

On Tail Probabilities of Continuous Probability Distributions with Heavy Tails

Seokhoon Yun^{a,1}

^aDepartment of Applied Statistics, University of Suwon

(Received July 30, 2013; Revised October 8, 2013; Accepted October 8, 2013)

Abstract

The paper examines several classes of probability distributions with heavy tails. An (asymptotic) expression for tail probability needs to be known to understand which class a given probability distribution belongs to. It is usually not easy to get expressions for tail probabilities since most absolutely continuous probability distributions are specified by probability density functions and not by distribution functions. The paper proposes a method to obtain asymptotic expressions for tail probabilities using only probability density functions. Some examples are given to illustrate the proposed method.

Keywords: Heavy tailed distributions, subexponential distributions, long tail, regularly varying tail.

1. 서론

요즘, 보험, 외환, 주식, 파생상품 등과 관련된 금융자료나, 대기오염, 수질오염 등을 포함한 환경자료 등의 모델링에 기존의 정규분포(normal distribution) 대신 보다 꼬리가 두꺼운 확률분포(heavy-tailed probability distribution)를 사용해야 한다는 주장이 많이 등장하고 있다. 일명 “Black Monday”라고 칭해지는, 1987년 10월 19일 월요일에 발생한 미국 증권시장 역사상 최악의 대폭락 사태 때는 하루 동안 다우존스 지수가 508.32가 폭락하였는데, 이는 22.6%의 폭락률에 해당하고 금액으로는 미국 증권시장 전체적으로 약 5,000억달러가 하루 사이에 사라지는 결과가 초래되었다. 이 대폭락 사태 이전까지만 해도 포트폴리오 매니저, 금융 의사결정권자, 증권 거래원들은 대부분의 유가 증권에서 수익률(혹은 손실률)의 분포는 정규분포를 따른다고 믿어 왔다 (Fergusson과 Platen, 2006). 그러나, 이후 극단 수익률(혹은 극단 손실률)의 발생은 정규분포에서 예상되는 것보다도 훨씬 자주 발생되고 있음이 자료 분석 결과 밝혀졌고, 오늘 날에는 대부분의 경제학자들이 주식, 외환에서의 수익률(혹은 손실률) 분포에 대한 정규분포 가정을 강하게 부정하고 있다.

천체 관측을 포함한 자연 현상의 관측에서 대부분 오차는 수반되게 마련이고 이러한 오차의 분포는 대칭적이며 종 모양의 정규분포를 따라야 정상(normal)이라는 19세기의 일반화된 생각이 현대 정보화 사회의 다양하고도 복잡한 자료를 분석하는 오늘 날의 통계 분석가에게는 더 이상의 표준이 되지 못한다. 오늘 날 여러 분야의 많은 사람들이 꼬리가 두꺼운 분포에 대해서 관심을 표명하는데, 20세기 통계 자료 모형화에 널리 사용된 바 있는 정규분포는

평균 $\pm 3 \times$ 표준편차

¹Professor, Department of Applied Statistics, University of Suwon, Suwon 445-743, Korea.
E-mail: syun@suwon.ac.kr

이내에 자료가 속하는 비율이 무려 99.73%나 되고 따라서 극단값들이 다수 포함되어 있는 현대의 복잡 한 자료를 모형화하기에 정규분포는 적합하지 않기 때문이다.

본 논문에서는 우선 두꺼운 꼬리를 갖는 확률분포들의 부류를 세분화해서 살펴본다. 어떤 확률분포가 이렇게 세분화된 여러 부류 중 어디에 속하는 지를 알아 내기 위해서는 해당 확률분포의 꼬리 확률에 대한 표현식 또는 점근 표현식을 알아야만 한다. 그러나, 대다수 절대 연속 확률분포(또는 확률 변수)들은 분포함수(distribution function)의 명시적인 표현식이 알려져 있지 않고, 대신 확률밀도함수(probability density function)만이 알려져 있다. 본 논문에서는 이러한 경우 확률밀도함수만을 이용하여 꼬리 확률의 점근 표현식을 쉽게 찾는 방법을 제안한다.

