김일성종합대학학보 주체110(2021)년
수 학 제67권 제3호

2차원타원형분포에 따르는 우연점으로부터 임의의 점까지의 거리의 분포

리경일, 손영순

다차원우연벡토르의 길이를 계산하는것은 통계학, 화상처리, 금융수학, 기계학습, 도형처리 등에서 중요한 의의를 가진다.[1, 2]

선행연구[3]에서는 평균이 0이고 척도행렬이 단위행렬의 상수배인 경우에 다차원타원형분포에 따르는 우연벡토르로부터 자리표원점까지의 L^p거리의 밀도함수와 모멘트들의 표시식을 계산하였다.

우리는 평균과 척도행렬이 일반적인 경우 2차원타원형분포에 따르는 우연벡토르로부터 평면의 임의의 점까지의 L^2거리의 분포에 대하여 밀도함수와 모멘트를 구하고 밀도함수를 얻는 수치모의실례를 론의한다.

정의 $(\Omega,\ \mathscr{F},\ P)$를 확률공간이라고 하고 $n \geq 1$을 자연수라고 하자. 우연벡토르 $\boldsymbol{X} = (X_1,\ X_2,\ \cdots,\ X_n)^{\mathrm{T}}$의 밀도함수가

$$f_X(x) = \frac{1}{\sqrt{|\Sigma|}} g\left(\frac{1}{2}(x-\mu)^{\mathrm{T}} \Sigma^{-1} (x-\mu)\right),\ x \in \mathbf{R}^n \tag{1}$$

일 때 $\boldsymbol{X}$는 평균이 μ이고 척도행렬이 Σ인 타원형분포에 따른다고 말하고 $\boldsymbol{X} \sim E_n(\mu,\ \Sigma,\ g)$로 표시한다. 여기서 $\mu \in \mathbf{R}^n$이고 Σ는 정의정값 n차행렬이며 $g(u),\ u \geq 0$은 조건

$$\int_0^{\infty} t^{n/2-1} g(t)dt < \infty \tag{2}$$

를 만족시키는 함수로서 밀도생성자라고 부른다.

$\mu = (\mu_1,\ \mu_2)^{\mathrm{T}} \in \mathbf{R}^2,\ \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \in \mathbf{R}^{2\times 2}$이고 $g(u),\ u \geq 0$을 조건 (2)를 만족시키는 밀도생성자라고 하자.

정리 1 $\boldsymbol{X} \sim E_2(\mu,\ \Sigma,\ g)$는 2차원타원형분포에 따르는 우연벡토르이고 $y = (y_1,\ y_2)^{\mathrm{T}} \in \mathbf{R}^2$을 평면의 임의의 점이라고 하자. 또한 $\boldsymbol{X}$의 자리표성분들을 $X_1,\ X_2$라고 하자. 이때 $\boldsymbol{X}$로부터 y까지의 L^2거리 $L = \sqrt{(X_1 - y_1)^2 + (X_2 - y_2)^2}$의 밀도함수는 다음과 같다.

$$f_L(l) = \frac{l}{\sqrt{|\Sigma|}} \int_0^{2\pi} h(l,\ \omega)d\omega \tag{3}$$

여기서

$$h(l,\ \omega) := g\left(\frac{1}{2|\Sigma|}[\sigma_2^2(l\cos\omega + y_1 - \mu_1)^2 - 2\sigma_{12}(l\cos\omega + y_1 - \mu_1)(l\sin\omega + y_2 - \mu_2) + \right.$$
$$\left. + \sigma_1^2(l\sin\omega + y_2 - \mu_2)^2]\right)$$

이다.

증명 L의 분포함수를 $F_L(l),\ l \ge 0$이라고 하면 $\boldsymbol{X} \sim E_2(\mu,\ \Sigma,\ g)$의 밀도함수가 식 (1)이라는데로부터 임의의 $l \ge 0$에 대하여

$$F_L(l) = P(L \le l) = P(\sqrt{(X_1 - y_1)^2 + (X_2 - y_2)^2} \le l) = \iint_{B_{y,l}} \frac{1}{\sqrt{|\Sigma|}} g\left(\frac{1}{2}(x-\mu)^{\mathrm{T}} \Sigma^{-1} (x - \mu)\right) dx_1 dx_2$$

가 성립한다. 여기서

$$B_{y,\,l} = \{x = (x_1,\ x_2)^{\mathrm{T}} \in \mathbf{R}^2 \mid \| x - y \|_2 = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \le l\}$$

을 의미한다. 우의 적분식에서 변수변환

$$\begin{cases} x_1 = t_1 + y_1 \\ x_2 = t_2 + y_2 \end{cases}$$

를 실시하면

$$F_L(l) = \iint_{B_l} \frac{1}{\sqrt{|\Sigma|}} g\left(\frac{1}{2}(t + y - \mu)^{\mathrm{T}} \Sigma^{-1} (t + y - \mu)\right) dt_1 dt_2$$

가 성립한다. 여기서 $B_l = \{t = (t_1,\ t_2)^{\mathrm{T}} \in \mathbf{R}^2 \mid \| t \|_2 = \sqrt{t_1^2 + t_2^2} \le l\}$ 이다.

