응용통계연구 (2009)
22(1), 171–179

곱 정규확률변수의 합에 대한 소표본 점근분포와 FSK 통신에의 응용

나종화[1] · 김정미[2]

[1]충북대학교 정보통계학과, [2]충북대학교 정보통계학과

(2008년 10월 접수, 2008년 10월 채택)

요 약

본 논문에서는 정규확률변수의 곱과 그의 합으로 표현되는 통계량의 분포에 대한 효과적인 근사법을 다루었다. 이차형식에 대한 안장점근사에 기초한 이 방법은 기존의 정규근사에 비해 매우 정확한 결과를 제공한다. 또한 이에 대한 응용으로 FSK 통신에서 발생하는 문제를 제시하고, 그 해결책으로 본 논문에서 제안한 안장점근사법을 사용하였다. 모의실험을 통해 제안된 근사법이 중심영역은 물론, 통신이론에서 주요 관심 영역인, 극단 꼬리부분의 확률 근사에도 매우 유용한 방법임을 확인하였다.

주요용어: 정규확률변수의 곱, 안장점근사, 이차형식, FSK 통신, 점근분포.

1. 서론

정규확률변수의 곱의 분포는 통신이론과 관련된 응용분야에서 중요하게 취급된다. 특히 시차 t를 두고 전달되는 신호의 인식과정에서 두 확률변수의 곱은 신호간의 상관관계를 표현하는 도구로 사용되며, 이들 곱에 대한 합의 분포로부터 신호에 대한 동기를 찾는데 활용하고 있다. 서로 독립이거나 상관된 두 정규확률변수의 곱에 대한 연구는 정규분포와 관련된 다른 연구에 비해 그리 활발하게 이루어지지 않았으나, 이와 관련된 고전적인 연구로 이미 1936년 Craig가 정규확률변수의 곱에 대한 적률생성함수(MGF)를 유도한 바 있다. Aroian 등 (1978)은 이들의 분포가 적절한 조건하에서 피어슨의 TypeⅢ 분포로 수렴함을 보인 바 있다. Conwell 등 (1978)은 이들 분포의 수치적 계산과정에 대해 다룬 바 있으며, 관련된 연구로 Conradie와 Gupta (1987)는 다변량 정규분포의 이차형식의 분포에 대해 연구한 바 있다.

일반적으로 정규확률변수의 곱에 대한 분포는 두 확률변수의 평균과 분산의 비(μ/σ)와 상관계수(ρ)의 영향을 크게 받으며, 이들 값에 따라 정규분포의 형태를 크게 벗어나기도 한다. 본 논문에서는 정규확률변수의 곱에 대한 분포함수의 추정문제를 다루었다. 기존의 분포이론을 통한 분포함수의 추정은 대부분 복잡한 수치적분의 과정을 포함할 뿐 아니라, 이들의 합의 분포에 대해서는 다시 합성곱(convolution)공식을 반복적으로 적용하여 구할 수는 있으나, 계산과정이 매우 복잡하여 실제의 응용문제의 해결에 효과적인 방법이라 말할 수 없다. 본 논문에서는 이차형식에 대한 안장점근사와 관련된 Na와 Kim (2005)의 결과를 이용하여, 이들 분포함수에 대한 정확하고도 편리한 근사과정을 보이고자 한다. 본

이 논문은 2008년도 충북대학교 학술연구지원사업의 연구비지원에 의하여 연구되었음.
[1]교신저자: (361-763) 청주시 흥덕구 개신동 12, 충북대학교 정보통계학과 & 기초과학연구소, 교수.
E-mail: cherin@chungbuk.ac.kr

논문의 2절에서는 정규확률변수의 곱과 관련된 몇 가지 기존의 결과들과 안장점근사의 과정에 필요한 성질들을 유도하고, 3절에서는 본 논문에서 다루게 될 이차형식의 분포함수에 대한 안장점근사와 FSK 통신에서의 응용을 다루기로 한다. 4절은 실제의 예를 통해 분포함수에 대한 안장점근사의 정확성을 확인하고, 5절은 결론으로 구성되었다.

