한 적용예를 들어 그 유용함을 보였고, Serfling(1980)은 서로 독립적이고 동일한 분포에 따른다는 가정하에서 위 세 통계량이 χ^2 분포에 따름을 증명하였다. Basawa 등(1984)는 AR 시계열 모형에서도 정규조건하에서 위 통계량들이 동일한 χ^2 분포를 가짐을 증명하였으나 본 연구는 계절성이 있는 승법 ARIMA 시계열모형에 대한 대표본검정으로 확장하고자 한다.

본 연구에서는 시계열모형의 계절성 검정을 위하여 Wald, 우도비 그리고 Rao통계량을 제안하고, 세 검정통계량의 극한분포가 χ^2 분포임을 유도한다. 이를 확인하기 위해 SAR(1,1)s모형의 계절성 모수와 표본수를 변화시킴에 따라 평균, 분산 그리고 유의확률이 어떻게 변하는지를 알아보고, 귀무가설하에서 χ^2 분포로 근사하는지를 모의실험한다. 계절성 모수에 대한 검정으로 일반적인 승법모형은 모형구조가 매우 복잡하기 때문에 본 연구에서는 계절성 모수를 가진 자기상관과정에 대한 연구로 범위를 제한한다.

---

1) 본 논문은 1997년도 충북대학교 학술연구재단 연구지원에 의하여 연구되었음.
2) (360-763) 충북 청주시 개신동 48, 충북대학교 자연과학대학 통계학과 부교수
3) (565-701) 전북 완주군 삼례읍 후정리 49, 우석대학교 자연과학대학 전산통계학과 조교수
4) (300-100) 대전시 동구 자양동 226-2 대전실업전문대학 전자계산과 조교수
5) (360-763) 충북 청주시 개신동 48, 충북대학교 자연과학대학 통계학과

2. 모형 및 가설 설정

2.1 승법계절자기회귀모형

승법 계절자기회귀모형은 일반화 승법 계절혼합모형에서 d=D=q=Q=0인 경우와 같고 $SARIMA(p\times0\times0)\times(P\times0\times0)_s$로 표기된다. 이 모형은 다음과 같이 차수가 (p,P)인 승법 계절자기회귀모형이 된다.(Box 와 Jenkins(1976) 참고)

$$\phi(B)\Phi(B^s)X_t = e_t$$

여기서 $\{e_t\}$는 백색잡음이고,

$$\phi(B) = 1-\phi_1B^1-\phi_2B^2-\cdots-\phi_pB^p$$
$$\Phi(B^s) = 1-\Phi_1B^{1s}-\Phi_2B^{2s}-\cdots-\Phi_PB^{Ps}$$

이며, $\phi(B)$와 $\Phi(B^s)$의 해가 모두 1을 넘는다고 가정한다.

이 식을 X_t에 대해서 다시 정리하면 다음과 같다.

$$\begin{aligned} X_t = {} & \phi_1X_{t-1}+\cdots+\phi_pX_{t-p}+\Phi_1X_{t-s}+\cdots+\Phi_PX_{t-Ps} \\ & -\phi_1\Phi_1X_{t-s-1}-\cdots-\phi_p\Phi_PX_{t-Ps-p}+e_t, \quad t=1,2,\cdots,n. \end{aligned} \tag{1}$$

X_t 의 기대값 $E(X_t)$와 시차 h의 자기 공분산 함수 $\gamma(h)$는 다음과 같다.

$$E(X_t) = E(\sum_{r=0}^{\infty}\Psi_r e_{t-r}) = 0$$
$$\gamma(h) = \sigma^2\sum_{r=0}^{\infty}\Psi_r\Psi_{r+h}, \quad h=0,1,\cdots$$

특히, $p=P=1$일 때 $\gamma(0)$는 다음과 같다.

$$\Psi_r = \phi^h[\phi^{(u+1)s}-\Phi^{(u+1)}]/(\phi^s-\Phi), \quad r=us+h,\ h=0,1,\cdots,s-1,\ h=0,1,\cdots$$

$$\gamma(0) = \sigma^2\sum_{r=0}^{\infty}\Psi_r^2 = \frac{\sigma^2(1+\phi^s\Phi)}{(1-\phi^2)(1-\Phi^2)(1-\phi^s\Phi)}$$

2.2 가설설정과 최우추정량

주어진 시계열 모형이 순수한 비계절성 자기회귀 모형이라는 가설을 검정하기 위해 귀무가설과 대립가설을 다음과 같이 설정한다.

귀무가설 H_0 : $\Phi_i = 0$, $i = 1, 2, \cdots, P$.

대립가설 H_1 : 최소한 Φ_i 중 하나는 0이 아니다.

