Historia Mathematica Vol. 9. No. 2, (1996), 15 - 21.

3차 적률(Moment)의 선형적 성질

한양대학교 수학과 동 경 화

요 약

적률법(method of moment)이란 변수 X의 멱승에대한 기대치를 이용하여 분포의 성질을 알아보는 방법이다. 여기서 적률법이 이용되어진 역사적 배경을 소개하고, 3차 적률들의 선형적 성질을 비교하였다. 먼저, Kagan이 입증한 표본평균에 관한 3차 표본적률의 선형적 성질과 Bayesian 경우에 3차 사후적률(posterior moment)과 사후평균(posterior)의 선형성을 소개하였다. 그리고, 자연지수족(natural exponential family)아래서도 표본평균에 관한 3차 표본적률의 선형성을 알아보기 위해 단순함수(simple function)의 형태로 유도하였으며, 정규분포인 경우에 적용시켜 보았다.

1. 배경

적률(Moment)이란 변수 X의 멱승에 대한 기대값을 지칭한다. 예컨대, 분포함수 $F(X)$를 갖는 확률변수 X에 대하여 $E\{(g(X))^r\} = \int_{-\infty}^{\infty}\{(g(x))^r\}dF(x)$를 $g(X)$의 r차적률(r-th moment)이라한다. 특히, $g(x) = x$일때 $m_r = E(X^r) = \int_{-\infty}^{\infty} x^r dF(x)$를 원점둘레의 r차 적률(r-th moment about origin)이라하고 $r=1$일때 $m_1 = E(X)$가 된다. 다른 한편 상수 a 에 대하여 $r_a = \int_{-\infty}^{\infty}(x-a)^r dF(x)$를 a둘레의 r차적률(r-th moment about a)이라한다. 이때 $a=E(X)$ 이면 $\mu_r = E\{(X-E(X))^r\} = \int_{-\infty}^{\infty}(x-E(X))^r dF(x)$를 평균 둘레의 r차적률(r-th moment about mean $E(X)$)이라한다.

통계학에서 적률을 이용한 방법은 모집단을 수식화시키는데 필요한 모수를 추정하거나 분포의 어떤 적률들의 값을 추정하는 방법으로 이용된다. 예를들어, 단일변수의 경우 분포

$f(x, \theta)$의 k개의 모수 $\theta_1, \theta_2, \cdots, \theta_k$ 가 $\{\theta^*\}$의 적률추정치를 아래와 같이 갖는다고 하자.

$$\mu_r'(\theta) = \int_a^b x^r f(x;\theta)\, dx, \quad r = 0, 1, \cdots, j \ ; \quad j \geq k$$

가 존재하고,

$$\mu_r'(\theta^*) = m_r', r = 1, 2, \cdots, k \ . \tag{1.1}$$

여기서, m_r'는 원점둘레에 관한 표본적률(r-th noncentral sample moment)이라한다.

적률법(method of moment)은 긴 역사를 갖고 있으며 여기에 관련된 논문들 역시 방대한 양에 달한다. 또한 표본을 추출하는 과정에서도 다른 추정과정들과 비교하여 이에 관련된 여러 부분들이 오랜 세월동안 연구되어 왔음에도 불구하고 적률법이 쉽게 이행될 수 있고 또한 폭넓은 일반성을 띄고 있기에 아직까지 효과적인 방법으로 여겨지고 있다. Karl Pearson(1857-1936)은 아래와 같은 두개의 성분을 갖는 복합 정규모델에서

$$P(x;\ p_1,\ p_2,\ \sigma_1,\ \sigma_2,\ \lambda_1,\ \lambda_2) = \sum_{r=1}^{2} p_i\ g_i(x) \quad ,$$

