

## 이산코사인변환을 이용한 단위근 검정

이고운<sup>1</sup> · 여인권<sup>2</sup>

<sup>1</sup>숙명여자대학교 통계학과, <sup>2</sup>숙명여자대학교 통계학과

(2010년 10월 접수, 2011년 1월 채택)

---

### 요약

이 논문에서는 이산코사인변환(discrete cosine transform) 계수를 이용하여 AR(1) 과정에서 단위근을 검정하는 방법을 소개한다. 제안하고자 하는 방법의 이론적 타당성을 보이기 위해 정상 AR(1) 과정과 확률보행과정 하에서의 이산코사인변환 계수의 통계적 성질을 비교한다. 비교 결과를 이용하여 단위근 여부를 확인할 수 있는 검정통계량을 제안하고 붓스트랩을 활용하여 검정통계량의 분포를 유도한다. 모의실험을 통해 기존 Dickey-Fuller 검정과 검정력 비교를 실시하였다.

주요용어: 모수적 붓스트랩, 확률보행과정, Dickey-Fuller 검정.

---

### 1. 서론

자기회귀이동평균(ARMA) 모형을 이용한 분석에서는 시계열자료가 관측 시점과 관계없이 통계적으로 유사한 성질을 가지는 정상성을 만족한다고 가정한다. 일반적으로 정상성은 시계열자료의 평균과 분산이 시점에 관계없이 동일하고 자기공분산은 시차에만 의존하는 약한 의미에서의 정상성을 의미한다. 자기회귀이동평균모형에서 정상성 만족여부는 AR의 구조에 의해 결정되며 확률적 추세를 가지는 비정상 시계열의 경우 시계열을 차분하여 정상성을 만족하게 만든 후 분석한다.

이 논문에서는 AR(1) 과정을 기본 모형으로 사용하여 단위근 여부를 확인하는 방법을 제안하고자 한다. AR(1) 과정은 현재시점  $t$ 에서의 관측값이 바로 앞 시점  $t-1$ 에서의 관측값에 영향을 받으며 다음과 같은 관계식을 가진다.

$$(1 - \phi B)X_t = \delta + \varepsilon_t, \quad t = 1, 2, \dots, n$$

여기서  $B$ 는 후향연산자를 의미하며  $\varepsilon_t$ 는 평균이 0이고 분산이  $\sigma^2$ 인 백색잡음을 나타낸다. 단위근은 확률보행과정과 같이 자기회귀모수가  $\phi = 1$ 인 것을 의미하며  $X_0 = 0$ 인 경우  $E[X_t] = t\delta$ 이고  $\text{Var}[X_t] = t\sigma^2$ 이 되어 비정상적 시계열이 된다. 정상성을 만족하지 않는 시계열자료는 미래시계열의 패턴이 현재까지의 시계열 패턴과 다를 수 있기 때문에 시계열분석의 궁극적인 목적인 예측이 어려워진다. 추세성분이나 계절성분과 같이 결정적 추세에 의해 비정상성이 나타나는 경우에는 이들 성분을 모형화하여 미래 값에 대한 예측을 할 수 있지만 확률보행과정과 같이 확률적 추세를 가지는 경우에는 분산이 지속적으로 커지기 때문에 예측에 대한 신뢰도가 떨어질 수 밖에 없다. 시계열 앙상블과 같이 동일한 시점에

본 연구는 숙명여자대학교 2009년도 교내연구비 지원에 의해 수행되었음.

<sup>2</sup>교신저자: (140-742) 서울특별시 용산구 효창원길 52, 숙명여자대학교 통계학과, 부교수.

E-mail: inkwon@sm.ac.kr

여러 시계열 관측값이 있는 경우에는 분산이 일정하지를 쉽게 파악할 수 있으나 대부분의 분석은 단일 시계열에 대한 분석이기 때문에 이를 파악하기 힘들다. 이러한 과정에서 시계열이 결정적 추세를 가지는지 확률적 추세를 가지는지를 확인하는 방법으로 단위근 검정이 사용된다.

