응용통계연구 제 18권 1호, 2005년, pp. 183-196

비동차 이차형식의 분포함수에 대한 안장점근사 *

나종화 [1] 김정숙 [2]

요 약

본 논문에서는 다변량 정규분포하에서 비동차(non-homogeneous) 이차형식의 분포함수에 대한 안장점근사법을 다루었다. 이는 Kuonen (1999)의 동차(homogeneous) 이차형식에 대한 안장점근사를 비동차의 경우로 확장한 것이다. 안장점근사의 적용을 위해 비동차 이차형식의 누율생성함수 및 관련 성질들을 유도하였다. 모의실험을 통해 안장점근사의 정도가 매우 뛰어남을 확인하였다.

주요용어: 안장점근사, 동차 이차형식, 비동차 이차형식, 누율생성함수.

1. 서론

다변량 정규분포를 따르는 확률변수들로 구성되는 이차형식(quadratic form)의 밀도함수나 분포함수와 관련된 연구는 지금까지 주로 동차 이차형식(homogeneous quadratic form)을 중심으로 진행되어 왔다. 동차 이차형식의 분포함수와 관련된 연구는 주로 동차 이차형식의 분포가 카이제곱이나 비심(non-central) 카이제곱 또는 이들의 선형결합의 분포를 따르기 위한 수학적 가정을 찾는 문제에 치우쳐 왔다.(Johnson & Kotz, 1970) 동차 이차형식의 분포함수에 대한 직접적인 근사를 다룬 논문으로는 Imhof (1961), Davies (1973) 등이 있으나, 이들은 모두 복잡한 수치적분과정을 요구하고 있어 계산과정에 많은 어려움을 포함하고 있다. 이러한 수치적분의 과정을 효과적으로 대체하기위한 근사적 방법에 대한 연구가 활발히 진행되고 있다.

최근 Kuonen (1999)은 동차 이차형식의 분포함수에 대한 안장점근사(saddlepoint approximation)를 실시하여, 밀도함수에 대한 Imhof (1961)의 근사를 부분적으로 개선한 바 있다. 본 논문에서는 Kuonen (1999)이 제시한 동차 이차형식에 대한 안장점근사를 비동차(non-homogencous) 이차형식의 경우로 확장하고자 한다. 다변량 정규분포하의 동차 또는 비동차 이차형식의 통계량은 모두 서로 상관된 non-iid 확률변수들의 함수형태를 따르며, 시계열 또는 다변량 분야에 폭 넓게 활용될 수 있다.

본 논문에서 다루게 될 비동차 이차형식은 다음과 같이 정의된다. 먼저 $X=(X_1,...,X_n)'$을 다변량 정규분포를 따르는 확률벡터라 하고, 이의 평균을 $\mu=(\mu_1,\cdots,\mu_n)'$, $n\times n$의 대

* 이 논문은 2004년도 충북대학교 학술연구지원사업의 연구비 지원에 의하여 연구되었음.
1) (361-763) 청주시 흥덕구 개신동 12, 충북대학교 정보통계학과 & 기초과학연구소 , 부교수
E-mail: cherin@cbu.ac.kr
2) (121-749) 서울시 마포구 염리동 168-9, 건강보험심사평가원, 책임연구원
E-mail: chastity@hiramail.net

칭(symmetric)인 공분산행렬을 Σ라고 하자. 비동차 이차형식(non-homogeneous quadratic form) $Q_N(X)$는 다음과 같다.

$$\begin{aligned} Q_N(X) &= X'AX + b'X + c \\ &= \sum_{1\le i\le j\le n} a_{ij}X_iX_j + \sum_{i=1}^{n} b_iX_i + c. \end{aligned} \tag{1.1}$$

여기에서 $A = (a_{ij})_{n\times n}$는 대칭행렬이고, $b = (b_1, \ldots, b_n)'$이고, c는 상수이다. 식 (1.1)에서 $b = 0$, $c = 0$ 인 경우에는 모든 항들의 X에 대한 차수가 2로 동일하므로 동차 이차형식이라 부른다. 본 논문의 2절에서는 식(1.1)의 분포함수에 대한 안장점근사를 소개하고, 이의 적용과정에 필요한 식(1.1)의 누율생성함수와 관련 성질들을 유도한다. 3절에서는 모의실험결과를 소개하고, 4절은 결론으로 구성되었다.

