New Measures of Skewness of a Probability Distribution

Ashok K. Singh[1], Laxmi P. Gewali[2], Jiwan Khatiwada[2]

[1]RGG Department, Harrah College of Hospitality, University of Nevada, Las Vegas, USA
[2]Department of Computer Science, University of Nevada, Las Vegas, USA
Email: Ashok.singh@unlv.edu, Laxmi.gewali@unlv.edu, Jiwan.khatiwada@unlv.edu

Abstract

Symmetry of the underlying probability density plays an important role in statistical inference, since the sampling distribution of the sample mean for a given sample size is more likely to be approximately normal for a symmetric distribution than for an asymmetric one. In this article, two new measures of skewness are proposed and the confidence intervals for true skewness are obtained via Monte Carlo simulation experiments. One advantage of the two proposed skewness measures over the standard measures of skewness is that the proposed measures of skewness take values inside the range $(-1, +1)$.

Keywords

Sample Moments, Quantiles, Computational Geometry, Symmetry, Robust Measure, Central Limit Theorem, Trapezoid Rule

1. Introduction

Many of the common statistical inference methods rely on the approximate normality of the sample mean via the Central Limit Theorem (CLT) for sufficiently large number of samples (n). A rule of thumb says that the CLT can be used for $n > 30$ [1] [2] [3] [4]. Singh, Lucas, Dalpatadu, & Murphy [5] showed that this rule of thumb may be inaccurate for highly skewed distributions. Veluchamy [6] developed a graphical approach based on bootstrap for verification of normality of the sample mean.

Skewness plays an important role in statistical analyses in almost all disciplines, and especially in finance. Johnson, Sen and Balyeat [7] applied a skewness adjusted binomial model to futures options pricing and derived the asymptotic skewness model properties. Their results showed that the futures options price, in the presence of skewness, depends not only on mean and standard deviation

(sd), but other parameters as well. Kun [8] investigated daily time series of four Shanghai Stock market indices and found inclusion of skewness in models to yield higher investor utility. Chateau [9] investigated the effects of skewness and kurtosis by starting with the Black's normal model for the European put values, replacing the Gaussian distribution by the Gram-Charlier and the Johnson distribution, and showed that both skewness and kurtosis have significant impact on the model results. The effects of skewness on stochastic frontier models are discussed in [10].

Several measures of skewness are available in statistical literature [11], but most of these are based on the sample moments or quantiles, and as such are adversely affected by the presence of a few outliers. Robust skewness measures such as medcouple have been proposed and investigated in the literature [12]; the medcouple measures of skewness are a function of sample quantiles and order statistics. A comparison of skewness and kurtosis measures is provided by [13]; a comparison of the standard t-test and a modified t-test for skewed distributions is available in [14].

Skewness of a probability distribution refers to the departure of the distribution from symmetry. A symmetric distribution has no skewness, a distribution with longer tail on the left is negatively skewed, and a distribution with longer tail on the right is positively skewed [15].

There are mainly three types of skewness measures available in the literature: Fisher-Pearson skewness, adjusted Fisher-Pearson skewness, and Pearson Type 2 skewness. Fisher-Pearson skewness measures are functions of the second and third central sample moments:

$$
m_k = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^k}{n-1}, \ k = 2,3
$$
$\bar{x}$ = sample mean, and

$\sqrt{m_2}$ = sample standard deviation. $\qquad$ (1)

The formulas for calculating Fisher-Pearson sample skewness used by popular statistical software packages [16] are shown below; the statistical software environment R [17] can be used to compute all of the three types.

Fisher-Pearson Skewness (Type 1):

$$
g_1 = \frac{m_3}{m_2^{3/2}}
$$
$\qquad$ (2)

Adjusted Fisher-Pearson Skewness (Type 2):

$$
G_1 = \frac{\sqrt{n(n-1)}}{(n-2)} g_1
$$
$\qquad$ (3)

Pearson Type 2 skewness is a simple measure that is calculated from the sample mean, standard deviation, and the sample median *m*:

$$
Sk_2 = 3\frac{(\bar{x} - m)}{s}
$$
$\qquad$ (4)

Hotelling and Solomon [18] have shown that $-3 \leq Sk_2 \leq 3$; a close look at the

proof shows that the "proof" is actually an intuitive argument for the population value of the Pearson Type 2 skewness, and not for the sample estimate, and hence Sk_2 may fall outside the range [−3, +3]. In this article, alternative measures of skewness are proposed that are based on nonparametric density estimates, and are compared to some of the commonly used skewness measures. A computational geometric measure of skewness is also introduced.

2. Proposed Measure of Skewness

Many introductory statistics text books include a rule of thumb regarding the relative positions of the mean, the median: for a positively skewed distribution, mean > median > mode, and for a negatively skewed distribution, mean < median < mode [19] [20] [21]. It was pointed out by von Hippel [22] that many violations of this rule exist, especially in the case of discrete probability distributions (see Figure 1(b), Figure 1(c)).

Letting $f(x)$ and $F(x)$ denote the population probability density and cumulative distributions functions of the random variable, with mean μ and median Q_2, the proposed skewness measure is defined as the area under $f(x)$ between μ and median Q_2 (Figure 2).

$$\text{Area skewness} = F(\mu) - F(Q_2).$$

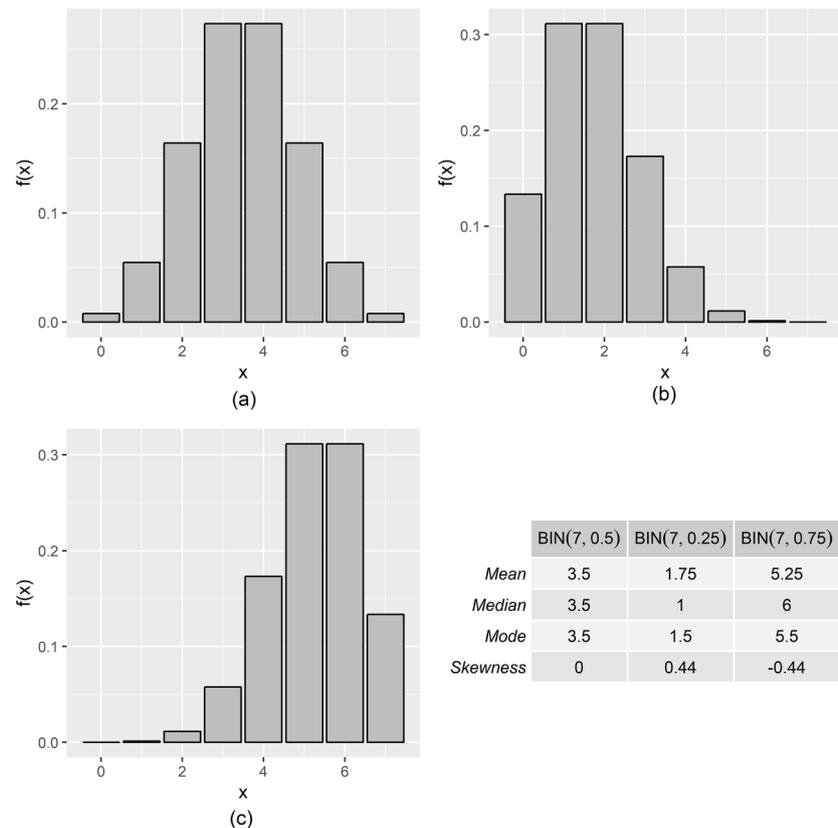|	BIN(7, 0.5)	BIN(7, 0.25)	BIN(7, 0.75)
Mean	3.5	1.75	5.25
Median	3.5	1	6
Mode	3.5	1.5	5.5
Skewness	0	0.44	-0.44