시계열 모형에서 계절성의 존재 유무에 대한 가설검정을 하기 위해 계절성모수가 0이라는 귀무가설을 만족하는 MLE와 아무런 가정이 없는 상태에서 구한 MLE는 다음과 같다. 먼저 계절성 모수가 0이라는 제약조건이 없고, 초기치 $X_0 = X_{-1} = \cdots = X_{-ps-p} = 0$인 조건하에서의 조건부 우도함수는 다음과 같다.

$$\begin{aligned} L = (2\pi\sigma^2)^{-n/2}\exp\Big\{-(2\sigma^2)^{-1}\sum_{t=1}^{n}(X_t-\phi_1X_{t-1}-\cdots-\phi_pX_{t-p}-\Phi_1X_{t-s}-\cdots- \\ \Phi_pX_{t-Ps}+\phi_1\Phi_1X_{t-s-1}+\cdots+\phi_p\Phi_PX_{t-Ps-p})^2\Big\} \end{aligned} \tag{2}$$

제약이 없는 MLE는 $\partial \ln L/\partial \phi = 0$, $\partial \ln L/\partial \Phi = 0$ 으로부터 구해지며,
계절성모수가 0이라는 귀무가설하에서 n개의 관찰치에 대한 우도함수는 다음과 같다.

$$L_H = (2\pi\sigma^2)^{-n/2} \exp\left\{-(2\sigma^2)^{-1}\sum_{t=1}^{n}(X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p})^2\right\}. \tag{3}$$

그러면 귀무가설하에서 ϕ_H의 최우 추정량 $\hat{\phi}_H$은 $\partial \ln L_h/\partial \phi = 0$의 해로서 $\hat{\phi}_H = (\hat{\phi}_{1H}, \cdots, \hat{\phi}_{pH})^T$이다. 즉 다음과 같이 구할 수 있다.

$$\hat{\phi}_H = D_n^{-1} e_n$$

여기서 e_n은 $(p \times 1)$벡터이고, 그리고 D_n은 $(p \times p)$행렬이다.

$$e_n^T = \left[\sum_{t=1}^{n} X_t X_{t-1} \quad \sum_{t=1}^{n} X_t X_{t-2} \quad \cdots \quad \sum_{t=1}^{n} X_t X_{t-p}\right]$$

$$D_n = \begin{bmatrix} \sum_{t=1}^{n} X_{t-1}^2 & \sum_{t=1}^{n} X_{t-2}X_{t-1} & \cdots & \sum_{t=1}^{n} X_{t-p}X_{t-1} \\ \sum_{t=1}^{n} X_{t-1}X_{t-2} & \sum_{t=1}^{n} X_{t-2}^2 & \cdots & \sum_{t=1}^{n} X_{t-p}X_{t-2} \\ \vdots & \vdots & & \vdots \\ \sum_{t=1}^{n} X_{t-1}X_{t-p} & \sum_{t=1}^{n} X_{t-2}X_{t-p} & \cdots & \sum_{t=1}^{n} X_{t-p}^2 \end{bmatrix}$$

3. 검정통계량

먼저 모수벡터 $\theta = (\phi, \Phi)^T$라 하면, θ는 $(p+P) \times 1$벡터이고 Wald 통계량은 다음과 같다.

$$Q_{1n} = n(\hat{\theta} - \theta)^T \{D G_n^{-1}(\hat{\theta}) D^T\}^{-1} (\hat{\theta} - \theta) \tag{4}$$

여기서 $\hat{\theta}$은 θ의 최우 추정량이고, $D = (I, O)$는 $p \times (p+P)$행렬, I는 $(p \times p)$ 단위행렬, $G_n(\hat{\theta})$는 $(p+P) \times (p+P)$ 행렬이다.

$$\begin{aligned} G_n(\hat{\theta}) &= -n^{-1}(\partial^2 \log L \,/\, \partial\theta_r\, \partial\theta_s)\,|_{\phi = \hat{\phi}, \Phi = \hat{\Phi}}, \quad r, s = 1, \cdots, p+P. \\ &= n^{-1}\sigma^{-2}\begin{bmatrix} A & C \\ C^T & B \end{bmatrix}_{\phi = \hat{\phi}, \Phi = \hat{\Phi}} \end{aligned} \tag{5}$$

여기서 A, B, C는 각각 $(p \times p)$, $(P \times P)$, 그리고 $(p \times P)$차 행렬들이다.
한편, 우도비 통계량은 다음과 같다.

$$\begin{aligned} Q_{2n} &= -2\log\{L(\hat{\phi}_H) \,/\, L(\hat{\phi}, \hat{\Phi})\} \\ &= \sigma^{-2}\sum_{t=1}^{n}(X_t - \hat{\phi}_{H1}X_{t-1} - \cdots - \hat{\phi}_{Hp}X_{t-p})^2 \\ &\quad - \sigma^{-2}\sum_{t=1}^{n}(X_t - \hat{\phi}_1 X_{t-1} - \cdots - \hat{\Phi}_1 X_{t-s} - \cdots \\ &\qquad + \hat{\phi}_1 \hat{\Phi}_1 X_{t-s-1} + \cdots + \hat{\phi}_p \hat{\Phi}_P X_{t-Ps-p})^2 \\ &= (b_t^T b_t - c_t^T c_t)\sigma^{-2} \end{aligned} \tag{6}$$