2. 비동차 이차형식에 대한 안장점근사와 누율생성함수

2.1. 분포함수에 대한 안장점근사

안장점근사(saddlepoint approximation)는 밀도함수나 분포함수에 대한 뛰어난 근사를 제공하는 방법으로 잘 알려져 있다. (Daniels (1954, 1987), Reid (1988), Barndorff-Nielsen & Cox (1979, 1989), McCullagh (1987) 참고.) 이는 통계량의 누율생성함수의 퓨리에 역변환식(Fourier inversion formula)에 대한 Laplace근사를 통해 유도되며, 기존의 정규근사 또는 Edgeworth근사에 비해 많은 장점을 가지고 있다. 특히 분포의 꼬리부분에서도 근사의 정도가 뛰어나며, 대표본은 물론 소표본의 경우에도 기존의 근사방법을 획기적으로 개선시킬 수 있음이 여러 연구에서 밝혀져 있다.

본 논문에서는 지금까지 안장점근사와 관련된 여러 연구 결과들 중 분포함수에 대한 Lugannani & Rice (1980)의 결과를 재 표현한 Daniels (1987)의 근사식을 소개하고, 비동차 이차형식 통계량에 대한 적용과정을 다루기로 한다. 먼저 식 (1.1)에서 소개된 비동차 이차형식 통계량 $Q_N(X)$의 누율생성함수를 $K(t)$라 할 때, Daniels (1987)에 의한 분포함수에 대한 안장점근사식은 다음과 같다.(이 근사식은 흔히 Lugannani와 Rice의 공식으로 불려진다.)

$$Pr\{Q_N(X) \le q\} \simeq \begin{cases} \Phi(\omega) + \phi(\omega)\{\frac{1}{\omega} - \frac{1}{\zeta}\}, & q \ne E(Q_N(X)) \\ \frac{1}{2} - \frac{K^{(3)}(0)}{6\sqrt{2\pi(K''(0))}}, & q = E(Q_N(X)) \end{cases} \tag{2.1}$$

위식에서 ω와 ζ는 다음과 같이 정의된다.

$$\omega = sgn(t_0)[2\{t_0q - K(t_0)\}]^{1/2},\ \zeta = t_0\{K''(t_0)\}^{1/2}.$$

여기서 안장점(saddlepoint)으로 불리는 t_0는 다음의 안장점방정식(saddlepoint equation)에 대한 수치해를 통해 구해진다.

$$K'(t_0) = q$$

또한 $\Phi(\cdot)$와 $\phi(\cdot)$는 표준정규분포의 분포함수와 밀도함수를 나타낸다.

다음으로 식 (2.1)과 함께 본 논문의 모의실험에서 고려되는 Jensen (1992, 1995)의 안장점근사식은 다음과 같다.

$$Pr\{Q_N(X) \le q\} \simeq \Phi\left\{\omega + \frac{1}{\omega}\log\left(\frac{\zeta}{\omega}\right)\right\}. \tag{2.2}$$

위 근사식의 정도(precision)는 iid 표본평균에 대해 근사식 (2.1)과 동일한 것으로 밝혀져 있으며(Jensen, 1992), 식 (2.1)의 근사식은 분포함수의 근사값이 음의 값을 취하는 경우가 발생할 수 있으나, 식 (2.2)의 근사식은 이러한 단점을 보완한 결과로써 항상 양의 값으로 주어지는 장점을 가진다.

2.2. 비동차 이차형식에 대한 누율생성함수와 그의 성질

이 절에서는 앞서 소개한 분포함수에 대한 안장점근사를 비동차 이차형식 통계량에 적용하기 위해 요구되는 통계량의 누율생성함수를 비롯한 관련 성질들을 유도하고자 한다. 먼저 1절에서 정의된 다변량 정규분포를 따르는 확률벡터 X는 다음과 같이 표현된다.

$$X = \mu + \Gamma Z. \tag{2.3}$$

위 식에서 Z는 평균이 0이고 공분산행렬이 단위행렬 I인 확률벡터이며, Γ는 Σ의 콜레스키분해(cholesky decomposition)를 통해 구해지는 하삼각행렬로써 $\Sigma = \Gamma\Gamma'$를 만족한다. 식 (2.3)을 (1.1)에 대입하면 $Q_N(X)$는 다음과 같이 표현된다.

