財務管理研究 第十九券 第一號(2002. 6), 1~32

자본시장과 구조변화

蜀道之難/難於上青天/側身西望長咨嗟*

이 일 균**

〈요 약〉

한국 종합주가지수는 한국증권거래소에 상장된 모든 기업들의 가치 가중치에 의한 포트폴리오의 가격의 시계열이라고 할 수 있다. 따라서 이 지수는 한국 경제의 현재의 활동과 미래의 활동에 대한 예상의 총합의 반영 또는 표상을 표현하는 정보라 할 수 있다.

본 논문에서는 시차변수가 독립변수로 도입될 수도 있으며, 시차변수가 아닌 변수가 독립변수로 도입되는 것이 허용되는 회귀모형을 통하여 구조변화의 회수와 구조변환점을 검정할 수 있는 통계량과 이 통계량의 확률분포를 분석하고 한국 종합주가지수에 적용하여 한국 종합주가지수의 일별수익률에 구조변화가 발생하였는지의 여부와 발생했다면 발생회수와 변환점들을 발견하는데 그 목적이 있다.

한국 종합주가지수의 일별수익률은 분산의 변화 그리고 평균 및 분산의 동시변화가 1997년 9월 27일에 발생하였다. 자기회귀모형에 의할 때 증권시장의 구조변화는 1999년 11월 16일에 이루어졌다. 평균과 분산의 변화가 일어나 구조변화의 단계를 시작하고 구조변화에 알맞는 환경조성에 2년이 소요된 후에 1999년 11월 16일에 구조변화가 정착되었다. 정착이 이루어진 후에야 비로서 이 두기간은 서로 다른 경제구조와 증권시장구조가 이루어지고 이에 입각하여 시장의 새로운 운동법칙이 전개되고 있다고 할 수 있을 것이다.

I. 서 론

증권시장의 운동법칙의 한 단면을 표상하고 있는 주가지수는 자본시장에 참여하여 투자활동을 전개하고 있는 투자자들에게는 시장의 움직임을 파악할 수 있는 주요한 지표가 되고 있다. 따라서 이 지표가 앞으로의 투자활동에 중요한 정보를 제공한다. 뿐만 아니라 주가지수는 경제활동의 선행지표로서의 역할을 담당하고 있어 경제정책의 형성에 중요한 정보를 공여하고 있고, 소비자와 생산자도 앞으로의 경제상황을 예측하는 지표로 사용하고 있다.

어느 한 기업의 현재의 경영성과와 미래의 경영성과가 공평하게 반영되어 형성되는

---

* 촉으로 가는 길 험난하다 / 푸른 하늘 오르는 것보다 어려우니 / 몸누여 서쪽 바라보며 긴 한숨 짓는다.(李白)

** 명지대학교 경영학과 교수

*** 본 논문의 질적향상에 도움을 주신 두분의 심사위원에게 감사를 드립니다.

것이 이 기업의 주가다. 주가지수는 이 지수를 계산하는데 산입되는 기업들의 주가들의 선형결합에 의하여 형성된다. 선형결합중에서도 볼록결합(convex combination)이 일반적이다. 볼록결합을 형성시키는 가중치는 대부분 가치가중치이다. 지수형성에 작용하는 가치가중치는 어느 기업의 시장가치를 증권시장 전체의 총가치로 나눈 값이다. 한국 종합주가지수는 한국증권거래소에 상장된 모든 기업들의 주가의 가치 가중치에 의하여 형성된 포트폴리오의 가격의 시계열이라고 할 수 있다. 따라서 한국 종합주가지수의 시계열은 증권시장에 참여된 기업들의 가치의 볼록결합의 집합이라 할 수 있을 것이다. 한국 종합주가지수는 한국 경제의 현재의 활동과 미래의 활동에 대한 예상의 총합의 반영 또는 표상을 표현하는 정보라 할 수 있다.

증권시장의 운동법칙을 요약하고 있으며 증권시장의 활동에 대한 종합적 정보를 제공하고 있는 한국 종합주가지수의 구조가 변화하고 있는지, 아니면 변화가 없는 상태를 유지하고 있는지를 파악하는 것은 정부의 자본시장 정책의 형성에 중요한 정보를 공여한다. 그리고 보다 직접적으로는 증권시장에 참여하여 투자활동을 전개하고 있는 투자자에게도 중요한 정보를 제공하고 있다. 증권시장이 구조변화를 겪어 변환점이 형성되고 이 변환시점 이전의 구조와 그 이후의 구조가 다르다면 증권시장의 운동법칙도 이 두기간에 상이할 것이다. 투자활동의 전개도 그 이전과 이후가 상이해야 할 것이다. 증권시장이 한 국민경제활동의 표상으로 파악된다면 증권시장의 구조변화여부는 곧 그 국민경제의 구조변화여부로 인식될 수도 있을 것이다.

어느 한 경제변수가 구조변화 없이 활동을 수행하고 있다면 이 함수는 매끄러운 함수(smooth function)로서 연속함수이다. 이 연속함수는 모든 곳에서 그 성질이 동일하다. 다시 말하여 성질들의 변화가 없는 상태가 유지된다. 경제변수가 동태적으로 움직이는 것은 사실이지만 각 시점에서의 행동은 그 성질이 동일하거나 동일하게 볼 수 있을 정도로 유사하다. 그러나 어떤 요인이나 충격이 발생하고 동시에 그 정도가 무척 크면 이 함수는 도약(jump)을 한다. 한 시점에서의 도약의 폭이 크면 함수는 불연속을 노정한다. 말하자면 도약이 발생하기 이전과 이후는 불연속이 된다. 이 불연속을 통하여 경제는 진보한다. 물론 불연속이 나쁜 방향으로 이루어지면 이 변수는 소멸하고 말 것이다. 진보와 발전은 불연속이 발생하는 지점에서 형성된다. 이 불연속이 발생한 지점이 구조변화가 형성되는 지점인 것이다. 따라서 어느 한 경제변수를 비롯하여 경제 전체를 표상하는 변수의 구조변화를 파악하는 것은 이처럼 중요한 것이라 할 수 있을 것이다.

경제의 구조변환 발생은 위에서 언급한 바와 같이 중요한 주제임에도 불구하고 연구가 활발하게 진행되고 있는 것 같지는 않다. 구조변화는 Hawkins(1977), Worsley(1979), Andrews(1993), Andrews와 Ploberger(1994)등이 다루고 있다. 이들은 주로 최대우도법을 이용하여 구조변화를 추정하는 식을 정립하였다. Perron(1989)은 시계열이 단위근을 갖는다는 귀무가설과 시계열과정이 추세 정상적(trend-stationary)이라는 대립가설을 설정하고 구조변화를 추정하고 있다. Bai와 Perron(1998)은 여러번에 결쳐 발생한 구조변화를 회귀식을 통하여 추정하는 방법을 정립하였다. Bai(1999)는 우도비율법에 의한 구조변화추정식을 도출하였다. Marriot와 Newbold(2000)는 베이즈 통계학에 의하여, Delgado와 Hidalgo(2000)은 비모수방법에 의하여, Hansen(2000)은 조건부 모형에 의하여 구조변화를 발견하는 검정모형을 정립하였다.

위에서 제시한 연구들은 어느 한 변수의 구조가 변화하지 않는다는 귀무가설과 어느 시점에서 변화가 발생한다는 대립가설을 세우고 이 가설을 검정하는 방법을 계량경제학적 관점에서 분석한 연구들이다. 그러나 한 경제변수를 확률변수로 보고 이 확률분포의 성질로부터 분포를 형성하고 있는 모수들의 변화 발생 여부를 도출하여 검정식을 유도한 논문도 많다. Chen과 Gupta(1999), Gupta와 Chen(1996), Sen과 Srivastava(1975)등은 정규확률변수의 평균의 변화, Inclán(1993), Chen과 Gupta(1997)등은 분산의 변화, Chen과 Gupta(1999)는 평균과 분산의 동시변화, Sen과 Srivastava(1973), Srivastava와 Worsley(1986) 등은 다변량 변수의 모수들의 변화, Hsu(1979)는 감마모형의 모수변화, Haccou등(1988)은 지수모형의 모수변화를 발견하는 검정방법을 개발하였다. Chen과 Gupta(2000)은 구조변화를 파악하는 여러 방법들을 정리하였다.

본 논문에서는 시차변수가 독립변수로 도입될 수도 있으며, 시차변수가 아닌 변수가 독립변수에로의 도입도 허용되는 회귀모형을 통하여 구조변화의 회수와 구조변환점을 검정할 수 있는 통계량과 이 통계량의 확률분포를 분석한다. 나아가 이를 한국 종합주가지수에 적용하여 한국 종합주가지수의 일별수익률에 구조변화가 발생하였는지의 여부와 발생했다면 발생회수와 변환점들을 발견하는데 그 목적이 있다.

이 논문의 진행은 다음과 같다. 제 2장에서는 구조변화를 파악할 수 있는 회귀식을 정립하고 이 회귀식을 통하여 구조변화 발생여부와 발생회수 및 변환점을 발견하는 검정통계량과 이 검정통계량의 확률분포를 천착한다. 아울러 한국 종합주가지수의 일별수익률을 사용하여 구조변화와 이에 관련된 문제들을 검정한다. 제 3장에서는 분산의 변화와 평균 · 분산의 동시변화를 정규분포모형을 사용하여 검정하는 방법을 분석한다. 그리고 한국 종합주가지수의 일별수익률이 분산과 평균 · 분산의 측면에서 변화가 발생하였는지의 여부를 검정한다. 제 4장에서는 한국증권시장과 한국경제의 구조변화를 야기시킨 원인들을 탐구한다. 제 5장에서는 결론을 제시한다.

Ⅱ. 구조변환점의 추정

어느 한 경제변수에 대한 회귀방정식 모형에는 선형모형이건 비선형모형이건 간에 종속변수와는 다른 외생변수들이 독립변수 또는 설명변수로 도입되는 경우가 있다. 계량경제학 분야에서는 이 같은 현상을 집중적으로 분석하고 있다. 그런데 설명변수로 도입되는 변수들이 종속변수의 시차변수들(lagged varables)로 도입되는 경우도 있다. 이 경우는 주로 시계열 분석에서 다루고 있다. 물론 계량경제학은 이 두 분야를 모두 아울러 대상으로 삼고 있음을 부인할 수 없다. 그런데 시차변수가 도입되는 시계열은 시계열의 비정상성(nonstationarity)등을 비롯한 여러 가지의 통계학적 문제가 야기되는 경우가 많다. Andrews(1993)등을 비롯한 여러 연구들은 종속변수의 시차변수가 설명변수로 도입되는 경우보다는 종속변수의 시차변수 이외의 변수들이 설명변수로 도입되는 회귀방정식 모형에 의하여 구조변화를 다루고 있다. 그런데 Bai(1999)는 종속변수의 시차변수가 도입될 때에도 강건성(robustness)을 갖는 구조변화 포착모형을 정립하였다. 그의 소론을 아래에 요약하고자 한다.

