

진행중인 시계열데이터에서 분산 변화점 탐지에 관한 연구*

최현석¹⁾ 강훈규²⁾ 송규문³⁾ 김태윤⁴⁾

요약

현재 발생중인 시계열 데이터에 분산변화가 일어날 경우 이동 분산비를 사용하여 분산 변화점을 빠른 시간 내에 탐지하는 문제를 다룬다. 이동 분산비의 분포로서 F 분포와 데이터에 의존하여 추정되는 실증적 분포를 제안한 후 상호비교를 통하여, 어느 방법이 시계열 데이터에서 분산의 변화점을 잘 탐지하는지 연구하였다.

주요용어: 분산 변화점, 이동 분산비, 실증적 분포

1. 서론

시간의 흐름에 따라 관측되는 시계열 자료는 갑작스런 외부요인의 변화에 의해 과거 자료와는 전혀 다른 형태로 변화하는 경우가 종종 있다. 즉 정책의 변경, 사회경제여건의 변화 등과 같은 외부요인에 의해 평균, 분산등과 같이 자료의 설명에 핵심적인 역할을 하는 모수의 변화가 흔히 발생한다. 이러한 현상은 특히 금융시계열에서 자주 발생하는데 금융경제정보가 빠르게 확산되는 요즘 시계열데이터내의 모수들이 자주 변화하는 경향을 보이고 있다. 본 연구에서는 시계열 데이터의 모수 변화가 일어나는 경우 이를 탐지하는 기법을 분산 변화점 탐지를 중심으로 연구한다.

지금까지 분산 변화점 탐지 문제는 과거에 관찰된 시계열 데이터를 분석하여 과거에 발생한 분산 변화점을 추정하는 문제가 주된 연구대상이었다. 이러한 문제에 대한 연구로서는 Kramer, Ploberger and Alt(1988), Tang and Macnail(1993), Inclan and Tiao(1994), Lee and Park(2001)을 들 수 있다. 본 논문에서는 현재에 계속 발생중인 시계열 데이터에 분산 변화가 일어날 경우 분산 변화점(Variance Change Point, VCP)을 빠른 시간 내에 탐지하는 문제를 다루고자 한다. 이와 관련 연구로는 금융시장의 변동성에 관한 연구(김태윤 등, 2004), 외환위기를 조기 경보하기 위한 연구(박원암, 2001; 김경수, 2004) 등이 있다.

* 본 연구는 산업자원부의 지역혁신 인력양성사업의 연구결과로 수행되었음.

1) (704-701) 대구광역시 달서구 신당동 1000번지, 계명대학교 자연과학대학 통계학과, 초빙전임강사

E-mail: chsuk1@kmu.ac.kr

2) (704-701) 대구광역시 달서구 신당동 1000번지, 계명대학교 자연과학대학 통계학과, 석사과정

E-mail: neogomusin@nate.com

3) (704-701) 대구광역시 달서구 신당동 1000번지, 계명대학교 자연과학대학 통계학과, 교수

E-mail: kms252@kmu.ac.kr

4) (교신저자)(704-701) 대구광역시 달서구 신당동 1000번지, 계명대학교 자연과학대학 통계학과, 교수

E-mail: tykim@kmu.ac.kr

최근 박경화 등(2005), 박윤성 등(2005)은 진행 중인 시계열 데이터에 대해 분산 변화점 탐지를 연구하였는데 그들은 분산 변화점 추정량으로써 이동 분산비를 제안하고 이동 분산비의 분포로써 F 분포를 사용하였다. 본 논문에서는 추정량의 분포로써 데이터에 의존하는 실증적(Empirical) 분포를 사용하는 실증적 기법을 제안하며 박경화 등(2005)의 F 분포에 근거한 기법과 상호비교를 통해 어느 기법이 분산 변화점을 효율적으로 탐지하는지 연구한다. 2절에서는 실증적 분산변화탐지 기법에 대하여 기술하고, 3절에서는 모의실험 절차와 그 결과에 대하여 기술하고 4절에서 결론을 맺는다.