여기서 $b_t = \begin{bmatrix} X_1 - \phi_{H1}X_{1-1} - \cdots - \phi_{Hp}X_{1-p} \\ X_2 - \phi_{H1}X_{2-1} - \cdots - \phi_{Hp}X_{2-p} \\ \vdots \qquad \vdots \qquad \vdots \\ X_n - \phi_{H1}X_{n-1} - \cdots - \phi_{Hp}X_{n-p} \end{bmatrix}_{\phi = \phi_H}$

그리고 Rao통계량은 다음과 같다.

$$Q_{3n} = \Delta_n^T(\hat{\phi}_H)G_n^{-1}(\hat{\phi}_H)\Delta_n(\hat{\phi}_H) \tag{7}$$

여기서 $\Delta_n(\phi_H) = n^{-\frac{1}{2}}(\partial \log L / \partial\theta)\,|_{\phi = \hat{\phi}_H, \Phi = 0}$

$= n^{-\frac{1}{2}}\sigma^{-2}[b_t^T d_{01}, \cdots, b_t^T d_{0p}, b_t^T b_{t-s}, \cdots, b_t^T b_{t-Ps}]$

이다.

$$G_n(\hat{\phi}_H) = -n^{-1}(\partial^2 \log L / \partial\theta_r \partial\theta_s)\,|_{\phi = \hat{\phi}_H, \Phi = 0}, \quad r, s = 1, \cdots, p+P. \tag{8}$$

$$= n^{-1}\sigma^{-2}\begin{bmatrix} D_n & E \\ E^T & B \end{bmatrix}_{\phi = \phi_H, \Phi = 0}$$

여기서 E 는 $(P \times p)$인 행렬이다.

$$E = \begin{bmatrix} b_{t-1s}^T d_{01} + b_t^T d_{11} & \cdots & b_{t-1s}^T d_{0P} + b_t^T d_{1P} \\ b_{t-2s}^T d_{01} + b_t^T d_{21} & \cdots & b_{t-2s}^T d_{oP} + b_t^T d_{2P} \\ & \vdots & \\ b_{t-ps}^T d_{01} + b_t^T d_{p1} & \cdots & b_{t-ps}^T d_{oP} + b_t^T d_{pP} \end{bmatrix}$$

4. 극한분포

세 통계량의 공통 극한분포를 구하기 위하여 우도 방정식 $S_n(\theta) = 0$의 해를 $\hat{\theta}$이라 하자. 그리고 실제 모수는 θ라고 가정하자. 먼저 $(\partial \log L/\partial\theta) = S_n(\theta)$는 다음과 같이 주어진다

$$S_n(\theta) = \sigma^{-2}[c_t^T a_{t-1}, \cdots, c_t^T a_{t-p}, c_t^T b_{t-s}, \cdots, c_t^T b_{t-Ps}]^T \tag{9}$$

Taylor 급수의 두 항을 생각하면

$$S_n(\hat{\theta}) = S_n(\theta) - \{W_n(\theta^*)\}(\hat{\theta} - \theta) \tag{10}$$

여기서 $W_n(\theta^*) = (-\partial^2 \log L/\partial\theta_r\partial\theta_s)\,|_{\theta = \theta^*}, r, s = 1, \cdots, p+P$는 $\hat{\theta}$대신 θ^*를 대입한 것으로, θ^*는 $\hat{\theta}$과 θ사이의 가능한 다른 점들로부터 구한다.

Crowder(1976)는 B_n 이 Fisher의 정보행렬 $B_n = E[-W_n(\theta)]$이고, $-B_N^{-1/2}W_n(\theta^*)$이 무한하게 I로 접근한다는 조건아래서 $\hat{\theta}$가 θ에 대하여 약하게 일관성이 있다는 것을 보여주었다. 따라서 $\hat{\theta}$의 근사정규성에 의해 조건은 다음과 같이 표현할수 있다.

$$-B_n^{-1}W_n(\theta) \xrightarrow{p} I$$

$G_n(\theta)$는 $(p+P) \times (p+P)$인 행렬이고 $G_n(\theta) = n^{-1}W_n(\theta)$이기 때문에

$$-\{E(-W_n(\theta))\}^{-1}W_n(\theta) = \{E(-nG_n(\theta)\}^{-1}\{-nG_n(\theta)\}$$
$$= G^{-1}(\theta)G_n(\theta) \xrightarrow{p} I_{(p+P)\times(p+P)}$$

이다. 여기서 $G_n(\theta)$와 $G(\theta)$는 식(5)와 (보조정리 1)에 의해서 정의되었다. 그러므로 우리의 모형은 Crowder에 의해 주어진 조건을 만족한다.