2. 정규확률변수의 곱의 합에 대한 분포와 관련성질

2.1. 곱의 합에 대한 분포

서로 상관된 정규확률변수의 곱에 대한 분포에 대해 생각하자. 즉, 확률변수 $X' = (X_1, X_2)$의 분포가

$$(X_1, X_2)' \sim N_2(\mu, \Sigma), \quad \mu = (\mu_1, \mu_2)',\ \Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

인 경우 두 확률변수의 곱 $Y = X_1X_2$의 밀도함수는

$$Y|X_2 = x_2 \sim N\left(x_2\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2^2 - \mu_2 x_2),\ x_2^2\sigma_1^2(1-\rho^2)\right), \quad X_2 \sim N(\mu_2, \sigma_2^2)$$

으로부터 다음의 식으로 나타낼 수 있다.

$$\begin{aligned} f(y) &= \int_{-\infty}^{\infty} f(y|x_2)f(x_2)dx_2 \\ &= \int_{-\infty}^{\infty} f(y, x_2)dx_2. \end{aligned}$$

위의 밀도함수는 X_1과 X_2의 분포 모수에 따라 다양한 형태로 주어지며, 가장 단순한 예로 X_1, X_2가 *iid* $N(0,1)$인 경우, $Y = X_1X_2$의 분포는 $\chi^2(1)$이 된다. 이 결과로부터 $Y = X_1X_2$의 분포함수는 다음 식을 수치 적분하여 구할 수 있다.

$$F(y) = \int_{-\infty}^{y} f(t)dt.$$

또한, Craig (1936)는 $Z = X_1X_2/(\sigma_1\sigma_2)$의 적률생성함수(MGF)를 유도하였으며, Aroian (1947)은 Z의 분포가, $\delta_1 = \mu_1/\sigma_1$또는 $\delta_2 = \mu_2/\sigma_2$가 ∞로 수렴할 때 정규분포로 근사함을 보였다. 또한 적절한 조건하에서, Z의 분포가 표준화된(standardized) 피어슨 TypeⅢ 분포로 수렴함을 보인 바 있다. 특히 Craig (1936)는 Z의 MGF를 다음의 식

$$M_Z(t) = \exp\left[\frac{2\delta_1\delta_2 t + (\delta_1^2 + \delta_2^2 - 2\rho\delta_1\delta_2)t^2}{2\{1-(1+\rho)t\}\{1+(1-\rho)t\}}\right] \Big/ \sqrt{\{1-(1+\rho)t\}\{1+(1-\rho)t\}}$$

으로 표현하였으며, 이를 $Y = X_1X_2$의 MGF로 표현하면 다음과 같다.

$$\begin{aligned} M_Y(t) &= M_Z(\sigma_1\sigma_2 t) \\ &= \exp\left\{\frac{\mu_1\mu_2 t + \left(\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 - 2\rho\mu_1\mu_2\sigma_1\sigma_2\right)t^2/2}{(1-\rho\sigma_1\sigma_2 t)^2 - \sigma_1^2\sigma_2^2 t^2}\right\} \Big/ \sqrt{(1-\rho\sigma_1\sigma_2 t)^2 - \sigma_1^2\sigma_2^2 t^2}. \end{aligned} \tag{2.1}$$

예를 들어, X_1, X_2가 *iid* $N(\mu, \sigma^2)$인 경우, Y의 MGF는 다음과 같이 주어진다.

$$M_Y(t) = \exp\left(\frac{\mu^2 t + \mu^2\sigma^2 t^2}{1-\sigma^4 t^2}\right) \Big/ \sqrt{1-\sigma^4 t^2}.$$

