

# 첨도와 뽀족함의 관계

## : Westfall의 논문과 시각적 확인을 바탕으로

목진휘, 김태오

July 3, 2022

### Abstract

Peter H. Westfall의 논문 "Kurtosis as Peakedness, 1905-2014. R.I.P."를 바탕으로 4차 중심 적률의 합수인 Kurtosis(첨도)가 어떤 분포의 뽀족함과는 관계가 없다는 저자의 주장을 정리하고, 대표적인 예시들을 재생산하며 비대칭분포에서의 예시를 보이는 것이 프로젝트의 목적이다.

## 1 Introduction

1905년 Karl Pearson은 임의의 분포에서 정상 부분의 평평한 정도를 측정하기 위한 척도로써 Kurtosis를 다음과 같이 정의하였다.

$$kurtosis = \frac{E[(X - \mu)^4]}{\sigma^4} - 3$$

이 때,  $X$ 가 정규분포를 따른다면, Kurtosis 값은 0이며, 정규분포를 기준으로 Kurtosis가 양수이면 "Leptokurtic", Kurtosis가 0이면 "Mesokurtic", Kurtosis가 음수이면 "Platykurtic"라고 지칭하였다. Kurtosis(이하 첨도)는 "볼록함" 혹은 "뽀족함"을 뜻하는 그리스어  $\kappaυρτος$  (kyrtos)에서 유래되었다. Pearson의 정의에 의하면 Leptokurtic은 정규분포보다 덜 뽀족함을, Mesokurtic은 정규분포와 같은 뽀족함을, Platykurtic은 정규분포보다 더 뽀족함을 나타낸다. 여러 교과서와 논문 또한 Pearson의 정의를 인용하고 있으며, 첨도와 뽀족함은 불가분의 관계가 되었다.

Lee(2013)에 의하면 첨도를 "뽀족함의 정도"로, Coolican(2013)는 "정점에서의 전반적인 모양"으로, Katz(2013)와 Taylor(2008)은 "정점이 정규분포보다 더 뽀족하거나 평평한 정도"로 정의하고있다. 논문 또한 비슷하게, Darlington(1970)은 뽀족함과 반대되는 개념으로 양봉형(Bimodal) 분포를 주장하고 있으며, Decarlo(1997)은 양의 첨도가 정규분포에 비해 두터운 꼬리와 뽀족함 나타남을 주장하고 있다. 이에 반해, Ali(1974)는 첨도가 오직 꼬리부분의 두터운 정도를 의미한다고 주장하며, Johnson(1980)은 첨도 값이 같음에도 그래프의 개형이 다른 여러 예시들을 제시함과 동시에 첨도가 뽀족함을 나타내는 것은 옳지 않음을 말하고 있다.

첨도와 뽀족함의 관계에 관한 이러한 혼동은 꼬리가 두터운 분포가 때때로 꼬리가 얇은 분포보다 peak(이하 정점)이 높은 것에서 기인한다. 하지만 꼬리가 두터운 분포가 항상 높은 정점을 갖진 않는다. Kaplansky(1945)는 leptokurtic한 분포와 platykurtic한 분포 모두에서 표준정규분포보다 정점이 높거나 낮을 수 있음을 보이고 있다. 본 레포트에서는 특히 대칭 분포에서, 분포의 정점이 있을 것으로 예상되는 중앙 부분의 값이 첨도에 미치는 영향이 적음을 수식과 함께 보이고, 여러 잘 알려진 분포의 경우를 표로 정리하여 나타낼 것이다. 또한, 첨도와 뽀족함이 관계가 없는 반례들을 그래프와 함께 표현 할 예정이다. 또한 첨도를  $E[\frac{(X-\mu)^4}{\sigma^4}]$ 로 정의하여 계산할 것이다.

## 2 Relationship between central part of distribution and kurtosis

### 2.1 Illustrating Example

꼬리가 두터운 분포를 가정하기 위해,  $Cauchy(0, 1)$ 을 따르는 1000개의 random sample을 생각해보자. Figure 1을 보았을 때, 대부분의 값들은 1 표준편차 안에 모여있지만,  $-834, -637, -185, -99$  와 같은 특이값 또한 존재하고 있다. 그림1의 데이터의 표본평균은  $m = -1.55$ , 표본표준편차는  $s = 34.85$  이다. 이 때,  $z_i = \frac{x_i - m}{s}$ 라고 하자. 그렇다면, 해당 분포의 표본 첨도는 다음과 같다.

$$\kappa = \frac{1}{1000} \sum z_i^4 = \frac{1}{1000} \sum_{|z_i| \leq 1} z_i^4 + \frac{1}{1000} \sum_{|z_i| > 1} z_i^4 = 0.0073 + 437.3798 = 437.3871$$

수식에서 볼 수 있듯이, 1 표준편차 이내에 있는 데이터들이 첨도에 미치는 영향은  $\frac{0.0073}{437.3871} = 0.000017$ 로, 단 0.0017%에 지나지 않음을 알 수 있다. 따라서 위의 예시를 통해, 첨도는 정점이 있을 것으로 예상되는 평균에서 1 표준편차 이내에 위치한 값에 영향을 받는 것이 아닌, 특이점에 매우 큰 영향을 받는다는 것을 알 수 있다. 이러한 점에 있어서, 오히려 첨도가 특이점을 진단하는 용도로 사용될 수도 있음을 Livesay(2007)은 주장한 바있다.