구조변화가 m번 발생하여 체제(regime)가 m+1개인 모형은 시계열 $\{y_t\}_{t=1}^{T}$에 대하여 다음과 같은 다중선형회귀식으로 정립할 수 있다.[1)]

$$
\begin{aligned}
y_t &= z_t'\delta_1 + u_t, \; t = 1,2,\cdots,k_1^0 \qquad (1)\\
y_t &= z_t'\delta_2 + u_t, \; t = k_1^0+1,\cdots,k_2^0\\
&\vdots\\
y_t &= z_t'\delta_{m+1} + u_t, \; t = k_{m+1}^0,\cdots,T
\end{aligned}
$$

위에서 y_t는 관찰된 종속변수이고 z_t는 $(q\times1)$ 벡터로 설명변수 또는 독립변수 벡터이다. 그리고 $\delta_i\,(i=1,\cdots,m+1)$는 회귀계수의 벡터로 $\delta_i \neq \delta_{i+1}$이다. u_t는 확률교란항(stochastic disturbance term)이다.

구조변화의 개수 m과 구조의 변환점 $k_1^0, \cdots, k_m^0$은 미지이다. 따라서 y_t와 z_t에 대한 관찰치가 T개 일 때 m개의 구조변환점과 이 구조변환점에 대응되는 회귀계수 δ_1,

1) 체제가 다르면 각 체제에 공통적인 성질들이 존재하고 이질적인 성질이 존재하며 각 체제가 서로 공유하고 있지 않은 성질들이 각 체제간의 서로다른 특성을 표출한다. 예컨대 각 체제에는 확률변수의 분포가 동일한 분포족(family of distributions)에 속하여 동질성을 구비하되 이 변수의 평균과 분산은 상이한 경우가 이에 해당된다고 할 수 있다. 물론 분포도 상이할 수 있다.

$\cdots, \delta_m$을 추정해야 한다. 구조변화의 개수 m과 m개의 구조변환점을 알고 있을 때에는 Perron(1987)이 수행한 것처럼 회귀계수의 추정이 용이하다. 식 (1)은 모든 회귀계수가 변화하고 있기 때문에 순수 구조변화모형(pure structural change model)이다.[2)]

전체표본의 분할 $(k_1, \cdots, k_m)$에 대응되는 회귀모형의 잔차들의 제곱의 합을 $S_T(k_1, \cdots, k_m)$이라 하자. 즉

$$S_T(k_1, \cdots, k_m) = \sum_{i=1}^{m} \inf_{\{\phi_i\}} \sum_{k_{i-1}+1}^{k_i} (y_t - z_t'\phi_i)^2 \tag{2}$$

위에서 $k_0 = 0$이고 $k_{m+1} = T$이다.

귀무가설은 l번 구조변화가 발생한다는 것이다. 즉

$$H_0 : m = l$$

이에 비하여 대립가설은 $l+1$회에 걸쳐 구조가 변화한다는 것이다. 따라서 대립가설은 다음과 같다.

$$H_1 : m = l+1$$

귀무가설 아래에서 구조변환점 $(k_1^0, \cdots, k_l^0)$의 추정치를 $(\widehat{k_1}, \cdots, \widehat{k_l})$이라 하자. 그리고 $l+1$개 변환점이 발생할 때 잔차의 자승의 합이 최소가 되는 점을 $(\widehat{k_1}^*, \cdots, \widehat{k}_{l+1}^*)$이라 하자. 그러면 검정통계량은 다음과 같다.

$$\sup LR_t\,(l+1 \mid l) = \frac{S_T(\widehat{k_1}, \cdots, \widehat{k_l}) - S_T(\widehat{k_1^*}, \cdots, \widehat{k^*}_{l+1})}{\widehat{\sigma}^2(l+1)} \tag{3}$$

위에서 $\sigma^2(1+1) = S_T(\widehat{k_1^*}, \cdots, \widehat{k}_{1+1}^*)/T$이다.

Bai(1999)는 식 (3)의 검정통계량이 (0, 1)위에서 독립적 Brownian bridge의 함수로

2) 순수구조변화모형 이외에 부분구조변화모형의 정립도 가능하며 이것은 다음과 같다.

$$y_t = x_t'\beta + z_t'\delta_i + u_t, \quad k_{i-1}^0 + 1 \le t \le k_i^0.$$

위에서 $i = 1, \cdots, m+1$이다. 여기에서 β는 표본전체를 통하여 일정한 상수이다. 따라서 우변의 첫항은 변하지 않는 부분이며 둘째항은 구조변화를 포착하는 항이다.

수렴하고 있음을 증명하였다.

식 (2)와 식 (3)을 좀 더 살펴보자. 식 (2)에 의하면 하나의 구조가 지속하고 있는 기간의 회귀잔차의 자승의 합을 계산할 수 있다. 여러번에 걸쳐 구조가 변하였을때 각 기간의 회귀잔차의 자승합을 계산하고 이 각각의 구조에 대하여 계산된 잔차의 자승합을 모두 합산한 것이 $S_T(\cdot)$이다. 각 구조마다 평균 또는 기대값이 상이할 것이다. 시계열 전체의 평균은 각 구조의 가중평균이다. 따라서 시계열 전체의 회귀잔차의 자승합은 각 구조의 잔차의 자승합의 합계보다 클 것이다.

식 (3)을 살펴보자. 실제에 있어서 구조변화가 $(l+1)$회 발생하였다고 하자. 그런데 구조변화가 실제보다 1회 적은 $m=l$회 발생한 것으로 착각되었다고 가정하자. 그러면 m개 구조가 존재하여 왔다고 착각되어 계산된 각 구조의 회귀잔차의 자승합의 합계가 실제로 $(l+1)$회 발생하였을때의 각 구조의 회귀잔차의 자승합의 합계보다 클 것이다. 이 경우에 있어서 이 두 합계의 차이는 클 것이다. 이 차이를 표준화하면 0보다 의미있게 (significantly) 클 것이다. 따라서 구조변화가 $l+1$회 발생하였다는 것이 파악될 것이고 각 구조마다 평균은 상이할 것이다. 이것이 식 (3)의 의미이다.

식 (2)와 식 (3)은 각 구조의 회귀잔차의 자승합을 모두 합계한 것을 기본으로 삼고 정립된 식들이다. 따라서 미세한 변화나 변동은 자승합들의 합산과정에서 흡수통합되고 만다. 말하자면 미시적(micro)인 구조변화는 합산과정에서 흡수통합되어 그 자취가 검정통계량을 계산하는 식 (3)의 분자의 첫항과 둘째항에 각각 동일하게 표출된다. 분자는 이 두항의 차를 계산하게 되므로 미시적 구조변화를 야기시키는 요소들은 상쇄된다. 반면에 거시적(macro)인 구조변화는 합산과정에서 흡수통합되지 않고 표출되므로 이 영향은 검정통계량에 그대로 반영된다.[3)]

이 점은 식 (3)의 검정통계량이 브라운 운동(Brownian motion)의 함수로 수렴하지 않고 그대신 Brownian bridge의 함수로 수렴하고 있음을 통하여 확인할 수 있다. 표준브라운 운동은 평균 또는 기대값이 0이다. 그러나 Brownian bridge의 기대값은 Dhrymes(1998, pp.386-387)가 제시한 바와 같이 $\mu(t)$이다.

식 (3)의 검정통계량이 Brownian bridge의 함수로 수렴하고 있다는 성질을 이용하면 식 (3)에 대한 임계값(critical values)는 다음에 의하여 구할 수 있다.

3) 이점을 지적해주신 심사위원께 이 자리를 빌어 감사를 드립니다. 심사위원께서는 Chow검정이 미시적 구조변화를 포착하고 있으므로 이 논문에서 제시된 구조변환점보다 많은 구조변환점을 발견할 수 있을 것이라고 주장하고 있다.

$$\lim_{T\to\infty} p[\sup L(l+1|l) > c] = 1 - \prod_{i=1}^{l+1} [1 - G_{i(c)}]$$

위에서

$$G_{i(c)} = \frac{c^{q/2}\exp(-c/2)}{2^{q/2-1}\Gamma(q/2)} \left[\left(1 - \frac{q}{c}\right)\log\frac{1-\eta_i}{\eta_i} + \frac{2}{c} + o(c^{-2}) \right] \qquad (4)$$

유의수준 α가 주어지면 이에 대응되는 임계값 c는 식 (4)에 의하여 구할 수 있다.[4)]시계열의 구조변화에 대한 설정과 검정식은 위에서 간략히 다루었는데 이에 대한 자세한 사항은 <부록 Ⅰ>로 할애한다.

위에서 제시한 검정방법에 의하여, 증권시장에서 구조변화가 발생하였는지 발생하였다면 몇 회에 걸쳐 발생하였으며 언제 발생하였는지를 추정하였다. 이를 위하여 사용한 데이터는 일별 종합주가지수의 수익률로 기간은 1980~2000년이다. 앞장에서 살펴본 추정방법에는 자기회귀모형의 사용이 허용되고 있어 증권시장의 구조변화의 회수와 구조변환점을 추정하기 위하여 자기회귀모형을 사용하였다. 자기회귀모형을 표본의 전기간에 걸쳐 살펴보았는 바 AR(3)이 데이터를 가장 잘 설명하는 것으로 판명되었다. 따라서 AR(3)를 사용하였다.

AR(3)에 의한 구조변화회수와 변환점에 대한 검정통계량을 제시하면 <표 1>과 같다. 이 표에 의하면 표본기간인 1980~2000년에 구조변화가 일회에 한하여 발생하였다. 구조변화가 발생한 일자는 1999년 11월 16일이다. 검정통계량은 20.09이고 유의수준 5%에서 임계값은 3.63이다. 따라서 구조변화가 발생하지 않았다는 귀무가설이 기각된다. 이것은 대체적으로 볼 때 1999년 11월에 구조변화가 발생하였다고 할 수 있을 것이다. <표 1>에 의하면 1999년 11월 16일의 검정통계량의 값과 유사한 검정통계량이 1999년 11월에 일어나고 있음을 알 수 있다. 말하자면 11월에 구조변화의 군집이 형성되어 있다. 구조변화가 하루에 형성된 것이 아니라 일정기간에 걸쳐 이루어졌다. 따라서 구조변화는 1999년 11월에 발생하였다고 할 수 있다.