단위근 검정에 대한 이론적 분석은 Dickey-Fuller (1979, 이하 DF)에 의해 제안되었으며 이를 일반화시킨 Augmented DF 검정, Phillips-Perron (1988) 검정 등이 많이 사용되고 있다. 기존 연구에서는 AR 모수에 대한 추론을 기반으로 단위근 여부를 검정하는 반면 이 논문에서 제안하는 검정법은 이산코사인변환(discrete cosine transform; DCT) 계수를 활용한 오차항의 분산 추정량과 회귀모형에서의 OLS 추정량에 의한 분산 추정량을 비교한다. 이에 대한 이론적 근거를 제시하기 위해 확률보행과정과 정상 AR(1) 과정에서의 이산코사인변환 계수를 이용한 분산 추정량의 통계적 성질을 유도하고 몬테카를로 모의실험을 통해 제안하는 검정법과 기존의 DF-검정과의 검정력을 비교한다. 제안 방법은 절편과 분산에 영향을 받기 때문에 표준화된 기각역을 구하기 어려워 이 논문에서는 모수적 붓스트랩을 통해 기각역 또는 유의확률을 계산한다.

## 2. 이산코사인변환 계수의 통계적 성질 비교

시계열  $\{X_t\}_{t=1}^n$ 에 대해 이산코사인변환 계수는 다음과 같이 정의된다.

$$F_j = w_j \sum_{t=1}^n X_t \cos \left\{ \frac{\pi}{2n} (2t-1)(j-1) \right\}, \quad j = 1, 2, \dots, n$$

여기서  $w_1 = 1/\sqrt{n}$ 이고  $j = 2, \dots, n$ 에 대해  $w_j = \sqrt{2/n}$ 이다. Ahmed 등 (1974)이 발표한 이후 이산코사인은 현재까지 정보통신분야 및 멀티미디어분야에서 다양하게 사용되고 있는 틀이다. 이산코사인변환은 이산푸리에변환(discrete Fourier transform)에 기초한 것으로 비슷한 점이 많다. 이산푸리에변환은 시계열자료를 실수와 복소수들의 선형결합으로 표현하여 시간영역에서의 자료를 주파수영역으로 변환하며 시간영역에서 알아내기 어려운 자료의 특징을 파악할 수 있는 장점을 가지고 있으나 해석이 쉽지 않은 단점도 있다. 그에 비해 이산코사인변환은 이산푸리에변환에서 실수 값만 가지는 변환으로 좌우대칭한 우함수성질을 가지고 있기 때문에 계산이 쉬우면서도 높은 설명력을 가진다. 또한 신호의 에너지성분 대부분이 저주파 일부에 집중되는 강력한 에너지 집중 특성을 가지고 있기 때문에 신호처리 및 영상처리에 많이 사용된다. Rao와 Yip (1990)에는 이산코사인변환의 주요 성질과 응용사례들이 나와 있다.

DC 성분이라고 불리는  $F_1$ 의 경우 다른 주파수 대역의 계수와 통계적 성질의 조금 다르기 때문에 이 논문에서는  $j \geq 2$ 인  $F_j$ 를 사용하고 이후에 언급되는  $F_j$ 는  $j \geq 2$ 인 경우에 한 한다. 표기의 편의를 위해 기저함수  $c_{jt} = \cos\{\pi/2n(2t-1)(j-1)\}$ 라고 할 때, 여인권 등 (2006)은 평균이  $E[X_t] = \mu$ 이고 자기공분산이  $\text{Cov}[X_i, X_j] = \gamma_{|i-j|}$ 인 정상시계열에서의 이산코사인변환 계수는 다음과 같은 적률을 가지는 것을 보였다.

$$E[F_j] = 0, \quad \text{Var}[F_j] = \gamma_0 + \frac{4}{n} \sum_{s < t} c_{js} c_{jt} \gamma_{t-s}.$$

시계열  $\{X_t\}_{t=1}^n$ 가 정상 AR(1) 과정을 따르면  $k \geq 0$ 에 대해  $\gamma_k = \phi^k \sigma^2 / (1 - \phi^2)$ 가 되기 때문에  $j \geq 2$ 인 이산코사인변환 계수의 평균과 분산은 다음과 같다.