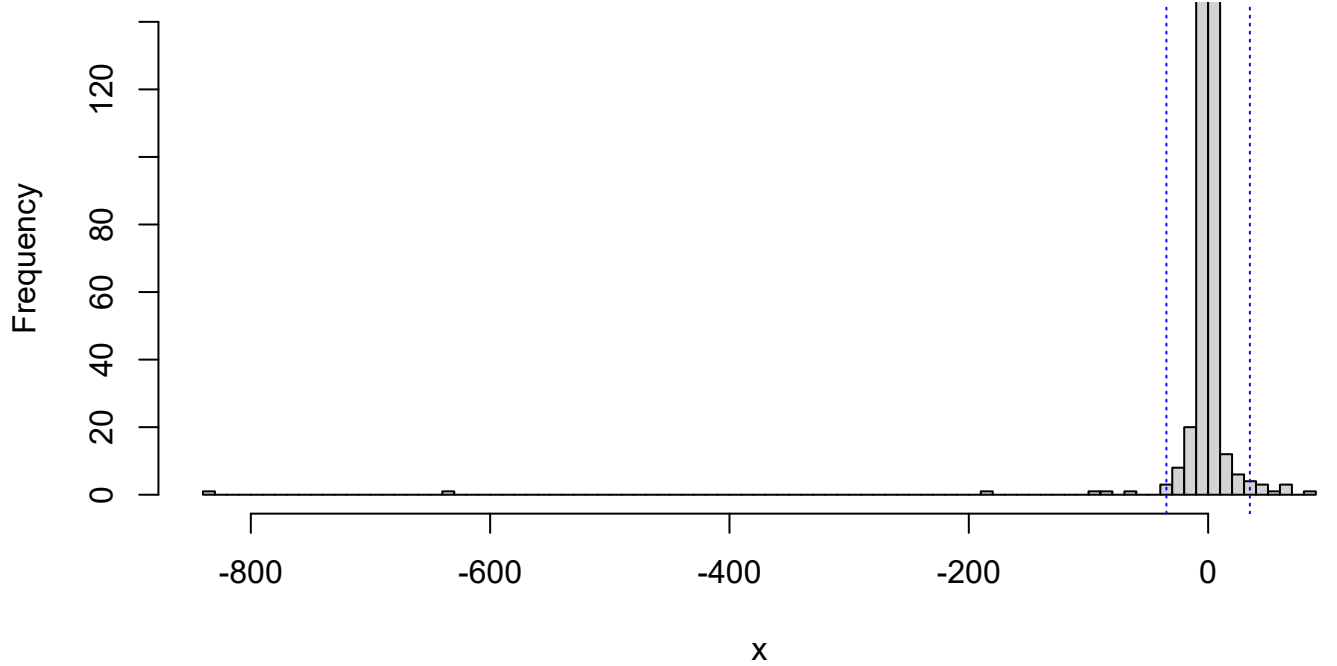


Figure 1: Histogram of a random sample of 1000 Cuachy random numbers. Dotted lines show mean  $\pm$  one standard deviation of the distribution

## 2.2 Theorem about Kurtosis and its proof

**Theorem 1** Define  $Z_X = \frac{X - \mu_X}{\sigma_X}$  for general random variable  $X$  and let  $\kappa(X) = E(Z_X^4)$  denote the kurtosis of  $X$ . Then for any distribution, the kurtosis for some interval  $[-h, h]$  is less or equal than  $h^4$ , i.e.  $0 \leq E[Z_X^4 \mid I(|Z_X| < h)] \leq h^4$ .

*pf)* Define central portion of distribution  $Z$  as  $A_h = \{z : |z| \leq h\}$ . Define the quantities  $Center_h = \int_{A_h} z^4 f(z) dz$  and  $Tail_h = \int_{A_h^c} z^4 f(z) dz$ . Then  $\kappa = Center_h + Tail_h$ . Since  $-h < z < h$ ,

$$Center_h = \int_{A_h} z^4 f(z) dz = \int_{-h}^h z^4 f(z) dz \leq \int_{-h}^h h^4 f(z) dz \leq \int_{-\infty}^{\infty} h^4 f(z) dz = h^4$$

Therefore  $0 \leq E[Z_X^4 \mid I(|Z_X| < h)] \leq h^4$ . When we think about  $h = 1$  case,  $0 \leq E[Z_X^4 \mid I(|Z_X| < h)] \leq 1$ .

**Theorem 2** Consider a sequence of random variables  $X_i$ ,  $i = 1, 2, \dots$  for which  $\kappa(X_i) \rightarrow \infty$ . Then  $\frac{E[Z_i^4 I(|Z_i| < 1)]}{\kappa(X_i)} \rightarrow 0$ .