여기서,

$g_i(x) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left\{\frac{-(x-\lambda_i)^2}{2(\sigma_i)^2}\right\}$이다. $(\sigma_i)^2$ (또는 $(\sigma_2)^2$)로 얻은 값이 음수이면 ; 복합비가 음이될 수 있으므로 이론상으로는 적어도 음이 아니라는(확률분포에서) 조건이 만족되지 않고서는 적률법이 해를 가질 수 없다고 하였다. Pearson 의 System에 따르면 (Pearson 은 분포를 12가지 형태로 분류하였음) Type IV는 4개의 적률들이 요구되며, Type III(with unknown origin)는 3개의 적률들을 포함하며, Type I (with know range limits)은 2개의 적률들을·갖는다고 밝혔다. (1.1)식의 경우에는 $k=4$일때 이용된다. 만일 Type I의 곡선이라면 양 끝점이 포함된다. Pearson System의 초기 계산치나 적률의 근사치는 Elderton(1960)에 의해 연구 되었으며, 후에 Elderton 과 Johnson(1969)에 의해 수정되어 졌다[표 1.1].

Tchouproff는(1918-1919)에 Bio-metrika에서

(1) noncentral이 central moments(표본과모집단)로 (현대용어상) 변환되는공식

(2) m_1'에서 8차까지의 적률들

(3) $E\{(m_r' - \mu_r')(m_s' - \mu_s')(m_t' - \mu_t')\}$;

(4) E_{m_r} , $r = 2, 3, \cdots, 6$;

(5) $E_{m_2}{}^r$, $r = 1, 2, 3, 4,$;

(6) $\mu_r(m_2)$, $r = 2, 3, 4.$ ($\mu_4(m_2)$는 order가 n^{-7}이며 38개의 계수를 갖는 형태로 나타남을 보였다).

Pearson 의 System

Type	Density	Support
I	$(1+x)^{m_1}(1-x)^{m_2}$	$-1\leq x\leq 1$
VI	$x^{m_2}(1+x)^{-m_1}$	$0\leq x< \infty$
IV	$(1+x^2)^{-m}\exp\{-\nu\tan^{-}(x)\}$	$-\infty< x< \infty$
Normal	$\exp(-\frac{1}{2}x^2)$	$-\infty< x< \infty$
II	$(1-x^2)^m$	$-1\leq x\leq 1$
VII	$(1+x^2)^{-m}$	$-\infty< x< \infty$
III	$x^m\exp(-x)$	$0\leq x< \infty$
V	$x^{-m}\exp(-x^{-1})$	$0\leq x< \infty$
VIII	$(1+x)^m$	$0\leq x\leq 1$
IX	$(1+x)^m$	$0\leq x\leq 1$
X	$\exp(-x)$	$0\leq x< \infty$
XI	x^{-m}	$1\leq x< \infty$
XII	$[(g+x)/(g-x)]^h$	$-g\leq x\leq g$

[표 1.1]

이에 앞서 Thiele는 Semi-invariants(or cummulants)를 이용하여 적률들의 표현을 단순화 시키고자 하는 모두의 바램을 이루었다 (즉, $E\prod_{s=1}^{4}(m_{\lambda_s}-\mu_{\lambda_s})$에서 n^{-2}항은 60개의 항으로 구성된다). 여기서, μ의 적률들에 관해 k 번 누적시키므로 $1+\mu_1'\alpha+\mu_2'\frac{\alpha_2}{2!}+\cdots=\exp\left\{k_1\alpha+k_2\frac{\alpha^2}{2}+\cdots\right\}$로 나타낼 수 있고, 정규분포에서는(단일변수이던지 아니던지) 2차이상 누적치는 0 이 된다. Thiele은 $k_2, k_3, \cdots$,는 원점둘레에서 자유로우며(free), $\left\{k_r/\ k_2^{\frac{r}{2}}\right\}$, $r\geq 2$는 원점과 scale에서 free하다고 언급하였다. 이후 1981년에 Hald는 Thiele의 결과를 Gram-Charlier의 급수를 이용하여 cummulants 를 표현하였다. 1925년 Church는 Tchouproff의 방법을 이용하여 m_2 (noncentral and central)에서 처음 4개의 적률들을 주었다. Sophister는 1927년에 Gamma모집단 ($y=y_0x^7\exp(-x)$, 5개와 20개의 표본을 가지고)으로부터 m_2의 분포를 다루었는데, 이때 Pearson의 Type VI곡선에서 4개의 적률들을 이용하였다.