2. 실증적 분산변화 탐지

실증적 분산 변화탐지 기법 소개를 위해 몇 가지 용어들을 먼저 정의한다.

- 분산 변화점

X_1, \dots, X_n 을 분산 변화점 ν 를 포함하고 있는 시계열데이터라고 하자.

$$\frac{\text{Var}(X_{i+1})}{\text{Var}(X_i)} = \alpha_{i+1}, \quad i = 1, \dots, n-1 \quad (2.1)$$

이라 할 때 만일 α_ν 가 1이 아니면 X_1, \dots, X_n 은 ν 시점에서 크기 α_ν 의 분산변화를 겪는다고 한다. 여기서 ν 는 1에서 n 사이의 어떤 정수이다.

- 이동 분산비(Moving Variance Ratio, MVR)

$v_{p,t} = \sum_{i=t-p+1}^t (X_i - \bar{X}_{p,t})^2$ 이고, $\bar{X}_{p,t} = \frac{1}{p} \sum_{i=t-p+1}^t X_i$ 일 때 t 시점의 MVR은 다음과 같이 정의한다.

$$r_{p,q,t} = \frac{v_{p,t}}{v_{p,t-q}} \quad (2.2)$$

여기서 p 는 $v_{p,t}$ 와 $v_{p,t-q}$ 계산에 사용된 데이터 수, q 는 $v_{p,t}$ 와 $v_{p,t-q}$ 간의 시차(lag)를 나타낸다.

박경화 등(2005)은 p 를 정보조절모수(information tuning parameter), q 를 시차 조절모수(lag tuning parameter)라 정의하고 이들의 역할에 대해 논의한 바 있다.

- 분산 변화점 탐지

분산변화가 없다는 가정 하에서 $r_{p,q,t}$ 의 분포를 $G_{p,q,t}$ (i.e., $G_{p,q,t}(x) = P(r_{p,q,t} \leq x)$)라 하자. 주어진 t_0 시점 에서 G_{p,q,t_0} 의 $100(1-\alpha)\%$ 기각역에 r_{p,q,t_0} 의 값이 속하게 되는 경우 t_0 시점 에서 분산 변화를 탐지하는 신호를 발령한다.

실험에서 분산 변화점 탐지 기법의 정확도를 측정하기 하기 위해 다음과 같은 측도를 사용한다.

• VT비(Valid to Total Ratio, VTR)

주어진 구간 $[1, T]$ 에 한 개의 분산 변화점 ν 를 포함하고 있다고 하자. 이 구간에서 VT비는 $VTR_{r,T} = \text{'유효한 신호수/총 신호수'}$ 로 정의된다. 여기서 총신호수는 $[1, T]$ 에서 발생한 총 신호수, 유효한 신호수는 구간 $[\nu - \gamma, \nu + \gamma] \subset [1, T]$ (여기서 γ 는 사전에 주어지는 작은 수)에서 발생한 신호수를 나타낸다.

박경화 등(2005), 박윤성 등(2005)은 $r_{p,q,t}$ 의 분포로써 자유도 (p, p) 인 F 분포($F(p, p)$)를 사용하여 이동분산비의 분산 변화점 탐지에 대하여 연구하였다. 그들의 방법은 이동분산비가 F 분포라는 가정에 기초한 것이므로 실제 상황에서 이러한 가정이 타당하지 않을 수 있다. 본 논문에서는 이러한 문제점을 해결하기 위해 실증적(Empirical) 기법을 통해 이동분산비의 분포를 추정한 후 그 분포에 기초를 둔 분산 변화점 탐지를 제안한다.

• 실증적 분산변화 탐지기법

주어진 데이터 X_1, \dots, X_t 로부터 $r_{p,q,t}$ 분포를 추정한 후 $(\hat{G}_{p,q,t})$ 그 분포로부터 $100(1 - \alpha)\%$ 기각역을 설정하여 $r_{p,q,t}$ 의 값이 기각역에 속할 때 t 시점에서 분산 변화가 일어났다는 신호를 발생시킨다. 여기서 주어진 데이터로부터 $r_{p,q,t}$ 분포를 추정하여 $100(1 - \alpha)\%$ 기각역을 설정하는 과정은 $\{r_{p,q,i} : 1 \leq i \leq t\}$ 의 값들로부터 $100(1 - \frac{\alpha}{2})\%$ 와 $100\frac{\alpha}{2}\%$ 백분위수를 찾음으로써 쉽게 해결된다.