*pf)* By theorem 1,  $0 \leq Center_h \leq h^4$ . Then

$$\frac{Center_1}{\kappa} \leq \frac{1^4}{\kappa} \rightarrow 0 \quad \text{as } \kappa \rightarrow \infty$$

**Corollary** Assume that  $X$  is continuous and that the density of  $Z_X^2$  is decreasing on  $[0, 1]$ . Then  $0 \leq E[Z_X^4 \mid I(|Z_X| < 1)] \leq 0.5$ . (Sharper bound than Theorem 1 when  $h = 1$  case)

*pf)* Let  $W = Z^2$  which is continuous and decreasing on  $[0, 1]$ . Also let  $Y = W^2$ . Then  $z \in (-\infty, \infty)$ ,  $w \in (0, \infty)$ ,  $y \in (0, \infty)$

$$F_Y(y) = P[Y \leq y] = P[W^2 \leq y] = P[0 \leq W \leq \sqrt{y}] = F_W(\sqrt{y})$$

Then,  $f_Y(y) = \frac{1}{2\sqrt{y}} f_W(\sqrt{y})$ ,  $y \in (0, \infty)$

$$E[W^2 \mid I(0 < W < 1)] = E[Y \mid I(0 < Y < 1)] = \int_0^1 y \frac{1}{2\sqrt{y}} f_W(\sqrt{y}) dy = \frac{1}{2} \int_0^1 \sqrt{y} f_W(\sqrt{y}) dy \leq \frac{1}{2} \int_0^1 \sqrt{y} f_W(y) dy$$

(Note that,  $f_W(w)$  is decreasing on  $[0, 1]$ ,  $0 < w < 1$ ,  $0 < f_W(w^2) < 1$ )

$$E[W^2 \mid I(0 < W < 1)] \leq \frac{1}{2} \int_0^1 \sqrt{y} f_W(y) dy = \frac{1}{2} \int_0^1 w f_W(w^2) dy = \frac{1}{2} w f_W(w^2) \leq \frac{1}{2}$$

$$\therefore E[Z^4 \mid I(-1 < Z_X < 1)] \leq 0.5$$

Theorem1, Theorem2, Corollary의 내용을 정리하면 다음과 같다. 임의의 분포 X에서 값을 표준화시킨 분포 Z를 생각하자. 이 때  $h = 1$ 인 경우, Theorem 1에 의해 1 표준편차 안의 값이 첨도에 기여하는 바는 절대적으로 1 보다 작다는 것을 알 수 있다. 가령, 첨도가 3인 표준정규분포를 가정할 경우, 1 표준편차안에 있는 값으로만 첨도를 구하면 기껏해야 1을 넘지 못한다는 것이다. 더 나아가, Theorem 2에 의하면 첨도가 무한히 커지는 상황을 가정했을 때, 1 표준편차 이내의 값이 첨도에 기여하는 바는 0에 수렴한다. 이는 첨도가 크면 그래프의 개형이 더욱 뾰족해질 것이라는 그동안의 예상과는 달리, 첨도가 클 때, 정점이 있을 것으로 예상되는 평균에서 1 표준편차 이내에 위치한 값은 첨도에 아무런 영향을 끼치지 못하며 첨도는 꼬리부분에 큰 영향을 받는다는 것을 의미한다. 또한 Collorary에 의해 특정 조건을 만족하는 분포는 Theorem1에서 정의한 Upper bound가 더 줄어 들 수 있음을 의미한다.

### 2.3 Table of central part to the kurtosis

아래의 표는 random sampling을 이용하여 흔히 잘 알려진 분포들의 첨도와, 중앙 부분이 첨도 계산에 기여하는 비율을 나타낸 것이다. 여섯 가지 분포 모두  $Center_1$ 의 값이 2.2 Theorem 1과 Corollary의 Upper bound 보다 훨씬 작은 값을 보이고 있다. 또한, 모분포의 평균이 정의되지 않으며, 이에 따라 첨도도 정의되지 않는 코시분포의 표본은 매우 큰 첨도 값을 보이고 있고,  $Center_1/\kappa$ 은 0에 수렴하는 것을 알 수 있다. 이러한 점에 있어서 Theorem 2의 내용도 확인해 볼 수 있다.

Distribution	$\kappa$	$Center_1$	$Center_1/\kappa$
Unif(0,1)	1.800	0.114	0.063
Normal(0,1)	2.992	0.112	0.037
Exp(1)	9.117	0.202	0.022
Cauchy(0,1)	186868	0.0003	2.01e-09
Poisson(5)	3.195	0.168	0.052
Bimodal (combination of Beta(2,2) dist)	1.586	0.159	0.100

Table 1: Proportion of the central part to the  $\kappa$  in  $(\mu - \sigma, \mu + \sigma)$  with sample version.  $10^6$  random sample with seed 412

## 3 Graphical Examples of "Kurtosis is not related to Peakedness"

### 3.1 Examples of Kaplansky(1945)

Kaplansky는 일찍이 첨도와 뾰족함이 관련 없음을 보인 바있다. 3.1절에서는 Kaplansky의 예시들을 그림으로 구현해보고자 한다. 다음과 같은 네 분포를 생각해 볼 수 있다.  $x \in (-\infty, \infty)$

$$P(x) = \frac{1}{3\sqrt{\pi}}(9/4 + x^4)e^{-x^2}, \quad S(x) = \frac{3\sqrt{3}}{16\sqrt{\pi}}(2 + x^2)e^{-3x^2/4}$$

$$Q(x) = \frac{3}{2\sqrt{2\pi}}e^{-x^2/2} - \frac{1}{6\sqrt{\pi}}(9/4 + x^4)e^{-x^2}, \quad R(x) = \frac{1}{6\sqrt{\pi}}(e^{-x^2/4} + 4e^{-x^2})$$

