김일성종합대학학보 JOURNAL OF KIM IL SUNG UNIVERSITY
(자연과학) (NATURAL SCIENCE)
주체106(2017)년 제63권 제10호 Vol. 63 No. 10 JUCHE106(2017).

지수형분포에서 J－분리도

한광룡, 전순영

위대한 수령 **김일성**동지께서는 다음과 같이 교시하시였다.

《통계는 사회주의건설에서 매우 중요한 역할을 합니다. 정확한 통계가 있어야 옳은 계획을 세울수 있고 모든 사업을 과학적으로 해나갈수 있습니다. 통계는 곧 사회주의입니다.》
(《**김일성**전집》 제43권 368페지)

선행연구[3－5]에서는 J－분리도, 분리도개념을 새롭게 론의하였으며 선행연구[2]에서는 두 파라메터를 가진 우연결음분포족에서 두 밀도함수사이의 J－분리도를 밝혔다.

선행연구[1]에서는 구체적인 몇가지 중요한 분포들에 대한 J－분리도를 연구하였다.

론문에서는 선행연구[1, 2]의 결과를 보다 일반화하여 지수형분포(족)에 대하여 두 분포사이의 J－분리도를 밝혔다.

1. 통계적모형과 J－분리도

확률분포를 원소로 하는 모임으로 이루어진 어떤 통계적모형을 다양체로 보면 이 다양체우에서 어떤 류형의 리만계량과 아핀접속이 자연스럽게 도입된다.[2－5]

모임 X 우에서 확률분포족 M 에 대하여 보기로 하자.

확률분포족 M 의 매 원소의 어떤 확률분포가 m 개의 실파라메터 $(\theta_1, \cdots, \theta_m)$ 으로서 $M = \{p_\theta = p(x\,;\theta) \mid \theta = (\theta_1, \cdots, \theta_m) \in z\}$ 로 표시될 때 M 을 X 우에서의 m 차원통계적모형 또는 간단히 모형이라고 부른다. 여기서 z 는 $\boldsymbol{R}^m$ 의 부분모임이며 $\theta \to P_\theta$ 는 1∶1넘기기이다.

통계적모형 M 에 대해서는 보통 다음과 같이 가정한다.

우선 파라메터 θ 에 대한 미분연산을 자유롭게 론의할수 있도록 z 를 $\boldsymbol{R}^m$ 에서의 열린모임으로 가정하며 $\forall x \in X$ 에 대하여 함수 $\theta \to p(x\,;\theta)$ $(z \to p)$ 는 C^∞ 급임을 가정한다.

이때 $\partial_i p(x\,;\theta),\ \partial_i \partial_j p(x\,;\theta)$ 들이 정의된다.

또한 미분과 적분의 순서도 교환가능하다고 가정한다.

실례로 $\int \partial_i p(x;\theta)dx = \partial_i \int p(x;\theta)dx = \partial_i = 0$ 인 등식을 자주 리용한다.

M 을 m 차원통계적모형이라고 할 때 주어진 점 $\theta(\in z)$ 에서 M 의 핏샤정보행렬은 다음의 식으로 정의된 $g_{ij}(\theta)$ 를 $(i,\ j)$ 원소로 하는 $m \times n$ 형행렬 $G(\theta) = (g_{ij}(\theta))$ 이다.

$$g_{ij}(\theta) = \mathrm{E}_\theta(\partial_i l_\theta,\ \partial_j l_\theta) = \int_{-\infty}^{+\infty} \partial_i l_{(x,\ \theta)} \partial_j l_{(x,\ \theta)} p_{(x,\ \theta)} dx$$

여기서 $l_\theta = l(x,\ \theta) = \ln p(x,\ \theta),\ \partial_i = \dfrac{\partial}{\partial \theta^i}$ 이며 E_θ 는 파라메터 θ 를 가진 분포 $p(x\,;\theta)$ 에 대한 기대값을 의미한다.

모형에 의하여 식 (1)이 발산할수 있지만 보통 임의의 θ, i, j에 대하여 $g_{ij}(\theta)$가 유한이며 $g_{ij}: z \to \boldsymbol{R}$가 C^{∞}급이라고 가정한다. 그리고 g_{ij}는 $g_{ij}(\theta) = -\mathrm{E}_{\theta}(\partial_i \partial_j l_{\theta})$로 표시할수 있다. 이것은 $\int \partial_i p(x;\theta)dx = \partial_i \int p(x;\theta)dx = \partial_i = 0$을 $\mathrm{E}_{\theta}(\partial l_{\theta}) = 0$으로 표시할 때 량변에 ∂_i를 실시하여 얻는다.