2. 두꺼운 꼬리를 갖는 확률분포

X 를 하나의 확률변수라고 하자. 만약

$$\lim_{x \rightarrow \infty} e^{\lambda x} P\{X > x\} = \infty, \quad \text{for all } \lambda > 0$$

이 성립하면, X 의 확률분포는 두꺼운 오른쪽 꼬리(heavy right tail)를 갖는다고 말하는데 (Asmussen, 2003), 이는 곧 X 의 확률분포가 지수분포(exponential distribution)보다 오른쪽 꼬리가 더 두껍다는 것을 의미한다. X 의 확률분포가 두꺼운 오른쪽 꼬리를 갖는다는 조건은 조건

$$E\left(e^{tX}\right) = \infty, \quad \text{for all } t > 0$$

과 동치임을 보일 수가 있는데 (Rolski 등, 1999), 따라서 X 의 확률분포가 두꺼운 오른쪽 꼬리를 갖는 경우 X 의 적률생성함수(moment generating function)가 0 근방(neighborhood)에서 정의되지 못하는 결과가 초래되고 이는 곧 고차 적률(high-order moments)의 존재성에 대한 의구심으로 이어질 수도 있다는 것을 의미한다. 20세기 통계 자료 모형의 대표 주자인 정규분포는 오른쪽 꼬리가 두껍지 않다. 반면에, 정보화 시대를 살고 있는 현대인들이 빈번하게 마주치는 오늘 날의 많은 통계 자료들은 오른쪽 꼬리가 두꺼운 확률분포의 모형화가 필수적이라는 주장이 여러 분야에서 제기되고 있는데, 따라서 이 경우 별 생각 없이 기존의 정규분포 모형화를 사용하게 되면 많은 문제점이 나타날 수도 있다.

확률분포의 왼쪽 꼬리에 대해서도 유사한 정의를 내릴 수가 있다. 만약

$$\lim_{x \rightarrow -\infty} e^{-\lambda x} P\{X < x\} = \infty, \quad \text{for all } \lambda > 0$$

이 성립하면, 혹은 동등하게

$$E\left(e^{tX}\right) = \infty, \quad \text{for all } t < 0$$

이 성립하면, X 의 확률분포는 두꺼운 왼쪽 꼬리(heavy left tail)를 갖는다고 말한다. 확률변수 X 의 확률분포가 두꺼운 왼쪽 꼬리를 갖는다는 것과 $-X$ 의 확률분포가 두꺼운 오른쪽 꼬리를 갖는다는 것은 동치이므로, 확률분포의 왼쪽 꼬리에 대한 두꺼움에 대해서는 더 이상 언급하지 않기로 한다.

확률분포의 오른쪽 꼬리에 대한 두꺼움은 좀 더 세분화해서 살펴볼 수도 있다. 첫째, 만약

$$\lim_{x \rightarrow \infty} P\{X > x + t | X > x\} = 1, \quad \text{for all } t > 0,$$

혹은 동등하게

$$P\{X > x + t\} \sim P\{X > x\} \text{ as } x \rightarrow \infty, \quad \text{for all } t \in \mathbb{R}$$

이 성립하면, X 의 확률분포는 긴 오른쪽 꼬리(long right tail)를 갖는다고 말한다 (Asmussen, 2003). X 의 확률분포가 긴 오른쪽 꼬리를 갖는다는 것은 결국 X 의 값이 어떤 높은 수준을 넘어서면 X 의 값이 보다 더 높은 수준을 넘어설 확률이 1에 가까워진다는 것이므로, 예를 들어 일상 생활에서 어떤 상황이 나쁘다는 것을 알 때, 실제로는 생각보다 훨씬 더 나빠져 있는 경우를 종종 접할 수가 있는데 이런 경우가 해당된다고 볼 수 있다.

둘째, 만약

$$\lim_{x \rightarrow \infty} \frac{P\{X_1 + X_2 > x\}}{P\{X > x\}} = 2,$$

혹은 동등하게, 모든 $n = 2, 3, \dots$ 에 대하여,

$$P\{X_1 + \dots + X_n > x\} \sim P\{\max(X_1, \dots, X_n) > x\} \sim nP\{X > x\} \text{ as } x \rightarrow \infty \quad (2.1)$$

가 성립하면, X 는 열지수분포(subexponential distribution)를 따른다고 말한다 (Teugels, 1975; Goldie와 Resnick, 1988; Klüppelberg, 1989). 여기서, X_1, \dots, X_n 은 X 의 독립적인 복제(independent copies)를 나타낸다. X 가 열지수분포를 따른다는 것은 결국 $X_1 + \dots + X_n$ 의 값과 $\max(X_1, \dots, X_n)$ 의 값이 높은 수준을 넘어설 확률은 서로 가까워진다는 것인데, 즉 달리 말하면 꼬리 확률은 n 개의 누적 값이 단일 최대값에 의해 좌우된다는 것으로서, 이를 종종 단일폭등의 원리(principle of the single big jump)라고 부른다.