Table 2에 의하면,  $P(x)$  와  $S(x)$ 의 표본평균과 표본 표준편차는 정규분포와 거의 차이가 나지 않는다. 하지만, 표본 첨도는 둘 다 정규분포 보다 작은 것을 알 수 있다. 흔히 알려져있는 첨도의 개념으로 이를 이해하면,  $P(x)$  와  $S(x)$  모두 정규분포보다 덜 뾰족한 형태여야한다. 하지만 Figure2에서 볼 수 있듯이,  $P(x)$ 의 첨도는 정규분포의 첨도보다 낮음에도 불구하고 더 뾰족한 형태임을 알 수 있다. 이는  $Q(x)$ 와  $R(x)$ 의 경우에서도 마찬가지이다.  $Q(x)$ 와  $R(x)$ 의 표본평균과 표본 표준편차는 정규분포와 차이가 나지 않으며, 표본 첨도는 둘 모두 정규분포보다 높다. 그동안 알려져 온 첨도의 관점에 의하면 두 그래프 모두 정규분포보다 더 뾰족한 형태여야 한다. 그러나,  $Q(x)$ 의 첨도는 정규분포의 첨도보다 높음에도 불구하고 덜 뾰족한 형태를 나타내고 있다. 위의 두 예시들 모두, 특히 일반적으로 생각할 수 있는 0을 기준으로 대칭인 분포임에도 불구하고, 첨도는 뾰족함을 나타내는 척도라는 정의에 대한 반례이다.

Distribution	$\mu$	$\sigma$	$\kappa$
P(x)	-0.009	0.999	2.746
S(x)	0.002	0.998	2.676
N(0,1)	0	1	3
Q(x)	0	1	3.129
R(x)	0	0.997	4.417

Table 2: Sample version of mean, sd and kurtosis from  $P(x)$ ,  $S(x)$ ,  $Q(x)$ ,  $R(x)$  with  $10^6$  random samples.

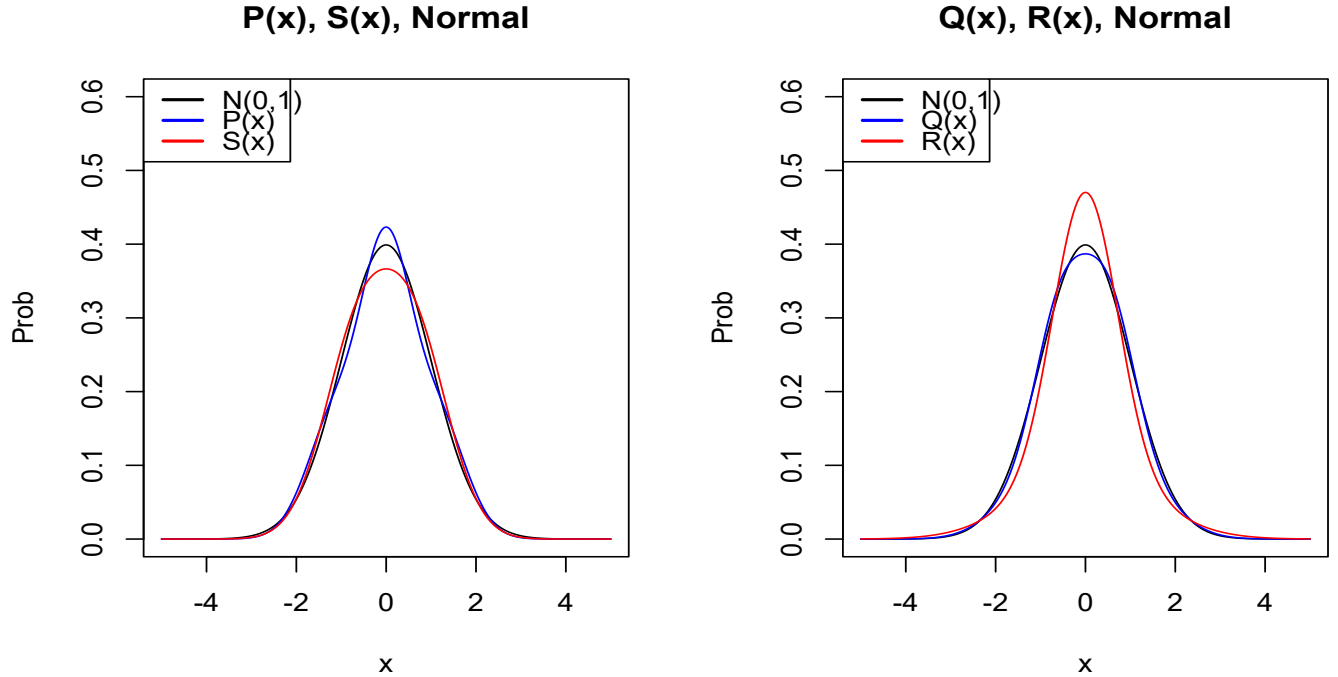


Figure 2: Graph of  $P(x)$ ,  $S(x)$  compared to Normal(0,1), and  $Q(x)$ ,  $R(x)$  compared to Normal(0,1)