$$\mu_j(\phi) = E[F_j] = 0, \quad \sigma_j^2(\phi) = \text{Var}[F_j] = \frac{\sigma^2}{1 - \phi^2} \left\{ 1 + \frac{4}{n} \sum_{s < t} c_{js} c_{jt} \phi^{t-s} \right\}.$$

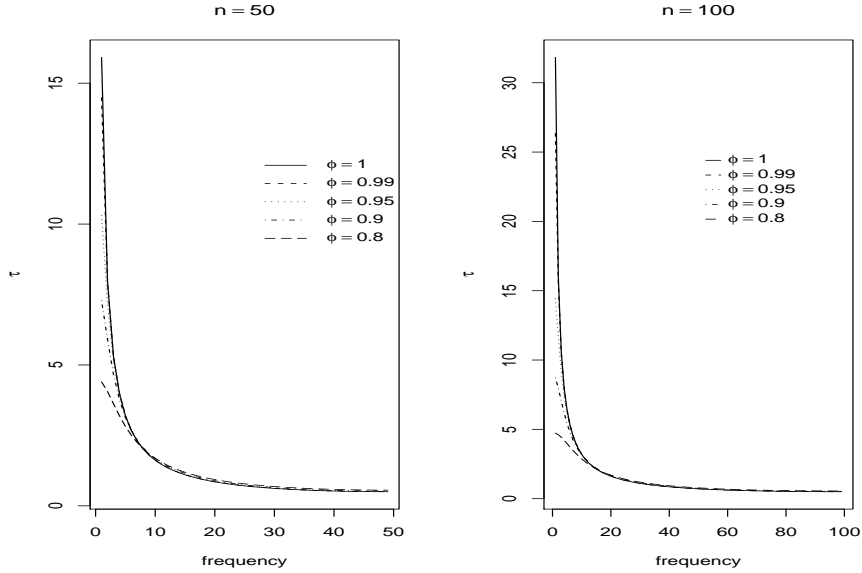


그림 2.1.  $\phi$ 의 변화에 따른  $\tau$ 의 비교

반면,  $X_0 = 0$ 인 확률보행과정인 경우  $E[X_t] = t\delta$ 이고  $s \leq t$ 에 대해  $\text{Cov}[X_s, X_t] = s\sigma^2$ 이기 때문에 확률보행과정에서의 이산코사인변환 계수의 적률은 다음과 같다.

$$\mu_j = E[F_j] = w_j \delta \sum_{t=1}^n tc_{jt} = \begin{cases} 0, & j = \text{홀수}, \\ \frac{\delta\sqrt{2}}{\sqrt{n}} \sum_{t=1}^n tc_{jt}, & j = \text{짝수}, \end{cases} \quad (2.1)$$

$$\sigma_j^2 = \text{Var}[F_j] = \frac{2\sigma^2}{n} \left\{ \sum_{t=1}^n tc_{jt}^2 + 2 \sum_{s < t} sc_{js}c_{jt} \right\}.$$

그림 2.1은  $n = 50$ 이고 100일 때 이산코사인변환 계수의 분산에서  $\sigma^2$ 을 제외한 나머지부분의 제곱근을  $j \geq 2$ 인 주파수에 따라 표시한 것으로 정상 AR(1)는

$$\tau_j(\phi) = \sqrt{\frac{1}{1-\phi^2} \left\{ 1 + \frac{4}{n} \sum_{s < t} c_{js}c_{jt} \phi^{t-s} \right\}}$$

이고 확률보행과정에서는

$$\tau_j = \sqrt{\frac{2}{n} \left\{ \sum_{t=1}^n tc_{jt}^2 + 2 \sum_{s < t} sc_{js}c_{jt} \right\}}$$

를 나타낸 그래프이다. 그림에서 보는 것과 같이  $\phi$ 값이 커질수록 저주파수 대역의  $\tau$ 값이 커지는 경향이 있으며 중주파수 대역과 고주파수 대역  $\tau$ 는 차이가 거의 없는 것으로 나타났다.