또한 $n \to \infty$일 때 $n^{-1}\{W_n(\theta^*) - W_n(\theta)\} \xrightarrow{d} 0$이고, 다음 관계를 얻을 수 있다.

$$\sqrt{n}(\hat{\theta} - \theta) \cong \{n^{-1}W_n(\theta)\}^{-1}\{n^{-1/2}S_n(\theta)\} = G_n^{-1}(\theta)(n^{-1/2}S_n(\theta))$$

여기서 $G_n(\theta)$와 $S_n(\theta)$는 앞에서 식(5)과 식(9)에 의해서 정의되었다.

<보조정리 1>

$\{X_t\}$, $t = 1, \cdots, n$는 식(1)을 만족하는 정상 시계열 모형이라 하자. 그러면 $n \to \infty$일 때 $G_n(\theta) \xrightarrow{p} G(\theta)$이다. 여기서 $G_n(\theta)$는 식 (5)과 같고 $G(\theta)$는 다음과 같은 행렬이다.

$$G(\theta)_{(p+P)\times(p+P)} = \sigma^{-2}\begin{bmatrix} G & H \\ H^T & F \end{bmatrix} \tag{11}$$

여기서 G, H 그리고 F는 각각 다음과 같다.

$$G_{(p\times P)} = \begin{bmatrix} g(0) & g(1) & \cdots & g(p-1) \\ g(1) & g(0) & \cdots & g(p-2) \\ \vdots & \vdots & & \\ g(p-1) & g(p-2) & \cdots & g(0) \end{bmatrix},$$

$$H_{(p\times P)} = \begin{bmatrix} h(s-1) & h(2s-1) & \cdots & h(Ps-1) \\ g(s-2) & h(2s-2) & \cdots & h(Ps-2) \\ \vdots & \vdots & & \\ h(s-p) & h(2s-p) & \cdots & h(Ps-p) \end{bmatrix}$$

$$F_{(p\times P)} = \begin{bmatrix} f(0) & f(s) & \cdots & f((P-1)s) \\ f(s) & f(0) & \cdots & f((P-2)s) \\ \vdots & \vdots & & \vdots \\ f((P-1)s) & f((P-2)s) & \cdots & f(0) \end{bmatrix}$$

여기서 자기공분산 함수들은 $u = 1, \cdots, p$, $v = 1, \cdots, P$ 에 대하여 $g(u) = E[a_{t-1}^T a_{t-u}]$, $h(vs-u) = E[a_{t-u}^T b_{t-vs} + c_t^T d_{uv}]$, $f(vs) = E[b_{t-s}^t b_{t-vs}]$이다.

(증명)

$X_{t,m} = \sum_{r=0}^{m} \Psi_r e_{t-r}$에서 $\{X_{t,m}\}$, $t = 1, 2, \cdots$는 m 개의 종속적인 과정이며, 대수의 법칙에 의해,

$$n^{-1}\sum_{t=1}^{n} X_t X_{t+h} \xrightarrow{p} \gamma(h) \quad \text{as} \quad n \to \infty \quad h=0, 1, \cdots$$

으로 에르고딕 정리에 의해 자명하다.

<보조정리 2>

$\{X_t\}$, $t = 1, 2, \cdots, n$, 은 식(1)을 만족하는 정상 시계열 모형이라 하자.

$$n^{-1/2}S_n(\theta) \xrightarrow{d} N_{p+P}(0, G(\theta)) \quad \text{as } n \to \infty$$

여기서 $S_n(\theta)$와 $G(\theta)$는 식(9)와 식(11)과 같다.

(증명) $X_{t,m}$은 보조정리 1에 의하고,

$$Y^i_{t,m} = X_{t-i,m} - \Phi_1 X_{t-s-i,m} - \cdots - \Phi_p X_{t-Ps-i,m}, \ i = 1, \cdots, p$$
$$Z^j_{t,m} = X_{t-js,m} - \phi_1 X_{t-js-1,m} - \cdots - \phi_p X_{t-js-p,m}, \ j = 1, \cdots, P$$

이다. 그러면, $\{Y^i_{t,m} e_t\}$와 $\{Z^j_{t,m} e_t\}$는 $(m+1)$개의 종속 과정들이다. 그리고 $r = 0, 1, \cdots, k$, $i = 1, \cdots, p$, $j = 1, \cdots, P$에 대하여 $E(e_{t-r}) = 0$, $E(Y^i_{t,m} e_t) = 0$, $E(Z^j_{t,m} e_t) = 0$이며 $S_n(\theta)$는 평균이 0인 마팅게일 차(martingale difference)이므로 Hall과 Heyde(1980, p.54)의 마팅게일 중심극한 정리에 의해 다음과 같은 결과를 얻는다.