<표 1>에서 볼 수 있는 바와같이 AR(2)도 계산하였다. AR(2)에 의하여서도 구조변화가 1회에 한하여 발생하였으며, AR(3)에 있어서 구조변화가 발생한 변환점의 검정통계량과 유사한 검정통계량의 군집과 AR(2)의 검정통계량의 군집이 서로 거의 동일하다.

이 계산치에 의하면 한국 종합주가지수의 일별수익률은 AR(3)과 AR(2)가 동시에 적

---

4) α가 적으면 c는 크다.

<표 1> 구조변환점

AR(3)		AR(2)	
일 자	통계량	일 자	통계량
1999. 11. 16	28.09	1999. 11. 16	25.88
2000. 01. 04	27.33	2000. 01. 04	25.47
1999. 11. 15	26.12	1999. 11. 15	24.11
1999. 11. 12	25.83	1999. 11. 12	23.81
1999. 11. 19	25.27	1999. 12. 28	23.52
1999. 12. 01	25.08	1999. 11. 18	23.31
1999. 12. 28	25.05	1999. 11. 19	23.21
1999. 11. 30	24.97	1999. 11. 17	23.20
1999. 11. 18	24.96	1999. 11. 11	23.16
1999. 11. 17	24.85	1999. 11. 09	22.99
1999. 11. 23	24.79	2000. 06. 07	22.91
1999. 11. 22	24.78	2000. 04. 17	22.79
1999. 11. 25	24.72	1999. 12. 01	22.72
1999. 11. 11	24.68	2000. 07. 10	22.68
1999. 11. 24	24.68	1999. 11. 30	22.59
1999. 11. 09	24.59	1999. 11. 10	22.50
1999. 11. 26	24.18	1999. 11. 04	22.50
1999. 11. 29	24.17	1999. 11. 05	22.47
1999. 11. 10	24.08	1999. 11. 25	22.44
1999. 12. 27	23.65	1999. 11. 23	22.41
1999. 12. 06	23.62	1999. 11. 22	22.39

AR(3)의 임계값 : 3.63 AR(2)의 임계값 : 1.97

용되는 자기회귀과정이라고 할 수 있을 것 같다.

1997년 말에 외환위기로 인하여 한국이 IMF로 갔고, 1998년 초 선거에 의한 새로운 정부가 국가운영을 시작하였다. 새 정부는 경제의 구조조정을 경제정책의 최우선 과제로 정립하고 정부의 모든 역량을 경주하였으며 구조조정을 위한 역할을 전담하는 부서로서 의 역할을 금융감독원이 전담하다시피 하였다. 경제계 역시 구조조정에 전념하였고 기업들이 외국인에게 팔리기도 하고 기업분할(divestiture)이 이루어지기도 하고 기업통합이 형성되기도 하였으며, 그룹내부간의 상호출자나 지급보증이 제한되기도 하였다. 대우가 도산하였고 현대가 분할되는 등 여러 가지의 구조조정 결과가 나타났다. 이러한 과정을 2년간 걸쳐서 1999년 11월에 그 동안의 구조조정 노력이 경제에 반영 정착

되고 이로 인하여 구조조정이 이룩되어 전과는 다른 새로운 구조가 형성된 것 같다. 말하자면 새로운 구조는 2년간의 조정기간을 거친 후 이루어진 것 같다.[5)]

<표 2> 구조변환점 전후기간의 평균과 분산

기 간	평 균	표준편차	최 소	최 대
전 기 간	0.00038	0.01509	-0.11633 (2000.4.17)	0.08504 (1998.6.17)
변환점이전	0.00049	0.01421	-0.08393 (1981.1.5)	0.08504 (1998.6.17)
변환점이후	-0.00200	0.02777	-0.11633 (2000.4.17)	0.08001 (2000.3.2)

* 변화일 1999. 11. 16

1997년 9월 27일을 구조변환일로 파악하고 1999년 11월 16일 역시 구조변환일로 인정한다면 구조변환의 조정기간은 존재하지 않을 것이다. 그러나 1997년 9월 27일을 평균과 분산이 동시에 변하여 구조변화의 발생을 예고하는 날로 파악한다면 2년간의 구조조정의 노력이 있었고, 그 결과로 1999년 11월 16일에 구조변화가 발생한 것으로 생각할 수도 있을 것이다. 그러나 이같은 것은 하나의 견해에 불과하며 이를 정당화시키기 위하여는 엄밀한 검정이 필요하다. 이것은 이 논문의 범위를 넘고 있어 별도의 논의가 필요한 것 같다.[6)]

증권시장은 우리나라의 대기업들의 경영성과가 평가를 받는 장소이다. 한국 종합주가지수는 가치가중 주가지수이다. 한국 종합주가지수의 일별수익률의 시계열은 한국기업의 매일매일의 경영성과와 더불어 미래에 형성될 기업성과의 집합이라 할 수 있을 것이다. 따라서 증권시장의 구조변화를 이룩하였다는 것은 한국경제가 구조변화를 이룩한 것이라고 볼 수 있을 것이다.

구조변환점이 1999년 11월 16일이므로 표본의 전기간을 비롯하여 1999년 11월 16일 이전과 이후의 표본들의 평균과 분산을 계산하였다. 이를 제시하면 <표 2>와 같다. 변환점 이전과 이후는 평균의 차이가 무척 심하고, 뿐만 아니라 표준편차도 이전의 시기보다 이후의 시기가 무척 높다. 평균과 표준편차의 측면에서 보아도 이 두기간에는 차이가 있음을 알 수 있다. <표 2>에 의하여 두 기간의 차이에 대한 검정을 용이하게 할 수 있

5) 제 4장에서 이점을 실증적으로 제시한다. 한국이 IMF로 가기 직전의 여러 가지의 복합적 경제상황은 구조변화를 형성해 나가는 시작단계로 볼 수 있으며 제 3장은 이점을 포착하기 위한 분석이고 제 4장을 통하여 이점을 확인한다.

6) 이점을 지적해주신 심사위원께 감사를 드립니다.

으므로 검정통계량은 할애하였으나 이 두기간의 모수들의 차이가 발생하였다는 귀무가설은 기각에 실패하고 있다.

Ⅲ. 평균과 분산의 동시변화

<표 2>에서 본 바와같이 구조변화 이전과 이후의 평균과 분산이 상이하다. 평균과 분산이 언제 변화하였는가를 아는 것도 구조변화를 이해하기 위한 중요한 정보라고 할 수 있을 것이다. 따라서 평균과 분산의 변환점을 찾는 방법을 모색해 보고자 한다.

평균의 변화는 한 시계열이 시간의 흐름에 걸쳐 운동을 진행하는 과정에서 일정기간의 행동과 다음기간의 행동, 그리고 그 다음기간의 행동이 평균의 측면에서 다르다는 것을 의미한다. 말하자면 이 시계열의 확률분포가 시계열의 전 기간에 있어서 동일하지 않고 두 개 기간 이상에 있어 서로 다르다는 것을 의미한다. 따라서 평균의 변환점은 분포가 서로 다르게 되는 점이라 할 수 있다. 전체표본을 여러개의 부분표본으로 분할하여 각 부분표본의 성질을 비교하면 분포의 모수들의 변화여부를 인식할 수 있다. 이점에 착안하여 분포의 모수들의 변화여부와 변환회수 및 변환점들을 발견하기 위한 일련의 연구들을 서론에서 언급한 바 있다. 그 논문들과 특히 Chen과 Gupta(2000)를 참조하여 분포의 모수들의 변환점들을 발견하는 방법을 다음에 요약하고자 한다. 보다 자세한 사항은 <부록 Ⅱ>로 미룬다.

독립적 정규확률변수의 수열이 $\{x_i\}_{i=1}^{T}$이고 각 $i=1, \cdots, T$에 대하여 각 x_i가 $x_i \sim N(\mu_i, \sigma_i^2)$ 이라 하자. 이때 변환이 1회에 한하여 발생할때 다음과 같은 귀무가설과 대립가설을 수립하고 있다.

$$H_0: \ \mu_1 = \mu_2 = \cdots = \mu_n; \ \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_T^2 \tag{5}$$

$$H_1: \ \mu_1 = \cdots = \mu_k \neq \mu_{k+1} = \cdots = \mu_T;$$

$$\sigma_1^2 = \cdots = \sigma_k^2 \neq \sigma_{k+1}^2 = \cdots = \sigma_T^2$$

평균과 분산이 동시에 변하는 시점이나 위치를 찾기 위한 분석은 Akaike의 정보기준을 사용하여 이루어질 수 있다. 이 정보기준에 의할 때 귀무가설하의 정보기준 SIC(T)는 다음과 같다.

$$\mathrm{SIC}(T) = T\log 2\pi + n\log\hat{\sigma}^2 + T + 2\log T \tag{6}$$

평균과 분산이 동시에 변화하는 시점 또는 위치를 k라 할 때 시계열의 처음부터 k까지의 수열과 k+1부터 T까지의 수열을 두 개의 집단으로 할 때 대립가설 H_1 아래에서 최대우도법에 의한 두 집단의 모수 추정치를 각각 $\widehat{\mu_1}$ 및 σ_1^2과 $\widehat{\mu_T}$ 및 σ_T^2라 하자. 그러면 $2 \leq k \leq T-2$에 대하여 대립가설 H_1 아래에서 정보기준에 의한 통계량은 최대우도함수에 Akaike 정보기준을 이용하면 다음을 얻는다.

$$SIC(k) = T \log 2\pi + k \log \hat{\sigma}_1^2 + (T-k) \log \hat{\sigma}_T^2 + T + 4 \log T \tag{7}$$

평균과 분산이 동시에 변화하는 변환점 k를 다음이 성립하는 $\widehat{k}$에 의하여 추정한다.

$$SIC(\widehat{k}) = \min_{2 \leq k \leq T-2} SIC(k) \tag{8}$$

검정통계량을 살펴보자. 먼저 다음과 같이 정의하자.

$$\begin{aligned} \triangle_T &= \min_{2 \leq k \leq T-2} [SIC(k) - SIC(T)] \\ &= - \max_{2 \leq k \leq T-2} [SIC(k) - SIC(T)] \\ &= \lambda_T^2 + 2 \log T \end{aligned}$$

위에서

$$\lambda_T^2 = \left[\max_{2 \leq k \leq T-2} \{T \log \hat{\sigma}^2 - k \log \hat{\sigma}_1^2 - (T-k) \log \hat{\sigma}_n^2\} \right]^{1/2}$$

그런데 λ_T은 $\lambda_T = (2 \log T - \triangle_T)^{1/2}$ 이므로 귀무가설 H_0 아래에서 모든 $x \in R$에 대하여 다음이 성립한다.

$$\lim_{T \to \infty} P\left[a(\log T)(2 \log T - \triangle_T)^{1/2} - b(\log T) \leq x \right] = \exp(-2e^{-x}) \tag{9}$$

위에서

$$a(\log T) = (2 \log \log T)^{1/2}$$
$$b(\log T) = 2 \log \log T + \log \log \log T$$

유의수준을 α, 이와 연관된 임계치를 c_α라 하자. $SIC(\widehat{T}) < \min_{2\leq k\leq T-2} SIC(k)$ 일때 귀무가설 H_0를 수락하는 것은 $SIC(\widehat{k}) = \min_{2\leq k\leq T-2} SIC(k) + c_\alpha$를 수락하는 것과 같다. 이때 c_α 와 α 와의 관계는 다음과 같다.

$$1-\alpha = P\left[SIC(T) < \min_{2\leq k\leq T-2} SIC(k) + c_\alpha | H_0 \right]$$

위에서 평균과 분산이 동시에 변화하는 것을 포착하는 방법을 살펴보았다. 이 방법을 한국 종합주가지수의 일별수익률에 적용하여 얻은 것이 <표 3>과 <표 4>이다.