본 논문에서 이런 많은 적률들의 성질 중에서 3차적률들의 선형적 성질을 비교하고자 한다. 표본평균에 관한 3차 표본 적률의 선형적인 성질을 Kagan이 밝혔으며, 그는 또한 Bayesian경우에 3차 사후적률(posterior moment)과 사후평균(posterior mean)에 대해 선형성을 입증하였다. 이와 연관지어 지수족 아래서 3차적률의 선형적 성질을 유도해보고, 정규분포와의 관계성을 보이고자한다

2. 3차 적률의 선형성 비교

첫째, 표본평균에대한 3차 평균적률을 생각해보자. $X_1, X_2, \cdots, X_n$은 분포함수 F 로부터 얻은 $n \geq 2$인 표본이라하자.

$$\overline{X} = (X_1 + \cdots + X_n)/n, \quad S^2 = M_2 = \sum_{j=1}^{n}(X_j - \overline{X})^2/n, \quad M_3 = \sum_{j=1}^{n}(X_j - \overline{X})^3/n.$$

Linnik(1970)에 따르면 $n \geq 3$이고 $E|X_j|^3 < \infty$ 이면,

$$E(M_3 \mid \overline{X}) = const \tag{2.1}$$

로 Gaussian 분포 F를 특성화시키는 것을 보였다. 그러나, Kagan(1990)은 아래와 같은 정리를 통해서 (2.1)식의 오른쪽 부분인 상수항을 $\overline{X}$ 에관해 선형적으로 나타냄으로 (2.1)식을 일반화시켰다.

Theorem 2.1 Let $X_1, X_2, \cdots, X_n$, $n \geq 3$ be an independent sample from a d.f. F with $E|X_j|^3 < \infty$. The relation $E(M_3 \mid \overline{X}) = a_0 + a_1 \overline{X}$ for some constants $a_0, a_1 \neq 0$ holds if and only if F is possibly shifted convolution of Poisson distributions $P_0(\Delta, \lambda_1)$ and $P_0(-\Delta, \lambda_2)$ concentrated on lattices $0, \Delta, 2\Delta, \cdots$, and, $-2\Delta, -\Delta, 0$ and having intensity parameters $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ respectively.

Proof. See Kagan(1990)

둘째, 기본적인 Bayesian 경우에 대해 생각해보자. $X = \theta + \varepsilon$ 이라하고 여기서 θ는 (location)모수이고 ε은 오차항이며 서로 독립이라하자. Diaconis 와 Ylvisaker는 ε의 특성함수가 0이 아니거나 그것의 적률들이 오차의 분포를 결정하지 않는다면, ε의 분포와 사후평균 $E(\theta \mid X)$의 선형성이 θ의 (사전)분포를 결정한다는 것을 밝혔다. 이를 발전시켜, ε와 θ의 모든 분포들은

(1) $E(\theta \mid X)$ 는 선형이고

(2) $Var(\theta \mid X)$는 qudratic 형이거나

(3) $E\{(\theta - E(\theta \mid X))^3 \mid X\}$는 선형이 되는 것을 보이고,

(2)와 (3)의 조건은 자연통계적 특성치들과 연관성이 있으며 $E(\theta \mid X)$는 θ 의 사전분포를 (ε의 분포를 모르더라도) 결정하기에 충분하다는 것을 의미한다.

Theorem 2.2 Suppose $E|\theta|^3 < \infty$ is satisfied. If

$$E(\theta \mid X) = a_0 + a_1 X, \quad E\{(\theta - (\theta \mid X))^3 \mid X\} = c_0 + c_1 X,$$

where $a_0, a_1 \neq \frac{1}{2}$, $c_0, c_1 \neq 0$ are constants, then Π and F are possibly shifted convolutions of Poisson distributions concentrated on lattices $0, \Delta, 2\Delta, \cdots,$ and $-2\Delta, -\Delta, 0$ respectively.