Figure 1. Plots of the binomial distribution with (a) BIN, $n = 7$ and $p = 0.5$; (b) BIN, $n = 7$ and $p = 0.25$ and (c) BIN, $n = 7$ and $p = 0.75$.
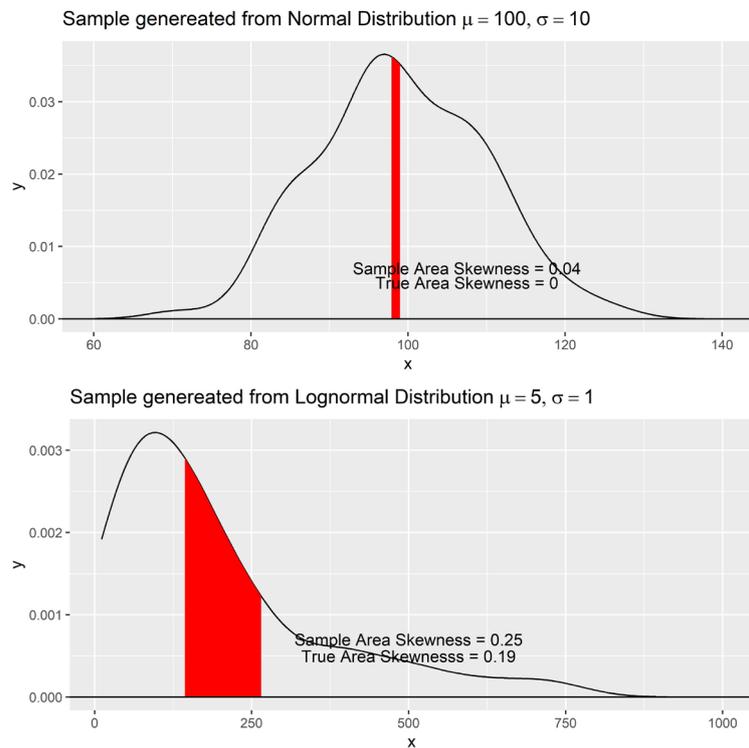
Figure 2. Examples showing area skewness computations.

Area skewness, the probability that the random variable falls inside the true mean μ and the median Q_2, can be computed in two steps:

Step 1. The probability density is estimated from the sample; in this article, a nonparametric density estimate [23] [24] is used, but a parametric density estimate can also be used.

Step 2: A numerical integration method can then be used to compute the area between the sample mean and sample median; the trapezoid rule is used in this article for computing area skewness.

Figure 2 shows two simulated examples of area skewness computation. Data from the first example (top graph) is simulated from a normal distribution with mean $\mu = 100$ and standard deviation $\sigma = 10$; the true area skewness, in this case, equals 0, and the area skewness computed for the samples is −0.004. The second example in **Figure 2** (bottom graph) is generated from the log-normal (LN) distribution which is defined as: Y is LN with parameters μ and σ if log(Y) is normally distributed with mean μ and standard deviation σ; here the log function is the natural log, *i.e.*, the base is *e*. The LN (μ, σ) distribution has population mean, standard deviation, and skewness given by [25]:

$$\text{Mean} = \exp\left(\mu + 0.5\sigma^2\right)$$

$$\text{Median} = \exp\left(\mu\right)$$

$$CV = \sqrt{\exp\left(\sigma^2\right) - 1}, \ CV = \text{Coefficient of Variation}$$

$$\text{Skewness} = \left(CV\right)^3 + 3\left(CV\right)$$

True population mean, median and area skewness for the LN ($\mu = 5$, $\sigma = 1$) distribution are:

$$\text{mean} = \exp(5.5) = 244.6919$$

$$\text{median} = \exp(5) = 148.4132$$

$$\text{standard skewness} = 6.1849$$

$$\text{area skewness} = F(244.6919) - F(148.4132) = 0.1915$$

The sample area skewness value for the generated sample is 0.2047, and the standard skewness estimate is 4.3192.

3. Monte Carlo Simulation for Comparison of Skewness Measures

Three probability distributions with varying degrees of skewness are used in simulation in this study:

N (μ, σ)—normal distribution with mean μ and standard deviation σ.

GAM (α, β)—gamma distribution with shape = α and scale = β, skewness = $2/\sqrt{\alpha}$.

Tr (a, b, c)—Triangular distribution with parameters a, b, c [26] [27] with probability density and cumulative distribution given by

$$f(x) = \begin{cases} \dfrac{2(x-a)}{(b-a)(c-a)}, & a \leq x \leq c \\ \dfrac{2(b-x)}{(b-a)(b-c)}, & c < x \leq b \end{cases}$$

$$F(x) = \begin{cases} \dfrac{(x-a)^2}{(b-a)(c-a)}, & a \leq x \leq c \\ 1 - \dfrac{(b-x)^2}{(b-a)(b-c)}, & c < x \leq b \end{cases}$$

The skewness of the triangular distribution Tr (a, b, c) is given by

$$g_1 = \frac{\sqrt{2}(a+b-2c)(2a-b-c)(a-2b+c)}{5(a^2+b^2+c^2-ab-ac-bc)^{3/2}}.$$

Triangular distribution is selected for this study as it can be used to model both positively skewed and negatively skewed distribution.

Table 1 shows the specific distributions and their skewness values used in this simulation, and Figure 3 shows plots of the two triangular distributions used in the simulations.

The simulation experiment used in this study is carried out in the following steps:

1) A random sample of size n is generated from the selected probability distribution.

2) Each of the five skewness coefficients (proposed area skewness, Pearson

Figure 3. Plots of the two triangular distributions used in the simulations.

Table 1. Probability distributions used in this simulation.

| | Normal | Gamma | Triangular Distribution Tr (a = 0, b = 1, c) | | |
	N ($\mu = 100$, $\sigma = 20$)	GAM ($\alpha = 2$, $\beta = 1$)	c = 0.5	c = 0.95	c = 0.05
Standard Skewness	0	1.4142	0	−0.5606	0.5606
Pearson Skewness	0	0.6823	0	−0.5217	0.5217
Area Skewness	0	0.0940	0	−0.0564	0.0564

skewness, and the sample-moments based Types 1-3 skewness coefficients are computed.

Steps (1) and (2) are repeated 10,000 times and the 90%, 95%, and 99% confidence intervals for true skewness are calculated from the 10,000 skewness values.

The simulation experiment was run for n = 25, 50, 75, 100, for each of the three probability models, for each of the two sets of parameter values. The samples sizes chosen represent moderate to a large number of samples, and the true skewness values selected cover a wide range of skewness. **Figures 4-23** show the histograms of the 10,000 skewness estimates and the confidence intervals.

Empirical Distribution of Skewness from 10,000 Simulations
Normal mean = 100, sigma = 20, sample size = 25

	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.093	0.092	-0.112	0.112	-0.149	0.146
skew.Pearson	-0.746	0.75	-0.884	0.862	-1.131	1.093
skew.Type1	-0.725	0.693	-0.864	0.851	-1.216	1.155
skew.Type2	-0.772	0.738	-0.92	0.906	-1.295	1.23
skew.Type3	-0.682	0.652	-0.812	0.8	-1.144	1.087

Figure 4. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 25$ samples from N (100, 20).

Empirical Distribution of Skewness from 10,000 Simulations
Normal mean = 100, sigma = 20, sample size = 50

	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.065	0.063	-0.077	0.076	-0.1	0.104
skew.Pearson	-0.522	0.506	-0.614	0.608	-0.784	0.792
skew.Type1	-0.552	0.531	-0.655	0.642	-0.88	0.866
skew.Type2	-0.569	0.548	-0.676	0.662	-0.908	0.894
skew.Type3	-0.535	0.515	-0.636	0.623	-0.854	0.841

Figure 5. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 50$ samples from N (100, 20).

Empirical Distribution of Skewness from 10,000 Simulations
Normal mean = 100, sigma = 20, sample size = 75