<표 3> 분산의 변화와 변환점

SIC(n)	최소 SIC(k)	변환일
-33588.26	-35393.70	1997.9.27

임계값 $\alpha = 1\%$, $c_\alpha = 15.416$; $\alpha = 2.5\%$, $c_\alpha = 11.067$.

<표 4> 평균과 분산의 동시변화

SIC(n)	최소 SIC(k)	변환일
-33579.55	-41421.68	1997.9.27

임계값 $\alpha = 1\%$, $c_\alpha = 14.451$; $\alpha = 2.5\%$, $c_\alpha = 9.463$.

먼저 분산의 변화가 발생한 적이 있는지 없는지를 검정하였다. 이를 제시하면 <표 3>과 같다. <표 3>에서 보는 바와 같이 전기간에 걸친 Akaike 정보기준 통계량은 -33588.26이다. 그런데 Akaike 정보기준에 의한 최소 통계량은 -35393.70이다. 그런데 임계값이 1%에서 15.416이고 0.25%에서 11.067이다. 최소통계량에 11.067이나 15.416을 더해도 $SIC(T) > \min SIC(\widehat{k})$보다 크다. 따라서 분산의 변화가 발생하였으며 SIC(k)의 값 중 가장 작은 즉 -35393.70이 발생한 일자가 1997년 9월 27일이다. 따라서 분산의 변화가 1997년 9월 27일 발생하였다.

평균과 분산이 동시에 변화하고 있는지 또는 하고있지 않은지를 검정하였으며 이를 제시하면 <표 4>와 같다. 전기간에 걸친 Akaike 정보기준에 의한 검정통계량은 -33579.55이고 최소가 이루어진 SIC(k)는 -41421.68이다. -41421.68 + c_α = -41421.68 + 14.45 = -41407.23이다. 이 값은 -33579.55보다 작다. 따라서 평균과 분산의 변환일은 1997년 9월 27일이다. 이 변환일은 분산의 변환일과 일치하고 있다. 종합주가지수는 모

든 상장기업들의 주가의 가중평균이다. 말하자면 한국 종합주가지수는 우리나라 경제의 총체적 표상이라 할 수 있다. 한국 종합주가지수의 수익률의 평균과 분산이 1997년 9월 27일 이전과 이후가 서로 다르다는 검정의 결과를 얻었다. 이것은 우리나라의 경제가 1997년 9월 27일 기준으로 평균과 분산의 관점에서 변환을 이루었다는 것을 의미한다.

1997년 말은 우리나라가 IMF 간 시기이다. 1997년 9월과 10월은 아시아의 금융위기가 발생하거나 발생할 징조가 도처에서 나타나기 시작했다. 우리나라도 여러 측면에서 금융위기의 징조가 보인 시기이다. 예컨대 1997년 9월말과 10월초에는 홍콩에서 원화선물이 1000원~1100원정도 발생된 시기였다. 우리나라가 실제로 IMF에 가기 전에 이미 한국경제의 구조가 평균과 분산의 측면에서 변화하였음을 알 수 있다. 이점을 좀더 살펴보기 위하여 SIC(K)의 최소값들의 집단을 구성하였으며 이를 게시하면 <표 5>와 같다.

<표 5>에 의하면 Akaike 정보기준에 의한 SIC(k)의 최소값들의 군집이 1997년 9월과 10월에 발생하였음을 알 수 있다. 뿐만 아니라 분산의 변화가 발생한 일자들과 평균 및 분산의 변화가 동시에 발생한 일자들이 일치하고 있다. 따라서 구조변화가 발생한 시기에 대한 검정이 강건성(robustness)을 갖고 있음을 알 수 있다.

한국의 IMF는 단순한 단기금융위기, 특히 단기 외국부채의 지급불능에 의해서 발생한 것이라고는 보기 어렵다. IMF가 단기외채의 지급불능 사태에 의하여 발생한 것은 사실이 다. 이것은 표면적 사건에 불과하다고 할 수 있다. 실물경제적으로는 이미 구조의 변화 또는 변혁이 이루어지는 시점에 도달해 있는 상황이었다. 그런데 마침 단기외채의 문제가 발생하여 구조변환 또는 변혁이 수면위로 또는 표면상으로 분명하고 명백하게 떠오른것에 불과하다고 할 수 있다. IMF가 발생한 후에는 정부가 취한 경제정책은 구조변화를 정착시키기 위한 일련의 경제조치로 간주할 수도 있을 것이다.

증권시장이 평균과 분산의 측면에서 구조 변화가 1997년 9월 27일에 발생하였으므로 두 기간의 평균과 분산이 실제로 상이한지 또는 동일한 지를 보기 위하여 두 기간의 평균과 분산을 추정하였다. 이것을 <표 6>에 제시한다. <표 6>에 의하면 1980~2000간의 전체표본기간의 평균과 표준편차는 각각 0.038%와 1.509%이다. 그러나 변환점 이전인 1997년 9월 27일 이전 기간의 표본의 평균과 표준편차는 각각 0.042%와 1.141%로 평균은 전체표본보다 크고 표준편차는 전체표본보다 작다. 변환점 이후의 표본의 평균과 표준편차는 각각 0.014% 와 2.864%이다. 이 평균이 전체표본보다 상당히 적고 변환점 이전의 평균에 비하여 3배나 적다. 반면 표준편차는 전체표본의 약 2배, 전환점 이전보다 2.5배가 크다. 따라서 1997년 9월 이전에 비하여 그 이후가 위험은 무척 증가한 반면 수익률은 상당히 감소하였다는 사실을 발견할 수 있다. 경제구조가 상당한 변환을 겪었다고 할 수 있을 것이다.

<표 5> SIC(k)의 최소값들

순 위	분산의 변화		평균 및 분산의 변화	
	SIC(k)	일 자	일 자	SIC(k)
1	-35393.70	19970927	19970927	-41421.68
2	-35392.39	19971004	19971004	-41420.40
3	-35392.21	19970926	19970926	-41420.19
4	-35391.96	19971013	19971013	-41419.93
5	-35391.68	19970925	19970925	-41419.68
6	-35391.08	19970924	19970924	-41419.10
7	-35391.07	19971011	19971011	-41419.02
8	-35391.00	19971006	19971006	-41418.98
9	-35390.62	19971002	19971002	-41418.64
10	-35390.46	19970923	19970923	-41418.50
11	-35390.16	19971010	19971001	-41418.14
12	-35390.15	19971001	19971010	-41418.12
13	-35390.10	19971009	19971009	-41418.05
14	-35389.88	19970919	19970919	-41417.98
15	-35389.78	19970920	19970920	-41417.85
16	-35389.48	19970913	19970913	-41417.63
17	-35389.43	19970930	19970930	-41417.45
18	-35389.43	19971014	19971918	-41417.38
19	-35389.26	19970918	19971014	-41417.37
20	-35388.64	19970922	19970922	-41416.68
21	-35388.26	19971008	19971008	-41416.21
22	-35387.68	19970929	19970912	-41415.82
23	-35387.67	19970912	19970929	-41415.70
24	-35386.44	19971007	19971007	-41414.40
25	-35385.84	19970911	19970911	-41414.00
26	-35384.19	19970910	19970910	-41412.32
27	-35382.60	19971017	19970909	-41410.65
28	-35382.52	19970909	19971017	-41410.51
29	-35380.77	19970908	19970908	-41408.89
30	-35380.77	19971016	19971016	-41408.69

<표 6> 평균과 분산

기 간	평 균	표준편차	최 소	최 대
전체표본	0.00038	0.01509	-0.11633 (2000.4.17)	0.08504 (1998.6.17)
변환이전표본	0.00042	0.01141	-0.08393 (1981.1.5)	0.05395 (1982.6.28)
변환이후표본	0.00014	0.02864	-0.11633 (2000.4.17)	0.08504 (1998.6.17)

* 변환점 : 1997. 9. 27

평균과 분산이 동시에 변한 것은 구조가 변한 것이라고 볼 수 있다. <표 3>과 <표 4>에서 볼 수 있는 바와 같이 평균 및 분산의 동시 변환점은 1997년 9월 27일에 형성되었다. 구조변환점이 발생한 1999년 11월 16일도 <표 2>에서 본 바와같이 평균과 분산이 크게 변화하였다. 이 관점에서 1997년 9월 27일도 구조변환점으로 볼 수 있을 것이다.[7)]

Ⅳ. 구조변화

제 2장에서 지적한 바와같이 이 논문에서 사용된 검정방법은 거시적 구조변화의 회수와 일자를 포착하는 방법이다. 거시적방법에 의하여 표본기간 동안에 구조변화가 1회 발생한 것으로 발견되었다. 재무관리에서 증권시장의 구조가 미시적 관점에서 변화한 시점을 포착하는 것도 중요하지만 그에 못지않게 거시적관점에서 구조변환점을 파악하는 것도 중요하다. 미시적 구조변화는 그동안 여러각도에서 해결점이 제시되었다. 자기회귀모형, 자기회귀 이동평균모형을 출발점으로 하여 분수적분 자기회귀 이동평균모형, 자기회귀 조건부이분산(ARCH) 계통의 여러모형과 분수적분 일반자기회귀 조건부이분산모형등을 대표적으로 들 수 있다. 이 모형들에서는 미시적 구조변화가 반영되고 이에 따라 모형의 모수가 변하고 있으므로 주가의 예측에 적극 활용할 수 있다. 그러나 거시적 관점에서의 구조변화에 대한 연구는 활발하게 진행되고 있지 않은 것 같다. 거시적 구조변화가 발생하면 위에서 제시한 모형들의 모수값들이 상당한 변화를 가지게 되는 경우가 발생할 수도 있다. 또는 위에 제시된 모형들의 모수 또는 일부가 구조가 변화된 주식시장에서는 효력(validity)을 갖지 못할 수도 있다. 이 같은 경우에는 새로운 주가결

7) 이 점을 지적해주신 심사위원께 이 자리를 빌어 감사의 말씀을 드립니다.