$G(\theta)$는 대칭행렬이며 반정값(또는 정값)행렬이다. 이것은 $(\partial_1 l_{\theta}, \cdots, \partial_m l_{\theta})$가 X 우의 함수로서 1차독립이라는것과 동등하며 $(\partial_1 l_{\theta}, \cdots, \partial_m l_{\theta})$가 1차독립성을 가진다는것과 동등하다.

M은 간단한 다양체구조를 가질뿐아니라 매 점이 확률분포를 표시하는 특성이 있다.

(θ^i)과 $(\theta^i + d\theta^i)$이 열린모임 z의 가까운 점이라고 하면 이 두 점사이의 무한소거리 ds는 $ds^2 = \sum_{i,\, j=1}^{m} g_{ij}(\theta) d\theta^i d\theta^j$과 같이 주어진다. 이때 통계다양체우의 두 점에 대응되는 분포밀도 $p = p(x;\theta)$, $q = (x;\, \theta + d\theta)$사이의 측지거리는 $s(p,\, q) = \int_p^q \sqrt{\sum_{i,\, j=1}^{m} g_{ij}(\theta) d\theta^i d\theta^j}$ 이다.

한편 아주 가까운 두 분포밀도함수 $p = p(x;\, \theta)$와 $q = p(x;\, \theta + d\theta)$사이의 J－분리도는 $J(p,\, q) = \int_{-\infty}^{+\infty} (p - q) \ln \frac{p}{q} dx$와 같이 정의된다.[1－4]

웃식을 테일러전개하면 $J(p,\, q) = \sum_{i,\, j=1}^{m} \left[\int_{-\infty}^{+\infty} \frac{\partial \ln p}{\partial \theta^i} \cdot \frac{\partial \ln p}{\partial \theta^j} p dx \right] d\theta^i d\theta^j = \sum_{i,\, j=1}^{m} g_{ij}(\theta) d\theta^i d\theta^j = ds^2$

즉 두 분포밀도 $p = p(x;\, \theta)$와 $q = p(x;\, \theta + d\theta)$사이의 J－분리도는 두 점 θ^i와 $\theta^i + d\theta^i$사이의 측지거리의 두제곱과 국부적으로 일치한다.

J－분리도는 클백크정보량 $I(p,\, q) = \int_{-\infty}^{+\infty} p(x) \ln \frac{p(x)}{q(x)} dx$에 기초하여 다음과 같이 표시된다.

$$J = J(p,\, q) = I(p,\, q) + I(q,\, p)$$

2. 지수형분포의 J－분리도

통계적추론문제의 해결에서 특별히 중요한것은 지수형분포(족)이다.

정의 밀도함수 또는 확률함수가

$$p(x;\theta) = d(x) \exp\left\{ \sum_{i=1}^{k} \theta_i T_i(x) + b(\theta) \right\} \quad (x \in \boldsymbol{R}^m) \tag{1}$$

로 주어지는 m차원분포를 k파라메터지수형분포라고 부른다.

일반적으로 크기가 n인 표본의 동시적분포도 지수형분포이다.

사실 동시적밀도(또는 확률)는

$$L(x;\theta) = \prod_{j=1}^{n} d(x_j) \cdot \exp\left\{ \sum_{i=1}^{k} \theta_i \cdot \sum_{j=1}^{n} T_i(x_j) + nb(\theta) \right\} \tag{2}$$

이므로 $D(x) = D(x_1, x_2, \cdots, x_n) = \prod_{j=1}^{n} d(x_j)$, $H_i(x) = H_i(x_1, x_2, \cdots, x_n) = \sum_{j=1}^{n} T_i(x_j)$로 놓으면 된다.

그리고 이때 식 (2)의 동시적밀도는 $L(x;\theta)=D(x)\cdot\exp\left\{\sum_{i=1}^{k}\theta_i H_i(x)+nb(\theta)\right\}$ 로 되므로 k 차원통계량 $(H_1(X),\ H_2(X),\cdots,\ H_k(X))$ 는 $\theta=(\theta_1,\theta_2,\cdots,\theta_n)$ 에 대한 충분통계량이다.