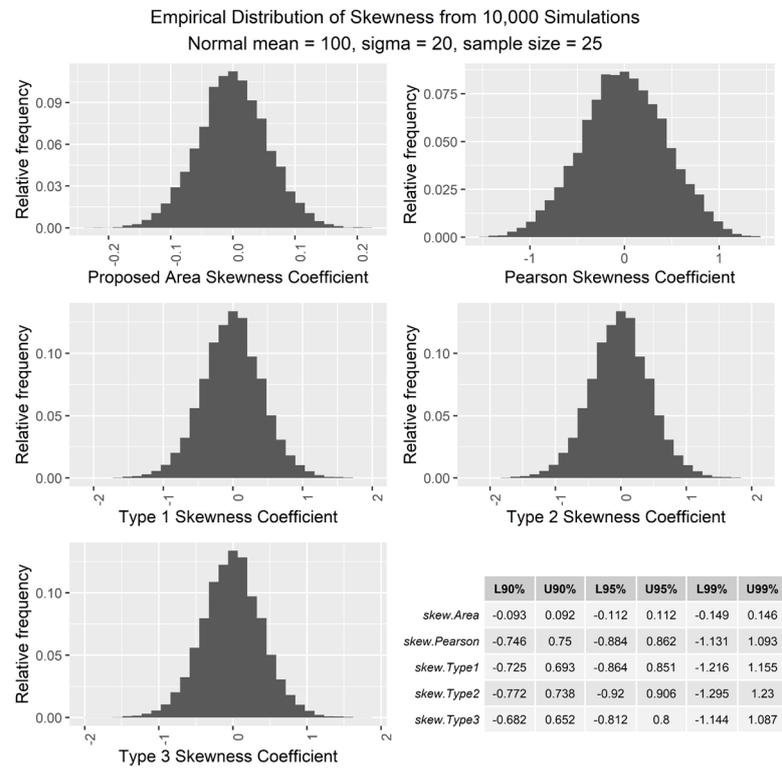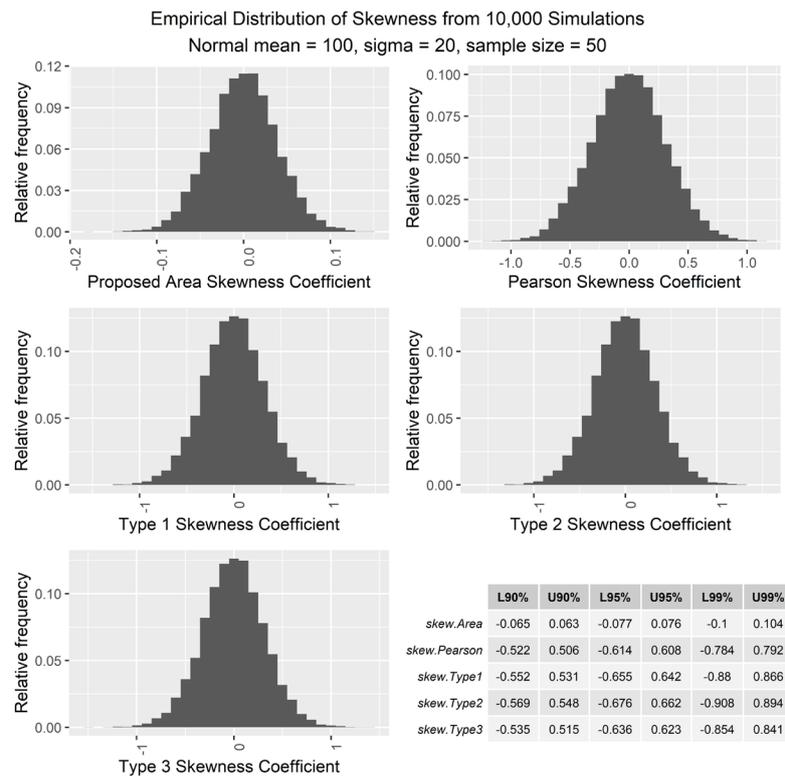|	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.055	0.053	-0.065	0.064	-0.087	0.085
skew.Pearson	-0.433	0.427	-0.522	0.503	-0.678	0.661
skew.Type1	-0.457	0.435	-0.553	0.527	-0.761	0.724
skew.Type2	-0.466	0.444	-0.565	0.538	-0.776	0.739
skew.Type3	-0.448	0.426	-0.542	0.517	-0.746	0.71

Figure 6. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 75$ samples from N (100, 20).

Empirical Distribution of Skewness from 10,000 Simulations
Normal mean = 100, sigma = 20, sample size = 100

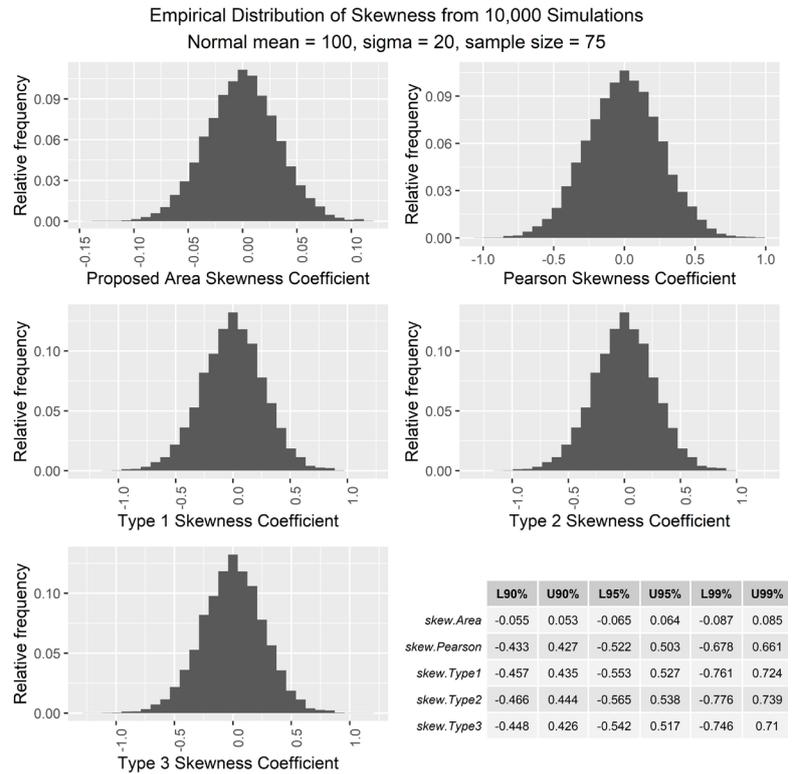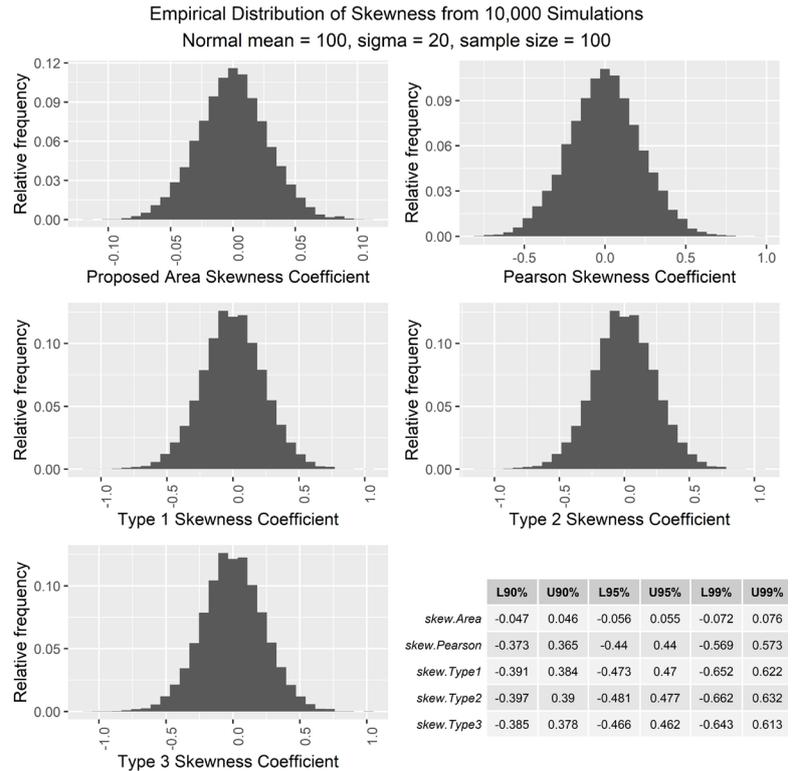|	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.047	0.046	-0.056	0.055	-0.072	0.076
skew.Pearson	-0.373	0.365	-0.44	0.44	-0.569	0.573
skew.Type1	-0.391	0.384	-0.473	0.47	-0.652	0.622
skew.Type2	-0.397	0.39	-0.481	0.477	-0.662	0.632
skew.Type3	-0.385	0.378	-0.466	0.462	-0.643	0.613

Figure 7. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 100$ samples from N (100, 20).

