

## 모형선택에서의 평활 적합도 검정<sup>1)</sup>

윤 용 화<sup>2)</sup>, 김 중 태<sup>3)</sup>

### 요약

본 연구의 목적은 평활 적합도 검정에 이용되는 모형선택 기법 중 AIC (Akaike information criteria) 기법과 BIC (Bayesian information criteria) 기법을 사용한 검정통계량들에 대한 검정력을 비교 분석함에 있다. 또한 이 두 가지 기법을 이용한 새로운 검정 통계량을 제시하고 기존의 검정 통계량들과 비교 분석하였다.

### 1. 서론

모형선택 기법은 그 자체적으로도 중요한 성질을 가질 뿐 아니라 밀도함수의 추정, 회귀함수의 추정, 베이지안 추론, 판별분석, 적합도 검정 등과도 깊은 관련이 있는 분야로 인식되고 있다.

최근 Eubank와 Ledwina를 비롯한 많은 통계학자들이 평활 적합도 검정에 대하여 모형선택 기법을 이용한 평활 검정 (smoothing test) 통계량을 제시하고 있다. 그 예들로서 Hart(1997), Eubank와 Hart(1992), Eubank와 Hart(1993), Eubank와 LaRiccia (1992), Ledwina(1994), Kim (1994), Kim(1994, 1998)등 많은 연구 논문들이 있다. 이러한 평활 적합도 검정 문제에 이용된 모형선택 기법은 크게 두 가지 유형으로 나누어진다. 그 첫 번째는 Akaike(1974)의 평균제곱합을 최소로 하는 AIC 유형의 기법이며 두 번째로 많이 사용되는 기법은 Schwartz(1978)의 베이지안 추론에서 유도된 BIC 모형선택 기법을 이용하는 방법이다. 그러나 이들의 모형선택 기법들 중에 평활 적합도 검정의 검정력을 높이는 문제에 있어서 어느 방법이 더 큰 검정력을 제시하는가에 대한 의문을 풀어주는 연구는 Lee(1996)의 웨이블릿 함수에서의 검정통계량 연구를 제외하고는 거의 없었다.

본 연구의 목적은 이들 두 모형선택 기법을 평활 적합도 검정 통계량에 각각 사용함으로써 검정에 대한 검정력들을 비교하여 AIC 기법과 BIC 기법 중 어느 기법을 이용한 검정 통계량들이 더 좋은 결과를 가지는가를 많은 대립 함수들을 사용하여 그 장단점들을 관찰하고 연구하는데 있다. 이 연구를 위하여 아래의 "효과 없음(no effect)" 모형에 대하여 위에서 언급한 두가지 모형선택 기법들을 이용한 평활 적합도 검정의 검정력을 비교하고자 한다. 제시된 모형은 다음과 같다.

$$y_j = f(t_j) + \varepsilon_j, \quad j = 1, \dots, n. \quad (1.1)$$

여기서  $f$ 는 미지의 함수이고, 오차  $\varepsilon_j$ 는 평균 0과 분산  $\sigma^2$ 을 가지는 독립적이고 동일한

1) 이 논문은 1998년도 대구대학교 학술 지원비에 의한 논문임  
2) 경북 경산시 진량면 대구대학교 통계학과, 712-714, 교수  
3) 경북 경산시 진량면 대구대학교 통계학과, 712-714, 조교수

분포를 가지는 확률변수들로 가정한다. 비록 모형 (1.1)은 귀무가설이 “효과 없음” - 즉,  $H_0: f = 0$  - 을 검정하기 위하여 비모수 회귀모형이나 선형모형에서의 가장 기본적으로 사용하는 모형이지만, 많은 다른 유형의 적합도 검정문제들로 변형, 적용할 수 있으며, 새로운 검정 통계량들의 형태와, 그에 대한 이론들과 정보를 제공하여 준다.