### 3.2 Examples of Westfall(2014)

Westfall은 Mixed Distribution을 이용하여 첨도와 뽀족함이 관계 없음을 보였다. Westfall은 "Triangular", "Devil's Tower", "Slip-dress"라는 임의의 세 분포에  $t(4 + 0.000001)$ 을 0.0001의 가중치를 부여해 혼합분포를 만들었고, 혼합 이전과 혼합 후 분포의 개형과 첨도를 비교하였다. 혼합 전과 혼합 후 모두 그래프의 개형은 차이하지 않았지만, 첨도는 상당한 차이가 있음을 확인할 수 있었다. 특히, 혼합분포의 첨도는 6000을 상회하여 매우 높게 나타나고 있다. 이는 첨도가 뽀족함을 나타내는 척도라는 정의에 대한 또다른 반례라고 할 수 있다. 3.2 절에서는 Westfall의 예시를 조금 변형하여 살펴볼 것이다. "Triangular", "Slip-dress", "Bimodal" 분포를  $t(1)$ 에 여러 가중치를 부여하여 혼합한 분포의 개형을 혼합 전 분포의 개형과 비교하고, 첨도를 비교해 볼 것이다.

Mixing Prob		Triangular	Mixed Triangular	Slip-Dress	Mixed Slip-Dress	Bimodal	Mixed Bimodal
0.001	$\kappa$	2.39	27261	2.40	73377	1.5	742185
0.0001	$\kappa$	2.39	578808	2.39	183320	1.58	753658
0.00005	$\kappa$	2.40	1684	2.40	604905	1.49	22451

Table 3: Kurtosis of Mixed with  $t(1)$  with seed (14, 4, 95) for Triangular, (3,123,12) for Slip-dress, (4,32,30) for Bimodal, for each mixing probability

Table 3에서 볼 수 있듯이,  $t(1)$ 를 여러 가중치로 혼합한 결과에 따르면, 혼합분포의 첨도는 급격하게 커지게 된다. 하지만, Fig3, Fig4, Fig5에서 볼 수 있듯이 원래 분포와 혼합분포의 개형은 크게 차이가 나지 않는다. 또한, 세가지 혼합 분포의 첨도가 Mixing Probability에 따라 일괄적으로 증가하거나 감소하지는 않는다는 점에서, 일정 수준 이상의 확률이면 혼합 분포의 첨도에 Mixing Probability가 큰 영향을 끼치는 것은 아니라고 생각할 수 있다. 결론적으로, Kaplansky와 접근 방식은 다르지만, 첨도가 매우 다름에도 불구하고 그래프의 개형이 차이가 나지 않는다는 점에서 첨도는 뽀족함과 연관이 없다. 더 나아가, 꼬리가 매우 두꺼운  $t(1)$  (= 코시 분포)와 혼합하였기 때문에 꼬리 부분에 있는 특이점이 추출되어 첨도에 큰 영향을 끼친 것으로 파악할 수 있다. 한편, Westfall의 경우 혼합에 사용하는 분포로  $t(4 + 0.000001)$ 을 사용하였지만, 이 분포의 경우 꼬리 부분이  $t(1)$  분포에 비해 두껍지 않기 때문에, 혼합 하였을 때 급격히 첨도가 증가하는 경우를 찾기 힘들었다.