Figure 8. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of *n* = 25 samples from GAM (2, 1); standard skewness = 1.41, pearson skewness = 0.68, area skewness = 0.09.
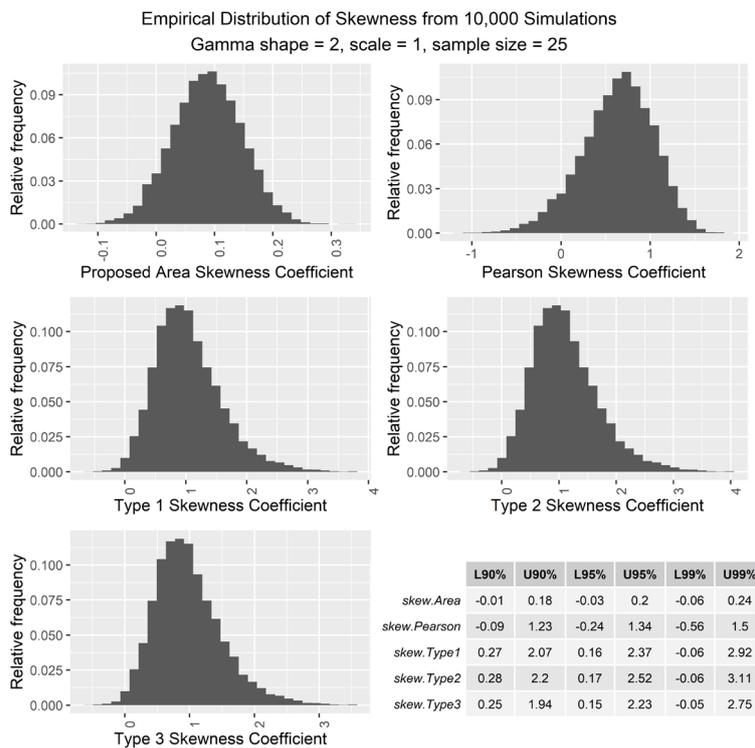

Figure 9. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of *n* = 50 samples from GAM (2, 1); standard skewness = 1.41, pearson skewness = 0.68, area skewness = 0.09.

Empirical Distribution of Skewness from 10,000 Simulations
Gamma shape = 2, scale = 1, sample size = 75

	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	0.03	0.15	0.02	0.16	0	0.18
skew.Pearson	0.27	1.02	0.18	1.08	0.01	1.19
skew.Type1	0.65	2.09	0.56	2.37	0.39	3.04
skew.Type2	0.66	2.13	0.57	2.41	0.4	3.11
skew.Type3	0.64	2.05	0.55	2.32	0.39	2.98

Figure 10. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 75$ samples from GAM (2, 1); standard skewness = 1.41, pearson skewness = 0.68, area skewness = 0.09.
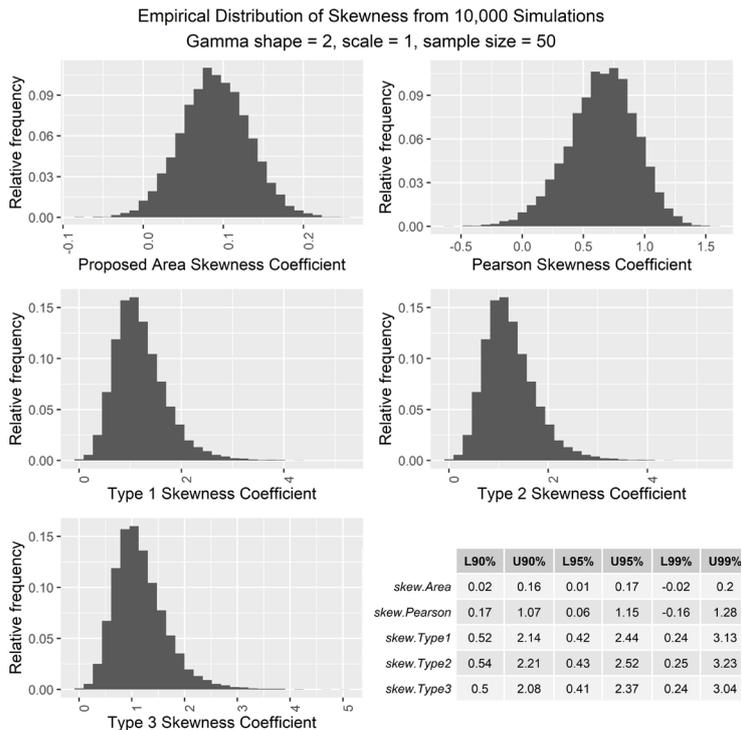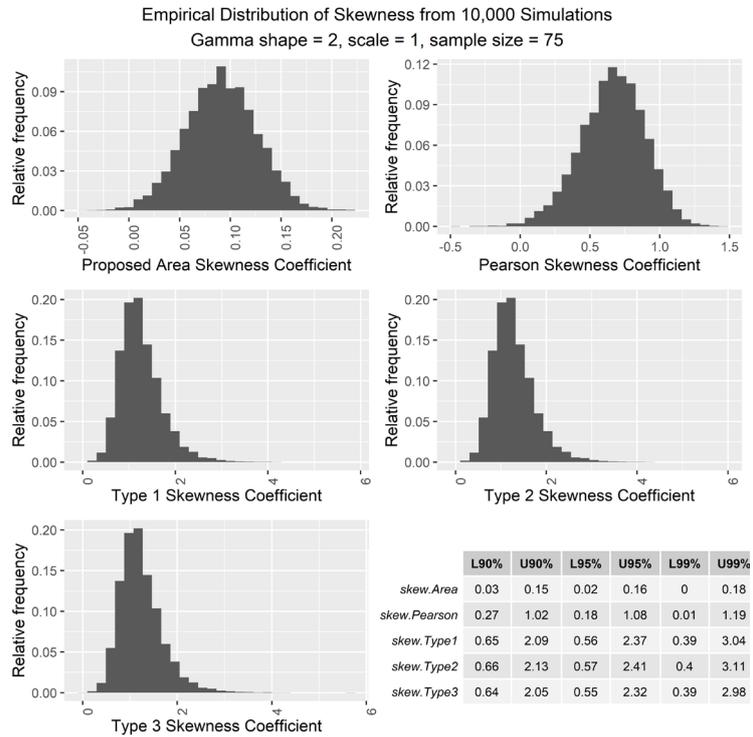
Empirical Distribution of Skewness from 10,000 Simulations
Gamma shape = 2, scale = 1, sample size = 100

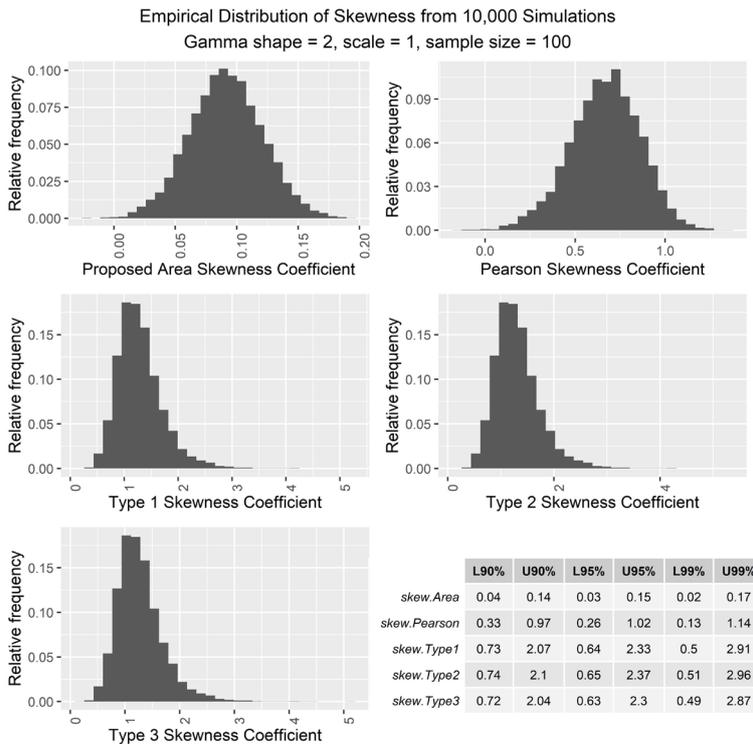|	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	0.04	0.14	0.03	0.15	0.02	0.17
skew.Pearson	0.33	0.97	0.26	1.02	0.13	1.14
skew.Type1	0.73	2.07	0.64	2.33	0.5	2.91
skew.Type2	0.74	2.1	0.65	2.37	0.51	2.96
skew.Type3	0.72	2.04	0.63	2.3	0.49	2.87