백색잡음  $\varepsilon_t$ 가 가우시안 백색잡음이면,  $F_j$ 는  $X_t$ 들의 선형결합이기 때문에 단위근 여부와 관계없이

표 3.1. DF검정

	실제모형	적합모형	귀무가설
Case I	$X_t = X_{t-1} + \varepsilon_t$	$X_t = \phi X_{t-1} + \varepsilon_t$	$\phi = 1$
Case II	$X_t = X_{t-1} + \varepsilon_t$	$X_t = \delta + \phi X_{t-1} + \varepsilon_t$	$\phi = 1, \delta = 0$
Case III	$X_t = \delta + X_{t-1} + \varepsilon_t$	$X_t = \delta + \phi X_{t-1} + \varepsilon_t$	$\phi = 1$
Case IV	$X_t = \delta + X_{t-1} + \varepsilon_t$	$X_t = \delta + \beta t + \phi X_{t-1} + \varepsilon_t$	$\phi = 1, \beta = 0$

$F_j$ 는 정규분포를 따른다. 확률보행과정에서는  $\phi = 1$ 이기 때문에  $F_j \sim N(\mu_j, \tau_j \sigma^2)$ 이므로

$$\frac{F_j - \mu_j}{\tau_j} \sim N(0, \sigma^2), \quad \left( \frac{F_j - \mu_j}{\tau_j} \right)^2 \sim \sigma^2 \chi_1^2$$

가 성립한다.  $F_j$ 들은 점근적으로 독립(asymptotically independent)이기 때문에 점근적으로 다음이 성립한다.

$$Y = \frac{1}{\sigma^2} \sum_{j=2}^n \left( \frac{F_j - \mu_j}{\tau_j} \right)^2 \sim \chi_n^2.$$

하지만  $\phi < 1$ 인 정상시계열인 경우에는

$$\frac{F_j - \mu_j}{\tau_j} = \frac{\sigma_j(\phi)}{\tau_j} \frac{F_j}{\sigma_j(\phi)} - \frac{\mu_j}{\tau_j} \sim N\left(-\frac{\mu_j}{\tau_j}, \frac{\sigma_j^2(\phi)}{\tau_j^2}\right)$$

이고

$$\left( \frac{F_j - \mu_j}{\tau_j} \right)^2 \sim \frac{\tau_j^2(\phi)}{\tau_j^2} \sigma^2 \chi_1^2(\lambda_j)$$

가 된다. 여기서  $\lambda_j = \mu_j^2 / \sigma_j^2(\phi)$ 는 비중심모수(non-centrality parameter)를 나타낸다. 절편이 없는 확률보행과정이나 절편이 있는 확률보행과정에서  $j$ 가 홀수이면  $\mu_j = 0$ 이기 때문에  $\lambda_j = 0$ 이 된다. 이것은  $|\delta|$ 가 상대적으로  $\sigma$ 보다 크면 클수록 확률보행과정과 정상 AR(1)과정에서 유도된  $Y$ 의 통계적 성질에 차이가 커지는 것을 의미한다.

### 3. 단위근 검정

Hamilton (1994)에서는 표 3.1과 같이 네 가지 케이스에 대해 DF검정의 검정통계량과 귀무가설 하에서의 유의수준에 대한 임계값을 제시하고 있다. 앞 절에서 언급한 것과 같이  $\delta = 0$ 인 경우에는 확률보행과정과 정상 AR(1)에서의  $Y$ 는 통계적으로 큰 차이가 없기 때문에 이 논문에서는 실제모형이 절편이 있는 확률보행과정에서의 단위근 검정에 대해 알아본다.

이 논문에서 귀무가설 하에서는 이산코사인변환 계수를 활용한 분산의 추정량과 적합모형  $X_t = \delta + \phi X_{t-1} + \varepsilon_t$  하에서  $\delta$ 와  $\phi$ 의 OLS 추정량  $\hat{\delta}$ 와  $\hat{\phi}$ 를 이용한 분산 추정량을 비교하는 검정통계량을 제안한다. 귀무가설 하에서 이산코사인변환 계수의 평균을 추정하기 위해 차분한 시계열  $W_t = X_t - X_{t-1}$ 를 구한 후 식 (2.1)의  $\delta$ 에  $\bar{W} = \{1/(n-1)\} \sum_{t=2}^n W_t$ 를 대입하여  $\mu_j$ 의 추정량  $\hat{\mu}_j$ 를 구한다. 이산코사인변환 계수를 이용한 분산추정량과 OLS 추정량을 이용한 분산추정량은 각각 다음과 같다.