정모형들의 정립이 요청될 것이다.

증권시장의 중요 모수인 분산의 변화와 평균 및 분산의 동시변화는 1997년 9월에 발생하고 증권시장의 구조는 1999년 11월에 변화가 발생하였다. 왜 이와같은 시차가 발생한 것인가?

1999년 9월은 한국이 IMF에 가기 직전이다. 기아사태를 비롯한 여러 가지의 실무적 측면에서의 불안정과 단기외채위기의 전조가 여러 면에서 노정되는 시기이다. 이 시기는 이미 한국경제의 구조변화의 실마리가 경제내에 뿌려지는 시기라고 할 수 있을 것이다. 이것을 평균과 분산이 포착하고 있다. 그래서 변화가 시작하는 출발점이었다고 할 수 있을 것이다. 그리고 평균과 분산이 변하면서 구조변화가 형성되기 위한 변환지속시기 또는 전이시기(transition period)가 2년정도 지속된 것 같다. 변화의 싹이 트면 착근의 시기가 필요하다. 착근을 확고하게 하기 위하여는 착근에 방해되는 요소는 제거되어야 한다. 이 착근의 방해요소를 제거하는 작업이 IMF 직후에 들어선 새정부에 의하여 실행된 여러 정책이었다. 먼저 제도적 측면에서 이루어졌다. 구조조정에 대한 원리들이 정립되었다. 이에따라 정부는 정부가 해야할 역할들을 수행하였고 기업은 기업이 해야 할 역할을 수행하였다. 기업간의 합병, 기업의 외국인에게의 매각, 대우같은 초대형 그룹의 도산과 해체, 현대그룹의 분할등 경제체제의 정비가 이루어졌다.

IT산업의 탄생과 발전이 이루어지고 있었다. 대기업이 한 차륜을 맡고 IT산업과 벤처산업을 중심으로 하는 중소기업이 한 차륜을 맡아 고속도로를 달리는 경제정책이 입안되고 실현되었다. 이러한 모든 것은 착근에 방해되는 요소의 제거와 착근을 도와주는 요소로 해석될 수 있다. 이와같은 기간이 2년정도 걸렸다고 볼 수 있다. 이 과정을 거친 후 새로운 증권시장구조, 즉 경제구조가 형성되었다고 볼 수 있을 것이다.

V. 결 론

증권시장이 구조변화를 하였는지, 구조변화가 발생하였다면 몇 회에 걸쳐 언제 발생하였는지를 검정하였다. 자기회귀모형을 정립하고 자기회귀모형에서 구조변화가 이루어진 회수와 구조변화 발생일을 찾아내는 검정통계량과 이 통계량에 대한 임계값을 구하는 식을 유도하고, 한국 종합주가지수의 일별 수익률에 적용하였다. 증권시장을 규명하는 모수들인 평균과 분산의 변화가 발생한 회수와 발생일을 찾고자 정규분포의 최대우도함수를 유도하고 이 우도함수에 Akaike 정보기준을 적용시켜 모수들의 변화회수와 변환점을 발견하는 검정통계량 추정식과 임계값 추정식을 살펴보았다.

한국 종합주가지수의 일별수익률은 분산의 변화 그리고 평균 및 분산의 동시변화가 1997년 9월 27일에 발생하였다. 변화는 1회 발생하였다. 자기회귀모형에 의할 때 증권시장의 구조변화는 1999년 11월 16일에 이루어졌다. 평균과 분산의 변화가 일어난 시기는 구조변화의 단계를 시작하는 시기이고 구조변화에 알맞는 환경조성에 2년이 소요된 후에 1999년 11월 16일에 구조변화가 정착되었다. 정착이 이루어진 후에야 비로서 이 두기간은 서로 다른 경제구조와 증권시장구조 하에서 발전이 이루어지고 있다고 할 수 있을 것이다.

참 고 문 헌

李逸均, "주가시계열의 성질과 특성 : 한미비교", 재무관리논총, 제7권 제1호, (2001. 2), 1-48.

李逸均, "증권의 日別收益率과 月別收益率의 特性에 관한 연구", 증권학회지, 제11집, (1989), 199-229.

Akaike, H., "Information Theory and an Extension of the Maximum Likelihood Principle," 2nd *International Symposium of Information Theory* (B.N.Petrov and E. Csaki, eds.), Akademiai Kiade, Budapest, (1973), 267-81.

Andrews, D. W. K., "Testing for Structural Instability and Structural Change with Unknown Change Point," *Econometrica* 61, (1993), 821-856.

Andrews, D. W. K. and W., "Ploberger Optimal Tests When a Nuisance Parameter is Present Only under the Alternative," *Econometrica* 62, (1994), 1383-1414.

Bai, J. and P. Perron., "Estimating and Testing Linear Models with Multiple Structural Changes," *Journal of Econometrics* 66, (1998), 47-78.

Bai, J., "Likelihood Ratio Tests for Multiple Structural Changes," *Journal of Econometrics* 91, (1999), 299-323.

Chen, J. and A. K. Gupta, "Likelihood Procedure for Testing Change Points Hypothesis for Multivariate Gaussian Model," *Random Operators and Stochastic Equations*, 3, (1995), 235-244.

Chen, J. and A. K. Gupta, "Testing and Locating Variance Change Points with Application to Stock Prices," *Journal of the American Statistical Association,* 92, (1997), 739-747.

Chen, J. and A. K. Gupta, "Change Point Analysis of a Gaussian Model," *Statistical Papers*, 40, (1999), 323-333.

Chen, J. and A. K. Gupta, *Parametric Statistical Change Point Analysis.* Birkháuser, (2000).

Chernoff, H. and S., "Zacks Estimating the Current Mean of a Normal Distribution Which is Subject to Changes in Time," *Annals of Mathematical Statistics,* 35, (1964), 999-1018.

CsöRgö, M. and L. Horváth, "Nonparametric Methods for the Change Point Problem," *Handbook of Statistics*, 7,(P. R. Krishnaiah and C. R. Rao, Eds.), John Wiley, New York, (1988), 403-425.

Delgado, A.M. and J. Hidalgo, "Nonparametric Inference on Structural Breaks," *Journal of Econometrics* 96, (2000), 113-144.

Diaz, J., "Bayesian Detection of a Change of Scale Parameter in Sequences of Independent Gamma Random Variables," *Journal of Econometrics*, 19, (1982), 23-29.

Dhrymes, P., *Time Series, Unit Roots, and Cointegration*, Academic Press, (1998).

Gupta, A. K. and J. Chen, "Detecting Changes of Mean in Multidimensional Normal Sequences with Application to Literature and Geology," *Computational Statistics*, 11, (1996), 211-221.

Haccou, P., E. Meelis, and S. Geer, "The Likelihood Ratio Test for the Change Point Problem for Exponentially Distributed Random Variables," *Stochastic Processes and Their Applications*, 27, (1988), 121-139.

Hansen, B. E. "Testing for Structural Change in Conditional Models," *Journal of Econometrics,* 97, (2000), 93-115.

Hawkins, D. M., "Testing a Sequence of Observation for a Shift in Location," *Journal of the American Statistical Association*, 72, (1977), 180-186.

Horváth, L., "The Maximum Likelihood Method for Testing Changes in the Parameters of Normal Observation," *Annals of Statistics*, 21, (1993), 671-680.

Hsu, D. A., "Tests for Variance Shifts at an Unknown Time Point," *Applied Statistics*, 26, No.3, (1977), 179-184.

Hsu, D. A., "Detecting Shifts of Parameter in Gamma Sequences with Applications to Stock Price and Air Traffic Flow Analysis," *Journal of the American Statistical Association*, 74, (1979), 31-40.

Inclán, C., "Detection of Multiple Changes of Variance Using Posterior Odds," *Journal of Business and Economics Statistics*, 11, (1993), 189-300.

Inclán, C. and G. C. Tiao, "Use of Sums of Squares for Retrospective Detection of Changes of Variance," *Journal of the American Statistical Association*, 89, (1994), 913-923.

James, B., K. L. James, and D. Siegmund, "Tests for a Change-Point," *Biometrika*, 74, (1987), 71-83.

Krishnaiah, P. R. and B. Q. Miao, "Review about Estimation of Change Points(P.R. Krishnaiah and C.R. Rao, Eds.)," *Handbook of Statistics*, 7(Elsevier, Amsterdam), (1988), 375-402

Lehmann, E. L., *Testing Statistical Hypotheses*, 2nd Edition. Wiley, New York, (1986).

Lucas, R., "Econometric Policy Evaluation : A Critique. In K. Brunner and A. Meltzer, Eds.," *The Phillips Curve and the Labor Market*. Amsterdam : North Holland. (2000), 1976.

Marriott, J, and P. Newbold, "The Strength of Evidence for Unit Autoregressive Roots and Structural Breaks," *Journal of Econometrics* 98, (2000), 1-25.

Perron, P., "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, (1989), 1361-1401.

Sen, A. K. and M. S. Srivastava, "On Multivariate Tests for Detecting Change in Mean," *Sankhya*, A35, (1973), 173-186

Sen, A. K. and M. S. Srivastava, "On Tests for Detecting Change in Mean," *Annals of Statistics*, 3, (1975), 98-108.

Sen, A. K. and M. S. Srivastava, "On Tests for Detecting Change in the Multivariate Mean. University of Toronto," *Technical. Report,* No.3, (1980).