이제 상관된 두 정규확률변수의 곱으로 표현되는 서로 독립이 아닌 확률변수들을 $Y_1, Y_2, \ldots, Y_n$이라 하고, 이들의 합에 대한 분포를 생각하자. 일반적으로 두 확률변수의 합의 밀도함수를 구하기 위해서는 잘 알려진 합성곱공식을 이용할 수 있다. 즉, $Z = Y_1 + Y_2$의 밀도함수는

$$f(z) = \sum_{y_2} f(z - y_2,\ y_2)$$

으로부터 구해지며, 이 과정을 반복적으로 적용하여 $\sum_{i=1}^{n} Y_i \, (n \geq 2)$의 밀도함수를 구할 수 있다. 이렇게 구해진 밀도함수에 대해 다시 수치적분을 수행하여 분포함수 값을 구할 수 있다. 그러나 이 과정은 n이 커짐에 따라 매우 복잡한 계산절차를 포함하고 있어 실제의 응용문제에 효과적으로 활용될 수 없다.

2.2. 관련성질

이 절에서는 서로 상관된 정규확률변수의 곱의 분포와 이들의 합의 분포와 관련된 몇 가지 성질들을 알아본다. 먼저 곱에 대한 적률생성함수 식 (2.1)로부터 $Y = X_1X_2$의 평균과 분산 그리고 왜도를 계산하면 다음과 같다.

$$\begin{aligned}
E(Y) &= \mu_1\mu_2 + \rho\sigma_1\sigma_2. \\
V(Y) &= \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2 + 2\rho\mu_1\mu_2\sigma_1\sigma_2 + \rho^2\sigma_1^2\sigma_2^2. \\
\alpha_3(Y) &= \frac{6\sigma_1\sigma_2\left\{\mu_1\mu_2\sigma_1\sigma_2\left(\rho^2+1\right) + \rho\left(\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2\right)\right\} + 2\rho\sigma_1^3\sigma_2^3\left(3+\rho^2\right)}{\left\{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2\left(1+\rho^2\right) + 2\rho\mu_1\mu_2\sigma_1\sigma_2\right\}^{\frac{3}{2}}}.
\end{aligned}$$

예를 들어, X_1과 X_2가 $iid\ N(\mu, \sigma^2)$경우,

$$E(Y) = \mu^2, \qquad V(Y) = \sigma^2\left(2\mu^2 + \sigma^2\right), \qquad \alpha_3(Y) = \frac{6\mu^2\sigma^4}{\left(2\mu^2\sigma^2 + \sigma^4\right)^{\frac{3}{2}}} \tag{2.2}$$

이 성립한다.

다음으로, 서로 독립$(\rho = 0)$인 정규확률변수

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}\right)$$

에 대해 $Y_1 = X_1X_2$와 $Y_2 = X_2X_3$의 공분산은

$$\begin{aligned}
\mathrm{Cov}(Y_1, Y_2) &= E\left[(X_1X_2 - \mu_1\mu_2)(X_2X_3 - \mu_2\mu_3)\right] \\
&= E(X_1X_2^2X_3) - \mu_1\mu_2E(X_2X_3) - \mu_2\mu_3E(X_1X_2) + \mu_1\mu_2^2\mu_3 \\
&= \mu_1\left(\mu_2^2 + \sigma_2^2\right)\mu_3 + \mu_1\mu_2^2\mu_3 - \mu_1\mu_2^2\mu_3 - \mu_1\mu_2^2\mu_3 \\
&= \sigma_2^2\mu_1\mu_3
\end{aligned}$$

이 되고, 이는 μ_1또는 μ_3가 영(zero)인 경우에는 역시 영이 되는 성질을 가지고 있다. 위의 결과로부터, 이들의 합 $Y = Y_1 + Y_2$에 대한 평균과 분산은 다음의 식

$$\begin{aligned}
E(Y) &= E(Y_1) + E(Y_2). \\
V(Y) &= V(Y_1) + V(Y_2) + 2\mathrm{Cov}(Y_1, Y_2)
\end{aligned}$$

을 이용하여 쉽게 계산될 수 있다. 예를 들어, X_1, X_2, X_3가 *iid* $N(\mu,\ \sigma^2)$인 경우,

$$E(Y) = 2\mu^2, \qquad V(Y) = 2\sigma^2(3\mu^2 + \sigma^2) \tag{2.3}$$

이 됨을 쉽게 확인할 수 있다.