$$Q_N(X) = Z'BZ + v'Z + u. \tag{2.4}$$

여기서, B, v, u는 각각 다음과 같이 정의된다.

$$\begin{aligned} B &= \Gamma' A \Gamma && (n \times n \text{ 대칭행렬}) \\ v &= \Gamma'(b + 2A\mu) && (n \times 1 \text{ 벡터}) \\ u &= \mu' A \mu + b'\mu + c && (1 \times 1\text{인 스칼라}) \end{aligned}$$

정리 2.1 비동차 이차형식 $Q_N(X)$의 적률생성함수 $M_{Q_N}(t)$는 다음과 같다.

$$M_{Q_N}(t) = |I - 2tB|^{-\frac{1}{2}} \exp(tu) \exp\left\{\frac{1}{2} tv'(I - 2tB)^{-1} tv\right\}.$$

증명:

$$
\begin{aligned}
M_{Q_N}(t) &= E(e^{tQ_N}) \\
&= E[\exp(tZ'BZ + tv'Z + tu)] \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} |2\pi I|^{(-1/2)} \exp\Big(-\frac{1}{2}z'z + tz'Bz + tv'z + tu\Big)dz \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} |2\pi(I-2tB)^{-1}|^{-\frac{1}{2}} \exp\Big\{-\frac{1}{2}z'(I-2tB)z\Big\} \\
&\quad \exp(tv'z)dz \times |I-2tB|^{-\frac{1}{2}} \exp(tu) \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \{\text{density of } \ N(0,(I-2tB)^{-1})\} \exp(tv'z)dz \\
&\quad \times |I-2tB|^{-\frac{1}{2}} \exp(tu) \\
&= \Big\{\text{mgf of } N(0,(I-2tB)^{-1}) \text{ evaluated at } tv'\Big\} \times |I-2tB|^{-\frac{1}{2}} \exp(tu) \\
&= |I-2tB|^{-\frac{1}{2}} \exp(tu) \exp\Big\{\frac{1}{2}tv'(I-2tB)^{-1}tv\Big\}.
\end{aligned}
$$

□

보조정리 2.1 A가 $n \times n$행렬이고 A의 고유치를 $\lambda_i (i = 1, \ldots, n)$라고 하고, $D = diag(\lambda_i)$라 하자. 정수 p와 정칙행렬(non-singular matrix) Q에 대해 다음이 성립한다.

$$(a)\ A = QDQ^{-1} \qquad (b)\ \ A^p = QD^pQ^{-1}$$

$$(c)\ tr(A^p) = tr(QD^pQ^{-1}) = tr(D^pQ^{-1}Q) = tr(D^p) = \sum_{i=1}^{n} \lambda_i^p$$

보조정리 2.2 A가 $n \times n$행렬이고 $A = QDQ^{-1}$을 만족한다. 만약 $f(A)$를 $f(A) = Qf(D)Q^{-1}$으로 정의할 때 다음이 성립한다.

$$(a)\ (I - A)^{-1} = \sum_{k=0}^{\infty} A_k \qquad (b)\ -\log(I - A) = \sum_{k=1}^{\infty} \frac{1}{k} A^k$$

증명: (a) 보조정리2.1로부터 $I - A = Q(I - D)Q^{-1}$가 성립한다. 따라서,

$$
\begin{aligned}
(I - A)^{-1} &= Q(I - D)^{-1}Q^{-1} \\
&= Q\ diag\{(1 - \lambda_i)^{-1}\}Q^{-1} \\
&= Q\ diag\{1 + \lambda_i + \lambda_i^2 + \cdots\}Q^{-1} \\
&= Q(I + D + D^2 + \cdots)Q^{-1} \\
&= I + A + A^2 + \cdots \ \text{ if } \ |\lambda_i| < 1 \text{ for } \ i = 1, \cdots, n.
\end{aligned}
$$

(b) (a)와 동일한 방법을 적용하여 쉽게 증명할 수 있다. □

정리2.1로부터 $Q_N(X)$의 누율생성함수 $K_{Q_N}(t)$는 다음과 같다.

$$K_{Q_N}(t) = -\frac{1}{2}\log|I-2tB| + tu + \frac{1}{2}t^2 v'(I-2tB)^{-1}v. \tag{2.5}$$

정리2.1에 제시된 비동차 이차형식에 대한 누율생성함수는 몇가지 유용한 다른 형태로 표현될 수 있다. 먼저 식 (2.5)에서 B는 대칭행렬로써 다음과 같이 표현될 수 있다.