셋째, 만약

$$\lim_{t \uparrow 1} \limsup_{x \rightarrow \infty} \frac{P\{X > tx\}}{P\{X > x\}} = 1,$$

혹은 동등하게

$$\lim_{t \downarrow 1} \liminf_{x \rightarrow \infty} \frac{P\{X > tx\}}{P\{X > x\}} = 1$$

이 성립하면, X 의 확률분포는 일관적으로 변하는 오른쪽 꼬리(consistently varying right tail)를 갖는다고 말한다 (Cline, 1994).

마지막으로, 만약 적당한 상수 $\alpha \geq 0$ 에 대하여,

$$\lim_{x \rightarrow \infty} \frac{P\{X > tx\}}{P\{X > x\}} = t^{-\alpha}, \quad \text{for all } t > 0 \quad (2.2)$$

이 성립하면, X 의 확률분포는 지수(index)가 $-\alpha$ 이며 정규적으로 변하는 오른쪽 꼬리(regularly varying right tail)를 갖는다고 말한다 (Resnick, 1987). 특별히, $\alpha = 0$ 이면 X 의 확률분포는 극도로 두꺼운 오른쪽 꼬리(super-heavy right tail)를 갖는다고 말하고 (Alves 등, 2009), $\alpha > 0$ 이면 X 가 멱분포(power-law distribution)를 따른다고 말하기도 한다. 여기서, 주의할 점은 $\alpha = 0$ 이면 X (X 가 음수를 허용할 때는 $\max(X, 0)$)의 모든 양의 차수의 적률이 ∞ 가 되고, 즉 모든 $k > 0$ 에 대하여 $E(X^k) = \infty$ 가 되고, 따라서 이 경우 경험적 적률(empirical moments)에 기초한 통계적 추론은 의미가 없어진다. 또한, $\alpha > 0$ 이면 오직 α 보다 작은 양의 차수에 대해서만 X 의 적률이 유한한 값을 갖게 되는데 (Embrechts 등, 1997), 따라서 예를 들어 $\alpha = 1$ 이라면 평균이나 분산에 기초한 통계적 추론은 역시 의미가 없게 된다.

앞에서 소개한, X 의 확률분포에 대한 오른쪽 꼬리의 두꺼움에 대한 여러가지 정의는 포함 관계가 소개한 순서의 역순이다. 즉, X 의 확률분포가 정규적으로 변하는 오른쪽 꼬리를 가지면, X 의 오른쪽 꼬

리는 역시 일관적으로 변하고, 이는 다시 X 가 열지수분포를 따르게 되는 것을 의미하며, 이는 또 다시 X 가 긴 오른쪽 꼬리를 갖는다는 것을 의미하게 되고, 이는 결과적으로 X 가 두꺼운 오른쪽 꼬리를 갖는다는 것을 뜻하게 된다.

절대 연속 확률변수 X 의 확률분포는 확률밀도함수나 분포함수로 나타내는 것이 일반적이다. 그러나 많은 경우 명시적인 분포함수의 표현식이 알려져 있지 않은데 이런 경우 확률밀도함수만으로 해당 확률변수의 확률분포를 나타내고 있다. 예를 들면, 정규분포, 감마분포(gamma distribution), t 분포(t distribution) 등이 여기에 해당된다. 다음의 Lemma는 이러한 경우 $x \rightarrow \infty$ 혹은 $x \rightarrow -\infty$ 일 때 해당 확률변수의 꼬리 확률 $P\{X > x\}$ 혹은 $P\{X < x\}$ 의 점근 표현식을 얻는 데 효과적으로 사용될 수 있다.

Lemma 2.1 X 를 하나의 절대 연속 확률변수라고 하고 f_X 를 X 의 확률밀도함수라고 하자. X 가 오른쪽 끝점으로 ∞ 를 가지면, 적당한 식 $a(x)$ 가 존재하여

$$P\{X > x\} \sim a(x) \text{ as } x \rightarrow \infty \Leftrightarrow f_X(x) \sim -a'(x) \text{ as } x \rightarrow \infty$$

이 성립하고, X 가 왼쪽 끝점으로 $-\infty$ 를 가지면, 적당한 식 $b(x)$ 가 존재하여

$$P\{X < x\} \sim b(x) \text{ as } x \rightarrow -\infty \Leftrightarrow f_X(x) \sim b'(x) \text{ as } x \rightarrow -\infty$$

이 성립한다.