3. 안장점근사와 FSK 통신에의 응용

3.1. 곱 정규변수의 합에 대한 안장점근사

먼저 $X = (X_1, \ldots, X_n)'$를 다변량 정규분포를 따르는 확률벡터로써, 평균을 $\mu = (\mu_1, \ldots, \mu_n)'$, 공분산행렬을 Σ라고 하자. 이 때, 곱 정규확률변수의 합은 다음과 같이 이차형식 $Q(X)$로 표현될 수 있다.

$$Q(X) = \sum_{1 \le i \le j \le n} a_{ij} X_i X_j = X'AX, \tag{3.1}$$

여기에서 행렬 $A = (a_{ij})_{n\times n}$는 일반성을 잃지 않고 대칭으로 가정할 수 있다.

식 (3.1)의 이차형식 $Q(X)$의 누율생성함수를 $K(t)$라 할 때, Lugannani와 Rice (1980)의 분포함수에 대한 안장점근사식은 다음과 같다 (Daniels, 1987).

$$\Pr\{Q(X) \le q\} \simeq \begin{cases} \Phi(\omega) + \phi(\omega)\left(\dfrac{1}{\omega} - \dfrac{1}{\zeta}\right), & q \ne E(Q(X)), \\ \dfrac{1}{2} - \dfrac{K^{(3)}(0)}{6\sqrt{2\pi\{K''(0)\}^3}}, & q = E(Q(X)). \end{cases} \tag{3.2}$$

위식에서 ω와 ζ는 다음과 같이 정의된다.

$$\omega = \operatorname{sgn}(t_0)\,[2n\{t_0 q - K(t_0)\}]^{\frac{1}{2}}, \qquad \zeta = t_0\{nK''(t_0)\}^{\frac{1}{2}},$$

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

$$K'(t_0) = q.$$

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

또한 식 (3.2)의 변형으로 분포함수에 대한 Jensen (1992, 1995)의 근사식은 다음과 같다.

$$\Pr\{Q(X) \le q\} \simeq \Phi\left\{\omega + \frac{1}{\omega}\log\left(\frac{\zeta}{\omega}\right)\right\}. \tag{3.3}$$

위의 두 가지 근사식의 적용과정에 사용될 이차형식의 누율생성함수와 관련 성질들은 다음과 같이 주어진다. 먼저 콜레스키 분해를 통해 양정치 공분산행렬(Σ)을 $\Sigma = \Gamma\Gamma'$으로 표현하자. 여기서 Γ는 비정칙 하삼각행렬이다. 또한 $B = \Gamma' A \Gamma$이라 하고, 행렬 B의 고유치 $\lambda_i\,(i = 1, \ldots, n)$들로 구성된 대각행렬을 $\Lambda = \operatorname{diag}(\lambda_i)$라고 하면, $B = P\Lambda P'$으로 표현된다. 여기서 P는 $P'P = I$를 만족하는 고유벡터들로 구성된 직교행렬이다.

정리 3.1 이차형식 $Q(X) = X'AX$의 누율생성함수는 다음과 같이 주어진다.

$$\begin{aligned} K_Q(t) &= -\frac{1}{2}\sum_{i=1}^{n}\log(1 - 2t\lambda_i) + \frac{1}{2}\sum_{i=1}^{n}\frac{\delta_i^2\lambda_i t}{1 - 2t\lambda_i} \\ &= -\frac{1}{2}\sum_{i=1}^{n}\log(1 - 2t\lambda_i) + \frac{1}{2}\sum_{i=1}^{n}\frac{\delta_i^2}{1 - 2t\lambda_i} - \frac{1}{2}\sum_{i=1}^{n}\delta_i^2, \end{aligned}$$

여기서 $\delta = P\Gamma^{-1}\mu$로 정의된다.