다음절에서는 푸리에 급수를 이용한 함수 추정 방법과 모형선택 기법인 AIC 기법과 BIC 기법을 소개하고, 3절에서는 2절에서의 기법들을 이용한 기존의 적합도 검정 통계량들을 조사하여 새로운 평활 적합도 검정 통계량들을 제시한다. 4절에서는 여러 가지 대립 함수들에 대하여 AIC 기법과 BIC 기법을 이용한 모의실험을 수행하여 검정 통계량들의 검정력들을 비교 분석하였다. 5절에서는 이 연구에 대한 결론을 다룬다.

## 2. 모형 추정과 모형 선택 기법들

모형 (1.1)의 미지함수  $f$ 를 추정하기 위하여, 직교성(orthogonality) 조건을 만족하는  $[0, 1]$  상에서 정의되는 함수  $U_{jn}(\cdot)$ ,  $j=1, 2, \dots, n$ , 을 다음과 같이 정의한다.

$$\sum_{k=1}^n U_{jn}(x_k) U_{ln}(x_k) = \sqrt{n} \delta_{jl}, \quad j, l=1, \dots, n \text{ 이고 } \sum_{k=1}^n U_{jn}(x_k) = 0, \quad j = 1, \dots, n.$$

위의 직교성 조건을 이용한 표본 코사인 푸리에 코사인 급수의 계수는 다음과 같이 정의된다.

$$a_{jn} = \frac{\sqrt{2}}{\sqrt{n}} \sum_{k=1}^n y_k \cos(j\pi x_k). \quad (2.1)$$

이때 표본 푸리에 코사인 급수의 계수  $a_{jn}$ 은 급수의 모수  $a_j = \int_0^1 f(x) \sqrt{2} \cos(j\pi x) / \sqrt{n} dx$ 에 수렴한다. 그러면 모형 (1.1)의 미지함수  $f$ 에 대한 푸리에 코사인 급수 추정량은 다음과 같다.

$$f_m(x) = \frac{\sqrt{2}}{\sqrt{n}} \sum_{j=1}^m a_{jn} \cos(j\pi x) \quad (2.2)$$

이고, 이때  $m$ 은  $1 \leq m \leq n$ 을 만족하는 상수이다.

귀무가설  $H_0: f = 0$ 을 검정하기 위하여  $f_m = 0$ 이 아닌 추정된 대립모형은 다음과 같다.

$$\hat{y}_r = f_m(x_r), \quad r=1, \dots, m.$$

모형 (1.1)에서의 오차분산  $\sigma^2$ 을 이미 안다고 가정하면 위험함수 (risk function) 혹은 평균제곱 오차는 다음과 같다.

$$r(m) = \sum_{r=1}^m (y_r - \sum_{j=1}^m a_{jn} \sqrt{2} \cos(j\pi x_r) / \sqrt{n})^2 + 2\sigma^2 m. \quad (2.3)$$

위의 평균제곱오차는 Rice(1984)의 연구를 참조하기 바란다. (2.3)에 있는 위험함수  $r(m)$  을 최소로 하는 값  $m$  을 찾는 방법은 다음의 식 (2.4)에 있는  $AIC(m)$  을 최대로 하는 값  $m$  을 찾는 방법과 동일하다.

$$AIC(m) = \sum_{j=0}^m a_{jn}^2 - 2\sigma^2 m \quad (2.4)$$

이때  $AIC(0) = 0$  이다. 식 (2.4)를 이용한 방법은 실제로 Akaike (1974)의 모형선택 방법과 동일한 기법이다. 식 (2.4)에서  $AIC(m)$  을 최대로 하는  $m$  을 선택하여 식 (2.2)에 있는  $f_m$  에 대입하면 모형 (1.1)에 있는 함수  $f$  의 추정량이 구해진다.