$$\begin{aligned}\lim_{m\to\infty} Var(Y^i_{t,m} e_t) &= \sigma^2 \lim_{m\to\infty} E(X_{t-i,m} - \Phi_1 X_{t-s-i,m} - \cdots - \Phi_p X_{t-Ps-i,m})^2 \\ &= \sigma^2\left\{\left[1 + \sum_{k=1}^{P} \Phi_k^2\right]\gamma(0) - 2\sum_{k=1}^{P} \Phi_k \gamma(ks) - 2\sum_{k<}^{P}\sum_{l}^{P} \Phi_k \Phi_l \gamma((k-l)s)\right\} \\ &= \sigma^2 g(0)\end{aligned}$$

<정리 1>

$\{X_t\}$, $t = 1, \cdots, n$는 식(1)을 만족하는 정상성 시계열 모형이라 하자. 그러면,

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N_{p+P}(0, \sigma^2 G^{-1}(\theta)) \quad \text{as } n \to \infty \tag{12}$$

여기서 $G_n(\theta)$는 앞에서 식(11)에 의해 정의되었다.

위 정리의 증명은 보조정리 1, 2에 Slutsky's Theorem을 이용하면 쉽게 얻을 수 있다.

이제 세 가지 검정통계량의 극한분포를 유도하고자 한다. Basawa등(1984)은 정규조건하에서 세 가지 검정통계량의 극한분포가 같다는 것을 보였다. 먼저 제한이 없는 최우추정량만을 사용하여 귀무가설 하에서 Wald 통계량의 극한분포를 유도하자. 그리고 그 결과 Basawa등(1984)을 이용하여 귀무가설과 대립가설하에서 세 가지 검정통계량의 극한분포를 유도하자.

<정리 2>

$\{X_t\}$, $t = 1, \cdots, n$는 식(1)을 만족하는 정상성 시계열 모형이라 하자. 귀무가설 $H: \Phi_i = 0$하에서 Wald 통계량 Q_{1n} 은 자유도가 P인 χ^2분포로 분포수렴한다.

(증명) 모수벡터 $\theta = (\phi, \Phi)^T$는 귀무가설 $H: \Phi_i = 0, i = 1, 2, \cdots, P$하에서 $\eta = (\phi, 0)$로 나타낼수 있다. 즉 $H^*: \eta_t = 0, i = 1, 2, \cdots, P$. 따라서, $n \to \infty$일 때 정리 1로부터 다음을 얻을수가 있다.

$$\sqrt{n}(\hat{\eta} - \eta) \xrightarrow{d} N_P(0, J(\theta))$$

여기서 $J(\theta) = \sigma^2 H G^{-1}(\theta) H^T$이고, $H = (\partial\eta/\partial\theta)^T = [I, O]$이다. 여기서 H는 $P \times (P+p)$

행렬, I는 (P×P) 항등행렬이다.

그러므로 귀무가설 H^*하에서, $n \to \infty$일 때 Wald 통계량은 다음과 같은 극한분포를 갖는다.

$$Q_{1n} = n(\hat{\eta} - \eta)^T J^{-1}(\hat{\theta})(\hat{\eta} - \eta) = n\hat{\eta} J^{-1}(\hat{\theta})\hat{\eta} \xrightarrow{d} \chi^2(P)$$

Basawa등(1984)에 의해 Wald통계량은 적당한 조건하에서 우도비 검정통계량 그리고 Rao검정통계량과 귀무가설 H하에서 동일한 극한분포를 가지며 Hwang 과 Basawa(1993)에 의하면 local alternative하에서 같은 검정력을 가짐을 보일수 있고 여러 가지 기준하에서 최적(asympototically optimal)임을 증명할 수 있다.

한편, Basawa등(1984)의 결과에 의해 세 통계량 Q_{in}, $i=1,2,3$의 극한분포는 동일하다. 귀무가설하에서 공통 극한분포는 자유도가 P인 χ^2분포를 가진다. 그리고 대립가설하에서 공통 극한분포는 자유도가 P이고, 비중심 모수 $\delta^2 = h^T J^{-1}(\theta_H)h$ 를 가지는 비중심 χ^2분포를 따른다. 여기서 $h^T = (h_1, \cdots, h_P)^T$이고, h_j는 임의의 실수이다.

5. 모의실험

계절성 검정을 위한 시계열 모형은 $\phi = 0.5$인 SAR(1,1)모형에 대해서 계절성 모수의 값을 0부터 0.7까지 변동을 주고, 표본의 수는 30, 50, 100, 500인 경우에 SAS를 이용하여 난수를 발생시켜 시계열자료를 형성한다. 생성된 자료를 이용하여 세 통계량을 각각 1000개씩 발생시키고, 귀무가설하에서 χ^2 분포로 근접하는지를 알아본다. 그리고 계절성 모수가 0이 아닌 경우에 대해서 Wald 통계량, 우도비통계량 그리고 Rao통계량의 검정력을 비교한다.