Srivastava, M. S, and K. J. Worsley, "Likelihood Ratio Tests for a Change in the Multivariate Normal Mean," *Journal of the American Statistical Association*, 81, (1986), 199-204

Worsley, K. J., "On the Likelihood Ratio Test for a Shift in Location of Normal Populations," *Journal of the American Statistical Association*, 74, (1979), 365-367

Worsley, K. J., The Power of Likelihood Ratio and Cumulative Sum Tests for a Change in a Binomial Probability, *Biometrika*, 70, (1983), 455-464

<부록 I> 구조변환점의 추정방법

어느 한 경제변수에 대한 회귀방정식 모형에는 선형모형이건 비선형모형이건 간에 종속변수와는 다른 외생변수들이 독립변수 또는 설명변수로 도입되는 경우가 있다. 계량경제학 분야에서는 이 같은 현상을 집중적으로 분석하고 있다. 그런데 설명변수로 도입되는 변수들이 종속변수의 시차변수들(lagged varables)로 도입되는 경우도 있다. 이 경우는 주로 시계열 분석에서 다루고 있다. 물론 계량경제학은 이 두 분야를 모두 아울러 대상으로 삼고 있음을 부인할 수 없다. 그런데 시차변수가 도입되는 시계열은 시계열의 비정상성(nonstationarity)등을 비롯한 여러 가지의 통계학적 문제가 야기되는 경우가 많다. Andrews(1993)등을 비롯한 여러 연구들은 종속변수의 시차변수가 설명변수로 도입되는 경우보다는 종속변수의 시차변수 이외의 변수들이 설명변수로 도입되는 회귀방정식 모형에 의하여 구조변화를 다루고 있다. 그런데 Bai(1999)는 종속변수의 시차변수가 도입될 때에도 강건성(robustness)을 갖는 구조변화 포착모형을 정립하였다. 그의 소론을 아래에 요약하고자 한다.

구조변화가 m번 발생하여 체제(regime)가 m+1개인 모형은 시계열 $\{y_t\}_{t=1}^{T}$에 대하여 다음과 같은 다중선형회귀식으로 정립할 수 있다.

$$
\begin{aligned}
y_t &= z_t'\delta_1 + u_t,\ \ t = 1,2,\cdots,k_1^0 \\
y_t &= z_t'\delta_2 + u_t,\ \ t = k_1^0 + 1,\cdots,k_2^0 \qquad (1\text{-}1) \\
&\ \vdots \\
y_t &= z_t'\delta_{m+1} + u_t,\ \ t = k_{m+1}^0,\cdots,T
\end{aligned}
$$

위에서 y_t는 관찰된 종속변수이고 z_t는 $(q\times 1)$ 벡터로 설명변수 또는 독립변수 벡터이다. 그리고 $\delta_i\,(i=1,\cdots,m+1)$는 회귀계수의 벡터로 $\delta_i \neq \delta_{i+1}$이다. u_t는 확률교란항(stochastic disturbance term)이다.

구조변화의 개수 m과 구조의 변환점 $k_1^0,\cdots,k_m^0$은 미지이다. 따라서 y_t와 z_t에 대한 관찰치가 T개 일때 m개의 구조변환점과 이 구조변환점에 대응되는 회귀계수 $\delta_1,\cdots,\delta_m$을 추정해야 한다. 구조변화의 개수 m과 m개의 구조변환점을 알고 있을 때에는 Perron(1987)이 수행한 것처럼 회귀계수의 추정이 용이하다. 식 (1)은 모든 회귀계수가 변화하고 있기 때문에 순수 구조변화모형(pure structural change model)이다.

이 구조변화모형에서는 $\delta_i \neq \delta_{i+1}$ $(i=1, \cdots, m)$이라는 가정아래에서 미지의 계수 $\delta_1, \cdots, \delta_{m+1}$을 먼저 추정해야 한다. 여기에서는 회귀함수가 변환점들에서 연속이라는 제약조건을 부과하지 않는다. 대신 구조가 이산적으로 변화한다는 가정하에서 모수의 추정과 구조변화개수 및 변환점들을 추정하고자 한다. 구조변화의 개수와 변환점들이 미지이다. 따라서 추정해야 할 대상이다. 그러나 먼저 구조변화의 개수 m과 변환점들을 알고 있다고 가정하여 추정하는 방법을 살펴보고 이 추정방법을 변용하여 미지의 구조변화 개수와 변환점들을 발견하는 방법을 고찰하도록 하겠다.

구조변화의 계수들은 최소자승법에 의하여 추정하는 것이 편리하다. 데이터의 집합의 크기가 T일 때 이 집합의 부분집합의 크기를 각각 $k_1^0, \cdots, k_{m+1}^0$이라 하면 $1 < k_1^0 < \cdots < k_{m+1}^0 = T$가 형성된다. 그러면 각 회귀식의 확률교란들의 제곱의 합을 최소화시키면 $\delta_i(i=1, \cdots, m)$를 추정할 수 있다. 즉 최소자승법에 의한 추정은 다음과 같다.

$$\min \sum_{i=1}^{m+1} \sum_{t=k_{i-1}}^{k_i} [y_t - z_t' \delta_z]^2$$

따라서 구조변환점들의 추정식은 목적함수의 비국부적인 전체영역에서의 최소화(global minimizers)를 형성시키는 방정식이다. 구조변화의 개수 m은 미지이므로 m의 값에 대한 추론(inference)이 중요시된다. l번 변화가 발생하고 있다는 귀무가설과 $(l+1)$번 변화가 발생하고 있다는 대립가설을 정립하고 이에 따른 검정통계량을 얻고 이 검정통계량의 확률분포가 도출되면 귀무가설의 검정이 가능하다.

식 (1)에서 u_t는 모든 t에 대하여 분산이 동일하다고 가정한다. 그리고 각 체제(regime)에서 z_t에 대하여 함수극한정리가 성립한다고 가정한다. 이 가정에 의하여 시계열상관오차(serially correlated disturbances)는 제거된다. 따라서 시차종속변수(lagged dependent variable)의 도입이 가능하다. 독립변수들은 각 체제에 있어서 점근적 정상성(asymptotic stationarity)이 성립하고 있다고 가정한다. 한 체제내의 iid 독립변수들의 분산 · 공분산 행렬이 양정치(positive definite)이다. 한 체제내의 모든 제 2계 정상과정은 이 가정을 만족하고 있으므로 $z_t z_t'$에 대하여 강대수율(strong law for large numbers)이 성립한다. 구조변화가 발생하는 자기회귀모형은 이 가정을 만족하고 있으므로 구조변화를 자기회귀모형에 의하여 검정할 수 있다. 끝으로 $i=1, \cdots, m$에 대하여 $\delta_i \neq \delta_{i+1}$, $k_i^0 = [T\tau_i^0]$이고 $\tau_i^0 \in [0, 1]$이며 $\tau_i^0 < \tau_{i+1}^0$이라고 가정한다. 따라서 $\widehat{k_i}$를

추정된 i번째 변환점의 위치라 하면 $\hat{\tau}_i = \hat{k}_i / T$이다. 즉 구조변화비율(break fraction)이다.

관찰치의 i 번째부터 j 번째까지의 부분집합을 부분표본 $[i, j]$라 하자. 전체표본을 m개 분할하면 m개 정수 $(k_1, \cdots, k_m)$벡터가 이루어진다. 이때 $1 < k_1 < k_2 \cdots < k_m < T$이다. 이 분할에 의하여 (m+1)개의 부분표본(subsample)이 형성된다. 양수 $\pi > 0$에 대하여 다음과 같이 정의하자.

$$\Lambda_\pi = \{(k_1, \cdots, k_m) : k_i - k_{i-1} \geq \pi T \; ; \; i = 1, \cdots, m+1\}$$

위에서 $k_0 = 0$이고 $k_{m+1} = T$이다. π는 작은 양수이다. 이와 같이 분할된 각 부분표본은 관찰치의 일부를 포함하고 있으며 어느 부분표본이나 공집합 ϕ가 아니다. 왜냐하면 모형설계자 또는 계량경제학자가 부분표본이 ϕ가 되지 않도록 전 표본의 분할을 설계하기 때문이다. π가 작은 양수이므로 $(k_1^0, \cdots, k_m^0) \in \Lambda_\pi$가 형성된다.

전체표본의 분할 $(k_1, \cdots, k_m)$에 대응되는 회귀모형의 잔차들의 제곱의 합을 $S_T(k_1, \cdots, k_m)$이라 하자. 즉

$$S_T(k_1, \cdots, k_m) = \sum_{i=1}^{m} \inf_{\{\phi_i\}} \sum_{k_{i-1}+1}^{k_i} (y_t - z_t' \phi_i)^2 \qquad (1\text{-}2)$$

위에서 $k_0 = 0$이고 $k_{m+1} = T$이다.

최적잔차 제곱을 다음과 같이 정의하자.

$$S_T(\hat{k}_1, \cdots, \hat{k}_m) = \min_{(k_1, \cdots, k_m)} S_T(k_1, \cdots, k_m)$$

위에서 Λ_π 위에서 $S_t(k_1, \cdots, k_m)$을 최소화하여 얻은 것을 $S_t(\hat{k}_1, \cdots, k_m)$이라 하자. Bai(1999)는 π가 0.05나 0.10이 적당한 값이라고 주장하고 있다. 구조변화가 m번 발생한다고 가정하자. 그러면 $(\hat{k}_1, \cdots, \hat{k}_m)$가 $(k_1^0, \cdots, k_m^0)$의 추정식(estimator)를 형성한다. $\hat{\tau}_i = \hat{k}_i / T$라 하자. 이 $\hat{\tau}_i$는 추정된 변화비율(break fraction)이다. Bai와 Perron(1998)은 $\hat{\tau}_i$가 $\hat{\tau}_i^0$에 대하여 T 일치성(T consistency)을 갖는다는 점을 밝힌 바 있다. 즉 $T(\hat{\tau}_i - \tau_i^0) = \hat{k}_i - k_t^0 = O_p(1)$이다.[1] 따라서 T 일치성이 성립한다.

귀무가설은 l번 구조변화가 발생한다는 것이다. 즉

$$H_0 : m = l$$

이에 비하여 대립가설은 $l+1$회에 걸쳐 구조가 변화한다는 것이다. 따라서 대립가설은 다음과 같다.

$$H_1 : m = l+1$$

검정통계량은 위에서 본 바와 같이 l개의 변환점이 존재할 때 얻게되는 최적잔차자승합이다. 이것은 $l+1$회에 걸쳐 구조변화가 발생할 때 얻게되는 최적 회귀잔차자승합의 차이에 의하여 도출할 수 있다. 귀무가설 아래에서 구조변환점 $(k_1^0, \cdots, k_l^0)$의 추정치를 $(\widehat{k_1}, \cdots, \widehat{k_l})$이라 하자. 그리고 $l+1$개 변환점이 발생할 때 잔차의 자승의 합이 최소가 되는 점을 $(\hat{k}_1^*, \cdots, \hat{k}_{l+1}^*)$이라 하자. 그러면 검정통계량은 다음과 같다.

$$\sup LR_t\,(l+1 \mid l) = \frac{S_T(\widehat{k_1}, \cdots, \widehat{k_l}) - S_T(\widehat{k_1^*}, \cdots, \widehat{k^*}_{l+1})}{\widehat{\sigma^2}\,(l+1)} \tag{1-3}$$

위에서 $\sigma^2(l+1) = S_T(\widehat{k_1^*}, \cdots, \hat{k}_{l+1}^*)/T$이다. 잔차가 iid 정규확률변수일 때 식 (3)의 검정은 우도비율검정이다. 정규성을 가정하지 않을 경우 이 통계량은 의사우도비율검정(pseudo-likelihood ratio test)이다. 식 (1-3)은 동태계획법에 의하여 용이하게 계산을 할 수 있다.