Schwarz(1978)는 베이저안 기법을 이용하여 다음의 BIC 판정 기준을 제시하였다.

$$BIC(m) = \sum_{j=1}^m a_{jn}^2 - \log(n)\sigma^2 m \quad (2.5)$$

식 (2.5)의 BIC 기법과 식 (2.4)에 있는 AIC 기법에서의 차이점은 2 대신에  $\log(n)$  을 사용하였음을 알 수 있고, AIC 기법은  $m$  의 값이 0부터 구하기 시작하는데 비하여 BIC 기법은  $m$  의 값을 1부터 찾아나간다. Eubank 와 Hart(1992), Hart(1997), Kim (1998)의 연구에서는 식 (2.4)의 AIC 기법에서의 상수 2 대신 일반적인 벌칙항(penalty term)  $C$ 를 사용하여 다음의 판정기준을 최대로 하는 값  $m$  자체를 검정통계량으로서 제시하였다.

$$EIC(m) = \sum_{j=1}^m a_{jn}^2 - Cm\sigma^2 \quad (2.6)$$

다음 절에서는 평활 검정 통계량들을 소개하고 위의 AIC 기법과 BIC 기법에서 선택된  $m$  을 적용하는 방법을 제시할 것이다.

### 3. 평활 적합도 검정통계량

Eubank와 LaRiccia(1992)는 모형 (1.1)에서의 귀무가설  $H_0 : f = 0$  을 검정하기 위하여 다음의 검정통계량을 제시하였다.

$$T(m) = \sum_{j=1}^m a_{jn}^2 / \hat{\sigma}^2. \quad (3.1)$$

이때 오차분산의 추정량은

$$\hat{\sigma}^2 = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (y_{i+1} - y_i)^2 \quad (3.2)$$

으로서 Rice(1984)에 의해 제시되었다. 그리고  $a_{jm}$  은 식 (2.1)에서 제시되었다. 식 (3.1)에 있는  $m$  은  $m \ll n$  의 값을 가지며,  $f$  가 이차 적분 도함수를 가지는 연속함수라면 표준 평활 실행에서  $m$  의 값은  $n^{1/3}$  의 비율보다 더 빠르지 않게 증가한다는 사실을 보였다.

두 번째 검정통계량으로서 Lee, Kim, Moon(1998)이 보인 가중치를 사용한 검정통계량은 다음과 같다. 이 검정통계량은 푸리에 급수 추정에서의 가중치인  $b_j$  를 첨가하여 구한 검정통계량을 정규화 (normalization) 시켜서 만든 통계량이다.

$$Z(m) = \frac{n \sum_{j=1}^m b_j^2 a_{jm}^2 - \sigma^2 \sum_{j=1}^m b_j^2}{\hat{\sigma}^2 (2 \sum_{j=1}^m b_j^4)^{1/2}} \quad (3.3)$$

이때  $a_{jm}$  은 식 (2.4)에서 제시되었고, 가중치

$$b_j = 1 - \frac{j}{m+1}, \quad j = 1, \dots, m,$$

인 값으로서  $m$  의 값이 증가함에 따라  $b_j$  의 값이 0으로 수렴하는 성질을 가지고 있다. 식 (3.3)에 있는  $\hat{\sigma}^2$  은 식 (3.2)에서 제시하였다.

모형 (1.1)을 다음과 같이 정의하자.

$$y_i = h(n) g\left(\frac{2i-1}{2n}\right) + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (3.4)$$

이때  $g$ 는 미지의 함수이고,  $h(n)$ 은 표본의 크기에 대한 함수로서  $h(n) \rightarrow 0, n \rightarrow \infty$ 를 만족한다. 그러면  $Z(m)$ 에 대한 점근적 분포는 다음의 정의에서 얻어진다.

**정리 3.1. (Lee, Kim, Moon (1998).)** 위의 모형 식 (3.4)에 대한  $H_0 : f = 0$  하에서,  $m \rightarrow \infty, n \rightarrow \infty$  일때,  $\sup_{1 \leq j \leq m} m^{1/2} \gamma(m, n) \rightarrow 0$ . 이때  $\gamma(m, n) = |a_{jn} - a_{jm}|$ . 그러면, 만약  $h(n) = m^{1/4}/\sqrt{n}$  이라면,  $Z(m) \xrightarrow{d} Z$ 이다. 여기서  $Z$ 는 분산이 1이며 평균  $\frac{\sqrt{5} \|g\|^2}{\sqrt{2}\sigma^2}$  를 가지는 정규확률변수이다.