지수형분포는 보다 일반적으로 다음과 같이 표시된다.

$$p(x;\theta)=d(x)\exp\left\{\sum_{i=1}^{k}a_i(\theta)T_i(x)+b(\theta)\right\},\quad \theta=(\theta_1,\theta_2,\cdots,\theta_k) \tag{3}$$

사실 $v_i=a_i(\theta)$ 로 놓으면 이것은 확률함수 (1)의 모양으로 된다.

주의할것은 $b(\theta)$ 가 $v=(v_1,v_2,\cdots,v_k)$ 의 함수로 되겠는가 하는것인데 식 (1)에서 $e^{b(\theta)}$ 은 $e^{b(\theta)}\int_{\boldsymbol{R}^n}d(x)\exp\left\{\sum_{i=1}^{k}\theta_i T_i(x)\right\}dx=1$ 을 만족시키게 하기 위한 표준화결수이며 따라서 식 (3)의 $b(\theta)$ 도 $a_i(\theta)$ 들의 함수로 된다는것을 알수 있다.

선행연구들에서 지수형분포족과 관련하여

$$p(x;\theta)=C(\theta)\exp\left[\sum_{i=1}^{k}\theta_i T_i(x)\right] \tag{4}$$

의 형태도 론의하고있는데 그것은 $p(x;\theta)=C(\theta)d(x)\exp\left\{\sum_{i=1}^{k}\theta_i T_i(x)\right\}\ (x\in\boldsymbol{R}^m)$ 형태의 지수형분포가 가측공간 $(X,\boldsymbol{B})$ 에서 σ－유한측도 μ 에 관한 밀도라고 하면 본래의 측도 $d(x)$ 를 측도 $\mu(x)$ 로 바꾸고 $d\nu(x)=h(x)d\mu(x)$ 를 생각하면 되기때문이다.

지수형분포 (4)의 오른변의 적분이 유한 즉 $\int_{\boldsymbol{R}^n}\exp\left\{\sum_{i=1}^{k}\theta_i T_i(x)\right\}d\nu(x)<\infty$ 이면 표준화정수 $C(\theta)$ 를 적당히 택하여 $p(x;\theta)$ 가 확률밀도로 되게 할수 있다. 이때 $\nu(x)$ 는 미리 주어진것으로 되여야 하며 파라메터 θ 에 무관계한 $X=\boldsymbol{R}^k$ 에서의 σ－유한측도로 되여야 한다. 이것이 성립되는 파라메터점 $\theta=(\theta_1,\cdots,\theta_k)$ 들의 모임 Θ 를 지수형분포 (4)의 자연파라메터공간이라고 부른다.

지수형분포족의 자연파라메터공간은 불룩모임이다.

보조정리 φ 가 $(X,\boldsymbol{B})$ 에서의 유계가측함수이면 다음의 사실들이 성립된다.

① 복소수변수 $\theta_j=\alpha_j+i\beta_j,\ j=1,\cdots,k$ 의 함수인 적분 $\int\varphi(x)\exp\left[\sum_{i=1}^{k}\theta_i T_i(x)\right]d\nu(x)$ 는 $(\alpha_1,\cdots,\alpha_k)$ 가 자연파라메터공간 Θ 의 내점으로 되는 구역 $\boldsymbol{R}$ 에서 이 매 변수에 관하여 해석함수이다.

② 우의 적분의 θ 에 관한 임의의 계수의 도함수는 적분기호밑에서 미분하여 계산할수 있다.

이제 이 결과를 리용하여 지수형분포의 제곱지수통계량 T 에 대한 다음과 같은 특성값을 얻기로 한다.