$$B = P\Lambda P'.$$

여기서 $\Lambda = diag(\lambda_i)$이며, λ_i는 B의 고유치(eigenvalues)이고, P는 $P'P = I$를 만족하는 고유벡터들(eigenvectors)로 구성된 직교행렬이다. 식 (2.5)의 첫째항은 보조정리2.2에 의해 다음과 같이 표현된다.

$$\begin{aligned}
-\frac{1}{2}\log|I-2tB| &= -\frac{1}{2}\log|P(I-2t\Lambda)P'| = -\frac{1}{2}\log|I-2t\Lambda| \\
&= -\frac{1}{2}\log\left[\prod_{i=1}^{n}(1-2t\lambda_i)\right] = -\frac{1}{2}\sum_{i=1}^{n}\log(1-2t\lambda_i) \\
&= \frac{1}{2}\sum_{i=1}^{n}\sum_{r=1}^{\infty}\frac{(2t\lambda_i)^r}{r} = \frac{1}{2}\sum_{r=1}^{\infty}\frac{t^r}{r}\cdot 2^r\cdot(\sum_{i=1}^{n}\lambda_i^r) \\
&= \frac{1}{2}\sum_{r=1}^{\infty}\frac{t^r}{r}\cdot 2^r\cdot tr(\Lambda^r) = \frac{1}{2}\sum_{r=1}^{\infty}\frac{t^r}{r}\cdot 2^r\cdot tr(P\Lambda^r P') \\
&= \frac{1}{2}\sum_{r=1}^{\infty}\frac{t^r}{r}\cdot 2^r\cdot tr(B^r) = \frac{1}{2}\sum_{r=1}^{\infty}tr\left[\frac{(2tB)^r}{r}\right] \\
&= \frac{1}{2}tr\left[\sum_{r=1}^{\infty}\frac{(2tB)^r}{r}\right] = \frac{1}{2}tr[-\log(I-2tB)].
\end{aligned}$$

위의 결과로부터 다음의 정리2.2를 얻을 수 있다.

정리 2.2 비동차 이차형식 $Q_N(X)$의 누율생성함수 $K_{Q_N}(t)$는 다음과 같다.

$$\begin{aligned}
K_{Q_N}(t) &= \frac{1}{2}tr\left[\sum_{r=1}^{\infty}\frac{(2tB)^r}{r}\right] + \frac{1}{2}t^2 v'(I-2tB)^{-1}v + tu \\
&= \frac{1}{2}tr\Big[-\log(I-2tB)\Big] + \frac{1}{2}t^2 v'(I-2tB)^{-1}v + tu.
\end{aligned} \tag{2.6}$$

정리 2.3 정리2.2의 누율생성함수는 다음과 같이 표현될 수 도 있다.

$$K_{Q_N}(t) = t\{tr(B)+u\} + \sum_{r=2}^{\infty}\frac{t^r}{r!}\left\{tr(B^r) + \frac{r}{4}v'B^{r-2}v\right\}\cdot 2^{r-1}(r-1)!. \tag{2.7}$$

증명: 식 (2.6)은 다시 보조정리2.2의 (a)에 의해 다음과 같이 표현된다.

$$\begin{aligned} K_{Q_N}(t) &= \sum_{r=1}^{\infty}\frac{t^r}{r}\cdot 2^{r-1}tr(B^r)+tu+\frac{1}{2}t^2v'\sum_{r=0}^{\infty}(2tB)^r v \\ &= \sum_{r=1}^{\infty}\frac{t^r}{r}\cdot 2^{r-1}tr(B^r)+tu+\sum_{r=0}^{\infty}2^{r-1}t^{r+2}v'B^r v \\ &= \sum_{r=1}^{\infty}\frac{t^r}{r}\cdot 2^{r-1}tr(B^r)+tu+\sum_{r=2}^{\infty}2^{r-3}t^r v'B^{r-2}v \\ &= t\{tr(B)+u\}+\sum_{r=2}^{\infty}\frac{t^r}{r}2^{r-1}\left\{tr(B^r)+\frac{r}{4}v'B^{r-2}v\right\} \end{aligned}$$

□

또한 식 (2.7)의 누율생성함수는 다음과 같이 생각될 수 있다.

$$K_{Q_N}(t)=t\cdot\{Q_N\text{의 1차 누율}\}+\sum_{r=2}^{\infty}\frac{t^r}{r!}\{Q_N\text{의 r차}(r\geq 2)\text{ 누율}\}.$$

이제 지금까지 유도된 비동차 이차형식 $Q_N(X)$의 누율생성함수를 바탕으로 안장점근사식 (2.1)과 (2.2)의 적용과정에 요구되는 누율생성함수에 대한 1차와 2차 미분식의 유도과정을 살펴보자.