세 번째 검정통계량은 모형 (1.1)의 검정을 위하여 Kim(1998)에서 제시한 검정통계량을 변형한 것으로서 다음과 같은 검정통계량을 고려 할 수 있다.

$$TK(m) = \sum_{j=1}^m a_{jn} / (m \hat{\sigma}^2) \tag{3.5}$$

이 검정 통계량은 식 (3.1)에 있는 Eubank와 Hart의 검정 통계량  $T(m)$ 에 대하여  $m$ 을 나눔으로서 평균값의 개념을 도입한 통계량이다. 모의실험에서의  $TK(m)$ 의 성질이나 검정력은 Eubank와 Hart(1993)의 검정통계량인  $T(m)$ 보다 더 좋은 검정력을 가지고 있음을 보여준다.

**정리 3.2.** 위의 모형 식 (3.4)에 대한 귀무가설 하에서,  $m \rightarrow \infty, n \rightarrow \infty$  일때,  $\sup_{1 \leq j \leq m} m^{1/2} \gamma(m, n) \rightarrow 0$ . 이때  $\gamma(m, n) = |a_{jn} - a_{jm}|$ . 그러면,  $TK(m) \xrightarrow{d} T_m$ 이다. 여기서  $T_m = \frac{1}{m} \sum_{j=1}^m Z_j^2$  이고,  $Z_1, Z_2, \dots$ 는 평균 0이고 분산이 1을 가지는 정규 확률변수이다.

**증명.** Kim, J.H. (1994)한 증명을 이용하면 쉽게 증명이 된다.

식 (3.1)에 있는 검정통계량들에 대한 점근적 성질이나 점근적 분포에 대한 이론과 증명은 그들의 연구에서 이미 밝혀져 있다. 본 연구에서는 다만 위의 평활 통계량들에 대한 AIC와 BIC의 기법들의 비교 연구에 중점을 둔다.

다음은 2절에서 소개한 AIC와 BIC의 기법을 사용하여 적절한  $\hat{m}$  값을 구한 다음 모의실험을 이용하여 이 두 모형선택 기법에 대한  $T(m), Z(m), TK(m)$  검정통계량들의 검정력을 비교할 것이다.

#### 4. 모의실험을 이용한 검정력 비교

2절에서 소개한 모형선택에서의 AIC 기법을 이용하여 구한  $\hat{m}$  값을  $\hat{m}_{AIC}$  라 두고, BIC 기법을 이용하여  $f$  함수를 추정하기 위해 구한 적절한  $\hat{m}$  값을  $\hat{m}_{BIC}$  로 두자. 이 두 개의 값을 구한 다음에, 귀무가설  $H_0 : f = 0$  를 검정하기 위한  $T(m), Z(m), TK(m)$  검정통계량에서  $m$  대신에  $\hat{m}_{AIC}$  과  $\hat{m}_{BIC}$  을 각각 대입하여 모의실험을 이용하여 구한 검정력들을 비교 분석한다.

모형 (1.1)에서의 귀무가설  $H_0 : f = 0$  에 대한 대립함수로서 다음의 모형 A, B의 각각 한가지씩의 대립함수와 모형 C에서의  $b$ 의 값의 변화에 따른 4가지 대립함수들을 고려하여 6개의 경우에 대하여 고려하였다. 이때 설계점은  $x_i = (i-1/2) / n, i=1, 2, \dots, n$  으로 두었다.

$$\text{모형 A : } f_1(x) = x + 0.5 \exp(-50(x-0.5)^2) . \tag{4.1}$$

$$\text{모형 B : } f_2(x) = \exp(-x/2) \cos(x) . \tag{4.2}$$

$$\text{모형 C : } f_3(x) = b \left( e^{4x} - \frac{e^4 - 1}{4} \right) \left( \frac{e^8 - 1}{8} - \frac{(e^4 - 1)^2}{4} \right)^{-1/2} . \tag{4.3}$$

모형 C 에서는  $b = i/4$ ,  $i = 1, 2, 3, 4$  인 경우에 대하여 조사하였다. <그림 1>은 모형 A, B와 모형 C의  $b = 0.25, 0.5, 0.75, 1.0$  인 4가지 함수들의 분포를 조사하였다.