증명: 로피탈 정리에 따라

$$\lim_{x \rightarrow \infty} \frac{P\{X > x\}}{a(x)} = \lim_{x \rightarrow \infty} \frac{-f(x)}{a'(x)}$$

이고

$$\lim_{x \rightarrow -\infty} \frac{P\{X < x\}}{b(x)} = \lim_{x \rightarrow -\infty} \frac{f(x)}{b'(x)}$$

이므로, Lemma의 주장이 성립한다. □

3. 예

본 절에서는 앞 절에 소개된 Lemma 2.1의 방법을 사용하여 명시적인 분포함수의 표현식이 알려져 있지 않은 몇몇 확률분포들에 대하여 꼬리 확률의 점근 표현식을 구해보고 해당 확률분포들의 꼬리가 어느 정도 두꺼운 지를 살펴보기로 한다. X 를 하나의 절대 연속 확률변수라고 하고 X 의 확률밀도함수를 f_X 로 나타내기로 하자.

3.1. 베타2분포

적당한 상수 $\alpha > 0$ 와 $\beta > 0$ 에 대하여, f_X 가

$$f_X(x) = \frac{1}{B(\alpha, \beta)} \cdot \frac{x^{\alpha-1}}{(1+x)^{\alpha+\beta}}, \quad x > 0$$

으로 표현되면, 확률변수 X 는 베타2분포(beta 2 distribution)를 따른다고 말하고, 이를 간단히 기호 $X \sim \text{Beta2}(\alpha, \beta)$ 로 나타낸다. 여기서, $B(\alpha, \beta)$ 는 베타함수(beta function)로서

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt$$

로 정의된다. 여기서 베타2분포는 통상의 베타분포(beta distribution)로 부터 유래된 것인데, $Y \sim \text{Beta}(\alpha, \beta)$ 이면 $Y/(1 - Y) \sim \text{Beta}2(\alpha, \beta)$ 가 성립함을 쉽게 보일 수가 있다. 또한, $x \rightarrow \infty$ 일 때는 f_X 가

$$f_X(x) \sim \frac{1}{B(\alpha, \beta)} x^{-\beta-1}$$

을 만족하므로 Lemma 2.1에 의하여

$$P\{X > x\} \sim \frac{1}{\beta B(\alpha, \beta)} x^{-\beta}$$

가 성립함을 알 수 있다. 따라서, 식 (2.2)로 부터 X 의 확률분포, 즉 $\text{Beta}2(\alpha, \beta)$ 분포는 지수가 $-\beta$ 이며 정규적으로 변하는 오른쪽 꼬리를 갖는다고 말할 수 있고, 이는 다시 β 보다 작은 양의 차수에 대해서만 X 의 적률이 유한한 값을 갖게 됨을 의미한다. 실제로,

$$E(X^k) = \begin{cases} \frac{\Gamma(\alpha + k)\Gamma(\beta - k)}{\Gamma(\alpha)\Gamma(\beta)}, & \text{if } 0 < k < \beta, \\ \infty, & \text{if } k \geq \beta \end{cases}$$

가 성립함을 보일 수가 있다. 여기서, $\Gamma(\alpha)$ 는 감마함수(gamma function)로서

$$\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt$$

로 정의된다.

3.2. 코쉬분포

적당한 상수 $\mu \in \mathbb{R}$ 와 $\sigma > 0$ 에 대하여, f_X 가

$$f_X(x) = \frac{1}{\pi\sigma \left[1 + \left(\frac{x - \mu}{\sigma}\right)^2\right]}, \quad x \in \mathbb{R}$$