(표 1)은 귀무가설하에서 세 통계량의 극한분포가 χ^2 분포를 따른다는 결과를 이용하여, 각각의 평균과 분산, 유의수준을 비교하였다. 여기서 n이 커짐에 따라 세 통계량이 극한분포에 잘 수렴해 가는 것을 확인할 수 있다. SAR(1,1)모형에서 계절성 모수가 0이라는 귀무가설하에서는 $\chi^2(1)$ 로 분포수렴하여야 하는데, 평균이 1, 분산이 2, 그리고 각 유의수준에 일치하면 분포수렴한다고 할 수 있다. 모의실험의 결과 n이 100정도되면 이론에 잘 적합되는 것을 알수 있다. 또한 Wald 통계량이 나머지 두 통계량에 비해 정도가 떨어지고 있음을 확인할 수 있고, 우도비통계량이 가장 잘 $\chi^2(1)$ 분포에 수렴하는 것을 확인할 수 있다. Rao통계량은 n이 커짐에 따라 과소적합되는 경향이 있다.

(표 2)는 자료 수 n을 30, 50, 100, 500으로 변화시키면서 $\phi = 0.5$, $\Phi = 0$인 SAR(1,1)모형으로부터 시계열 자료를 1000번 추출하여 Wald통계량, 우도비통계량, Rao통계량의 구간별 확률을 구하고, $\chi^2(1)$분포에서 20000개의 χ^2 값을 랜덤하게 추출하여 구간별확률을 비교한 결과이다. 여기서 n이 커짐에 따라 세 통계량이 귀무가설하에서 χ^2 분포로 잘 분포수렴하고, Rao통계량이 나머지 통계량보다 더 좋은 결과를 가짐을 알수 있다.

계절성 검정을 위해 제안한 통계량의 검정력을 (표 3)과 같이 n이 30, 50, 100, 500인 경우와 계절성 모수를 0부터 0.7까지 변화하면서 $\chi^2(1)$ 의 유의수준을 기준으로 하여 계산하였다.

(표 3)에서 유의수준 0.1을 기준으로 하여 세 통계량의 검정력을 비교한 결과, n이 커짐에 따라 검정력이 현저하게 높아짐을 알 수 있다. 그리고 계절성 모수가 커짐에 따라서 검정력이 높아짐을

확인할수 있다. 또, Wald 통계량은 귀무가설하에서 유의수준 관리가 제대로 이루어지지 않는 것으로 보여지며, 우도비 통계량과 Rao통계량은 각각의 경우에서 거의 유사한 검정력을 보여주었다. 유의수준 0.05에서 세 통계량의 검정력을 비교한 결과는 유의수준 0.1에서와 거의 유사한 결과를 보였다. 유의수준 0.01인 경우에도 같은 결과를 보여주고 있다.

6. 결 론

본 연구를 통하여 얻은 결과는 다음과 같다.

첫째, 계절성 검정을 위한 Wald, 우도비통계량은, 귀무가설하에서 $\chi^2(1)$분포를 따름을 확인할수 있었다. 대표본 검정통계량이므로 n=30, 50, 100, 500으로 증가할수록 세 통계량이 χ^2분포에 잘 근사함을 구간별 확률 그리고 평균과 분산을 통해서 알 수 있었다. 제안된 통계량 중에서 우도비 통계량이 분포수렴에 가장 좋은 결과를 보였고 Rao통계량은 n이 커짐에 따라 다소 과소적합되는 경향을 보였다. 그리고 Wald통계량은 두 통계량과 비교하여 볼 때 가장 정도가 떨어지는 것으로 나타났다. 그러나 n이 커짐에 따라 세 통계량의 평균, 분산이 $\chi^2(1)$분포의 평균, 분산에 점근적으로 근사하였다.

둘째, 귀무가설하에서 세 통계량의 점근적인 극한분포가 χ^2분포로 근사하는 정도는, 우도비통계량, Rao통계량이 Wald통계량보다 더 정확하고 빨리 χ^2분포에 근사함을 보였다.

셋째, 세 통계량의 검정력비교에서 n이 커짐에 따라 검정력의 정도가 높아짐을 확인할 수 있었다. 또한 계절성 모수가 변화함에 따라 검정력의 차이가 있고, 계절성 모수가 커짐에 따라 검정력이 높아짐을 알수 있었다. 그리고 Wald통계량은 귀무가설하에서의 유의수준이 다른 통계량과 비교하여 볼 때 차이가 있어서, 우도비통계량과 Rao통계량을 비교한 결과 두 통계량의 검정력은 각각의 경우에서 거의 유사한 검정력을 가지고 있음을 보여주었다.