Bai(1999)는 식 (1-3)의 검정통계량이 (0, 1)위에서 독립적 Brownian bridge의 함수로 수렴하고 있음을 증명하였다. 이 성질을 이용하면 식 (3)에 대한 임계값(critical values)는 다음에 의하여 구할 수 있다.

$$\lim_{T \to \infty} p[\sup L\,(l+1 | l) > c] = 1 - \prod_{i=1}^{l+1} [1 - G_{i(c)}]$$

1) 확률을 P라 하자. $U_n \to 0$이면 $U_n = o_p(1)$이다. 모든 n에 대하여 $P[\,|U_n| \geq M\,] \leq \varepsilon$이 성립하도록 $\varepsilon > 0$에 대하여 $M < \infty$이 존재하면 $U_n = O_p(1)$이다. $|U_n|/|V_n| = o_p(1)$이면 $U_n = o_p(V_n)$이다. $|U_n|/|V_n| = O_p(1)$이면 $U_n = O_p(V_n)$이다.

위에서

$$G_i(c) = \frac{c^{q/2}\exp(-c/2)}{2^{q/2-1}\Gamma(q/2)}\left[\left(1-\frac{q}{c}\right)\log\frac{1-\eta_i}{\eta_i}+\frac{2}{c}+o(c^{-2})\right] \qquad (1\text{-}4)$$

유의수준 α가 주어질 때 이에 대응되는 임계값(critical value) c는 위 식에 의하여 쉽게 구할 수 있다. c가 크고 η_i가 작으면 식 (1-4)의 네모 괄호안의 첫 항이 $G_i(c)$의 값을 주도하게 되어 2/c는 무시할 수 있다. π값은 통제가능한 값이지만 $\eta_i = \pi(\tau_i^0 - \tau_{i-1}^0)$은 미지이므로 추정해야 한다. 이 추정은 $i=1,\cdots,l$에 대하여 $\widehat{\eta}_i = \pi(\widehat{\tau}_i - \widehat{\tau}_{i-1})$이다. 이때 $\widehat{\tau}_0 = 0$이고 $\widehat{\tau}_{l+1} = 1$이다. 귀무가설 아래에서 $\widehat{\tau}_i$는 T율로 $\widehat{\tau_i^0}$ 으로 수렴한다. 따라서 $\widehat{\eta}_i$가 η_i에 대한 적절하고 정당한 추정값이 된다.

위에서 제시된 검정과정(test procedure)은 일치성을 가진다. $m > l+1$이라는 대립가설 아래에는 $\sup LR_t(l+1 \mid l)$이 확률 1로 $\sup LR_t(l+1 \mid l) \rightarrow \infty$이다. 따라서 l보다 큰 구조변화가 존재할 때 l개 구조변화가 발생한다는 귀무가설은 확률 1로 기각하게 된다. Bai(1999)는 귀무가설 아래에서 $\sup LR_t(l+1 \mid l) \rightarrow O_p(T)$임을 증명하였다. l개 구조변화만이 발생한 경우에는 적어도 하나의 구조변환점은 일치성을 갖고 추정할 수 없다. 즉 $S_T(\widehat{k}_1, \cdots, \widehat{k}) - \Sigma_{t=1}^T u_t^2 = O_p(T) \rightarrow \infty$이다. 반면 $(l+1)$개 구조변환점이 추정에서 허용되면 $\Sigma_{t=1}^T u_t^2 - S_T(\widehat{k_1}^*, \cdots, \widehat{k}_{l+1}) = O_p(1)$이다 따라서 검정통계량의 분자는 $O_p(T)$이고 분모는 $O_p(1)$이다. 그러므로 대립가설 아래에서 검정통계량은 $O_p(T)$이다.

구조변화를 발견하기 위한 추정은 다음과 같이 수행할 수 있다. 먼저 구조변화가 발생하지 않는다는 귀무가설과 1회만 변화가 일어난다는 대립가설을 설정하고 검정을 수행한다. 이것이 기각되어 변화가 1회 발생하였다는 결과를 얻으면 변화가 1회 발생한다는 귀무가설과 2회 발생한다는 귀무가설을 검정한다. 이러한 절차로 진행되다가 $\sup LR_t(l+1 \mid l)$이 추가적 변화가 발생하지 않는다는 귀무가설을 기각하는데 실패하게 되면 이때 최적구조변화의 개수와 변환점을 얻게된다.

위에서 제시한 가정들이 충족된다고 하자. α_T가 충분하게 완만한 속도로 0으로 수렴하게 되면 $P(\widehat{m} = m_0) = 1$이 된다는 정리를 Bai(1999)가 증명하였다. 이때 P는 확률 척도(probability measure)이다.

α를 검정의 크기라 하고 c를 이 정리에 의한 극한분포로부터 계산된, 대응되는 임계값(critical value)이라 하자. 모든 큰 c에 대하여 $1-\Pi_{i=1}^{l+1}[1-G_i(c)] \leq \Sigma_{i=1}^{l+1} G_i(c)$ $\leq kc^{q/2}$ $\exp(-c/2)\exp(-c/3)$이므로 $\alpha \leq \exp(-c/3)$이다. 따라서 α가 적거나 또는 c가 크면 $c \leq -3\log(\alpha)$이다. α_T를 $\alpha_T = 1/T$로 잡으면 $c \leq 3\log T$이다. 대립가설 아래에서 $LR_t(l+1|l)$은 차수 $O(T)$이다. 따라서 귀무가설을 확률 1로 기각하게 된다. 이것은 표본이 클 때 변환점의 개수가 과소 추정되지 않다는 것을 의미한다. α_T가 천천히 0에 접근해가면, 즉 1/T이상으로는 접근해가지 않으면, 변화의 개수에 대한 일치추정이 형성된다. Bai(1999)는 simulation을 통하여 $\alpha = 0.05$가 적절함을 제시하고 있다.

<부록 II> 평균과 분산의 동시변화

독립적 확률분수의 수열을 x_1, x_2, …, x_T라 하고 이 수열의 확률분포함수를 각각 F_1, F_2, …, F_T라 하자. 그러면 변환점들(change points)은 다음의 가설을 검정하여 찾을 수 있다.

H_0: $F_1 = F_2 = \cdots = F_T$

H_A: $F_1 = \cdots = F_{k_1} \neq F_{k_1+1} = \cdots = F_{k_2} \neq F_{k_2+1} = \cdots = F_{k_q} \neq F_{k_q+1} = \cdots = F_T$.

위에서 H_0은 귀무가설이고 H_A는 대립가설이다. 그리고 $1 < k_1 < k_2 \cdots > k_q$이고 q는 미지의 변환점 개수이며 k_1, k_2, …, k_q는 각각 미지의 변환점들로서 추정해야 할 모수이다.

그런데 분포함수 F_1, F_2, …, F_T가 공통의 母數族 (parametric family) $F(\theta)$, $(\theta \in R^p)$에 속하는 경우에는 모집단 모수 θ_i $(i=1, \cdots, T)$에 대한 가설을 검정하여 변환점을 찾아 낼 수 있다.

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_T$$

위에서 q와 k_1, k_2, …, k_q를 추정해야 한다.

위에서 변환점을 찾기 위한 일반적 방법을 제시하였다. 이제 여기에서는 확률변수가 정규분포를 따를 때 발생되는 변환점을 발견하기 위한 추정방법을 살펴보도록 한다.

독립적 정규분포의 수열을 $\{x_i\}_{i=1}^{T}$이라 하고 이 수열의 모수를 (μ_i, σ_i^2) $(i=1, \cdots, T)$이라 하자. 그러면 변화가 발생하지 않는다는 안정성의 가설, 즉 귀무가설은 다음과 같다.

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_n = \mu$$

이에 비하여 변환점이 한번만 발생하였다는 가설을 정립할 수 있는데, 이것이 대립가설이며 이 가설은 다음과 같다.

$$H_1 : \mu_1 = \cdots = \mu_k \neq \mu_{k+1} \cdots = \mu_T$$

위에서 k는 변환점의 위치이다.

분산의 변화에 대하여 살펴보자. 독립적 정규확률변수의 수열이 $\{x_i\}_{i=1}^{T}$이며 $x_i \sim N(\mu, \sigma_i^2)$ $(i=1, \cdots, T)$이라 하자. 이때 귀무가설과 대립가설은 다음과 같다.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots, = \sigma_T^2 = \sigma^2$$

$$H_A : \sigma_1^2 = \cdots = \sigma_{ki}^2 \neq \sigma_{ki}^2 = \cdots = \sigma_{k2}^2 \neq \sigma_{k2+1}^2 \neq \cdots \neq \sigma_{kq+1}^2 = \cdots = \sigma_T^2$$

위에서 q는 미지의 변환점 개수이고, $1 \leqq k_1 \leqq k_2 < \cdots < k_q < T$는 각각 미지의 변환점들의 위치이다. 변환이 한번 발생한 경우에는 대립가설은 다음과 같다.

$$H_1 : \sigma_1^2 = \cdots = \sigma_k^2 \neq \sigma_{k+1}^2 = \cdots = \sigma_T^2$$

위에서 $1 < k < T$ 이다. 표본전체의 평균과 분산의 최대우도추정식은 잘 알려진 바와 같이 각각 다음과 같다.2)

$$\hat{\mu} = \frac{\sum_{i=1}^{T} x_i}{T}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{T} (x_i - \mu)^2}{T}$$

귀무가설 아래에서의 최대우도함수는 다음과 같다.

$$\log L_0(\hat{\sigma}^2) = -\frac{T}{2}\log 2\pi - \frac{T}{2}\log \hat{\sigma}^2 - \frac{T}{2}$$

그리고 변환점 k까지의 표본과 변환점 k이후부터 T까지의 표본에 대한 평균과 분산의 최대우도 추정식은 다음과 같다.

$$\hat{\mu}_1 = \frac{\sum_{i=1}^{k} x_i}{k}$$

$$\hat{\sigma}_1^2 = \frac{\sum_{i=1}^{k} (x_i - \hat{\mu}_1)^2}{k}$$

2) 대수우도함수는 다음과 같이 정립된다.

$$\log L_0(\sigma^2) = -\frac{T}{2}\log 2\pi - \frac{T}{2}\log \sigma^2 - \frac{\sum_{i=1}^{T}(x_i - \mu)^2}{2\sigma^2}.$$

$$\hat{\mu}_T = \frac{\sum_{i=k+1}^{T} x_i}{T-k}$$

$$\widehat{\sigma_T} = \frac{\sum_{i=k+1}^{T} (x_i - \hat{\mu}_T)^2}{T-k}$$

대립가설 H_1 아래에서 최대 대수우도함수는 다음과 같다.3)

$$\log L_1(\hat{\sigma}_1^2, \hat{\sigma}_T^2) = -\frac{T}{2}\log 2\pi - \frac{k}{2}\log \hat{\sigma}_1^2 - \frac{T-k}{2}\log \hat{\sigma}_n^2 - \frac{T}{2}$$

Akaike(1973)는 정보기준(information criterion)에 의하여 데이터에 적합한 모형을 선택할 수 있는 방법을 개발한 바 있다. Akaike(1973)에 의하면 모형선정을 위한 정보기준은 다음에 의하여 얻는다.

$$AIC(k) = -\log L(\hat{\theta}_k) + 2k, \quad k = 1, 2, \cdots, K$$

위에서 $L(\hat{\theta}_k)$는 모형 k에 대한 최대우도 함수이다. AIC를 최소화하는 모형이 고려하고 있는 데이터에 가장 적합한 모형으로 간주된다.