위의 대립모형 A와 B에서의 검정력 비교에 있어서 모형 (1.1)의 오차에 대한 평균 0과 분산  $\sigma^2 = 0.1$  인 경우의 표본의 크기  $n = 10, 20, 30, 40, 50$  에 대하여 실험하였다. 또한 표본의 크기  $n = 30$  인 경우에 오차분산  $\sigma^2 = 0.01, 0.05, 0.1, 0.5, 1.0$  으로 두어 분산의 변화에 따른 검정력의 변화를 살펴보았다. 대립모형 C에 있어서는 오차분산의 값이  $\sigma^2 = 1.0$  인 경우에서 각각의  $b = 0.25, 0.5, 0.75, 1.0$  값에 따라 표본의 크기  $n = 10, 20, 30, 40, 50$  인 경우에 대하여 검정 통계량의 값의 변화를 조사하였다.

모의실험의 횟수는 1000번으로 하였으며 귀무가설  $H_0 : f = 0$  에 대한 각각의 검정통계량들에 대한 기각값은 유의수준  $\alpha = 0.05$  에서 모의실험을 이용하여 구하였다.

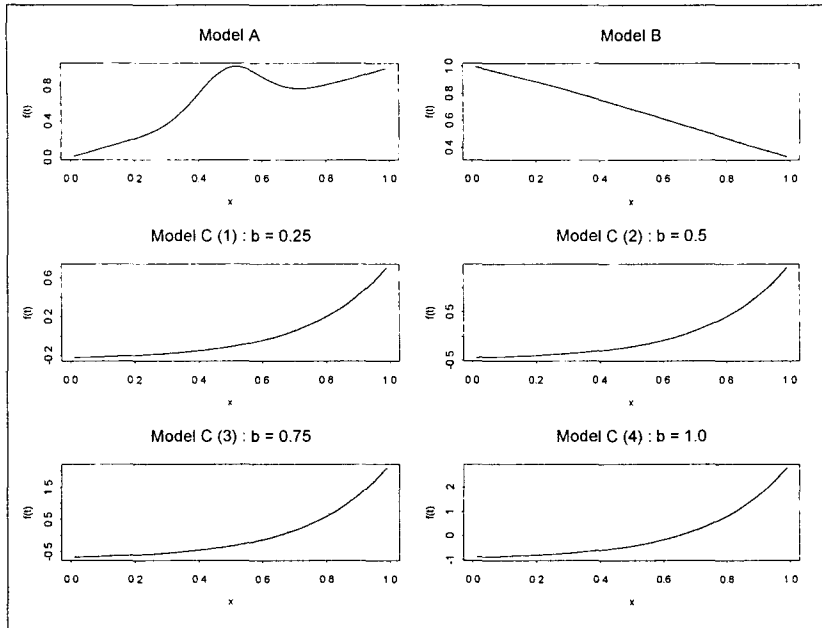
<그림 2>와 <그림 3>은 모형 A와 모형 B의 검정력 값을 비교한 것이다. AIC와 BIC 기법을 이용한 검정력의 비교 결과는 다음과 같다.

<그림 2> (a)와 <그림 3> (a)의 경우는 오차분산  $\sigma^2 = 0.1$  인 경우  $n$ 의 값에 따른 검정력 값의 변화를 조사하였는데  $n = 10$  인 경우에는  $Z(m)$ 과  $TK(m)$ 의 검정력 값의 비교에 있어서는 AIC 기법이 BIC 기법보다 조금 더 좋은 검정력을 가지는 것으로 나타나 보이지만  $n$  값이 20 이상인 경우에  $n$  값이 증가함에 따라서 6개의 검정통계량들의 검정력 값이 증가하는데,  $Z(m)$  통계량에서는 AIC와 BIC 기법을 이용한 두 검정력들의 값에 대한 차이가 거의 없었다.