주목할 점은 이러한 과정을 통해 현재 데이터 X_t 까지 사용하여 $r_{p,q,t}$ 의 분포를 최신화(update)한다는 점이다.

3. 모의 실험

본 절에서는 F 분포에 기초한 기법과 실증적 기법의 비교를 모의실험을 통해 연구하며 그 결과를 통해 어느 기법이 시계열 데이터에서 분산 변화점을 효율적으로 탐지할 수 있는지 살펴본다.

3.1. 실험 데이터 설명

시뮬레이션을 위해 $\rho = -0.7, -0.3, 0.3, 0.7$ 일 때 식 (3.1), (3.2)와 같은 다른 분산을 가지는 2개의 자기회귀 (autoregressive, AR) 시계열 데이터를 생성한다.

$$X_t = \rho \times X_{t-1} + \sqrt{1 - \rho^2} \times e_t, e_t \sim N(0, \sigma_1^2), t = 1, \dots, 190, \sigma_1^2 = 1 \quad (3.1)$$

$$X_t = \rho \times X_{t-1} + \sqrt{1 - \rho^2} \times e_t, e_t \sim N(0, \sigma_2^2), t = 191, \dots, 200, \sigma_2^2 \neq \sigma_1^2 \quad (3.2)$$

즉 $X_1, \dots, X_{190} \sim N(0, \sigma_1^2)$ 이고 $X_{191}, \dots, X_{200} \sim N(0, \sigma_2^2)$ 인 시계열 데이터가 생성되며, $\nu = 191$ 에서 분산변화가 일어난다 (여기서 σ_2^2 으로 3, 5, 10을 사용하며 따라서 $\alpha_{191} =$

$\sigma_2^2/\sigma_1^2 = 3, 5, 10$ 이 된다. 앞으로 본 논문에서는 α_{191} 대신 α 라는 표현을 사용한다). 본 모의실험에서는 분산 변화점 $t = \nu = 191$ 을 탐지하기 위해 두 기법(F 분포에 기초한 기법과 실증적기법)을 사용하였으며 각 기법에 대해 α 값의 변화에 따른 조절모수 p, q 의 최적 선택 혹은 최적 p, q 값의 존재를 확인하기 위해 다양한 p, q 값에 대한 모의실험을 수행하였다. 즉 $\alpha = 3, 5, 10$ 에 대해 $p = 5, 10, 15, 20, 25, 30$ 과 $q = 1, 5, 10$ 을 사용하였다. 기각역으로는 $F(p, p)$ 와 $\hat{G}_{p, q, t}$ 의 95백분위수를 각각 사용하였다. 위의 내용을 요약하면 표 3.1과 같다.

표 3.1: 모의실험 절차

▷ $\rho = -0.7, -0.3, 0.3, 0.7$
▷ $\alpha = 3, 5, 10$
▷ $p = 5, 10, 15, 20, 25, 30$
▷ $q = 1, 5, 10$
▷ $\nu = 191$ 에서 분산변화
▷ 총신호수 : $t = 101$ 에서 $t = 200$ 까지 분산 변화 신호가 발생한 회수
▷ 유효한 신호수 : $t = 186$ 에서 $t = 196$ 까지 발생한 신호수
▷ F 분포에 기초한 기법의 기각역 : $F(p, p)$ 의 95백분위수
▷ 실증적 기법의 기각역 : $\hat{G}_{p, q, t}$ 의 95백분위수

여기서 비교를 위한 모의실험은 $101 \leq t \leq 200$ 에서 실행하였으며 $1 \leq t \leq 100$ 구간은 훈련 과정 혹은 burning과정으로 사용되었다. 참고로 박경화 등(2005), 박윤성 등 (2005)에서는 모의실험 데이터 수를 50개로 사용하였다.