Akaike의 정보기준을 위에서 정립한 우도 함수 $L(\hat{\theta}_k)$에 적용할 때 AIC(k)가 최소화되는 k를 얻을 수 있다. 이 점이 분산의 변환점이다. 이것을 SIC(k)라 하자. 그러면 $SIC(\hat{k})$가 최소가 되는 $\hat{k}$를 찾으면 변환점 k의 위치를 추정하게 된다. 귀무가설 H_0 아래에서 이 식은 위에서 정립한 우도함수에 Akaike 정보기준을 적용할 때 다음과 같은 검정식을 얻는다.

$$SIC(T) = T\log 2\pi + n\log \hat{\sigma}^2 + T + \log T \qquad (2\text{-}1)$$

대립가설 H_1 아래에서는 다음과 같다.

$$SIC(k) = T\log 2\pi + k\log \widehat{\sigma_1}^2 + (T-k)\log \widehat{\sigma_T}^2 + T + 2\log T$$

---

3) 대립가설 H_1 아래에서 우도함수는 다음과 같다.

$$\log L_1(\sigma_1^2, \sigma_n^2) = -\frac{T}{2}\log 2\pi - \frac{k}{2}\log \sigma_1^2 - \frac{T-k}{2}\log \sigma_T^2 - \frac{\sum_{i=1}^{k}(x_i-\mu)^2}{2\sigma_1^2} - \frac{\sum_{i=k+1}^{T}(x_i-\mu)^2}{2\sigma_T^2}$$

따라서 변환점들은 두 번째 위치와 (T-2)번째 위치 사이에서만 찾아내게 된다. 정보기준에 의하면 다음이 성립하면 귀무가설 H_0를 수락한다.

$$SIC(T) < \min_{2 \leq k \leq T-2} SIC(k) \tag{2-2}$$

그리고 어떤 k에 대하여 다음이 성립하면 대립가설 H_1을 수락한다.

$$SIC(T) > SIC(k)$$

이때 변환점의 위치 $\hat{k}$는 다음에 의하여 확정한다.

$$SIC(\hat{k}) = \min_{2 \leq k \leq T-2} SIC(k) \tag{2-3}$$

Chen과 Gupta(1997)는 식 (Ⅱ-3)에 의하여 추정된 변환점의 위치 $\hat{k}$가 강한 일치성(consistency)을 가지고 있다는 점을 증명하였다. 유의수준 α와 이에 관련된 임계치(critical value)인 $c_\alpha(c_\alpha \geq 0)$를 고려하면 다음이 성립할 때 귀무가설 H_0를 수락한다.

$$SIC(T) \leq \min_{2 \leq k \leq T-2} SIC(k) + c_\alpha$$

이때 c_α는 다음에 의하여 얻는다.

$$c_\alpha \approx \left\{ -\frac{1}{a(\log T)} \log\log\left[1-\alpha+\exp\left(-2e^{b(\log T)}\right]^{-1/2} + \frac{b\log(T)}{a\log(T)} \right\}^2 \tag{2-4}$$

위에서

$$a(\log T) = (2\log\log T)^{1/2}$$
$$b(\log T) = 2\log\log T + \log\log\log T$$

앞에서는 평균은 고정되어 있고 분산만이 변하는 경우를 살펴보았다. 그러나 평균과 분산이 동시에 변하기도 한다. 여기에서는 이 점을 분석하고자 한다.

독립적 정규확률변수의 수열이 $\{x_i\}_{i=1}^{T}$이고 각 i=1, …, T에 대하여 각 x_i가 $x_i \sim N(\mu_i, \sigma_i^2)$이라 하자. 이때 Chen과 Gupta(1999)는 변환이 1회에 한하여 발생할때 다음과 같은 귀무가설과 대립가설을 수립하고 있다.

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_n \ ; \ \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_T^2$$

$$H_1 : \ \mu_1 = \cdots = \mu_k \neq \mu_{k+1} = \cdots = \mu_T;$$

$$\sigma_1^2 = \cdots = \sigma_k^2 \neq \sigma_{k+1}^2 = \cdots = \sigma_T^2$$

귀무가설 H_0 아래에서 최대대수우도함수는 다음과 같다.[4)]

$$\log L_0(\hat{\mu}, \ \hat{\sigma}^2) = -\frac{T}{2}\log 2\pi - \frac{T}{2}\log \ \hat{\sigma}^2 - \frac{T}{2}$$

최대우도법에 의한 각 모수의 추정식은 다음과 같다. 그리고 대립가설 H_1아래에서의 최대대수우도함수는 다음과 같다.[5)]

$$\log L_1(\ \hat{\mu}_1, \ \hat{\mu}_T, \ \hat{\sigma}_1, \ \hat{\sigma}_T) = -\frac{T}{2}\log 2\pi - \frac{k}{2}\log \ \hat{\sigma}_1{}^2 - \frac{T-k}{2}\log \ \hat{\sigma}_T{}^2 - \frac{T}{2}$$

평균과 분산이 동시에 변하는 시점이나 위치를 찾기 위한 분석은 Akaike의 정보기준을 사용하여 이루어질 수 있다. 이 정보기준에 의할 때 귀무가설하의 정보기준 SIC(T)는 다음과 같다.

$$SIC(T) = T\log 2\pi + n\log\hat{\sigma}^2 + T + 2\log T \qquad (2\text{-}5)$$

평균과 분산이 동시에 변화하는 시점 또는 위치를 k라 할 때 시계열의 처음부터 k까지의 수열과 k+1부터 T까지의 수열을 두 개의 집단으로 할 때 대립가설 H_1 아래에서 최대우도법에 의한 두 집단의 모수 추정치를 각각 $\hat{\mu}_1$ 및 σ_1^2과 $\hat{\mu}_T$ 및 σ_T^2라 하자. 이 추정식은 이미 앞에서 제시한 바 있다. 그러면 $2 \leq k \leq T\text{-}2$에 대하여 대립가설 H_1 아래에서 정보기준에 의한 통계량은 위에서 제시된 최대우도함수에 Akaike 정보기준을 이용하면 다음을 얻는다.

$$SIC(k) = T\log 2\pi + k\log\hat{\sigma}_1^2 + (T-k)\log\hat{\sigma}_T^2 + T + 4\log T \qquad (2\text{-}6)$$

---

4) 우도함수는 다음과 같다.

$$\log L_0(\mu, \sigma^2) = -\frac{T}{2}\log 2\pi - \frac{T}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{T}(x_i - \mu)^2$$

5) 대립가설 H_1 아래에서 대수우도함수는 다음과 같다.

$$\log L_1(\sigma_1^2, \sigma_T^2) = -\frac{T}{2}\log 2\pi - \frac{k}{2}\log\sigma_1^2 - \frac{T-k}{2}\log\sigma_T^2 - \frac{\sum_{i=1}^{k}(x_i - \mu)^2}{2\sigma_1^2} - \frac{\sum_{i=k+1}^{T}(x_i - \mu)^2}{2\sigma_T^2}$$

평균과 분산이 동시에 변화하는 변환점 k를 다음이 성립하는 $\hat{k}$ 에 의하여 추정한다.

$$\mathrm{SIC}(\hat{k}) = \min_{2 \le k \le T-2} \mathrm{SIC}(k) \qquad (\text{II-7})$$

검정통계량을 살펴보자. 먼저 다음과 같이 정의하자.

$$\begin{aligned} \triangle_T &= \min_{2 \le k \le T-2} [\mathrm{SIC}(k) - \mathrm{SIC}(T)] \\ &= - \min_{2 \le k \le T-2} [\mathrm{SIC}(k) - \mathrm{SIC}(T)] \\ &= \lambda_T^2 + 2 \log T \end{aligned}$$

위에서

$$\lambda_T^2 = \left[\max_{2 \le k \le T-2} \{ T \log \hat{\sigma}^2 - k \log \hat{\sigma}_1^2 - (T-k) \log \hat{\sigma}_n^2 \} \right]^{1/2}$$

그런데 λ_T은 $\lambda_T = (2 \log T - \triangle_T)^{1/2}$ 이므로 귀무가설 H_0 아래에서 모든 $x \in R$ 에 대하여 다음이 성립한다.

$$\lim_{T \to \infty} P\left[a(\log T)(2 \log T - \triangle_T)^{1/2} - b(\log T) \le x \right] = \exp(-2e^{-x})$$

위에서

$$\begin{aligned} a(\log T) &= (2 \log \log T)^{1/2} \\ b(\log T) &= 2 \log \log T + \log \log \log T \end{aligned}$$

유의수준을 α, 이와 연관된 임계치를 c_α라 하자. $\mathrm{SIC}(\hat{T}) < \min_{2 \le k \le T-2} \mathrm{SIC}(k)$ 일때 귀무가설 H_0를 수락하는 것은 $\mathrm{SIC}(\hat{k}) = \min_{2 \le k \le T-2} \mathrm{SIC}(k) + c_\alpha$ 를 수락하는 것과 같다. 이때 c_α 와 α 와의 관계는 다음과 같다.

$$1 - \alpha = P\left[\mathrm{SIC}(T) < \min_{2 \le k \le T-2} \mathrm{SIC}(k) + c_\alpha | H_0 \right]$$