보조정리 2.3 A가 $n\times n$행렬이고 t의 함수로써 $A=A(t)$와 같다고 하면 다음의 식을 만족한다. 여기서 $\dot{A}=dA(t)/dt$을 의미한다.

$$tr\left(\frac{d}{dt}A^p\right)=p\ tr[A^{p-1}\dot{A}].$$

증명: 연쇄규칙(chain rule)으로부터

$$\frac{d}{dt}A^p=\frac{d}{dt}(A\cdots A)=\dot{A}A^{p-1}+A\dot{A}A^{p-2}+\cdots+A^{p-1}\dot{A}$$

이고, 따라서 다음의 식이 성립된다.

$$\begin{aligned} tr\left(\frac{d}{dt}A^p\right) &= tr(\dot{A}A^{p-1})+tr(A\dot{A}A^{p-2})+\cdots+tr(A^{p-1}\dot{A}) \\ &= tr(A^{p-1}\dot{A})+tr(A^{p-2}A\dot{A})+\cdots+tr(A^{p-1}\dot{A}) \\ &= p\cdot tr(A^{p-1}\dot{A}). \end{aligned}$$

□

보조정리 2.4 $A(t)$의 역행렬을 $B=B(t)$라고 놓으면 다음의 식이 성립한다.

$$\dot{B}=-B\dot{A}A^{-1}=-B\dot{A}B.$$

여기서 $\dot{A}=dA(t)/dt$ 이고 $\dot{B}=dB(t)/dt$이다.

증명: $B(t) = A(t)^{-1}$로부터 $B(t)A(t) = I$가 성립한다. 양변을 t에 대해 미분하면 $\dot{B}(t)A(t) + B(t)\dot{A}(t) = 0$이 된다. 따라서 $\dot{B}(t)$는 다음 식을 만족한다.

$$\dot{B}(t) = -B(t)\dot{A}(t)A(t)^{-1} = -B(t)\dot{A}(t)B(t).$$

□

예를 들어,

$$\begin{aligned}\frac{d}{dt}(I-2tB)^{-1} &= -(I-2tB)^{-1}(-2B)(I-2tB)^{-1} \\ &= 2(I-2tB)^{-1}B(I-2tB)^{-1}\end{aligned}$$

이다.

보조정리2.3과 2.4로부터 누율생성함수의 1, 2차 미분에 대한 다음의 정리2.4와 2.5의 결과를 얻을 수 있다.

정리 2.4 비동차 이차형식 $Q_N(X)$의 누율생성함수의 1차 미분식은 다음과 같다.

$$\begin{aligned}\frac{d}{dt}K_{Q_N}(t) &= tr\{(I-2tB)^{-1}B\} \\ &\quad + tv'(I-2tB)^{-1}v + t^2v'(I-2tB)^{-1}B(I-2tB)^{-1}v + u.\end{aligned}$$

증명: 먼저 식 (2.6)의 첫째항의 미분식은 다음과 같다.

$$\begin{aligned}&\frac{d}{dt}[tr\{-\log(I-2tB)\}] \\ &= tr\left[\frac{d}{dt}\{-\log(I-2tB)\}\right] \\ &= tr\left[\frac{d}{dt}C + \frac{1}{2}\frac{d}{dt}(C^2) + \frac{1}{3}\frac{d}{dt}(C^3) + \cdots\right] \qquad \text{보조정리 2.2(b)에 의해} \\ &= tr\left(\frac{d}{dt}C\right) + \frac{1}{2}tr\left(\frac{d}{dt}C^2\right) + \frac{1}{3}tr\left(\frac{d}{dt}C^3\right) + \cdots \\ &= tr(\dot{C}) + tr(C\dot{C}) + tr(C^2\dot{C}) + \cdots \qquad \text{보조정리 2.3에 의해} \\ &= tr(\dot{C} + C\dot{C} + C^2\dot{C} + \cdots) \\ &= tr\{(I-C)^{-1}\dot{C}\} \qquad \text{보조정리 2.2(a)에 의해} \\ &= 2tr\{(I-2tB)^{-1}B\}.\end{aligned}$$