Proof. See Kagan(1992, p36)

셋째, 자연지수족(Natural exponential families)에서 지수함수의 특성을 이용하여 평균에 대한 3차 적률을 간단한 형태의 함수로 (즉, 특성함수를 유도하여)표현한다면 이는 곧 선형적 성질을 가짐을 알 수 있다. F를 $\int \exp(\theta x) dF(x) < \infty$를 만족하는 R^1위에서 Borel측도라하자. 여기서 $\theta \in \Theta$, $0 \in \text{int}\ \Theta$이며 F는 아래와 같은 $\{F_\theta, \theta \in \Theta\}$를 생성한다고 한다.

$$dF_\theta = \exp(\theta x - \phi(\theta)) dF \tag{2.2}$$

여기서 , $\phi(\theta) = \log \int \exp(\theta x) dF(x)$.

만일 X가 분포 (2.2)를 갖는 확률변수라면 $\mu(\theta) = E_\theta X = \phi'(\theta)$, $Var_\theta X = \phi''(\theta)$임은 널리 아는 바이다. 그러면 $E_\theta (X - \mu(\theta))^3$도 $\mu(\theta)$의 simple 함수로 표현 되어질 수있다.

Theorem2.3 The convolution of two Poisson distributions $F = P_0(\Delta, \lambda_1) * P_0(-\Delta, \lambda_2)$ where $\Delta > 0$, $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ is the only distribution that generates on NEF with

$$E_\theta (X - \mu(\theta))^3 = E_\theta (X - \phi'(\theta))^3 = a_0 + a_1 \phi'(\theta) \tag{2.3}$$

where a_0, a_1 are arbitrary constants and

$$\phi(\theta) = \frac{c_0}{\sqrt{a_1}} \exp(\sqrt{a_1}\theta), \quad -\frac{c_1}{\sqrt{a_1}} \exp(-\sqrt{a_1}\theta) - \frac{a_0}{a_1}\theta, \ (c_0, c_1\text{: constants}).$$

Proof. We rewrite the third central moment of mean as following $E_\theta (X - E_\theta X)^3 = E_\theta [X^3 - 3X^2 \phi'(\theta) + 3X(\phi'(\theta))^2 - (\phi'(\theta))^3]$. Using the property of p.d.f. under (2.2), we know

$$\int \exp(\theta x) d\nu = \exp \phi(\theta). \tag{2.4}$$

We differentiate each side of (2.4) with respect to θ, then

$\int xexp(\theta x)d\nu = \phi'(\theta)\exp(\phi(\theta))$. Now, we get expectation of X

$$E_\theta(X) = \int xexp(\theta x - \phi(\theta))d\nu = \phi'(\theta). \qquad (2.5)$$

We differentiate each side of (2.5) with respect to θ, then

$$\int (x-\phi'(\theta))x \cdot \exp(\theta x - \phi'(\theta))d\nu = \phi''(\theta).$$

So, we obtain the expectation of X^2 as following

$$E_\theta(X^2) = \int x^2\exp(\theta x - \phi(\theta))d\nu = \phi''(\theta) + (\phi'(\theta))^2. \qquad (2.6)$$

We differentiate each side of (2.6) one more time with respect to θ, we get

$$\int (x-\phi'(\theta))x^2\exp(\theta x - \phi(\theta))d\nu = \phi'''(\theta) + 2\phi'(\theta)\phi''(\theta).$$

Thus, we get the expectation of X^3

$$\begin{aligned}\int x^3\exp(\theta x - \phi(\theta))d\nu &= \phi'''(\theta) + 2\phi'(\theta)\phi''(\theta) + \phi'(\theta)(\phi''(\theta) + \phi'(\theta)^2) \\ &= \phi'(\theta) + 3\phi'(\theta)\phi''(\theta) + (\phi'(\theta))^3.\end{aligned} \qquad (2.7)$$