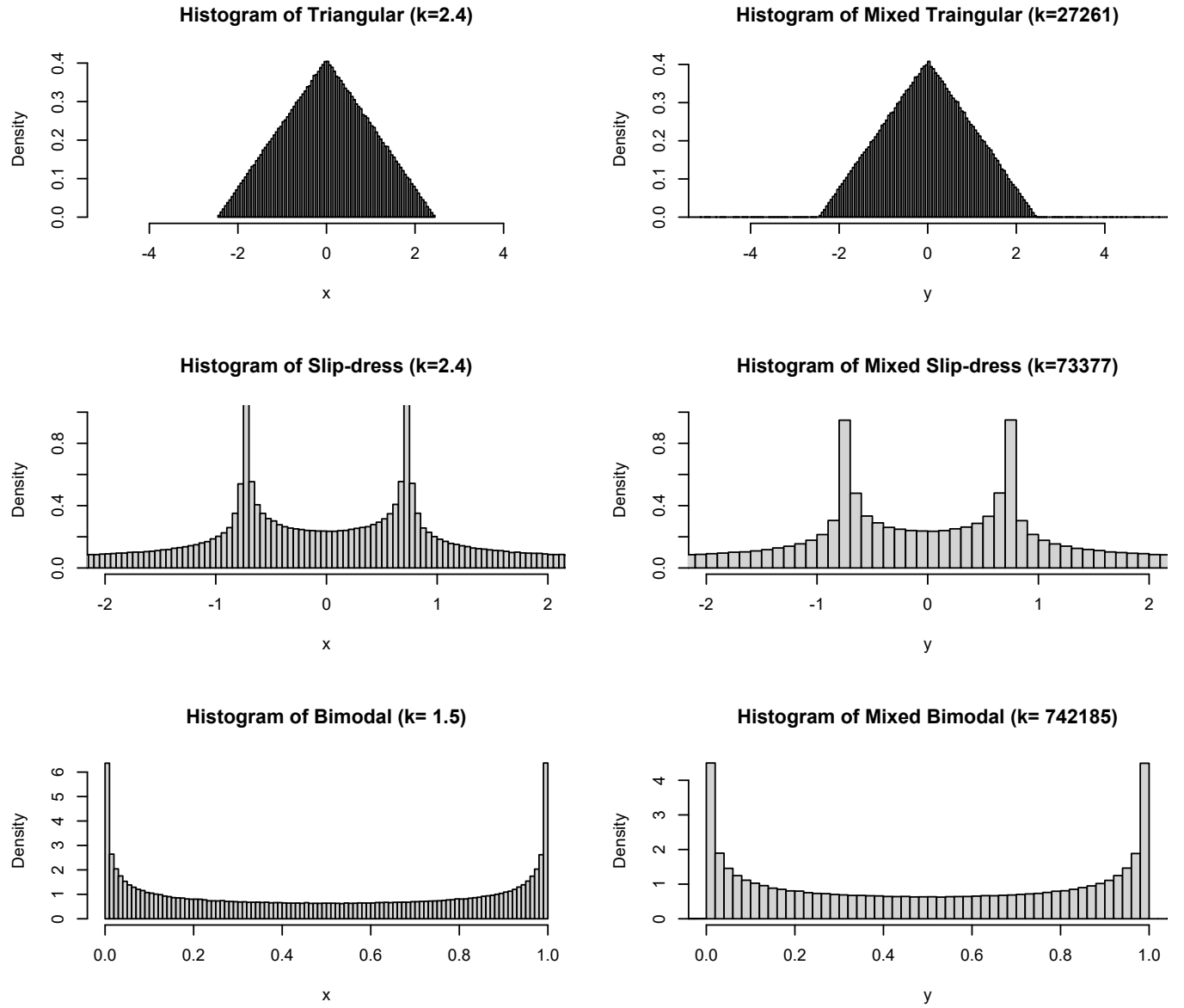


Figure 3: Histogram of Triangular, Slip-dress, Bimodal with mixing probability 0.001

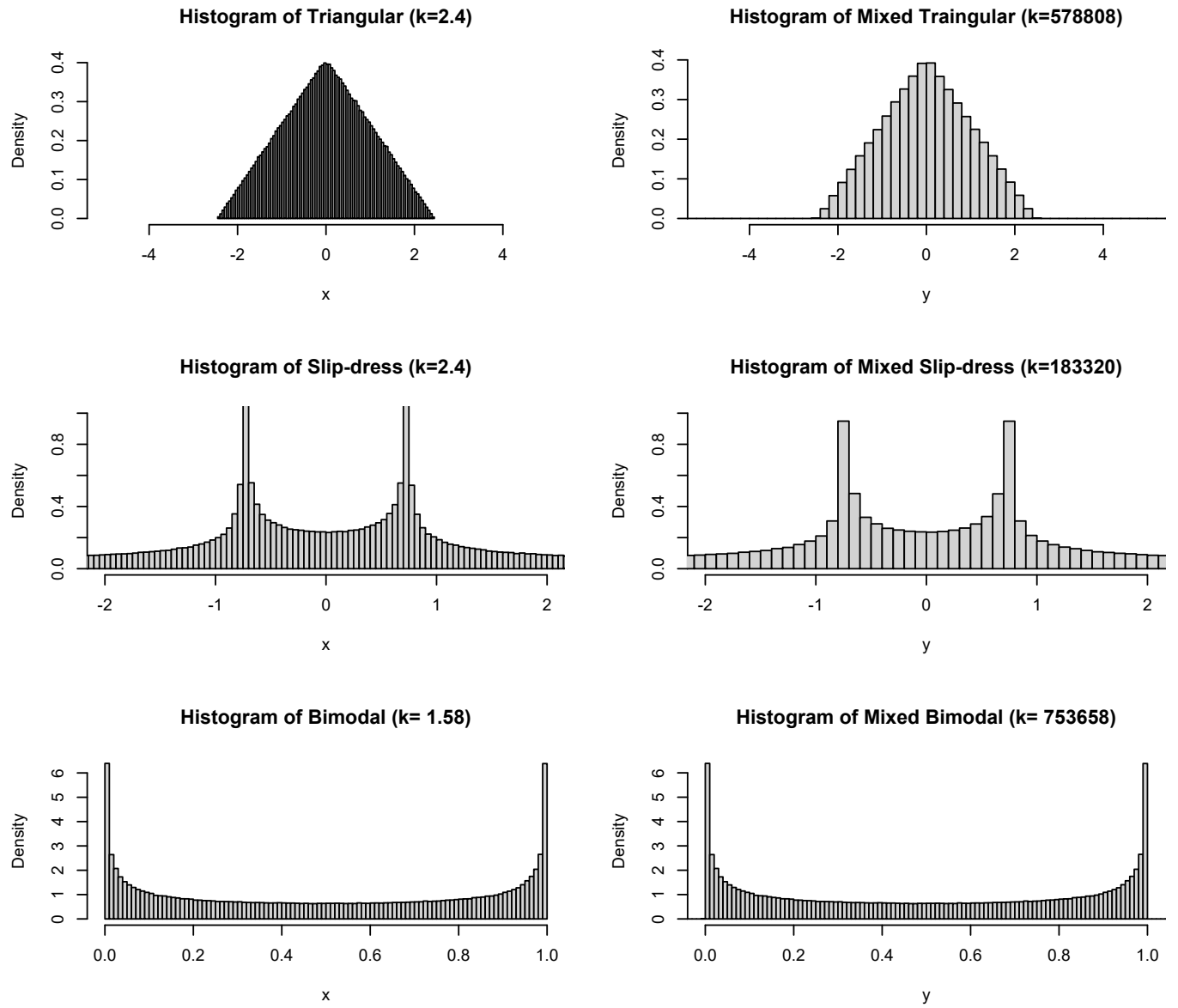


Figure 4: Histogram of Triangular, Slip-dress, Bimodal with mixing probability 0.0001