$|\Sigma| = \sigma_1^2 \sigma_2^2 - \sigma_{12}^2,\ \Sigma^{-1} = \dfrac{1}{|\Sigma|}\begin{pmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{pmatrix}$ 이라는것을 고려하면

$$F_L(l) = \frac{1}{\sqrt{|\Sigma|}} \iint_{B_l} g\left(\frac{1}{2|\Sigma|}[\sigma_2^2 (t_1 + y_1 - \mu_1)^2 - \right.$$

$$\left. - 2\sigma_{12}(t_1 + y_1 - \mu_1)(t_2 + y_2 - \mu_2) + \sigma_1^2 (t_2 + y_2 - \mu_2)^2\,]\right) dx_1 dx_2$$

가 성립된다. 이제 극자리표변환 $\begin{cases} t_1 = r\cos\omega \\ t_2 = r\sin\omega \end{cases}$ 를 실시하자.

$$F_L(l) = \frac{1}{\sqrt{|\Sigma|}} \int_0^l \int_0^{2\pi} r g\left(\frac{1}{2|\Sigma|}[\sigma_2^2 (r\cos\omega + y_1 - \mu_1)^2 - \right.$$

$$\left. - 2\sigma_{12}(r\cos\omega + y_1 - \mu_1)(r\sin\omega + y_2 - \mu_2) + \sigma_1^2 (r\sin\omega + y_2 - \mu_2)^2\,]\right) d\omega dr$$

$$f_L(l) := \frac{l}{\sqrt{|\Sigma|}} \int_0^{2\pi} g\left(\frac{1}{2|\Sigma|}[\sigma_2^2 (l\cos\omega + y_1 - \mu_1)^2 - \right.$$

$$\left. - 2\sigma_{12}(l\cos\omega + y_1 - \mu_1)(l\sin\omega + y_2 - \mu_2) + \sigma_1^2 (l\sin\omega + y_2 - \mu_2)^2\,]\right) d\omega$$

로 놓자. 그러면 f_L은 부가 아닌 적분가능한 가측함수이고 임의의 $l \ge 0$에 대하여

$$F_L(l) = \int_0^l f_L(r) dr$$

가 성립하므로 f_L은 L의 밀도함수이다. 따라서 정리의 결론이 성립한다.(증명끝)

정리 2 $k\geq 1$을 자연수라고 하자. 우연량 L의 k차모멘트는 다음과 같다.

$$\mathbf{E}[L^k]=\int_0^{\infty}\int_0^{2\pi}\frac{l^{k+1}}{\sqrt{|\Sigma|}}h(l,\ \omega)d\omega dl$$

증명 k차모멘트의 정의와 정리 1로부터

$$\mathbf{E}[L^k]=\int_0^{\infty}l^k f_L(l)dl=\int_0^{\infty}\frac{l^{k+1}}{\sqrt{|\Sigma|}}\int_0^{2\pi}h(l,\ \omega)d\omega dl=\int_0^{\infty}\int_0^{2\pi}\frac{l^{k+1}}{\sqrt{|\Sigma|}}h(l,\ \omega)d\omega dl$$

이다.(증명끝)

이제 가우스－르쟝드르구적법을 리용하여 우연량 L의 밀도함수 $f_L(l)$을 수치적으로 근사시켜보자. 먼저 가우스－르쟝드르구적법에 대하여 간단히 론의한다.

$q\in\mathbf{N}$에 대하여 $(c_j^q)_{j=1,\cdots,q}\subset[-1,\ 1]$을 q차르쟝드르다항식

$$L^q(x)=\frac{1}{2^q q!}\frac{d^q}{dx^q}[(x^2-1)^q]$$

의 뿌리들이라고 하자. 적분 $\int_a^b g(t)dt$를 q차가우스－르쟝드르구적법으로 근사시킬 때 구적점 $(t_j)_{j=1,\cdots,q}$와 구적무게 $(w_j)_{j=1,\cdots,q}$는 다음과 같이 주어진다.

$$t_j=\frac{c_j^q(b-a)+(a+b)}{2},\ \ w_j=\int_a^b\left[\prod_{\substack{i\in\{1,\cdots,q\}\\ t_i\neq t_j}}\frac{2x-(b-a)c_i^q-(a+b)}{2t_j-(b-a)c_i^q-(a+b)}\right]dx$$

이때 근사오차에 관하여 다음의 식이 성립한다.

$$\int_a^b g(t)dt-\sum_{j=1}^q w_j g(t_j)=\frac{[q!]^4(b-a)^{2q+1}}{(2q+1)[(2q)!]^3}g^{(2q)}(\xi),\ \xi\in[a,\ b]$$

우의 식에서 볼수 있는것처럼 르쟝드르구적법은 적은 개수의 구적점만으로도 적분근사오차를 충분히 작게 할수 있는 우점을 가지고있다.

$\widetilde{f}_L^q(l)$을 q점가우스－르쟝드르구적법에 의한 $f_L(l)$의 근사라고 하자. 이때

$$\widetilde{f}_L^q(l)=\frac{l}{\sqrt{|\Sigma|}}\sum_{j=1}^q w_j h(l,\ t_j),\ t_j=\pi(c_j^q+1),\ w_j=\int_0^{2\pi}\left[\prod_{\substack{i\in\{1,\cdots,q\}\\ t_i\neq t_j}}\frac{x-\pi c_i^q-\pi}{t_j-\pi c_i^q-\pi}\right]dx$$

이다.

참 고 문 헌

[1] Den Dekker, A. J. Sijbers; J. Phys. Medica., 30, 725, 2014.

[2] A. E. Gomes et al.; J. Stat. Comput. Simul., 84, 290, 2014.

[3] T. Shushi; Statist. Probab. Lett., 153, 104, 2019.

주체110(2021)년 3월 5일 원고접수

Distribution of the Distance from a Random Point with Two-Dimensional Elliptical Distribution to any Point

Ri Kyong Il, Son Yong Sun

In this paper, we obtain the distribution and moments of the distance from a random point with two-dimensional elliptical distribution to any point in the case of general location and scale parameter.

Keyword: elliptical distribution