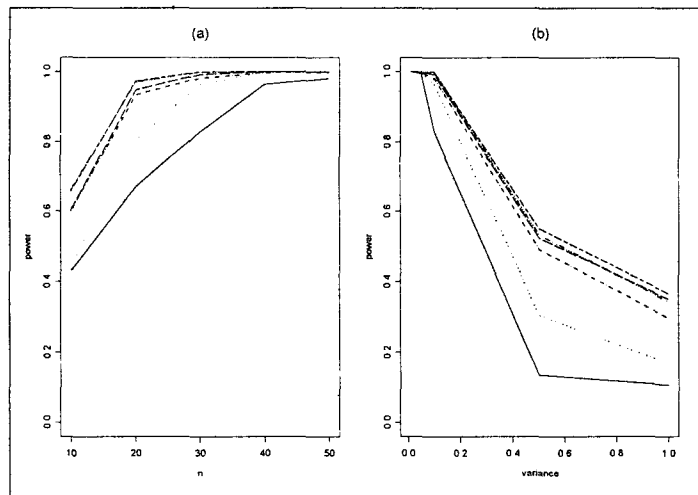
$TK(m)$  통계량에서는 BIC 기법이 AIC 기법보다 더 좋은 결과를 보였다.  $T(m)$  통계량에 있어서는 BIC 기법이 AIC 기법 보다 더 많은 차이를 가지고 좋은 검정력의 우위를 가짐을 알 수 있다. 그러나 전체적인 면에서 검정통계량들의 비교우위를 조사하여 보면  $Z(\hat{m}_{BIC})$ ,  $Z(\hat{m}_{AIC})$ ,  $TK(\hat{m}_{BIC})$ ,  $TK(\hat{m}_{AIC})$ ,  $T(\hat{m}_{BIC})$ ,  $T(\hat{m}_{AIC})$ 의 순서로  $Z(\hat{m})$ 이 가장 좋음을 알 수 있다.

<그림 2> (b)와 <그림 3> (b)는  $n = 30$  인 경우에 오차분산  $\sigma^2$  값들에 따른 6개의 검정통계량들에 대한 검정력 값들을 비교한 것이다. <그림 2> (a)와 <그림 3> (a)에서는  $n$ 의 값의 증가할 수록 검정력 값이 증가하였지만, <그림 2> (b)와 <그림 3> (b)에서는 오차분산의 값이 적을 수록 검정력의 값이 증가함을 보이고 있다. 검정통계량들의 검정력 값들에 대한 비교 우선순위는  $Z(\hat{m}_{BIC})$ ,  $Z(\hat{m}_{AIC})$ ,  $TK(\hat{m}_{BIC})$ ,  $TK(\hat{m}_{AIC})$ ,  $T(\hat{m}_{BIC})$ ,  $T(\hat{m}_{AIC})$  순서로서  $Z(\hat{m}_{BIC})$ 이 보다 강력한 검정력을 가짐을 알 수 있다.

아래의 <그림 4>는 식 (4.3)에 있는 모형 C에 대하여 오차분산  $\sigma^2 = 1.0$  인 경우에 있어서  $n$  값들의 변화에 따른 검정통계량들의 검정력 값들을 비교한 것이다.  $b = 0.25, 0.5, 0.75, 1.0$ 의 값의 변화에 따른 검정력 값들의 변화를 <그림 4>의 (a), (b), (c), (d)의 순서로 나타내었다. <그림 1>의 모형 C (1)에서 모형 C (4)까지의  $b$  값의 변화의 영향은 식 (4.3)에 있는 모형 C의  $f_3$  함수 분포의 폭의 변화를 증가시킴을 알 수 있다. <그림 4>에서 보듯이  $b$  값이 증가함에 따라 - 즉, 분포 값들의 폭이 커질수록 - 각 검정통계량들의 검정력 값들이 증가함을 알 수 있다. 전체