정리 1 여러파라메터지수형분포 $p(x;\theta)=C(\theta)D(x)e^{\sum_{i=1}^{k}v_i(\theta)T_i(x)}$, $\theta=(\theta_1,\cdots,\theta_k)$ 에서 통계량 T_i 의 수학적기대값과 공분산은 다음과 같다.

$$\mathrm{E}_\theta T_i(X) = -\frac{\partial \ln C(\theta)}{\partial \theta_i}\frac{1}{v_i'(\theta)},\ \ \mathrm{E}T_i(X)T_j(X) = \left(-\frac{C_{ij}''}{C} + \frac{C_i'C_j'}{C^2}\right)\frac{1}{v_i'(\theta)v_j'(\theta)},\ \ i,\ j = 1, \cdots, k$$

$$\mathrm{cov}(T_i(X),\ T_j(X)) = -C_{ij}''(\theta)/[C(\theta)\cdot v_i'(\theta)v_j'(\theta)],\ i,\ j = 1, \cdots, k$$

따름 여러파라메터지수형분포 $p(x;\theta) = C(\theta)\exp\left[\sum_{i=1}^{k}\theta_i T_i(x)\right]$ 에서 통계량 T_i 의 수학적기대값과 공분산은 각각 다음과 같다.

$$\mathrm{E}T_i(X) = -\frac{\partial \ln C(\theta)}{\partial \theta_i} = -\frac{C_i'(\theta)}{C(\theta)}\ \ (i = \overline{1,\ k})$$

$$\mathrm{cov}(T_i(X),\ T_j(X)) = -\frac{\partial^2 \ln C(\theta)}{\partial \theta_i \partial \theta_j} = -\frac{C_{ij}''(\theta)}{C(\theta)}\ \ (i,\ j = \overline{1,\ k})$$

정리 2 $p(x;\theta) = C(\theta)D(x)e^{-\theta T(x)}$ 에 대한 J –분리도는 다음과 같다.

$$J(p,\ q) = [\theta_1 - \theta_2]\left[\frac{\partial \ln C(\theta_2)}{\partial \theta_2} - \frac{\partial \ln C(\theta_1)}{\partial \theta_1}\right],\ \ p = p(x;\theta_1),\ \ q = p(x;\theta_2)$$

증명 클백크정보량은

$$I(p,\ q) = \int(\ln p - \ln q)p(x\,;\,\theta_1)dx = \ln\frac{C(\theta_1)}{C(\theta_2)} + (\theta_1 - \theta_2)\int T(x)p(x\,;\,\theta_1)dx$$

와 같다. 여기서 지수형분포의 제곱지수 $T(x)$ 에 대하여 정리 1에서의 수학적기대값이 $\mathrm{E}_\theta T(X) = \int_{-\infty}^{\infty} T(x)p(x;\theta)dx = -\frac{\partial \ln C(\theta)}{\partial \theta}$ 이라는것과 밀도함수의 적분이 1이라는것을 리용한다.

따라서 J –분리도는 $J(p,\ q) = I(p,\ q) + I(q,\ p) = [\theta_1 - \theta_2]\left[\frac{\partial \ln C(\theta_2)}{\partial \theta_2} - \frac{\partial \ln C(\theta_1)}{\partial \theta_1}\right]$ 과 같이 표시된다.(증명끝)

정리 3 일반적인 파라메터지수형분포 $p(x;\theta) = C(\theta)D(x)e^{-v(\theta)T(x)}$ 에 대한 J –분리도는 $J(p,\ q) = [v(\theta_1) - v(\theta_2)]\left[\frac{\partial \ln C(\theta_2)}{\partial \theta_2} - \frac{\partial \ln C(\theta_1)}{\partial \theta_1}\right],\ \ p = p(x;\theta_1),\ \ q = p(x;\theta_2)$ 와 같다.(증명생략)

정리 4 여러파라메터지수형분포 $p(x;\theta) = C(\theta)D(x)e^{\sum_{i=1}^{k} v_i(\theta)T_i(x)}$, $\theta = (\theta_1, \cdots,\ \theta_k)$ 에 대한 J –분리도는 $J(p,\ q) = \sum_{i=1}^{k}[v_i(\theta_1) - v_i(\theta_2)]\left[\frac{\partial \ln C(\theta_2)}{\partial \theta_{2i}}\frac{1}{v'(\theta_{2i})} - \frac{\partial \ln C(\theta_1)}{\partial \theta_{1i}}\frac{1}{v'(\theta_{1i})}\right]$ 과 같다.

증명 클백크정보량은 $I(p,\ q) = \ln\frac{C(\theta_1)}{C(\theta_2)} + \sum_{i=1}^{k}(v_i(\theta_1) - v_i(\theta_2))\int T_i(x)p(x;\theta_1)dx$ 와 같다.