Figure 11. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 100$ samples from GAM (2, 1); standard skewness = 1.41, pearson skewness = 0.68, area skewness = 0.09.
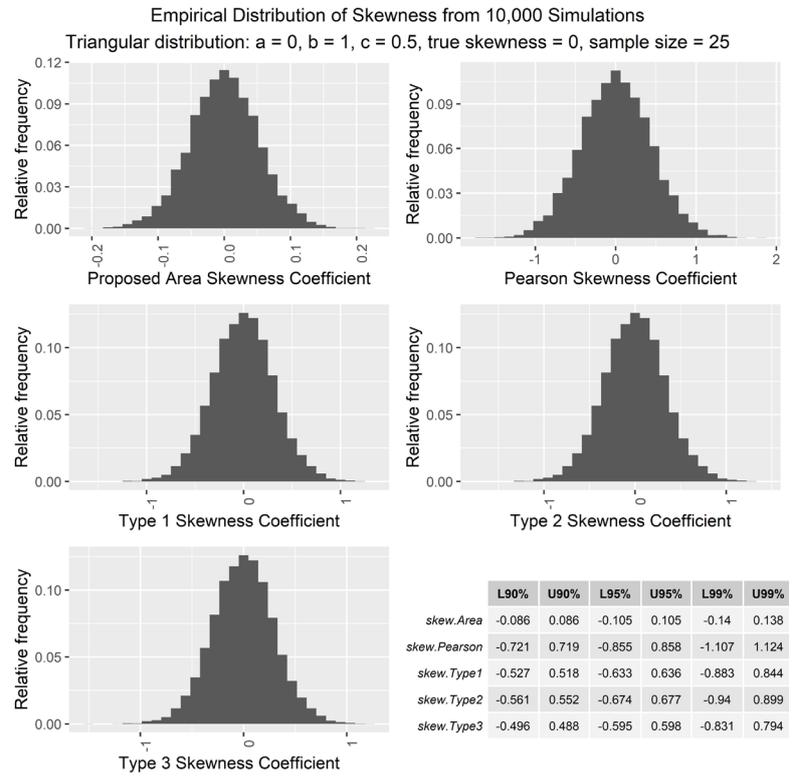
Empirical Distribution of Skewness from 10,000 Simulations
Triangular distribution: a = 0, b = 1, c = 0.5, true skewness = 0, sample size = 25

	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.086	0.086	-0.105	0.105	-0.14	0.138
skew.Pearson	-0.721	0.719	-0.855	0.858	-1.107	1.124
skew.Type1	-0.527	0.518	-0.633	0.636	-0.883	0.844
skew.Type2	-0.561	0.552	-0.674	0.677	-0.94	0.899
skew.Type3	-0.496	0.488	-0.595	0.598	-0.831	0.794

Figure 12. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 25$ samples from Tr (0, 0.5, 1).
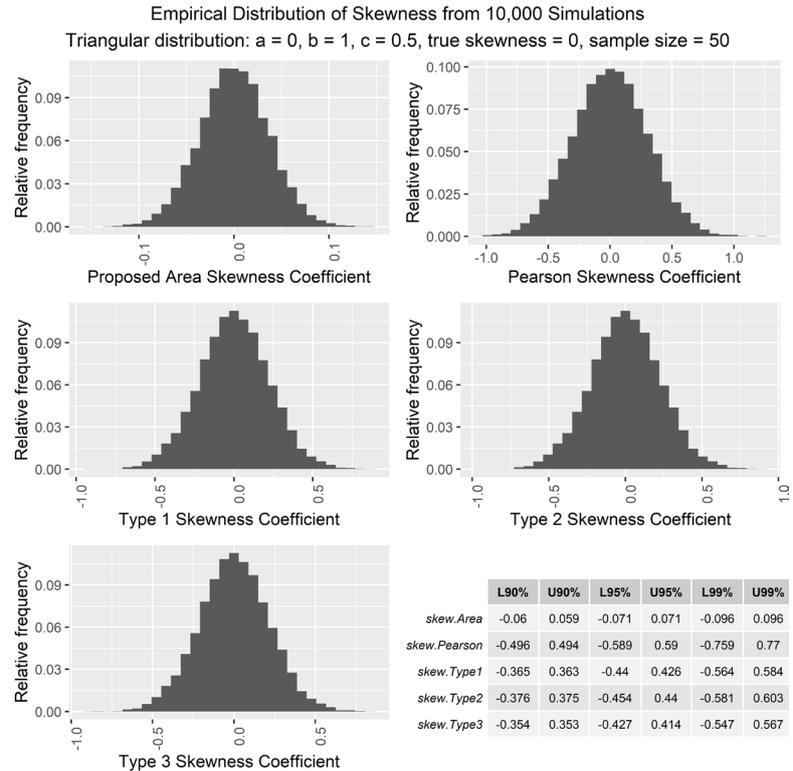
Empirical Distribution of Skewness from 10,000 Simulations
Triangular distribution: a = 0, b = 1, c = 0.5, true skewness = 0, sample size = 50

	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.06	0.059	-0.071	0.071	-0.096	0.096
skew.Pearson	-0.496	0.494	-0.589	0.59	-0.759	0.77
skew.Type1	-0.365	0.363	-0.44	0.426	-0.564	0.584
skew.Type2	-0.376	0.375	-0.454	0.44	-0.581	0.603
skew.Type3	-0.354	0.353	-0.427	0.414	-0.547	0.567

Figure 13. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 50$ samples from Tr (0, 0.5, 1).
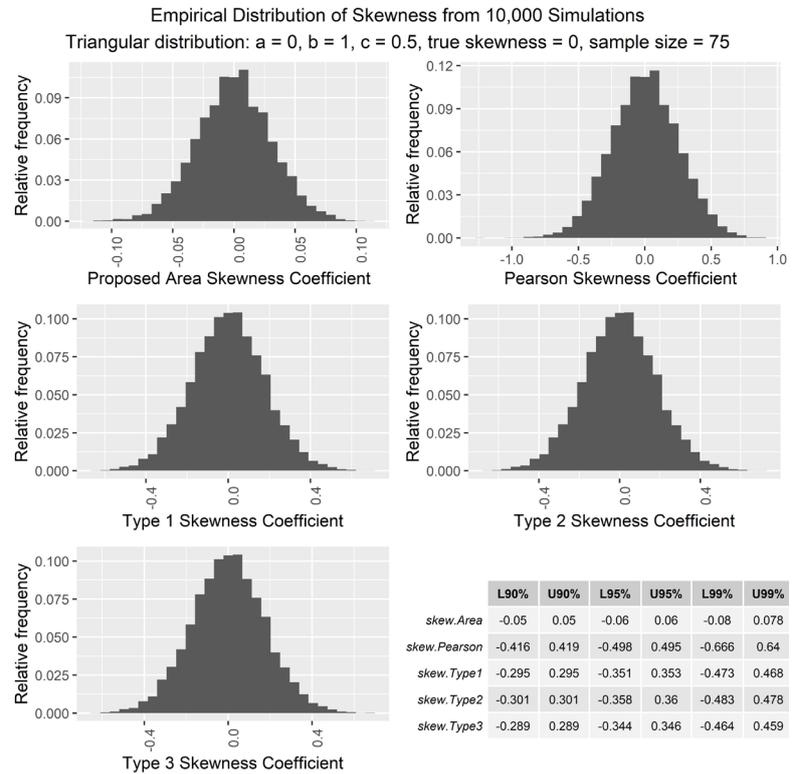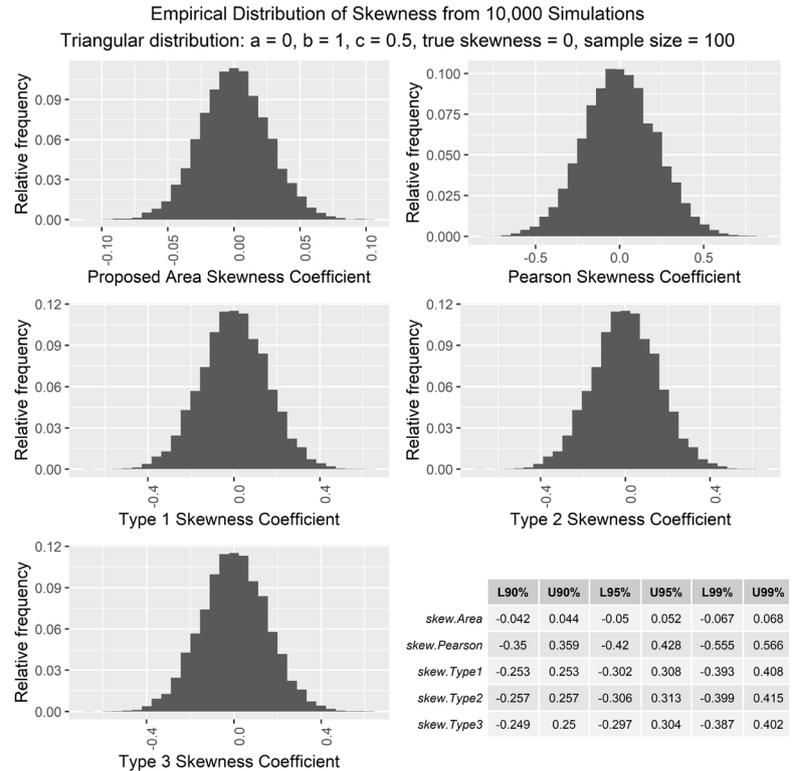
Empirical Distribution of Skewness from 10,000 Simulations
Triangular distribution: a = 0, b = 1, c = 0.5, true skewness = 0, sample size = 75