여기서 $C = 2tB$이고 $\dot{C} = dC(t)/dt = 2B$이다. 다음으로 식 (2.6)의 두 번째 항의 미분식은 다음과 같다.

$$\frac{d}{dt}\left[\frac{1}{2}t^2 v'(I-2tB)^{-1}v\right]$$
$$= tv'(I-2tB)^{-1}v + \left(\frac{1}{2}t^2\right)v'\left[\frac{d}{dt}(I-2tB)^{-1}\right]v$$
$$= tv'(I-2tB)^{-1}v + \left(\frac{1}{2}t^2\right)v'\left[2(I-2tB)^{-1}B(I-2tB)^{-1}\right]v$$

보조정리 2.4에 의해

$$= tv'(I-2tB)^{-1}v + t^2v'(I-2tB)^{-1}B(I-2tB)^{-1}v.$$

□

정리 2.5 비동차 이차형식 $Q_N(X)$의 누율생성함수의 2차 미분식은 다음과 같다.

$$\begin{aligned}\frac{d^2}{dt^2}K_{Q_N}(t) &= 2\,tr\{B(I-2tB)^{-1}B(I-2tB)^{-1}\} + v'(I-2tB)^{-1}v \\ &\quad +4tv'(I-2tB)^{-1}B(I-2tB)^{-1}v \\ &\quad +4t^2v'(I-2tB)^{-1}B(I-2tB)^{-1}B(I-2tB)^{-1}v.\end{aligned}$$

증명: 정리2.4로부터 2차 미분식은 다음과 같다.

$$\frac{d^2}{dt^2}K_{Q_N}(t)$$
$$= tr\left\{\frac{d}{dt}(I-2tB)^{-1}B\right\} + v'(I-2tB)^{-1}v + tv'\left\{\frac{d}{dt}(I-2tB)^{-1}\right\}\nu$$
$$+2tv'(I-2tB)^{-1}B(I-2tB)^{-1}v + t^2v'\left\{\frac{d}{dt}(I-2tB)^{-1}\right\}B(I-2tB)^{-1}v$$
$$+t^2v'(I-2tB)^{-1}B\left\{\frac{d}{dt}(I-2tB)^{-1}\right\}v.$$

위 식에 보조정리2.4를 적용하면 쉽게 결과를 얻을 수 있다. □

3. 모의실험

이 절에서는 모의실험을 통한 비동차 이차형식 $Q_N(X) = X'AX + b'X + c$의 분포함수에 대한 안장점근사의 결과를 제시하였다. 모의실험에서는 먼저 $X = (X_1, X_2)'$의 평균과 분산이 아래와 같이 주어지는 이변량 정규분포에 대해

$$(a)\ \mu' = (1,1),\quad \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \qquad (b)\ \mu' = (0,0),\quad \Sigma = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

다음으로부터 정의되는 비동차 이차형식에 대한 모의실험을 수행하였다.

$$A = \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix},\quad b = \begin{pmatrix} 5 \\ 4 \end{pmatrix},\quad c = 3,$$

또한 $X = (X_1, X_2, X_3)'$의 분포가 평균과 분산이

$$(c)\ \mu' = (2, 5, 10), \quad \Sigma = \begin{pmatrix} 1 & 4.5 & 9 \\ 4.5 & 25 & 49 \\ 9 & 49 & 100 \end{pmatrix}$$

인 3변량 정규분포에서 다음으로부터 정의되는 비동차 이차형식에 대한 모의실험을 수행하였다.

$$A = \begin{pmatrix} 1 & 3 & 0 \\ 3 & -1 & -2 \\ 0 & -2 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 4 \\ -2 \end{pmatrix}, \quad c = 3$$

아래의 표4.1 $\sim$ 표4.3은 각각 (a)~(c)에 대한 모의실험 결과이다. 각 표에서 제시되어 있는 S.Exact(Simulated Exact)는 반복횟수 10만 번의 난수에 기초하여 모의실험을 수행한 결과이며, Normal은 근사식 (2.1)의 첫째항만을 이용한 정규근사이며, Saddle.LR과 Saddle.JN은 각각 식 (2.1)과 (2.2)의 안장점근사의 결과이다. 이 결과들을 살펴보면 안장점근사의 결과가 대단히 뛰어난 것을 확인할 수 있으며, 특히 통계적 추론의 관심 영역인 꼬리부분에서도 그 정확도가 유지되는 것을 확인할 수 있다.