정리 1에서 얻은 수학적기대값 $\mathrm{E}_\theta T_i(X) = \int_{-\infty}^{\infty} T_i(x)p(x;\theta)dx = -\frac{\partial \ln C(\theta)}{\partial \theta_i}\frac{1}{v_i'(\theta)}$ 과 밀도함수의 적분이 1이라는 사실을 리용하면 J –분리도는

$$J(p,\ q) = I(p,\ q) + I(q,\ p) = \sum_{i=1}^{k}[v_i(\theta_1) - v_i(\theta_2)]\left[\frac{\partial \ln C(\theta_2)}{\partial \theta_{2i}}\frac{1}{v'(\theta_{2i})} - \frac{\partial \ln C(\theta_1)}{\partial \theta_{1i}}\frac{1}{v'(\theta_{1i})}\right]$$

로 표시된다.(증명끝)

정리 2, 3, 4에서 밝힌 지수형분포의 J－분리도를 리용하면 선행연구[1]에서 얻은 다음과 같은 몇가지 중요한 분포의 J－분리도를 얻을수 있다.

$$J(p,\ q)=(1/(2\sigma_2^2)-1/(2\sigma_1^2))[(\mu_1^2+\sigma_1^2)-(\mu_2^2+\sigma_2^2)]+$$
$$+(\mu_1/\sigma_1^2-\mu_2/\sigma_2^2)(\mu_1-\mu_2)\ \ (\text{정규분포}\ \ N(\mu,\ \sigma^2))$$
$$J(p,\ q)=(\lambda_1-\lambda_2)^2/(\lambda_1\lambda_2)\ \ (\text{지수분포}\,e(\lambda))$$
$$J(p,\ q)=(p_1-p_2)\ln\{p_1(1-p_2)/[p_2(1-p_1)]\}\ \ (2\text{항분포}\ \ Bi(1\,;\,p))$$
$$J(p,\ q)=(1/p_1-1/p_2)\ln[(1-p_1)/(1-p_2)]\ \ (\text{기하분포}\ \ Ge(p))$$
$$J(p,\ q)=(\lambda_1-\lambda_2)+\ln(\lambda_1/\lambda_2)\ \ (\text{뽜쏭분포}\ \ Po(\lambda))$$
$$J(p,\ q)=k(D_1-D_2)^2/(D_1-D_2)\ \ (\text{질점들사이의}\ \ k\,\text{차거리분포}\ \ H_n(k,\ D))$$

마찬가지로 선행연구[2]에서 얻은 우연결음분포족에 대한 분리도를 정리 4로부터 얻을수 있다. 그것은 우연결음분포의 밀도함수

$$f(x;\ \mu,\ \lambda)=\sqrt{\lambda/(2\pi x)}\exp\{-\lambda x/2+\lambda/\mu-\lambda/(2\mu^2x)\},\ \ x,\ \lambda,\ \mu>0$$

을

$$C(\lambda,\ \mu)=\sqrt{\lambda/(2\pi)}e^{\lambda/\mu},\ \ D(x)=1/\sqrt{x},\ \ T_1(x)=x,\ \ T_2(x)=1/x,\ \ \nu_1(\lambda,\ \mu)=-\lambda/2,\ \nu_2(\lambda,\ \mu)=-1/(2\mu^2)$$

로 놓으면 지수형분포 $f(x;\theta)=C(\theta)D(x)e^{\sum_{i=1}^{k}\nu_i(\theta)T_i(x)}$, $\theta=(\mu,\ \lambda)$로 되기때문이다.

참 고 문 헌

[1] **김일성**종합대학학보(자연과학), 56, 5, 6, 주체98(2009).

[2] H. Nassar et al.; Chaos Solitions Fractals, 15, 161, 2003.

[3] E. Hayaman et al.; Tensor(N. S.), 57, 282, 1996.

[4] R. Ivanova et al.; Tensor(N. S.), 57, 300, 1996.

[5] O. Calin et al.; Geometric Modeling in Probability and Statistics, Springer, 3～375, 2014.

주체106(2017)년 6월 5일 원고접수

The J－Divergence from Exponential Type Distribution

Han Kwang Ryong, Jon Sun Yong

We first found the mathematical expectation and covariance formulas of power exponent T in exponential type distribution.

On the basis of it the J-divergence formulas for exponential type distribution are obtained.

Key words: exponential type distribution, divergence