$$\hat{\sigma}_1^2 = \frac{1}{n-2} \sum_{j=2}^n \left( \frac{F_j - \hat{\mu}_j}{\tau_j} \right)^2, \quad \hat{\sigma}_2^2 = \frac{1}{n-3} \sum_{t=2}^n \left( X_t - \hat{\delta} - \hat{\phi} X_{t-1} \right)^2.$$

표 3.2. DF검정과 DCT검정의 검정력비교

$\delta$	$\sigma$	$\phi$	$n = 50$				$n = 100$				$n = 200$			
			1%		5%		1%		5%		1%		5%	
			DF	DCT	DF	DCT	DF	DCT	DF	DCT	DF	DCT	DF	DCT
1	1	0.99	0.010	0.078	0.050	0.288	0.011	0.099	0.056	0.326	0.011	0.112	0.052	0.368
		0.95	0.012	0.098	0.056	0.350	0.017	0.184	0.083	0.548	0.039	0.454	0.169	0.852
		0.9	0.017	0.155	0.077	0.496	0.043	0.452	0.176	0.846	0.229	0.939	0.575	0.999
		0.8	0.043	0.403	0.169	0.825	0.204	0.944	0.534	0.999	0.894	1.000	0.990	1.000
		0.7	0.094	0.736	0.304	0.976	0.544	0.999	0.865	1.000	0.999	1.000	1.000	1.000
		0.6	0.186	0.937	0.487	0.998	0.841	1.000	0.981	1.000	1.000	1.000	1.000	1.000
		0.5	0.310	0.991	0.666	1.000	0.959	1.000	0.998	1.000	1.000	1.000	1.000	1.000
1	3	0.99	0.011	0.030	0.054	0.138	0.009	0.054	0.051	0.236	0.013	0.088	0.056	0.327
		0.95	0.015	0.036	0.061	0.173	0.017	0.107	0.086	0.397	0.045	0.379	0.174	0.810
		0.9	0.016	0.063	0.077	0.261	0.045	0.281	0.176	0.735	0.236	0.906	0.568	0.997
		0.8	0.040	0.193	0.156	0.582	0.207	0.850	0.536	0.995	0.895	1.000	0.992	1.000
		0.7	0.095	0.470	0.301	0.873	0.554	0.997	0.868	1.000	0.999	1.000	1.000	1.000
		0.6	0.186	0.768	0.489	0.980	0.835	1.000	0.979	1.000	1.000	1.000	1.000	1.000
		0.5	0.313	0.937	0.656	0.998	0.963	1.000	0.998	1.000	1.000	1.000	1.000	1.000
3	1	0.99	0.011	0.088	0.050	0.301	0.008	0.104	0.052	0.339	0.011	0.126	0.057	0.382
		0.95	0.013	0.116	0.058	0.383	0.018	0.193	0.080	0.542	0.044	0.476	0.180	0.864
		0.9	0.015	0.186	0.078	0.534	0.043	0.466	0.173	0.858	0.232	0.947	0.568	0.999
		0.8	0.040	0.455	0.154	0.850	0.217	0.947	0.540	0.999	0.897	1.000	0.990	1.000
		0.7	0.091	0.777	0.304	0.978	0.549	0.999	0.868	1.000	0.999	1.000	1.000	1.000
		0.6	0.187	0.954	0.490	0.999	0.833	1.000	0.980	1.000	1.000	1.000	1.000	1.000
		0.5	0.309	0.994	0.661	1.000	0.964	1.000	0.998	1.000	1.000	1.000	1.000	1.000
3	3	0.99	0.011	0.077	0.052	0.278	0.011	0.101	0.054	0.329	0.013	0.124	0.060	0.385
		0.95	0.012	0.102	0.061	0.357	0.017	0.183	0.080	0.529	0.046	0.471	0.180	0.857
		0.9	0.017	0.161	0.078	0.506	0.043	0.448	0.175	0.850	0.230	0.945	0.572	0.999
		0.8	0.040	0.417	0.158	0.831	0.211	0.939	0.530	0.998	0.896	1.000	0.991	1.000
		0.7	0.092	0.757	0.308	0.974	0.548	0.999	0.865	1.000	0.998	1.000	1.000	1.000
		0.6	0.185	0.934	0.485	0.997	0.838	1.000	0.978	1.000	1.000	1.000	1.000	1.000
		0.5	0.312	0.992	0.645	1.000	0.962	1.000	0.998	1.000	1.000	1.000	1.000	1.000