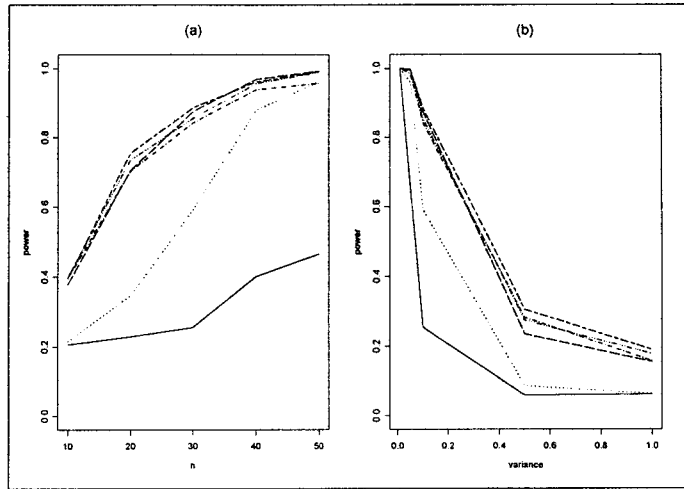


<그림 1.> 대립함수들에 대한 분포



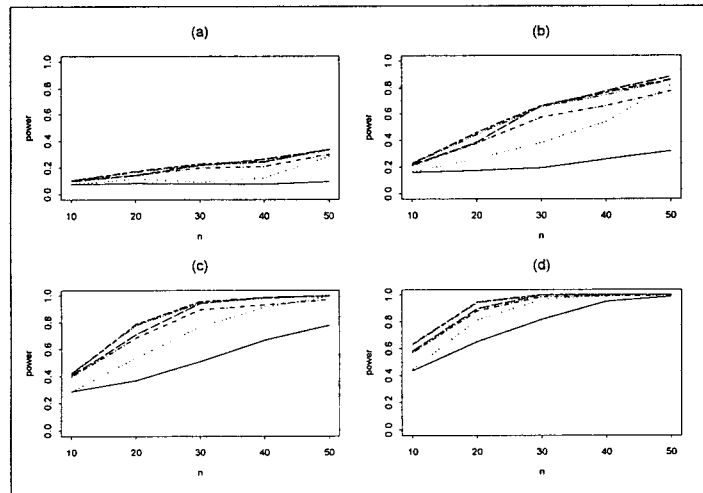
<그림 2.> 모형 A에 대한 검정력 비교

※ (a)  $\sigma^2=0.1$ 인 경우와 (b)  $n=30$ 인 경우, 검정력 값이 큰 것부터 작은 것의 순서는  $Z(\hat{m}_{BIC})$ ,  $Z(\hat{m}_{AIC})$ ,  $TK(\hat{m}_{BIC})$ ,  $TK(\hat{m}_{AIC})$ ,  $T(\hat{m}_{BIC})$ ,  $T(\hat{m}_{AIC})$  이다.



<그림 3.> 모형 B에 대한 검정력 비교

※ (a)  $\sigma^2=0.1$ 인 경우와 (b)  $n=30$ 인 경우, 검정력 값이 큰 것부터 작은 것의 순서는  $Z(\hat{m}_{BIC})$ ,  $Z(\hat{m}_{AIC})$ ,  $TK(\hat{m}_{BIC})$ ,  $TK(\hat{m}_{AIC})$ ,  $T(\hat{m}_{BIC})$ ,  $T(\hat{m}_{AIC})$  이다.



<그림 4.> 모형 C에 대한 검정력 비교

※  $\sigma^2=0.1$ 인 경우, 검정력 값이 큰 것부터 작은 것까지의 순서는  $Z(\hat{m}_{BIC})$ ,  $Z(\hat{m}_{AIC})$ ,  $TK(\hat{m}_{BIC})$ ,  $TK(\hat{m}_{AIC})$ ,  $T(\hat{m}_{BIC})$ ,  $T(\hat{m}_{AIC})$  이다.

적인 통계량들의 비교 순위는  $Z(\hat{m}_{BIC})$ ,  $Z(\hat{m}_{AIC})$ ,  $TK(\hat{m}_{BIC})$ ,  $TK(\hat{m}_{AIC})$ ,  $T(\hat{m}_{BIC})$



$T(\hat{m}_{AIC})$  순서로  $Z(\hat{m}_{BIC})$ 이 가장 우수함을 알 수 있다. 그러나 실제로  $Z(\hat{m})$ 와  $TK(\hat{m})$ 의 AIC 기법과 BIC 기법에 따른 검정력 차이는 모든  $b$  값에서 비교적 작게 나타나지만  $T(\hat{m})$ 은  $b$  값이 증가할수록 BIC 기법의 검정력이 두드러지게 커짐을 알 수 있다.