	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.05	0.05	-0.06	0.06	-0.08	0.078
skew.Pearson	-0.416	0.419	-0.498	0.495	-0.666	0.64
skew.Type1	-0.295	0.295	-0.351	0.353	-0.473	0.468
skew.Type2	-0.301	0.301	-0.358	0.36	-0.483	0.478
skew.Type3	-0.289	0.289	-0.344	0.346	-0.464	0.459

Figure 14. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 75$ samples from Tr (0, 0.5, 1).

Empirical Distribution of Skewness from 10,000 Simulations
Triangular distribution: a = 0, b = 1, c = 0.5, true skewness = 0, sample size = 100

	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.042	0.044	-0.05	0.052	-0.067	0.068
skew.Pearson	-0.35	0.359	-0.42	0.428	-0.555	0.566
skew.Type1	-0.253	0.253	-0.302	0.308	-0.393	0.408
skew.Type2	-0.257	0.257	-0.306	0.313	-0.399	0.415
skew.Type3	-0.249	0.25	-0.297	0.304	-0.387	0.402

Figure 15. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 100$ samples from Tr (0, 0.5, 1).
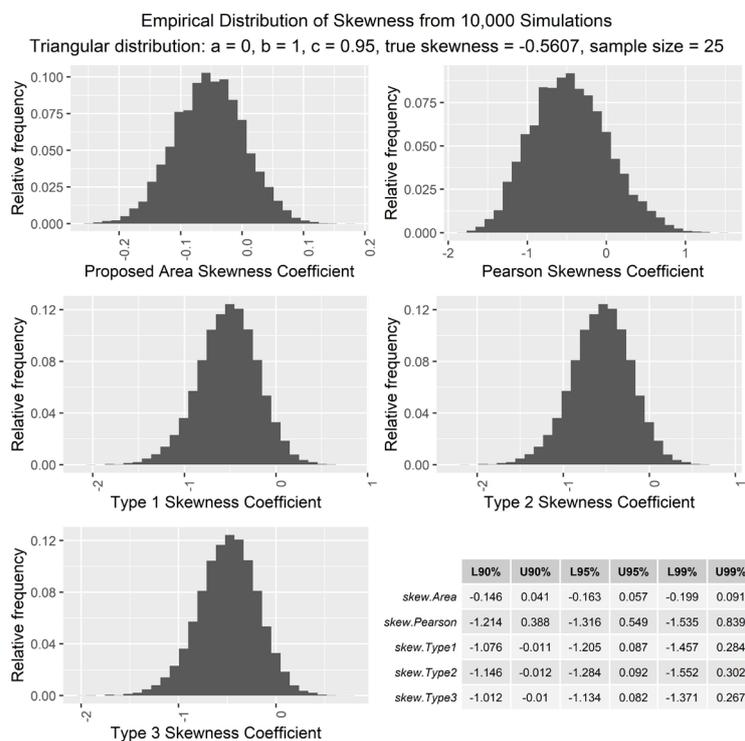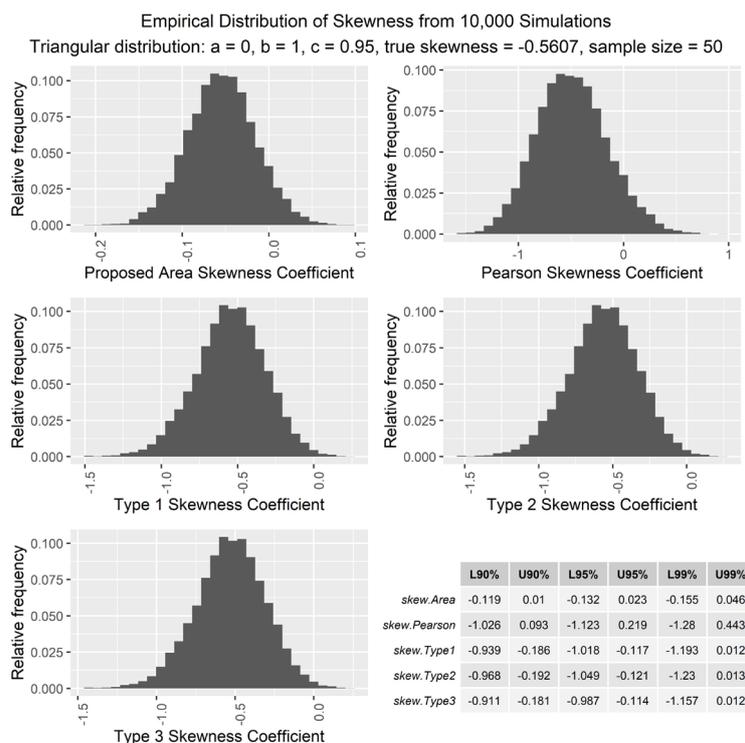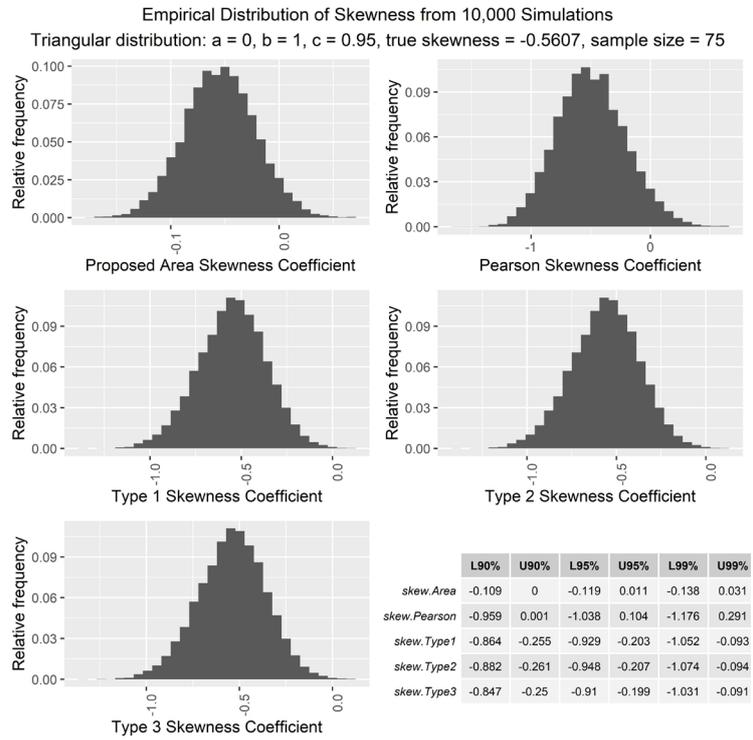
Figure 16. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 25$ samples from Tr (0, 0.95, 1); standard skewness = −0.56, Pearson skewness = −0.52, area skewness = −0.06.
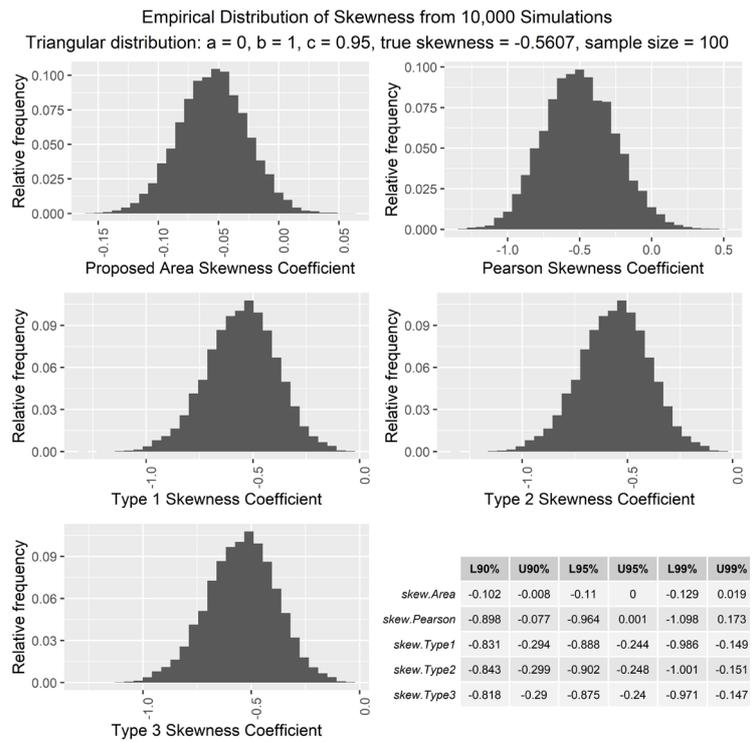

Figure 17. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 50$ samples from Tr (0, 0.95, 1); standard skewness = −0.56, Pearson skewness = −0.52, area skewness = −0.06.