제안하고자 하는 검정통계량은 다음과 같은 두 분산추정량의 비로 정의한다.

$$T = \frac{(n-2)\hat{\sigma}_1^2}{(n-3)\hat{\sigma}_2^2} = \frac{\sum_{j=2}^n \left( \frac{F_j - \hat{\mu}_j}{\tau_j} \right)^2}{\sum_{t=2}^n (X_t - \hat{\delta} - \hat{\phi}X_{t-1})^2},$$

문제는  $\hat{\sigma}_1^2$ 와  $\hat{\sigma}_2^2$ 가 독립이 아니기 때문에 검정통계량  $T$ 의 확률분포를 유도하기 어렵다는 것이다.

귀무가설 하에서 제안된 검정통계량의 분포 및 기각역 계산을 위해 먼저 몬테칼로 모의실험을 실시하였다. 또한 이 방법의 성능을 비교하기 위해 제안 방법과 DF 방법의 검정력을 비교하였다. 모의실험에서는  $n = (50, 100, 200)$ ,  $\delta = (1, 3)$ ,  $\sigma = (1, 3)$ 가 사용되었으며 각 설정에서 확률보행과정을 따르는 시계열을 10만 번 반복 발생하여  $T$ 의 분포를 생성하고 1%와 5%에 해당하는 임계값을 구했다. 동일한  $n$ ,  $\delta$ 와  $\sigma$  조건 하에서  $\phi = (0.99, 0.95, 0.9, 0.8, 0.7, 0.6, 0.5)$ 에 대해 한번 반복하여 귀무가설을 기각시킨 비율을 계산하였다. DF 방법은 R의 fUnitRoots 패키지에서 제공하는 unitrootTest 함수를 이용하였으며 선택사항에서 lags = 1, trend = "ct"로 설정하여 표 3.1에서 Case IV로 분석하였다.

표 3.2는 DF 검정과 DCT를 이용한 검정의 검정력 결과를 보여준다. 표에서 보는 것과 같이 제안된 DCT 검정의 검정력은 모든 설정에 대해 기존 DF 검정보다 비교할 수 없을 정도로 높은 것으로 나타났다. 그러나 DF 검정의 경우  $\delta$ 와  $\sigma$ 의 값에 관계없이 같은  $n$ 인 경우 비슷한 검정력을 가지는 것으로 나타난 반면 DCT 검정의 경우  $\delta/\sigma$ 가 크면 클수록 높은 검정력을 가지는 것으로 나타났으며  $\delta/\sigma$ 가 같으면 기각역과 검정력이 유사한 것을 추가 모의실험을 통해 확인했다.

모의실험과 2절에서 본 것과 같이 귀무가설 하에서의  $T$ 의 분포는  $\delta$ 와  $\sigma$ 에 영향을 받기 때문에 표준화된 기각역을 구하기 어렵다. 이 논문에서는 이러한 문제를 해결하기 위해 다음과 같이 모수적 붓스트랩을 이용하여 기각역 또는 유의확률을 계산하는 방법을 제안한다. 먼저 차분한 시계열을  $W_t = X_t - X_{t-1}$ 를 이용하여  $H_0 : \delta = 0$ 인지를 검정하여 제안된 방법의 사용여부를 결정한다. 가설  $H_0 : \delta = 0$ 가 기각되면 다음과 같은 방법으로 기각역이나 유의확률을 계산한다.

단계 1. 시계열로부터  $\hat{\sigma}_1^2$ 와  $\hat{\sigma}_2^2$ 을 구하고 검정통계량  $T$ 를 계산한다.