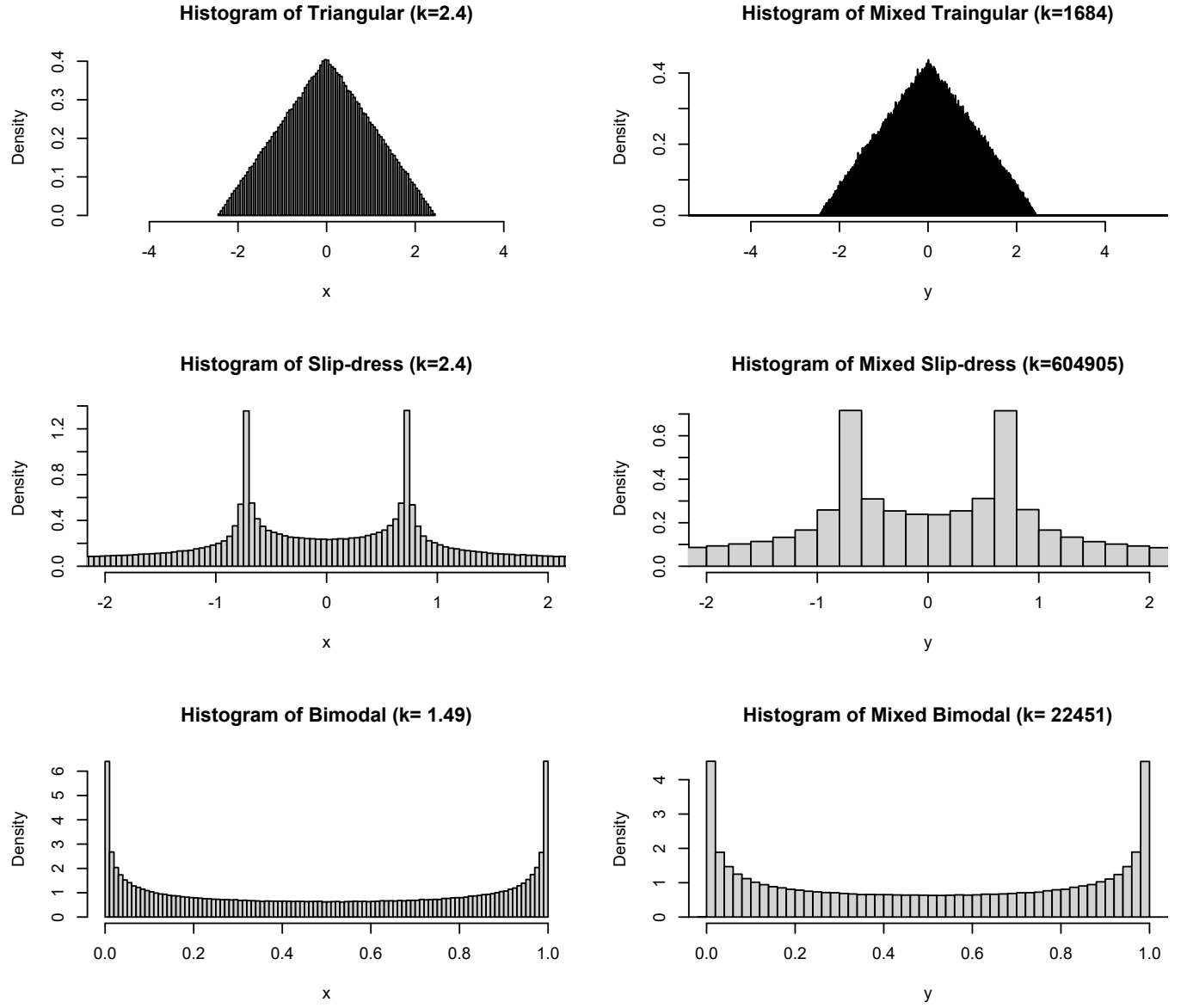


Figure 5: Histogram of Triangular, Slip-dress, Bimodal with mixing probability 0.0005

### 3.3 Think about Asymmetric case

기존의 논문들에서 대표적으로 나타나는 예시는 symmetric한 분포에 대해 주로 다루고 있기에, 본 레포트에서는 그 이외의 경우인 asymmetric한 분포에 대한 탐구가 이루어졌다. 이는 가장 잘 알려진 비대칭 분포인 감마분포를 활용하여 주제에 걸맞게 첨도값과 중앙부분의 뾰족한 정도를 비교하며 반례를 찾고자 하였음을 밝힌다.