3.2. 모의실험 결과 설명

정확한 VT비 값을 얻기 위해 주어진 (p, q) 조합에 대해 100번 반복한 모의실험의 평균 값이 표 3.2 ~ 표 3.5에 주어져 있다.

F 분포에 기초한 방법과 실증적 기법들에 대하여 $\alpha = 3, 5, 10$; $p = 5, 10, 15, 20, 25, 30$; $q = 1, 5, 10$ 인 경우의 VT비의 값($VTR_{5, 100}$)을 보여주고 있으며 각 케이스 별로 최고 VT비 값들은 굵게 표시되어 있다. VT비의 최고값을 표시한 이유는 VT비는 '유효한 신호수/총신호수'로 정의되므로 값이 클수록 바람직하며 그 때의 p, q 값들이 최적의 값들로 간주될 수 있기 때문이다.

모의실험 결과를 통해 다음과 같은 사실들을 확인할 수 있다(표 3.6 참조).

- (i) 전반적으로 실증적 기법이 F 분포에 기초한 기법보다 개선된 결과를 보여 준다.
- (ii) 실증적 기법이 F 분포에 기초한 기법보다 p 의 크기에 민감하지 않는 것으로 보인다. 즉 실증적 기법은 대부분의 α 값에서 최적의 p 가 15인 반면 F 분포에 기초한 기법은 최적의 p 가 5와 20 사이에서 변한다.
- (iii) $q = 5$ 에서 탐지기가 효율적인 결과를 보여준다.
- (iv) $|\rho|$ (의존도)가 증가함에 따라 신호비가 감소하는 경향이 있다.

표 3.2: 모의실험 결과($\alpha = -0.7$)

F분포에 기초한 기법					실증적 기법				
α	p	1	5	10	α	p	1	5	10
3	5	0.1637	0.1788	0.1682	3	5	0.1432	0.1926	0.1596
	10	0.0700	0.1618	0.1537		10	0.2224	0.1676	0.1396
	15	0.0300	0.1912	0.1431		15	0.1958	0.2800	0.1088
	20	0.0300	0.1438	0.1220		20	0.2238	0.1521	0.1444
	25	0.0300	0.1876	0.1125		25	0.1946	0.1582	0.1830
	30	0.0000	0.1880	0.1415		30	0.0000	0.1880	0.1415
5	5	0.1267	0.2244	0.1536	5	5	0.2220	0.2176	0.2032
	10	0.1050	0.2551	0.1595		10	0.2575	0.2727	0.1347
	15	0.1200	0.2479	0.1953		15	0.2805	0.2360	0.1390
	20	0.0800	0.2616	0.1470		20	0.3114	0.3035	0.1532
	25	0.0900	0.2710	0.1378		25	0.2816	0.3208	0.2078
	30	0.0800	0.2755	0.1372		30	0.2561	0.2187	0.1516
10	5	0.2283	0.2620	0.2193	10	5	0.2470	0.3871	0.2754
	10	0.2950	0.3446	0.2261		10	0.3289	0.4059	0.2521
	15	0.2850	0.3599	0.2254		15	0.3203	0.3881	0.1931
	20	0.2050	0.3596	0.2315		20	0.3315	0.4066	0.2466
	25	0.2800	0.3845	0.2604		25	0.3714	0.3556	0.2554
	30	0.2200	0.4012	0.2548		30	0.3859	0.3930	0.2636

표 3.3: 모의실험 결과($\alpha = -0.3$)