로 표현되면, 확률변수 X 는 코쉬분포(Cauchy distribution)를 따른다고 말하고, 이를 간단히 기호 $X \sim \text{Cauchy}(\mu, \sigma)$ 로 나타낸다. 여기서, $x \rightarrow \infty$ 일 때, f_X 가

$$f_X(x) \sim \frac{\sigma}{\pi} x^{-2}$$

을 만족하므로 Lemma 2.1에 의하여

$$P\{X > x\} \sim \frac{\sigma}{\pi} x^{-1}$$

가 성립함을 알 수 있다. 따라서, 식 (2.2)로 부터 X 의 확률분포, 즉 $\text{Cauchy}(\mu, \sigma)$ 분포는 지수가 -1 이며 정규적으로 변하는 오른쪽 꼬리를 갖는다고 말할 수 있고, 이는 다시 1보다 작은 양의 차수에 대해서만 X 의 적률이 유한한 값을 갖게 됨을 의미한다. 코쉬분포는 대칭적으로 생겼으므로 왼쪽 꼬리 확률은 오른쪽 꼬리 확률과 같다. 실제로, 코쉬분포는 평균이 정의되지 않고 분산이 ∞ 의 값을 갖는, 즉 양쪽 꼬리가 매우 두꺼운 분포로 널리 알려져 있다.

3.3. 로그감마분포

적당한 상수 $\alpha > 0$ 와 $\beta > 0$ 에 대하여, f_X 가

$$f_X(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{-\frac{1}{\beta}-1} (\log x)^{\alpha-1}, \quad x > 1$$

로 표현되면, 확률변수 X 는 로그감마분포(log-gamma distribution)를 따른다고 말하고, 이를 간단히 기호 $X \sim \text{Log}\Gamma(\alpha, \beta)$ 로 나타낸다. 로그감마분포는 통상의 감마분포로부터 유래된 것으로서 $Y \sim \Gamma(\alpha, \beta)$ 이면 $\exp(Y) \sim \text{Log}\Gamma(\alpha, \beta)$ 가 성립한다. Lemma 2.1을 적용하면, $x \rightarrow \infty$ 일 때,

$$P\{X > x\} \sim \frac{1}{\Gamma(\alpha)\beta^{\alpha-1}} x^{-\frac{1}{\beta}} (\log x)^{\alpha-1}$$

이 성립함을 알 수 있다. 따라서, 식 (2.2)로부터 X 의 확률분포, 즉 $\text{Log}\Gamma(\alpha, \beta)$ 분포는 지수가 $-1/\beta$ 이며 정규적으로 변하는 오른쪽 꼬리를 갖는다고 말할 수 있고, 이는 다시 $1/\beta$ 보다 작은 양의 차수에 대해서만 X 의 적률이 유한한 값을 갖게 됨을 의미한다. 실제로,

$$E(X^k) = \begin{cases} (1 - \beta k)^{-\alpha}, & \text{if } 0 < k < \frac{1}{\beta}, \\ \infty, & \text{if } k \geq \frac{1}{\beta} \end{cases}$$

이 성립함을 보일 수가 있다.

3.4. 로그정규분포

적당한 상수 $\mu \in \mathbb{R}$ 와 $\sigma > 0$ 에 대하여, f_X 가

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi x}} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right], \quad x > 0$$

으로 표현되면, 확률변수 X 는 로그정규분포(log-normal distribution)를 따른다고 말하고, 이를 간단히 기호 $X \sim \text{LogN}(\mu, \sigma^2)$ 으로 나타낸다. 로그정규분포 역시 통상의 정규분포로부터 유래된 것으로서 $Y \sim N(\mu, \sigma^2)$ 이면 $\exp(Y) \sim \text{LogN}(\mu, \sigma^2)$ 이 성립한다. Lemma 2.1을 적용하면, $x \rightarrow \infty$ 일 때,

$$P\{X > x\} \sim \frac{\sigma}{\sqrt{2\pi} \log x} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right]$$

이 성립함을 알 수 있다. 여기서,

$$\lim_{x \rightarrow \infty} \frac{P\{X > tx\}}{P\{X > x\}} = \begin{cases} \infty, & \text{if } 0 < t < 1, \\ 1, & \text{if } t = 1, \\ 0, & \text{if } t > 1 \end{cases}$$

과 같이 계산되고, 따라서 식 (2.2)는 어떠한 상수 $\alpha \geq 0$ 에 대해서도 성립하지 않음을 알 수 있다. 실제로 로그정규분포는 오른쪽 꼬리가 정규적으로 변하는 정도로 두껍지는 않지만 식 (2.1)이 만족되어 열지수분포 중의 하나임이 알려져 있는데, 따라서 X 의 적률생성함수는 0 근방에서 정의되지 못한다. 반면에, X 의 모든 양의 차수의 적률이 유한한 값을 갖는데, 정규분포의 적률생성함수를 이용하면

$$E(X^k) = \exp\left(k\mu + \frac{k^2\sigma^2}{2}\right), \quad k > 0$$

이 성립함을 쉽게 보일 수 있다.