Distribution	Normal(0,1)	Gamma(8, 0.5)	Gamma(10, 0.5)
$\kappa$	3	3.75	3.6

Table 4: Kurtosis of Normal and Gamma Distribtuion with scale parameter

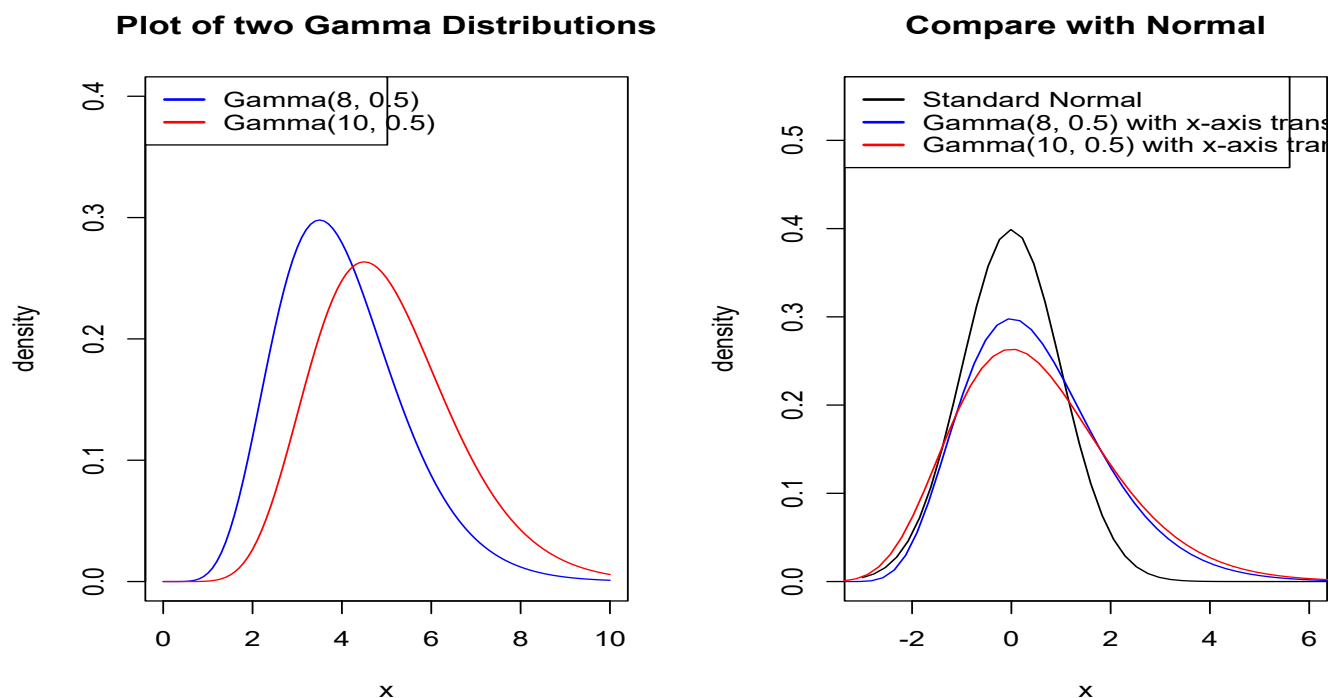


Figure 6: Graph of Gamma(8, 0.5), Gamma(10,0.5) with x-axis translation compared to Normal

Table 4에 의하면, 두 감마분포 모두 첨도가 표준정규분포 보다 크며,  $Gam(8, 0.5)$ 의 첨도가  $Gam(10, 0.5)$  보다 크다. 또한,  $Gam(\alpha, \beta)$ 의 첨도는  $6/\alpha + 3$ 으로 알려져있기 때문에, 어떠한 감마분포든 항상 정규분포보다 첨도가 크다는 것을 알 수 있다. 또한, Figure 6의 첫번째 그래프를 확인해보았을 때, 첨도가 더 큰  $Gam(8, 0.05)$ 가  $Gam(10, 0.05)$  보다 더 뾰족한 것을 알 수 있다. 하지만, 더 정밀한 비교를 위해 두 그래프를 정규분포와 같이 위치시켜보았다. 이 때, 직관적인 이해를 위해, 두 감마분포를 각각  $x$ 축에서  $-3.5, -4.5$  만큼 평행이동하였다. 두 감마분포의 첨도는 3보다 크기 때문에, 표준정규분포 보다 더 뾰족해야하지만, 오히려 세 분포 중 가장 뾰족한 것은 표준정규분포이다. 따라서 이 또한 첨도가 뾰족함의 척도라는 정의에 반례가 된다. 감마분포 뿐만 아니라 이를 이용한 F분포, 혹은 또다른 비대칭 분포를 활용하여 여러 다양한 반례를 찾을 수 있을 것이라 예상한다.

## 4 Conclusion

본 레포트에서는 4차적률값인 첨도가 분포 중심부분의 뾰족함을 나타내는 척도로 해석되어 온 것이 잘못됨과 더불어 실제로 중심부분에 대한 설명력이 결코 충분하지 않음을 발표된 논문들을 재현 및 가공하여 확인하였다. 특히 Westfall의 논문에 나와 있는 기존의 “첨도는 중심부분의 뾰족함을 나타내는 척도이다”에 대한 반례로, 대칭분포의 경우만 나와 있었기 때문에, 비대칭 분포에서의 예시에 대한 탐구와 더불어 혼합분포의 가중치와 분포의 자유도를 달리하여 다양한 반례들을 모색하였음을 밝힌다. 더하여 분포 중앙의 정점에 대한 설명력이 부족한 Kurtosis는 Peakedness와 연관이 없음을 확인하며 보고서를 마친다.