F분포에 기초한 기법					실증적 기법				
α	p	1	5	10	α	p	1	5	10
3	5	0.1450	0.2361	0.1701	3	5	0.1743	0.2175	0.1846
	10	0.0600	0.1969	0.1731		10	0.2101	0.2368	0.1556
	15	0.0500	0.2241	0.0980		15	0.2485	0.3129	0.1986
	20	0.0000	0.1362	0.1505		20	0.2217	0.2623	0.1424
	25	0.0400	0.1822	0.1720		25	0.2434	0.2796	0.1624
	30	0.0300	0.1223	0.0757		30	0.2776	0.2610	0.1987
5	5	0.1845	0.2299	0.2594	5	5	0.1768	0.2376	0.2910
	10	0.1800	0.3450	0.2213		10	0.2768	0.3142	0.2626
	15	0.1800	0.3615	0.2541		15	0.3138	0.3850	0.2140
	20	0.1400	0.3729	0.2185		20	0.3016	0.3753	0.2447
	25	0.1500	0.3207	0.1972		25	0.2992	0.3759	0.2671
	30	0.1400	0.3104	0.2041		30	0.3213	0.3697	0.2472
10	5	0.2782	0.3473	0.2799	10	5	0.2419	0.4004	0.3407
	10	0.3500	0.4658	0.2632		10	0.3183	0.4533	0.3295
	15	0.3517	0.4690	0.3247		15	0.3764	0.5258	0.3762
	20	0.4700	0.6028	0.3345		20	0.3757	0.4865	0.3520
	25	0.3400	0.5589	0.3380		25	0.3688	0.4901	0.3496
	30	0.2300	0.4664	0.3466		30	0.4077	0.5132	0.3933

표 3.4: 모의실험 결과($\alpha = 0.3$)

F분포에 기초한 기법					실증적 기법				
α	p	1	5	10	α	p	1	5	10
3	5	0.1220	0.2010	0.1802	3	5	0.1470	0.2463	0.1903
	10	0.1100	0.2277	0.1627		10	0.1951	0.2261	0.2163
	15	0.0800	0.1980	0.1943		15	0.2137	0.3192	0.1459
	20	0.0400	0.2256	0.1195		20	0.2163	0.2521	0.1562
	25	0.0400	0.1966	0.1474		25	0.2240	0.2854	0.1637
	30	0.0100	0.1159	0.0688		30	0.2352	0.2698	0.1736
5	5	0.1890	0.2780	0.2078	5	5	0.1840	0.2243	0.2184
	10	0.2617	0.4056	0.1805		10	0.2639	0.3663	0.2227
	15	0.1350	0.3707	0.2483		15	0.2942	0.4322	0.2561
	20	0.1350	0.4099	0.2019		20	0.3060	0.3832	0.2388
	25	0.1400	0.3501	0.2261		25	0.3050	0.3859	0.2402
	30	0.0900	0.2669	0.1842		30	0.3231	0.3340	0.2394
10	5	0.2692	0.2930	0.2483	10	5	0.2135	0.4343	0.3884
	10	0.3783	0.4445	0.2434		10	0.3464	0.4963	0.3282
	15	0.4250	0.5038	0.3279		15	0.3636	0.5102	0.3468
	20	0.4000	0.5679	0.3797		20	0.3830	0.5172	0.3637
	25	0.4100	0.4709	0.3490		25	0.3662	0.4966	0.3525
	30	0.2650	0.5614	0.3739		30	0.3977	0.5170	0.3620

표 3.5: 모의실험 결과($\alpha = 0.7$)

F분포에 기초한 기법					실증적 기법				
α	p	1	5	10	α	p	1	5	10
3	5	0.1166	0.1652	0.1843	3	5	0.579	0.2034	0.1663
	10	0.0983	0.1873	0.1417		10	0.1970	0.2217	0.1621
	15	0.1000	0.2316	0.0871		15	0.2029	0.1479	0.1686
	20	0.0700	0.1909	0.1153		20	0.2230	0.2003	0.1368
	25	0.0200	0.1889	0.1272		25	0.2371	0.1789	0.1138
	30	0.0000	0.2208	0.1145		30	0.2178	0.1887	0.1270
5	5	0.1830	0.2387	0.2034	5	5	0.1709	0.2663	0.2012
	10	0.2600	0.2523	0.0969		10	0.2567	0.2544	0.2139
	15	0.1700	0.2784	0.1919		15	0.2434	0.2971	0.2032
	20	0.1400	0.2947	0.1750		20	0.2492	0.2791	0.1847
	25	0.0900	0.2270	0.1583		25	0.2809	0.2270	0.1930
	30	0.0900	0.2019	0.1811		30	0.2960	0.2950	0.1650
10	5	0.2218	0.3164	0.2699	10	5	0.2254	0.3947	0.3274
	10	0.3133	0.3552	0.2343		10	0.2933	0.4157	0.3016
	15	0.3317	0.3434	0.2200		15	0.3348	0.4539	0.2185
	20	0.3250	0.3516	0.2523		20	0.3271	0.3818	0.2384
	25	0.2400	0.4080	0.2983		25	0.3684	0.4288	0.2137
	30	0.2450	0.4304	0.2361		30	0.3856	0.4024	0.2360