4. 결론

본 논문에서는 Kuonen (1999)의 동차 이차형식의 분포함수에 대한 안장점근사를 비동차의 경우로 확장하였다. 이를 위해 비동차 이차형식의 누율생성함수를 유도하고, 이로부터 안장점근사의 적용과정에 요구되는 여러 성질들을 구체적으로 유도하였다. 이는 뛰어난 정확도에도 불구하고 누율생성함수와 관련된 이론적 제약으로 인하여, 특정한 통계량의 형태에만 제한적으로 적용되어온 안장점근사의 범위를 좀 더 넓은 영역의 일반적 통계량으로 확장한 것이다. 모의실험의 결과를 통해 안장점근사의 결과가 대단히 뛰어남을 알 수 있으며, 기존의 정규근사와는 달리 통계적 추론의 주된 관심 영역인 꼬리부분에서도 정확한 근사가 가능하여, non-iid 경우를 비롯한 많은 통계적 응용문제로의 활용이 기대된다.

참고문헌

Barndorff-Nielsen, O. E. and Cox, D. R. (1979). Edgeworth and saddlepoint approximations with statistical applications (with discussion), *Journal of Royal Statistical Society B*, 41, 279-312.

Barndorff-Nielsen, O. E. and Cox, D. R. (1989). *Asymptotic Techniques for Use in Statistics*, Chapman and Hall, London.

Daniels, H. E. (1954). Saddlepoint approximations in statistics, *The Annals of Mathematical Statistics*, **25(4)**, 631-650.

Daniels, H.E. (1987). Tail probability approximations, *International Statistical Review*, **55(1)**, 37-48.

Davies, R. B. (1973). Numerical inversion of a characteristic function, *Biometrika*, **60**, 415-417.

Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables, *Biometrika*, **48**, 419-426

Jensen, J. L. (1992). The modified signed loglikelihood statistic and saddlepoint approximations, *Biometrika*, **79**, 693-703.

Jensen, J. L. (1995). *Saddlepoint Approximations*, Oxford Statistical Science Series 16, Oxford, Clarendon Press.

Johnson, N. L. and Kotz, S. (1970). *Distributions in Statistics: Continuous univariate distributions-2*, New York, Wiley.

Kuonen, D. (1999). Saddlepoint approximations for distributions of quadratic forms in normal variables, *Biometrika*, **86(4)**, 929-935.

Lugannani, R. and Rice, S. O. (1980). Saddlepoint approximation for the distribution of the sum of independent random variables, *Advances in Applied Probability*, **12**, 475-490.

McCullagh, P. (1987). *Tensor Methods in Statistics*, Chapman & Hall, London.

Reid, N. (1988). Saddlepoint expansions and statistical inference(with discussion.), *Statistical Science*, **3**, 213-238.

[2004년 8월 접수, 2004년 12월 채택]

표 4.1: 비동차 이차형식의 분포함수에 대한 근사: (a)의 경우

$Pr\{Q_N(X) \le q\}$에 대한 근사				
q	S.Exact	Normal	Saddle.LR	Saddle.JN
-10.0	0.0384	0.0694	0.0425	0.0428
-6.5	0.0587	0.0959	0.0652	0.0654
-3.0	0.0957	0.1313	0.1005	0.1007
4.0	0.2612	0.2346	0.2289	0.2289
11.0	0.4365	0.3737	0.4126	0.4125
18.0	0.5893	0.5203	0.5800	0.5798
25.0	0.7133	0.6507	0.7099	0.7096
32.0	0.8084	0.7555	0.8052	0.8050
39.0	0.8741	0.8342	0.8725	0.8723
46.0	0.9187	0.8906	0.9183	0.9182
53.0	0.9505	0.9294	0.9487	0.9486
60.0	0.9682	0.9553	0.9683	0.9683
67.0	0.9807	0.9722	0.9807	0.9807
70.5	0.9847	0.9782	0.9850	0.9850
74.0	0.9885	0.9830	0.9884	0.9884
77.5	0.9912	0.9867	0.9911	0.9910
81.0	0.9931	0.9897	0.9931	0.9931
84.5	0.9946	0.9920	0.9947	0.9947
88.0	0.9957	0.9938	0.9959	0.9959