증명: Na와 Kim (2005)의 정리 2.1을 참고할 것. □

보조정리 3.1 이차형식 $Q(X)$의 1, 2차 미분식은 다음과 같이 주어진다.

$$K'_Q(t) = \frac{d}{dt}K_Q(t) = \sum_{i=1}^{n} \frac{\delta_i^2\lambda_i}{(1-2t\lambda_i)^2} + \sum_{i=1}^{n} \frac{\lambda_i}{1-2t\lambda_i}.$$

$$K''_Q(t) = \frac{d^2}{dt^2}K_Q(t) = \sum_{i=1}^{n} \frac{4\delta_i^2\lambda_i^2}{(1-2t\lambda_i)^3} + \sum_{i=1}^{n} \frac{2\lambda_i^2}{(1-2t\lambda_i)^2}.$$

증명: 정리 3.1로부터 쉽게 증명할 수 있다. □

3.2. FSK 통신에의 응용

FSK(Frequency Shift Keying)는 통신이론에서 주파수 신호를 정확히 잡아내기 위해, 특정 신호에 대해 고의적으로 잡음(noise)을 주어 신호를 송출하고 이를 받아서 해독하는 일련의 과정을 말한다. 서로 다른 인코딩 원(original) 신호 $s(t)$는 보통 가우지안 노이즈 $\omega(t)$가 더해져 보내지는데, 이 때 수신된 신호 $y(t)$는 다음과 같이 보내진 원 신호와 잡음의 합으로 표현될 수 있다.

$$y(t) = s(t) + \omega(t).$$

이 때, $s(t)$와 $\omega(t)$의 분산의 비를 "신호 대 잡음비(signal to noise ratio)"라 한다. 원 신호 $s(t)$의 추정값을 $\hat{s}(t)$라 할 때, 수신기의 성능은 $s(t)$와 $\hat{s}(t)$사이의 오차확률에 의해 결정된다. 이제 N개의 수신된 신호를 다음과 같이 표현하자.

$$y = \begin{bmatrix} y_{-\xi+1} \\ y_{-\xi+2} \\ \vdots \\ y_t \\ \vdots \\ y_{-\xi+N} \end{bmatrix} = \begin{bmatrix} s_{-\xi+1} + \omega_{-\xi+1} \\ s_{-\xi+2} + \omega_{-\xi+2} \\ \vdots \\ s_t + \omega_t \\ \vdots \\ s_{-\xi+N} + \omega_{-\xi+N} \end{bmatrix} = s + \omega.$$

위 식에서 ξ는 양의 정수이다. 이 때, 오차거리(error metric)를 다음과 같이 정의하자.

$$M(s,\hat{s}) = 2\sum_{t=1}^{N\xi}(y_{R,t}y_{R,t-\xi}b_{R,t} + y_{R,t}y_{I,t-\xi}b_{I,t} - y_{I,t}y_{R,t-\xi}b_{I,t} + y_{I,t}y_{I,t-\xi}b_{R,t}). \tag{3.4}$$

위 식에서 $y_{R,\cdot}$과 $y_{I,\cdot}$는 2차원 공간상의 실수축 좌표값과 허수축 좌표값을 나타내며, $b_{R,t}$와 $b_{I,t}$는 실수값으로, 미리 정해지는 상수이다. 또한 가해지는 잡음은 다음을 만족하도록 주어진다.

$$\begin{bmatrix} y_{R,t} \\ y_{I,t} \\ y_{R,t-\xi} \\ y_{I,t-\xi} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{R,t} \\ \mu_{I,t} \\ \mu_{R,t-\xi} \\ \mu_{I,t-\xi} \end{bmatrix}, \begin{bmatrix} \sigma_\omega^2 & 0 & 0 & 0 \\ 0 & \sigma_\omega^2 & 0 & 0 \\ 0 & 0 & \sigma_\omega^2 & 0 \\ 0 & 0 & 0 & \sigma_\omega^2 \end{bmatrix} \right).$$