표 3.6: VT비의 최고값

(,)는 p, q 값

α	방법	ρ			
		-0.7	-0.3	0.3	0.7
3	실증적 기법	0.2800 (15,5)	0.3129 (15,5)	0.3192 (15,5)	0.2371 (25,1)
	F분포에 기초한 기법	0.1912 (15,5)	0.2361 (15,5)	0.2256 (20,5)	0.2316 (15,5)
5	실증적 기법	0.3208 (25,5)	0.3850 (15,5)	0.4322 (15,5)	0.2971 (15,5)
	F분포에 기초한 기법	0.2755 (30,5)	0.3729 (20,5)	0.4099 (20,5)	0.2947 (20,5)
10	실증적 기법	0.4066 (20,5)	0.5258 (15,5)	0.5172 (20,5)	0.4539 (15,5)
	F분포에 기초한 기법	0.4012 (30,5)	0.6028 (20,5)	0.5679 (20,5)	0.4304 (30,5)

(v) α 가 증가함에 따라 VT비가 증가하는 경향이 있다. 즉, 분산변화가 클수록 변화점 탐지가 쉽다는 것을 나타낸다.

주목할 만한 결과는 (i)과 (ii)로써 이를 구체적으로 살펴보면 다음과 같다. 첫째 전반적으로 실증적 기법이 우월하다는 사실은 $r_{p,q,t}$ 의 분포 추정을 위해서는 비모수 기법(실증적 기법)이 모수적 기법(F분포에 기초한 기법)보다 우월함을 보여주고 있다. 특히 실제 상황에서 $r_{p,q,t}$ 의 분포가 의존성 등으로 인해 이론적으로 도출하기가 쉽지 않다는 점과 모집단의 X_t 분포가 수시로 변할 수 있다는 점 등을 고려할 때 실증적 비모수적 접근 방식이 진행 중인 시계열의 변화점 탐지 문제에서는 바람직스러운 것으로 판단된다. 둘째, 모수적 기법에 비해 실증적 기법의 정보조절 모수 p 의 선택이 α 의 크기 변화에 따라 변하지 않는 것은 실증적 기법의 안정성과 효율성을 보여 주고 있다. 즉 실증적 기법에서 $p = 15$ 의 선택이 대부분의 α 에서 좋은 결과를 기록한다는 사실은 정보 조절 모수 p 의 선택이 외부값 변화에 영향을 상대적으로 받지 않으면서 $r_{p,q,t}$ 분포 추정 관점에서 안정적으로 이루어지고 있음을 뜻한다. F분포에 기초한 기법의 경우 p 의 선택이 외부값 α 변화에 많은 영향을 받는 것으로 판단된다.

위에서 관찰된 결과 중 (iii), (iv)와 (v)는 박경화 등(2005)과 박윤성 등(2005)에 의해서 관찰된 결과들이다. 이 중 특히 실증적 기법이 $q = 1$ 이 아닌 $q = 5$ 에서 좋은 결과를 보이고 있는 것은 실증적 기법에서도 이동 분산들($v_{p,t}$ 와 $v_{p,t-q}$)간의 독립성 확보가 변화점 탐지의 효율성을 제고 할 수 있는 요인임을 확인해 주고 있다.

4. 결론

진행중인 시계열데이터에서 변화점 탐지를 위해 기존의 모수적 기법(F분포에 기초한 기법)과는 구별되는 비모수적인 실증적 기법을 제안하여 그 효율성을 기존의 모수적 기법과 비교하였다. 그 결과 성능뿐만 아니라 안정성에서도 실증적 비모수적 기법이 F분포에 기초한 기법보다 우월함을 확인하였다. 이를 요약하면 표 4.1과 같다.