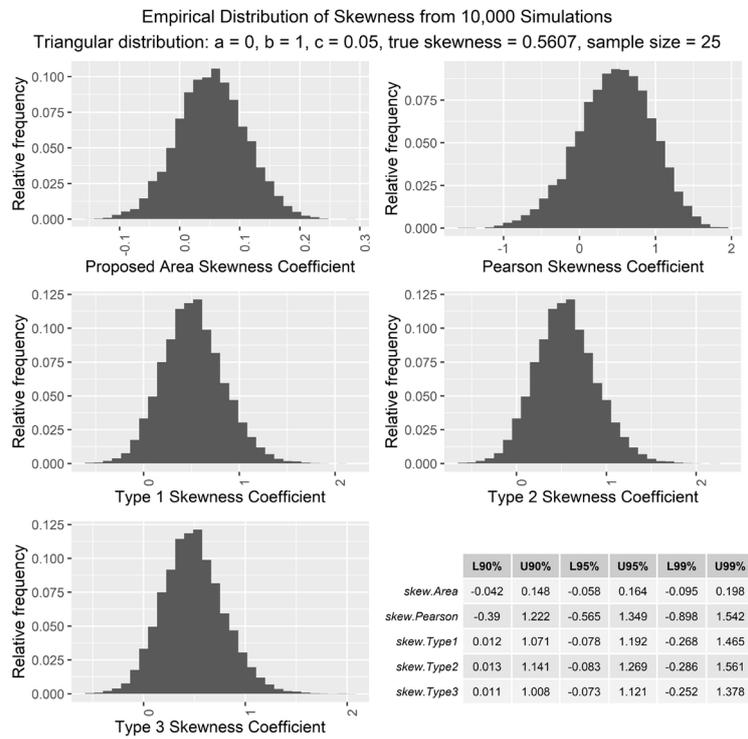**Empirical Distribution of Skewness from 10,000 Simulations**
Triangular distribution: a = 0, b = 1, c = 0.95, true skewness = -0.5607, sample size = 75

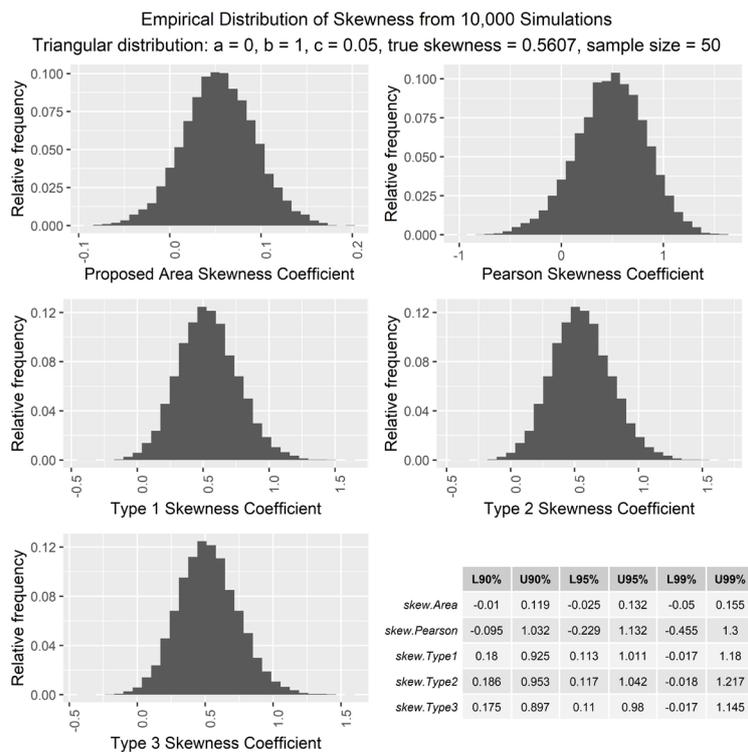|	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.109	0	-0.119	0.011	-0.138	0.031
skew.Pearson	-0.959	0.001	-1.038	0.104	-1.176	0.291
skew.Type1	-0.864	-0.255	-0.929	-0.203	-1.052	-0.093
skew.Type2	-0.882	-0.261	-0.948	-0.207	-1.074	-0.094
skew.Type3	-0.847	-0.25	-0.91	-0.199	-1.031	-0.091

Figure 18. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of n = 75 samples from Tr (0, 0.95, 1); standard skewness = −0.56, Pearson skewness = −0.52, area skewness = −0.06.

Empirical Distribution of Skewness from 10,000 Simulations
Triangular distribution: a = 0, b = 1, c = 0.95, true skewness = -0.5607, sample size = 100

	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.102	-0.008	-0.11	0	-0.129	0.019
skew.Pearson	-0.898	-0.077	-0.964	0.001	-1.098	0.173
skew.Type1	-0.831	-0.294	-0.888	-0.244	-0.986	-0.149
skew.Type2	-0.843	-0.299	-0.902	-0.248	-1.001	-0.151
skew.Type3	-0.818	-0.29	-0.875	-0.24	-0.971	-0.147

Figure 19. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of n = 100 samples from Tr (0, 0.95, 1); standard skewness = −0.56, Pearson skewness = −0.52, area skewness = −0.06.
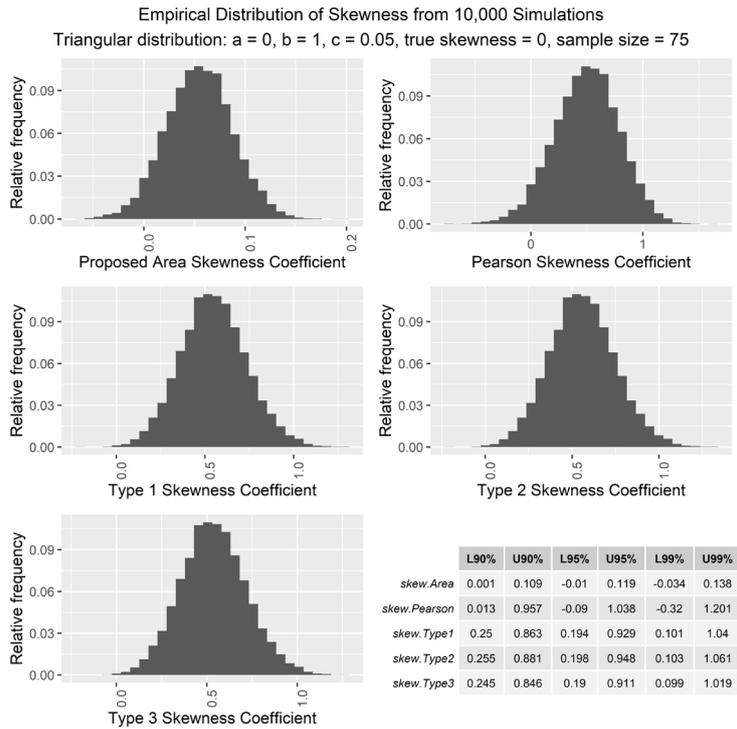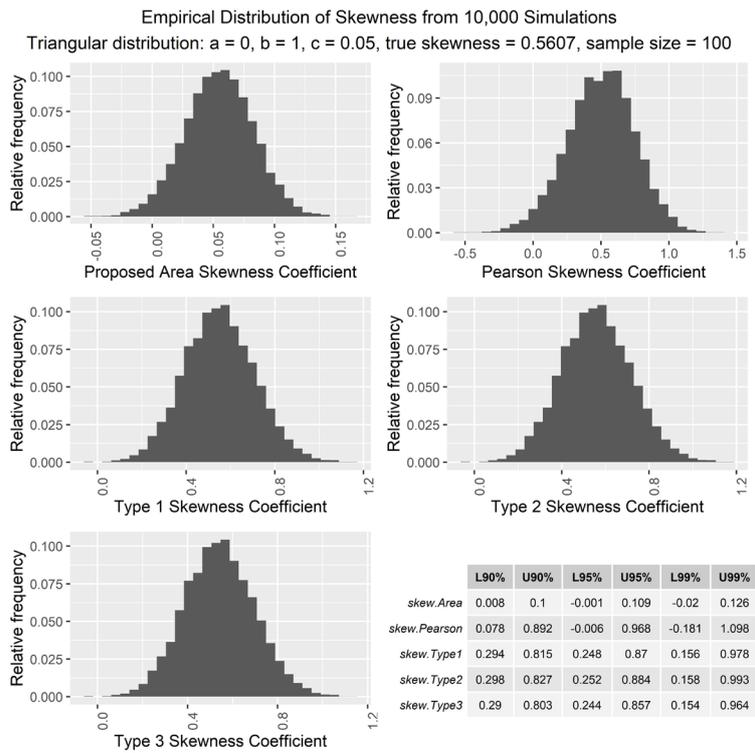
Empirical Distribution of Skewness from 10,000 Simulations
Triangular distribution: a = 0, b = 1, c = 0.05, true skewness = 0.5607, sample size = 25