Using (2.5),(2.6) and (2.7), we obtain (2.3) as below

$$\begin{aligned}E_\theta(X-E_\theta)^3 &= \{\phi'''(\theta) + 3\phi'(\theta)\phi''(\theta) + (\phi'(\theta))^3\} \\ &\quad -3\{\phi''(\theta) + (\phi'(\theta))^2\}\phi'(\theta) \\ &\quad +3(\phi'(\theta))^2\phi'(\theta) - (\phi'(\theta))^3 \\ &= \phi'''(\theta)\end{aligned}$$

Therfore, $\phi'''(\theta) - a_1\phi'(\theta) = a_0$.

By differential equation, we get the form of solution as following

$$\phi'(\theta) = c_0\exp(\sqrt{a_1}\theta) + c_1\exp(-\sqrt{a_1}\theta) - \frac{a_0}{a_1},$$

$$\phi(\theta) = \frac{c_0}{\sqrt{a_1}}\exp(\sqrt{a_1}\theta) - \frac{c_1}{\sqrt{a_1}}\exp(-\sqrt{a_1}\theta) - \frac{a_0}{a_1}\theta,$$

where a_0, a_1, c_0, c_1 are constants.

3.결론

2장에서 살펴본 바와 같이 여러경우에 3차적률의 선형적 성질이 Robust하다는 것을 이용한다면 모집단의 분포를 추정하는데 보다 쉽게 접근할 수 있을 것이며, 이는 $X_j = \theta + \varepsilon_j$, $j = 1, \cdots, n$.(n은 임의의 수)에서와 같은 모형에도 적용하면

$E\{(\theta - E(\theta \mid \overline{X}))^3 \mid \overline{X}\}$가 선형성을 갖는 것도 밝힐 수 있을 것이다. 이를 활용한다면 정규모집단의 경우에 대칭적 성질을 이용하여 메디안도 평균을 이용한 특성함수로 표현할 수 있을 것이다. 더 나아가 비대칭형의 경우에 연구를한다면 모집단의 분포를 추정하는 또 하나의 새로운 접근 방법으로 많은 활용이 되리라 사료되는 바이다.

참고문헌

[1]. Church, A. E. R.(1925). *Biometrika*, 17, 79-83.

[2]. Elderton, W. P. (1960). *Frequency curves and Correlation*, Cambridge University. Press, Cambridge, England.

[3]. Elderton, N. P. and Johnson, N. L. (1969). *Systems of Frequency Curves*. Cambridge University Press, cambridge, England.

[4]. Hald, A.(1981). *Int. Statist. Rev*., 49, 1-20.

[5]. Kagan, A. M. (1990). Linearity of regression of third sample moment on the sample average. *Metron*. Vol. XLVIII-N. 1-4.

[6]. Kagan, A. M. (1992). Linearity of Posterior mean for for location parameter families when Sample size equals one. *Mathematical Methods of Statistics*. p28-38.

[7]. Perason, K. (1894). *Phil, Trans. R. Soc. Land.* A., 185-1-40.

[8]. Samuel, K. & Norman, L. J. (1981). *Encyclopedia of Statistical Sciences*. Vol. 5-6.

[9]. Tchouproff, A. A. (1918). *Biometrika*, 12, 140-169.

[10]. Tchouproff, A. A. (1919). *Biometrika*, 12, 185-210.

[11]. Thiele, T. N. (1903). *Theory of Observations*. C. and E. Layton, London.

효과적으로 가르치기 위해서 교사는 반드시 수학에 대해 호감을 가져야 한다. 교사가 스스로 수학의 활력을 느끼지 못한다면, 학생에게 열정을 불어넣을 수 없다. 교사가 어떤 내용이 중요하다고 강조하려면, 반드시 그것이 중요하다고 스스로 느껴야 한다.

- George Ploya, *Mathematical Discovery*(1981)