넷째, 비계절 자기회귀모형에 대해서 계절성모수가 0이라는 귀무가설에 대해서 세 검정통계량을 제안했고, 통계량의 극한분포가 $\chi^2(P)$인 분포임을 보였다. 연구된 결과를 이용하여 순수한 계절자기회귀모형에 대해서 자기회귀모수가 0이라는 귀무가설에 본 연구에서 제안한 통계량을 사용할 수 있고, 검정통계량의 극한분포가 $\chi^2(p)$인 분포임을 추론할수 있다.

향후의 연구과제로서 자기회귀모수의 변화에 계절성 검정통계량의 값들이 영향을 받을지에 대한 연구가 필요하다. 또한 실증적인 자료를 통하여 제안한 검정통계량을 이용하여 계절성을 찾고, 계절성 검정결과를 이용하여 모형설정과 예측에 응용할 필요가 있다.

참 고 문 헌

[1] Basawa, I.V., Billard, L. and Srinivasan R. (1984). Large sample tests of homogeneity for time seires models, *Biometrika*, 71, 203-206.

[2] Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis : Forecasting and Control, 2nd ed., Holden-Day.

[3] Buse, A.(1982). The LR Ratio, Wald, and Largange Multiplier Tests : An Expository Note,

The American Statistician, 36, 153-157.

[4] Crowder, M.J.(1976). Maximum LR estimation for dependent observations, *Journal of Royal Statistical Society, B, 38, 45-53.*

[5] Serfling(1980). Approximation theorems of mathematical statistics, John Wiley and Sons.

[6] Hall, P.G. and Heyde, C.C. (1980). Martingale Limit Theory and Its Applications. Academic Press, New York.

[7] Hwang, S.Y. and Basawa, I.V.(1993). Asymptotic optimal inference for a class of nonlinear time series. *Stochastic Process and Their Application*, Vol 46, 91-113.

<표 1> 귀무가설하에서의 세 통계량의 평균과 분산, 유의수준 비교

		평균	분산	1%	5%	10%
χ^2		1.000	2.000	0.01	0.05	0.10
Wald	30	3.301	58.476	0.118	0.193	0.284
	50	1.936	8.739	0.056	0.161	0.215
	100	1.530	6.626	0.041	0.105	0.160
	500	1.257	3.149	0.018	0.069	0.134
LR	30	1.362	3.061	0.014	0.082	0.152
	50	1.213	2.082	0.016	0.066	0.117
	100	1.100	2.528	0.016	0.052	0.110
	500	1.001	1.916	0.013	0.047	0.088
Rao	30	1.085	2.705	0.013	0.067	0.116
	50	1.037	1.371	0.011	0.051	0.093
	100	0.978	2.263	0.013	0.047	0.091
	500	0.941	1.844	0.012	0.043	0.081

<표 2> 자료 수 n을 30, 50, 100, 500으로 변화시키면서 ϕ = 0.5, Φ =0인 SAR(1,1)모형으로부터 시계열 자료를 1000번 추출하여 Wald 통계량, 우도비 통계량, Rao 통계량 그리고 $\chi^2(1)$분포에서의 구간별확률을 비교한 결과.

	n=30				n=50				n=100				n=500			
	Wald	LR	Rao	$\chi^2(1)$	Wald	LR	Rao	$\chi^2(1)$	Wald	LR	Rao	$\chi^2(1)$	Wald	LR	Rao	$\chi^2($
0.5	51.151	61.649	67.167	68.355	53.053	61.649	66.866	68.355	60.060	68.295	70.270	68.545	62.200	67.826	69.800	68.235
1.5	14.114	16.353	16.416	15.685	17.517	16.353	15.315	15.685	17.717	15.789	15.315	16.065	17.900	18.260	17.300	16.040
2.5	10.310	11.432	6.106	7.600	9.609	11.432	9.609	7.600	7.907	7.393	7.107	7.390	8.300	6.847	6.300	7.285
3.5	5.805	4.920	4.404	3.735	5.305	4.920	3.203	3.735	4.204	3.508	2.902	3.695	5.200	3.369	3.200	3.675
4.5	3.303	2.894	2.802	1.960	4.104	2.894	2.102	1.960	2.702	1.127	1.201	2.075	1.700	1.521	1.400	2.005
5.5	2.302	0.868	1.201	1.045	3.303	0.868	1.601	1.045	2.302	1.127	1.001	0.940	1.900	0.434	0.400	1.220
6.5	1.501	1.447	0.800	0.700	1.701	1.447	0.800	0.700	1.301	1.253	1.001	0.580	1.000	0.434	0.400	0.645
7.5	1.501	0.289	0.500	0.425	0.700	0.289	0.400	0.425	0.800	0.250	0.200	0.290	0.100	0.543	0.500	0.430
8.5	1.301	.	.	0.225	0.800	.	.	0.225	0.400	0.877	0.700	0.165	0.400	0.434	0.400	0.185
9.5	0.800	.	0.100	0.130	1.101	.	.	0.130	0.900	.	.	0.085	0.400	0.326	.	0.095
10.5	0.900	0.144	0.100	0.065	0.700	0.144	.	0.065	0.200	0.125	0.100	0.065	0.500	.	0.300	0.095
11.5	1.001	.	.	0.035	0.500	.	0.100	0.035	0.400	.	.	0.055	0.400	.	.	0.035
12.5	0.200	.	0.300	0.020	0.300	.	.	0.020	0.300	0.250	0.200	0.025	.	.	.	0.020
13.5	0.300	.	.	0.015	0.300	.	.	0.015	0.100	.	.	0.015	.	.	.	0.020
14.5	0.100	.	.	0.005	.	.	.	0.005	0.100	.	.	0.005	.	.	.	0.005
15.5	0.400	.	.	.	0.400	.	.	.	0.100	.	.	.	.	.	.	0.010
16~	5.000	.	0.100	.	0.600	.	.	.	0.500	.	.	0.005	.	.	.	.