데이터의 변화점 탐지는 과학 기술의 발전과 더불어 그 수요가 크게 증대되고 있는 분야이다. 특히 최근 들어 전자 메일 등 통신 기술의 발전은 많은 부문에서 실시간 데이터(현

표 4.1: 모의실험 결과 요약

	F분포에 기초한 기법	실증적 기법
α 가 증가	VT비 증가	VT비 증가
α 가 증가함에 따라 최적 데이터수 p	민감	안정적
$ \rho $ (의존도) 증가	효율감소	효율감소
분자와 분모간의 시차 q 의 크기	$q > 1$ 이 바람직	$q > 1$ 이 바람직

재 진행 중인 시계열 데이터의 일종) 획득을 가능하게 하였으며 따라서 이러한 실시간 데이터의 분석 기술 개발이 앞으로 더욱 중요해질 전망이다. 본 논문에서 다룬 문제는 이러한 실시간 데이터의 변화점 탐지 문제에 적용될 수 있다.

참고문헌

- 김경수 (2004). 효율적 외환위기 예측시스템의 구축방안: 신호접근모형을 중심으로, <금융조사보고서>, 한국금융연구원.
- 김태윤, 송규문, 도종두 (2004). A study on development of economic instability index, *Journal of Koran Data & Information Science Society*, **15**, 355–365.
- 박경화, 김태윤, 송규문, 최종재 (2005). Quick detection of variance change point for I.I.D. Data, *Journal of Koran Data & Information Science Society*, **16**, 173–183.
- 박윤성, 박경화, 최성환, 김태윤 (2005). Quick variance change point detection for time series in process, *Journal of Korean Data & Information Science Society*, **16**, 289–300.
- 박원암 (2001). 한국외환위기의 조기경보모형, <국제경제연구>, **7**, 55–79.
- Inclan, C. and Tiao, G.C. (1994). Use of cumulative sums for retrospective detection of changes of variance. *Journal of American Statistical Association*, **89**, 913–923.
- Kramer, W., Ploberger, W. and Alt, R. (1988). Testing for structural change in dynamic models. *Econometrica*, **56**, 1355–1369
- Lee, S. and Park, S. (2001). The cusum squares test for scale changes in infinite order moving average process, *Scandinavian Journal of Statistics*, **28**, 625–644.
- Tang, S. M. and Macnail, I. B. (1993). The effect of serial correlation on tests for parameter change at unknown time, *Annals of Statistics*, **21**, 552–575.

[2006년 2월 접수, 2006년 5월 채택]

A Study on Variance Change Point Detection for Time Series Data in Progress*

Hyun Seok Choi¹⁾ Hoon Kyu Kang²⁾ Gyu Moon Song³⁾ Tae Yoon Kim⁴⁾

ABSTRACT

This paper considers moving variance ratio (MVR) for variance detection problem with time series data in progress. For testing purpose, parametric method based on F distribution and nonparametric method based on empirical distribution are compared via simulation study.

Keywords: Variance change detection, Moving variance ratio, Empirical distribution

* This research was supported by the Program for the Training of Graduate Students in Regional Innovation which was conducted by the Ministry of Commerce Industry and Energy of the Korean Government.

1) Full time lecturer, Department of Statistics, Keimyung University, 1000 Sindang-Dong, Dalseo-Gu, Daegu, 704-701, Korea.

E-mail: chsuk1@kmu.ac.kr

2) Graduate Student, Department of Statistics, Keimyung University, 1000 Sindang-Dong, Dalseo-Gu, Daegu, 704-701, Korea.

E-mail: neogomusin@nate.com

3) Professor, Department of Statistics, Keimyung University, 1000 Sindang-Dong, Dalseo-Gu, Daegu, 704-701, Korea.

E-mail: kms252@kmu.ac.kr

4) (Corresponding author) Professor, Department of Statistics, Keimyung University, 1000 Sindang-Dong Dalseo-Gu, Daegu, 704-701, Korea.

E-mail: tykim@kmu.ac.kr