단계 2. 귀무가설 하에서  $\delta$ 와  $\sigma^2$ 의 추정량을 다음과 같이 구한다.

$$\tilde{\delta} = \bar{W}, \quad \tilde{\sigma}^2 = \frac{1}{n-2} \sum_{t=2}^n (W_t - \bar{W})^2.$$

단계 3. 단계 2에서 계산된 추정량을 이용하여 다음과 같은 시계열을 생성하고,

$$X_t^* = \tilde{\delta} + X_{t-1}^* + \varepsilon_t^*, \quad \varepsilon_t^* \sim N(0, \tilde{\sigma}^2)$$

이 시계열을 이용하여 적합모형에서의 분산추정량  $\hat{\sigma}_2^{2*}$ 를 추정한다.

단계 4.  $T^* = (n-2)\hat{\sigma}_1^2 / ((n-3)\hat{\sigma}_2^{2*})$ 를 계산한다.

단계 5. 단계 3과 4를 반복하여  $T^*$ 의 붓스트랩분포를 생성하고 기각역 또는 단계 1에서 계산한  $T$ 에 대한 유의확률( $p$ -값)을 추정한다.

#### 4. 실증분석

Dickey 등 (1986)는 1948부터 1980년까지의 20-24세 미국 여성 1000명당 출생수를 이용하여 단위근 검정을 실시하였다. 이 논문에서 통계청 사회통계국 인구동향과에서 발표한 1970년부터 2009년까지의 ‘인구동태건수 및 동태율 추이’ 자료를 이용하여 출생아 수(명), 이혼건수(건), 기대수명(출생시 기대여명) 자료에 대해 제안한 방법을 이용하여 단위근 검정을 실시하였다. 그림 4.1은 이들 지표에 대한 연도별 추이를 보여주고 있는데 출생아 수는 지속적으로 감소하는 형태를 가지는 반면 이혼건수와 기대수명은 증가하는 형태를 가진다.

표 4.1은 붓스트랩에서 사용된 통계값과 만 번의 재표집과정을 통해 얻어진 기각역의 임계값과 유의확률을 정리한 것으로 출생아 수, 이혼건수, 기대수명의 시계열에는 모두 단위근을 있는 것으로 분석되었으며 이들 지표에 대한 분석에서는 차분된 시계열을 이용해야 하는 것으로 나타났다.

#### 5. 결론

Davies (2001)는 귀무가설이 정상과정이고 대립가설이 누적과정인 Harvey (1989)의 모형과 귀무가설이 단위근인 Dickey 등 (1986)의 모형에서 이산코사인변환을 이용하여 단위근 여부를 검정하는 방법을 소개하였다. 그의 검정통계량의 기본 구조는 단위근일때 저주파 대역에 에너지가 집중된다는 사실을 이