보통 N의 값은 2 또는 3으로, ξ는 $2^k, k \geq 0$으로 주어지며, $b_{R,t}$와 $b_{I,t}$는 1의 값으로 주어진다. 여기서 우리의 관심사는 식 (3.4)의 오차거리의 분포와 함께 이 측도가 음의 값을 가질 확률 즉,

$$P\{M(s,\hat{s}) < 0\} \tag{3.5}$$

이라 말할 수 있다. 본 논문에서는 앞 절에서 소개한 이차형식에 대한 안장점근사를 이용하여 정규확률 변수의 곱들의 합의 형태로 표현되는 오차거리의 분포를 추정하기로 한다.

4. 예제를 통한 비교

앞서 소개된 FSK 통신에서의 예를 통해 식 (3.5)의 확률에 대한 안장점근사를 실시하고, 이 근사의 정도(precision)에 대해 알아보기로 한다. 먼저 $N = 2$, $\xi = 1$이고, $b_{R,i} = b_{I,i} = 1$, $i = 1, 2$일 때,

$$\begin{bmatrix} y_{R,0} \\ y_{I,0} \\ y_{R,1} \\ y_{I,1} \\ y_{R,2} \\ y_{I,2} \end{bmatrix} \sim N \left(\begin{bmatrix} \mu \\ \mu \\ \mu \\ \mu \\ \mu \\ \mu \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma^2 \end{bmatrix} \right)$$

에 대한 오차거리는, 상수배를 무시하면,

$$\begin{aligned} M(s,\hat{s}) &= \sum_{t=1}^{2} (y_{R,t}y_{R,t-1}b_{R,t} + y_{R,t}y_{I,t-1}b_{I,t} - y_{I,t}y_{R,t-1}b_{I,t} + y_{I,t}y_{I,t-1}b_{R,t}) \\ &= y_{R,0}y_{R,1} + y_{R,1}y_{I,0} - y_{R,0}y_{I,1} + y_{I,0}y_{I,1} + y_{R,1}y_{R_2} + y_{R,2}y_{I,1} - y_{R,1}y_{I,2} + y_{I,1}y_{I,2} \end{aligned} \quad (4.1)$$

이 되고, 위 식은 다시 다음과 같은 이차형식으로 표현될 수 있다. 즉,

$$\begin{aligned} M(s,\hat{s}) &= (Y_1, Y_2, \ldots, Y_6) \begin{bmatrix} 0 & 0 & 1/2 & -1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 1/2 & -1/2 \\ -1/2 & 1/2 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & -1/2 & 1/2 & 0 & 0 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_6 \end{bmatrix} \\ &= Y' AY. \end{aligned}$$

위 식에서 $Y' = (Y_1, Y_2, \ldots, Y_6) = (y_{R,0}, y_{I,0}, y_{R,1}, y_{I,1}, y_{R,2}, y_{I,2})$이다. 이제 이차형식으로 표현된 식 (4.1)의 분포함수에 대해, 앞 절에서 소개한 안장점근사를 실시함으로써 분포함수에 대한 근사를 수행할 수 있다.

본 논문에서는 모의실험을 통해 안장점근사의 정확도를 알아보았다. (a) $\mu = 1$, $\sigma^2 = 0.1(0.9)0.2$, (b) $\mu = 2$, $\sigma^2 = 0.5$, $1(4)1$, (c) $\mu = 3$, $\sigma^2 = 1(9)2$의 경우에 대해 관심있는 확률 $P\{M(s,\hat{s}) < 0\}$에 대한 근사를 실시하고, 그 결과를 표 4.1에 나타내었다. 표 4.1에서 Sim. Exact(Simulated Exact)는 모의실험 1,000,000번에 기초한 추정값이며, Saddle I 과 Saddle II 는 각각 식 (3.2)와 (3.3)의 안장점근사의 결과이며, Normal은 정규근사의 결과를 나타낸다. 여기서 정규근사는 식 (2.2)를 식 (4.1)에 적용하여 구해지는 오차거리 $M(s,\hat{s})$의 평균과 분산

$$\begin{aligned} E[M(s,\hat{s})] &= 2N\xi\mu^2 = 4\mu^2 \\ V[M(s,\hat{s})] &= 4N\xi\sigma^2(\sigma^2 + 2\mu^2) = 8\sigma^2(\sigma^2 + 2\mu^2) \end{aligned}$$