표 4.2: 비동차 이차형식의 분포함수에 대한 근사: (b)의 경우

$Pr\{Q_N(X) \leq q\}$에 대한 근사				
q	S.Exact	Normal	Saddle.LR	Saddle.JN
-8.0	0.0522	0.0918	0.0604	0.0606
-4.5	0.1033	0.1504	0.1259	0.1260
-1.0	0.2264	0.2346	0.2632	0.2632
6.0	0.5907	0.4280	0.5747	0.5715
13.0	0.7357	0.5826	0.7276	0.7241
20.0	0.8227	0.6960	0.8160	0.8133
27.0	0.8814	0.7788	0.8732	0.8713
34.0	0.9144	0.8393	0.9117	0.9104
41.0	0.9426	0.8833	0.9382	0.9373
48.0	0.9583	0.9155	0.9566	0.9559
55.0	0.9707	0.9388	0.9694	0.9690
62.0	0.9794	0.9558	0.9784	0.9781
69.0	0.9855	0.9681	0.9848	0.9845
76.0	0.9905	0.9770	0.9892	0.9891
79.5	0.9916	0.9804	0.9909	0.9908
83.0	0.9927	0.9834	0.9924	0.9923
86.5	0.9940	0.9859	0.9936	0.9935
90.0	0.9951	0.9880	0.9946	0.9945
93.5	0.9959	0.9899	0.9955	0.9954

표 4.3: 비동차 이차형식의 분포함수에 대한 근사: (c)의 경우

$Pr\{Q_N(X) \leq q\}$에 대한 근사				
q	S.Exact	Normal	Saddle.LR	Saddle.JN
-40	0.0357	0.0719	0.0425	0.0428
-20	0.0652	0.1186	0.0839	0.0841
0	0.0339	0.1904	0.1699	0.1700
20	0.3078	0.2882	0.3181	0.3181
40	0.4975	0.3987	0.4770	0.4765
60	0.6228	0.5050	0.6011	0.6001
80	0.7147	0.5992	0.6941	0.6931
100	0.7834	0.6793	0.7652	0.7643
120	0.8392	0.7457	0.8201	0.8193
140	0.8706	0.7999	0.8625	0.8619
160	0.9080	0.8436	0.8952	0.8947
180	0.9311	0.8784	0.9203	0.9200
200	0.9481	0.9060	0.9396	0.9393
220	0.9630	0.9276	0.9544	0.9542
240	0.9712	0.9445	0.9656	0.9654
260	0.9772	0.9576	0.9741	0.9740
280	0.9836	0.9677	0.9806	0.9805
300	0.9869	0.9755	0.9855	0.9854
320	0.9915	0.9814	0.9891	0.9891
340	0.9940	0.9860	0.9919	0.9919
360	0.9958	0.9894	0.9940	0.9939
380	0.9959	0.9920	0.9955	0.9955
400	0.9964	0.9940	0.9967	0.9966
500	0.9995	0.9986	0.9992	0.9992

Saddlepoint Approximations to the Distribution Function of Non-homogeneous Quadratic Forms *

Jong-Hwa Na [1)] Jeong-Sook Kim [2)]

Abstract

In this paper we studied the saddlepoint approximations to the distribution of non-homogeneous quadratic forms in normal variables. The results are the extension of Kuonen's which provide the same approximations to homogeneous quadratic forms. The CGF of interested statistics and related properties are derived for applications of saddlepoint techniques. Simulation results are also provided to show the accuracy of saddlepoint approximations.

Keywords: Saddlepoint approximation, Homogeneous quadratic form, Non-homogeneous quadratic form, Cumulant generating function

* This work was supported by Chungbuk National University Grant in 2004.

1) Associate Professor, Dept. of Information and Statistics & Institute for Basic Science Research, Chungbuk National University, 12 Gaeshin-dong, Heungduk-gu, Cheongju, Chungbuk, 361-763, Korea.
E-mail: cherin@cbu.ac.kr

2) Principal Researcher, Information & Communication Dept., Health Insurance B/D 168-9, Yeomri-Dong, Mapo-ku, Seoul, 121-749, Korea.
E-mail: chastity@hiramail.net