## 5. 결론

본 연구에서의 AIC와 BIC 기법에 대한 검정 통계량들의 검정력 비교에 있어서는 BIC 기법이  $n = 10$ 인 경우처럼 표본의 크기가 적은 경우를 제외할 때, 전반적으로 AIC 기법을 이용하는 것 보다 우수한 것으로 나타난다. 우리는 AIC와 BIC 기법에 대한 비교를 위해 4절에서 제시한 3개의 대립함수 외에 이 논문에서 제시하지 않은 또 다른 6개의 대립함수들을 사용하여 분산의 변화에 따른 검정력의 비교와 표본의 크기에 따른 검정력의 비교 등 많은 모의실험을 하였다. 그러나 거기에서 얻은 결론도 전반적으로 볼 때 BIC 기법이 AIC 기법 보다 큰 검정력을 가진다는 결론을 얻었다.

그러면 여기서 우리는 왜 BIC 기법이 AIC 기법 보다 더 좋은 검정력을 가지는가에 대한 의문을 갖지 않을 수 없다. Ledwina (1994)는 자신의 연구에서 “BIC 기법이 AIC 기법 보다 좋을 것이다”라고 제시하였지만 또 다른 통계학자들은 반대로 “AIC 기법이 BIC 기법보다 더 우수할 것이다”라고 주장한다. 이러한 상반된 두 모형선택 기법에 대한 주장들의 시비를 가리기 위해 무척 많은 모의 실험을 하였지만 우리들이 발견한 것은 거의 대부분의 모형에 있어서 BIC 기법이 AIC 기법보다 더 큰 검정력을 가진다는 것이다.

## 참고문헌

- [1] Akaike, H. (1974). A new look at statistical model identification, *I.E.E.E. Transactions on Automation Control*, Vol. 19, 716-723.
- [2] Eubank, R.L. and Hart, J.D. (1992). Testing goodness-of-fit in regression via Order Selection Criteria, *The Annals of Statistics*, Vol. 20, 1412-1425.
- [3] Eubank, R.L. and Hart, J.D. (1993). Commonality of Cusum, von Neumann and smoothing-based goodness-of-fit tests, *Biometrika*, Vol. 80, 89-98.
- [4] Eubank, R, and LaRiccia, V. N. (1992)., Asymptotic comparison of Cramer-von Mises and nonparametric function estimation techniques for testing goodness-of-fit, *Annals of Statistics*, Vol. 20, 2071 - 2086.
- [5] Hart, J.D. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests*, Springer-Verlag, New York.
- [6] Kim, J.H. (1994). Test for change in a mean function when data are dependent. Ph.D. dissertation, *Department of statistics*, Texas A&M University.
- [7] Kim, J.T. (1994). Goodness of fit test based on smoothing parameter selection criteria, *The Korean Communications in Statistics*, Vol. 2, 122 - 136.
- [8] Kim, J.T. (1998). Testing goodness-of-fit via order section criteria, *Journal of American*

*Statistical Association*, be in progress.

- [9] Ledwina, T. (1994). Data-driven version of Neyman's smooth test of fit, *Journal of American Statistical Association*, Vol. 89, 1000-1005.
- [10] Lee, G.H. (1996). A statistical wavelet approach to model selection and data driven Neyman smooth tests, *Ph.D. dissertation, Department of statistics, Texas A&M University*.
- [11] Lee, S.H., Kim, J.T., Moon, G.A. (1998), Testing Goodness-of-Fit for No Effect Models, *The Korean Communications in Statistics*, be in progress.
- [12] Rice, J. (1984). Bandwidth choice for nonparametric kernel regression, *The Annals of Statistics*, Vol. 12, 1215-30.
- [13] Schwarz, G. (1978). Estimating the dimension of model, *The Annals of Statistics*, Vol. 6, 461-464.