표 4.1. FSK에서의 관심확률 추정

μ	σ^2	$P\{M(s,\hat{s})<0\}$의 근사			
		Sim.Exact	Normal	Saddle I	Saddle II
1	0.1	2.4E-05	1.0E-03	2.3E-05	2.3E-05
1	0.3	0.0177	0.0443	0.0176	0.0176
1	0.5	0.0673	0.1029	0.0664	0.0664
1	0.7	0.1198	0.1518	0.1173	0.1173
1	0.9	0.1648	0.1907	0.1612	0.1612
2	0.5	1.6E-04	3.0E-03	1.7E-04	1.7E-04
2	1.0	0.0091	0.0296	0.0091	0.0091
2	2.0	0.0676	0.1030	0.0664	0.0664
2	3.0	0.1315	0.1624	0.1290	0.1290
2	4.0	0.1834	0.2071	0.1802	0.1802
3	1.0	6.1E-05	1.8E-03	6.3E-05	6.3E-05
3	3.0	0.0249	0.0544	0.0246	0.0246
3	5.0	0.0827	0.1176	0.0810	0.0810
3	7.0	0.1380	0.1680	0.1353	0.1353
3	9.0	0.1843	0.2071	0.1802	0.1802

표 4.2. 분포함수 $P\{M(s,\hat{s})<m\}$에 대한 근사

m	$P\{M(s,\hat{s})<m\}$의 근사			
	Sim.Exact	Normal	Saddle I	Saddle II
0.0	0.0673	0.1029	0.0664	0.0664
0.5	0.1075	0.1342	0.1019	0.1019
1.0	0.1583	0.1714	0.1495	0.1495
1.5	0.2173	0.2146	0.2079	0.2079
2.0	0.2835	0.2635	0.2739	0.2739
3.0	0.4219	0.3759	0.4149	0.4148
5.0	0.6707	0.6241	0.6684	0.6682
6.0	0.7657	0.7365	0.7639	0.7638
7.0	0.8379	0.8286	0.8371	0.8370
8.0	0.8910	0.8970	0.8907	0.8906
9.0	0.9287	0.9431	0.9284	0.9283
10.0	0.9548	0.9711	0.9541	0.9541
11.0	0.9712	0.9866	0.9712	0.9712
12.0	0.9823	0.9943	0.9822	0.9822
13.0	0.9892	0.9978	0.9892	0.9892
13.5	0.9917	0.9987	0.9916	0.9916
14.0	0.9935	0.9992	0.9935	0.9935
14.5	0.9951	0.9996	0.9950	0.9950
15.0	0.9962	0.9997	0.9962	0.9962

을 이용하여, 다음의 정규근사식으로부터 구해진 값이다.

$$P\{M(s,\hat{s})<m\} \simeq \Phi\left\{\frac{m-4\mu^2}{\sqrt{8\sigma^2(\sigma^2+2\mu^2)}}\right\}.$$

표 4.1의 결과를 살펴보면, 정규근사의 경우 극단꼬리 영역에서 정확도가 크게 떨어지는 것을 확인할 수 있다. 반면 안장점근사의 경우 그 정확도는 두 경우(Saddle I , Saddle II) 모두 모의실험을 통한 정확한 값과 큰 차이가 없으며, 극단 꼬리 영역에서도 정도가 유지되는 것을 확인할 수 있다.