3.5. 로그파레토분포

적당한 상수 $\alpha > 0$ 와 $\beta > 0$ 에 대하여, f_X 가

$$f_X(x) = \alpha\beta^\alpha x^{-1}(\log x)^{-\alpha-1}, \quad x > e^\beta$$

로 표현되면, 확률변수 X 는 로그파레토분포(log-Pareto distribution)를 따른다고 말하고, 이를 간단히 기호 $X \sim \text{LogPar}(\alpha, \beta)$ 로 나타낸다. 로그파레토분포 역시 통상의 파레토분포로부터 유래된 것으로서 $Y \sim \text{Par}(\alpha, \beta)$ 이면 $\exp(Y) \sim \text{LogPar}(\alpha, \beta)$ 가 성립한다. 로그파레토분포의 분포함수는 잘 알려져 있는데, 이로 부터

$$P\{X > x\} = \beta^\alpha (\log x)^{-\alpha}, \quad x > e^\beta$$

가 됨을 알 수 있다. 따라서, 식 (2.2)로부터 X 의 확률분포, 즉 $\text{LogPar}(\alpha, \beta)$ 분포는 오른쪽 꼬리가 지수 0을 갖으며 정규적으로 변하는 것을 알 수 있는데, 이는 앞에서 예로 든 네 가지 분포의 그 어떤 것보다도 오른쪽 꼬리가 가장 두꺼운 분포임을 의미한다. 또한, 이 경우 X 의 모든 양의 차수의 적률은 ∞ 가 됨을 뜻한다.

References

- Alves, I. F., de Haan, L. and Neves, C. (2009). A test procedure for detecting super-heavy tails, *Journal of Statistical Planning and Inference*, **139**, 213–227.
- Asmussen, S. (2003). *Applied Probability and Queues*, Springer, Berlin.
- Cline, D. B. H. (1994). Intermediate regular and II variation, *Proceedings of the London Mathematical Society*, **68**, 594–616.
- Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, Berlin.
- Fergusson, K. and Platen, E. (2006). On the distributional characterization of daily log-returns of a world stock index, *Applied Mathematical Finance*, **13**, 19–38.
- Goldie, C. M. and Resnick, S. (1988). Distributions that are both subexponential and in the domain of attraction of an extreme-value distribution, *Advances in Applied Probability*, **20**, 706–718.
- Klüppelberg, C. (1989). Subexponential distributions and characterisations of related classes, *Probability Theory and Related Fields*, **82**, 259–269.
- Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*, Springer, New York.
- Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J. (1999). *Stochastic Processes for Insurance and Finance*, John Wiley & Sons, Chichester.
- Teugels, J. L. (1975). The class of subexponential distributions, *Annals of Probability*, **3**, 1000–1011.

두꺼운 꼬리를 갖는 연속 확률분포들의 꼬리 확률에 관하여

윤석훈^{a,1}

^a수원대학교 통계정보학과

(2013년 7월 30일 접수, 2013년 10월 8일 수정, 2013년 10월 8일 채택)

요약

본 논문에서는 두꺼운 꼬리를 갖는 확률분포들의 여러 부류에 대해서 살펴본다. 주어진 하나의 확률분포가 이들 중 어떤 부류에 속하는 지를 알려면 해당 분포의 꼬리 확률에 대한 (점근) 표현식을 알아야만 한다. 그러나 대다수의 절대 연속 확률분포들은 분포함수가 아닌 확률밀도함수로 명시되기 때문에 통상적으로 이들의 꼬리 확률에 대한 표현식을 얻는 작업은 그리 쉬운 일이 아니다. 본 논문에서는 이러한 경우 확률밀도함수만을 이용하여 꼬리 확률에 대한 점근 표현식을 쉽게 얻을 수 있는 하나의 방법을 제안한다. 또한 제안한 방법을 설명하기 위하여 몇가지 예를 첨부한다.

주요용어: 두꺼운 꼬리를 갖는 분포, 열지수분포, 긴 꼬리, 정규적으로 변하는 꼬리.

¹(445-743) 경기도 화성시 봉담읍 와우리 산2-2, 수원대학교 통계정보학과, 교수. E-mail: syun@suwon.ac.kr