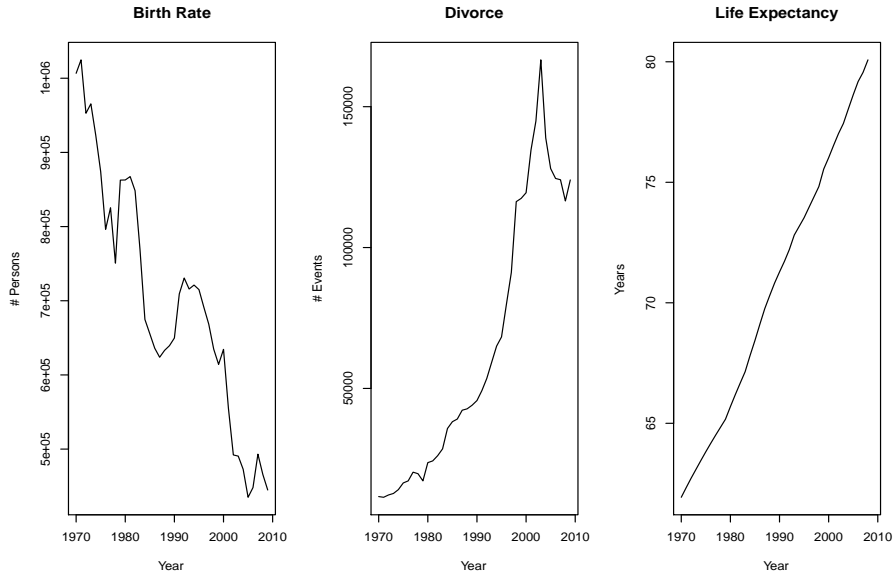


그림 4.1. 출생아 수, 이혼건수, 기대수명의 연도별 추이

표 4.1. 인구통계자료에 대한 단위근 검정 결과

통계값	출생아 수	이혼건수	기대수명
$\hat{\delta}$	-14405.03	2881.64	0.4776
$\hat{\sigma}$	41737.91	8351.57	0.1076
$(n-2)\hat{\sigma}_1^2$	$66198 \times 10^5$	$26504 \times 10^5$	0.42829
$(n-3)\hat{\sigma}_2^2$	$63819 \times 10^5$	$26409 \times 10^5$	0.38276
$T$	1.0373	1.0036	1.1189
95백분위수	1.6035	1.6065	1.5871
99백분위수	1.9222	1.9627	1.9176
유의확률	0.459	0.404	0.6116

용해 기본적으로  $\sum_{j=2}^n F_j^2 C_{jt}^2$ 과  $\sum_{j=2}^n F_j^2$ 를 비교하거나  $\sum_{j=2}^k F_j^2$ 와  $\sum_{j=k+1}^n F_j^2$ 을 비교한 것으로 우리가 제안한 방법과는 차이가 있다. 또한 Dickey 등 (1986) 모형에서 제안방법이 훨씬 큰 검정력을 가지는 것을 확인하였다.

제안된 방법에서는 절편  $\delta$ 와 표준편차  $\sigma$ 의 값에 검정력이 영향을 받으며 동일조건 하에서  $|\delta|/\sigma$ 의 값이 커지면 커질수록 검정력이 커지는 것으로 나타났다. 실제 적용에 있어 제안한 검정통계량의 분포를 유도하기 어렵고  $\delta$ 와 표준편차  $\sigma$ 를 모르기 때문에 이 논문에서는 붓스트랩을 이용하여 기각역의 임계값 또는 유의확률을 계산하는 방법을 알아 보았다. 향후 독립이 아닌 이차형식의 통계량 비(ratio)에 대한 적률과 분포에 대한 연구를 이용하여 귀무 가설 하에서의 검정 통계량의 분포를 근사적 또는 점근적으로 유도할 예정이다.

참고문헌

여인권, 윤화형, 조신섭 (2006). 시계열분석을 위한 주파수 공간상에서의 재표집 기법, <응용통계연구>, 19, 121-134.

- Ahmed, N., Natarjan, T. and Rao, K. R. (1974). Discrete cosine transform, *IEEE Transactions on Computers*, **23**, 90–93.
- Davies, R. B. (2001). Integrated processes and the discrete cosine transform, *Journal of Applied Probability*, **38A**, 701–717.
- Dickey, D. A., Bell, W. R. and Miller, R. B. (1986). Unit roots in time series models: Tests and implications, *The American Statistician*, **40**, 12–26.
- Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, **74**, 427–431.
- Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University Press.
- Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.
- Phillips, P. C. B. and Perron, P. (1988). Testing for unit roots in time series regression, *Biometrika*, **75**, 335–346.
- Rao, K. R. and Yip, P. (1990). *Discrete Cosine Transform: Algorithms, Advantages, Applications*, Academic Press, New York.



# A Unit Root Test via a Discrete Cosine Transform

Go-un Lee<sup>1</sup> · In-Kwon Yeo<sup>2</sup>

<sup>1</sup>Department of Statistics, Sookmyung Women's University

<sup>2</sup>Department of Statistics, Sookmyung Women's University

(Received October 2010; accepted January 2011)

---

## Abstract

In this paper, we introduce a unit root test via discrete cosine transform in the AR(1) process. We first investigate the statistical properties of DCT coefficients under the stationary AR(1) process and the random walk process in order to verify the validity of the proposed method. A bootstrapping approach is proposed to induce the distribution of the test statistic under the unit root. We performed simulation studies for comparing the powers of the Dickey-Fuller test and the proposed test.

**Keywords:** Dickey-Fuller test, parametric bootstrap, random walk.

---

---

This research was supported by the Sookmyung Women's University Research Grants 2009.

<sup>2</sup>Corresponding author: Associate Professor, Department of Statistics, Sookmyung Women's University, Hyochangwon-gil, 52, Yongsan-gu, Seoul 140-742, Korea. E-mail: inkwon@sm.ac.kr