위의 표 4.1에서는 FSK 통신에서 관심확률인 오차거리가 영보다 작은 확률을 추정하였다. 이제, 영의 값을 벗어난 분포의 전 영역에 걸쳐 안장점근사의 정도를 알아보기로 한다. 본 논문에서는 위의 표 4.1에서 3번째 경우에 해당하는 $\mu = 1, \sigma^2 = 0.5$인 경우에 대해, m의 값의 변화에 따른 오차거리의 분포함수 $P\{M(s,\hat{s}) < m\}$에 대한 근사결과를 표 4.2에 제시하였다. 표 4.2의 결과에서 알 수 있듯이 본 논문에서 제시한 안장점근사는 분포의 전 영역에 걸쳐 그 정확도가 뛰어남을 알 수 있다. 또한 이 근사법은 뛰어난 정확도와 더불어 복잡한 분포이론이나 수치적분과정을 포함하지 않는다는 점에서 많은 장점을 가진다.

5. 결론

정규확률변수의 곱에 대한 분포는 두 확률변수의 평균과 분산의 값에 따라 다양한 형태를 취할 수 있다. 또한 곱의 합으로 표현되는 통계량의 분포는 다양한 응용분야에서 중요하게 취급된다. 본 연구에서는 이들 통계량을 지금까지 시도된 바 없는 이차형식의 관점에서 이해하고, Na와 Kim (2005)의 안장점근사에 기초한 새로운 근사법을 제시하였다. 이 근사법은 통계량의 분포함수를 유도하기위한 분포이론이나 분포함수의 근사를 위한 밀도함수의 수치적분 등의 계산과정이 불필요하며, 기존의 정규근사와는 달리 꼬리영역 또는 극단 꼬리영역에서도 근사의 정확도가 유지되는 장점을 가지고 있다. 본 연구에서는 FSK 통신과 관련하여 제시된 근사법이 기존의 정규근사에 비해 매우 뛰어난 결과를 제공함을 확인하였다.

참고문헌

Aroian, L. A. (1947). The probability function of the product of two normally distributed variables, *The Annals of Mathematical Statistics*, **18**, 265–271.

Aroian, A. L., Taneja, V. S. and Cornwell, L. W. (1978). Mathematical forms of the distribution of the product of two normal variables, *Communications in Statistics- Theory and Methods*, **7**, 165–172.

Conradie, W. and Gupta, A. (1987). Quadratic forms in complex normal variates: Basic results, *Statistica*, **47**, 73–84.

Cornwell, L. W., Aroian, L. A. and Taneja, V. S. (1978). Numerical evaluation of the distribution of the product of two normal variables, *Journal of Statistical Computation and Simulation*, **7**, 123–131.

Craig, C. C. (1936). On the frequency function of xy, *Annals of Mathematical Statistics*, **7**, 1–15.

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

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.

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

Na, J. H. and Kim, J. S. (2005). Saddlepoint approximations to the distribution function of non-homogeneous quadratic forms, *The Korean Journal of Applied Statistics*, **18**, 183–196.

Small Sample Asymptotic Distribution for the Sum of Product of Normal Variables with Application to FSK Communication

Jong-Hwa Na[1] · Jung-Mi Kim[2]

[1]Dept. of Information and Statistics, Chungbuk National University;
[2]Dept. of Information and Statistics, Chungbuk National University

(Received October 2008; accepted October 2008)

Abstract

In this paper we studied the effective approximations to the distribution of the sum of products of normal variables. Based on the saddlepoint approximations to the quadratic forms, the suggested approximations are very accurate and easy to use. Applications to the FSK(Frequency Shift Keying) communication are also considered.

Keywords: Product of normal variables, saddlepoint approximation, quadratic form, FSK communication, asymptotic distribution.

This work was supported by the research grant of the Chungbuk National University in 2008.
[1]Corresponding author: Professor, Dept. of Information and Statistics Institute for Basic Science Research, Chungbuk National University, 12 Gashin-dong, Heungduk-gu, Cheongju, Chungbuk 361-763, Korea. E-mail: cherin@chungbuk.ac.kr