<표 3> 계절성 검정을 위해 제안한 통계량의 검정력을 n이 30, 50, 100, 500인 경우에 대하여,계절성 모수를 0부터 0.7까지 변화하면서 $\chi^2(1)$ 의 유의수준을 기준으로 하여 계산.

통계량	n	유의수준	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7
Wald	30	0.1	0.284	0.316	0.403	0.551	0.707	0.829	0.922	0.962
		0.05	0.193	0.236	0.316	0.463	0.616	0.732	0.868	0.925
		0.01	0.118	0.151	0.215	0.308	0.459	0.582	0.758	0.849
	50	0.1	0.215	0.310	0.498	0.701	0.860	0.971	0.986	0.999
		0.05	0.161	0.231	0.400	0.615	0.795	0.952	0.973	0.994
		0.01	0.056	0.115	0.250	0.441	0.640	0.845	0.951	0.983
	100	0.1	0.160	0.326	0.685	0.926	0.985	1.000	1.000	1.000
		0.05	0.105	0.219	0.575	0.852	0.978	1.000	1.000	1.000
		0.01	0.041	0.109	0.362	0.711	0.953	0.997	0.999	1.000
	500	0.1	0.134	0.772	0.999	1.000	1.000	1.000	1.000	1.000
		0.05	0.069	0.684	0.994	1.000	1.000	1.000	1.000	1.000
		0.01	0.018	0.493	0.980	1.000	1.000	1.000	1.000	1.000
LR	30	0.1	0.152	0.201	0.303	0.549	0.708	0.844	0.939	0.981
		0.05	0.082	0.142	0.222	0.419	0.630	0.779	0.905	0.966
		0.01	0.014	0.056	0.099	0.209	0.399	0.604	0.790	0.924
	50	0.1	0.117	0.247	0.425	0.665	0.855	0.966	0.995	1.000
		0.05	0.066	0.147	0.320	0.569	0.794	0.944	0.980	0.996
		0.01	0.016	0.055	0.126	0.350	0.616	0.852	0.949	0.993
	100	0.1	0.110	0.242	0.623	0.880	0.990	1.000	1.000	1.000
		0.05	0.052	0.143	0.508	0.833	0.976	1.000	1.000	1.000
		0.01	0.016	0.056	0.263	0.630	0.921	0.986	0.998	1.000
	500	0.1	0.088	0.711	0.997	1.000	1.000	1.000	1.000	1.000
		0.05	0.047	0.589	0.992	1.000	1.000	1.000	1.000	1.000
		0.01	0.013	0.356	0.964	1.000	1.000	1.000	1.000	1.000
Rao	30	0.1	0.116	0.172	0.255	0.513	0.716	0.866	0.959	0.984
		0.05	0.067	0.105	0.178	0.391	0.600	0.791	0.923	0.976
		0.01	0.013	0.041	0.083	0.192	0.395	0.626	0.816	0.929
	50	0.1	0.093	0.190	0.382	0.666	0.865	0.969	0.995	1.000
		0.05	0.051	0.116	0.284	0.561	0.789	0.948	0.986	1.000
		0.01	0.011	0.045	0.113	0.350	0.634	0.861	0.958	0.993
	100	0.1	0.091	0.217	0.621	0.886	0.989	1.000	1.000	1.000
		0.05	0.047	0.127	0.505	0.841	0.976	1.00	1.00	1.000
		0.01	0.013	0.050	0.263	0.640	0.923	0.986	0.998	1.000
	500	0.1	0.081	0.711	0.997	1.000	1.000	1.000	1.000	1.000
		0.05	0.043	0.592	0.992	1.000	1.000	1.000	1.000	1.000
		0.01	0.012	0.360	0.964	1.000	1.000	1.000	1.000	1.000