	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.042	0.148	-0.058	0.164	-0.095	0.198
skew.Pearson	-0.39	1.222	-0.565	1.349	-0.898	1.542
skew.Type1	0.012	1.071	-0.078	1.192	-0.268	1.465
skew.Type2	0.013	1.141	-0.083	1.269	-0.286	1.561
skew.Type3	0.011	1.008	-0.073	1.121	-0.252	1.378

Figure 20. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 25$ samples from Tr (0, 0.05, 1); standard skewness = 0.56, Pearson skewness = 0.52, area skewness = 0.06.

Empirical Distribution of Skewness from 10,000 Simulations
Triangular distribution: a = 0, b = 1, c = 0.05, true skewness = 0.5607, sample size = 50

	L90%	U90%	L95%	U95%	L99%	U99%
skew.Area	-0.01	0.119	-0.025	0.132	-0.05	0.155
skew.Pearson	-0.095	1.032	-0.229	1.132	-0.455	1.3
skew.Type1	0.18	0.925	0.113	1.011	-0.017	1.18
skew.Type2	0.186	0.953	0.117	1.042	-0.018	1.217
skew.Type3	0.175	0.897	0.11	0.98	-0.017	1.145

Figure 21. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 50$ samples from Tr (0, 0.05, 1); standard skewness = 0.56, Pearson skewness = 0.52, area skewness = 0.06.

Figure 22. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 75$ samples from Tr (0, 0.05, 1); standard skewness = 0.56, Pearson skewness = 0.52, area skewness = 0.06.

Figure 23. Histograms and confidence intervals of skewness coefficients from 10,000 simulations of $n = 100$ samples from Tr (0, 0.05, 1); standard skewness = 0.56, Pearson skewness = 0.52, area skewness = 0.06.

4. A Computational Geometric Measure of Skewness

The probability density function estimated from the data can be modeled by a simple polygon P as shown in **Figure 24** (thin solid line). Let I_m be the vertical line segment at the sample mean (thick vertical line). Let Ch_1 and Ch_2 denote polygonal chains to the right and left of I_m. By taking I_m as a mirror we can consider the reflected images of Ch_1 and Ch_2 denoted by I_1 and I_2, respectively. I_1 and I_2 are drawn as dashed lines in **Figure 24**. Chains I_1 and I_2 form a simple polygon P^*, which we call *image polygon*. The overlay of P and P^* results in two types of areas: (i) *Overlap Area* O_A, and (ii) *Spilled Area* S_A. In the figure spilled area components are labeled as A, B, C, and D. For a symmetric distribution, spilled area will be small. If the distribution is asymmetric then the portion of spilled area will be large. This motivates us to use the proportion of spilled area as a measure of skewness.

An algorithm for computing spilled area can be developed by using the data structures for representing simple polygon from computational geometry. A sketch of the algorithm for computing spilled areas is shown below. Efficient implementation of Step 5 and Step 6 needs techniques from computational geometry. For this, the input polygon is represented in a doubly connected edge list data structure as reported in [28]. By navigating through this data structure, the intersection points corresponding to the overlay of P and P' can be computed in linear time.

Algorithm 1: Computing Spilled Area.

Input: A simple polygon P constructed from samples points.

Output: Spilled Area S_A.

Step 1: Find the mean vertical line segment I_m.

Step 2: Find polygonal chains Ch_1 and Ch_2 implied by I_m from input polygon P.

Step 3: Determine corresponding image chains I_1 and I_2.

Step 4: Construct image polygon P^* by combining I_1 and I_2.

Step 5: Compute Overlap Area $O_A = \cap \left(P, P^* \right)$.

Step 6: Compute Union Area $U_A = \cup \left(P, P^* \right)$.

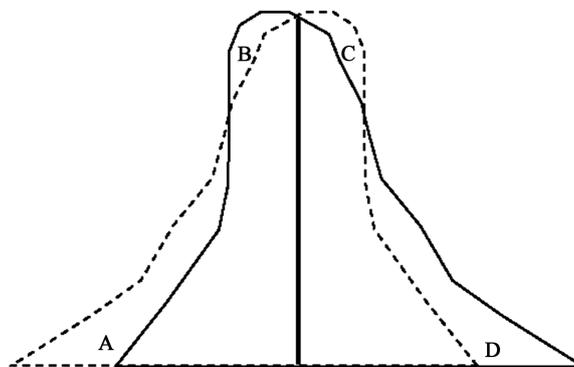**Step 7**: Spilled Area $S_A = U_A - O_A$.

Figure 24. Construction of an Image Polygon.

We implemented the algorithm in python programming environment. For illustration purposes, two different samples were generated from different normal distributions. The true geometric skewness measure for any normal distribution is 0, since the normal distribution is symmetric. The results for the two samples are presented below.

The input polygon computed from the first sample is shown in **Figure 25**, and the overlap area is shown in **Figure 26**.
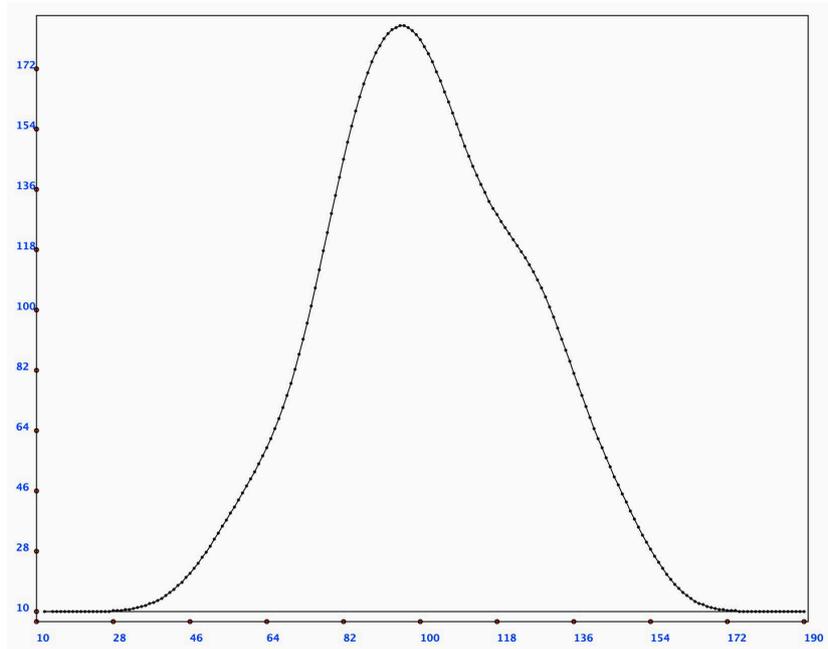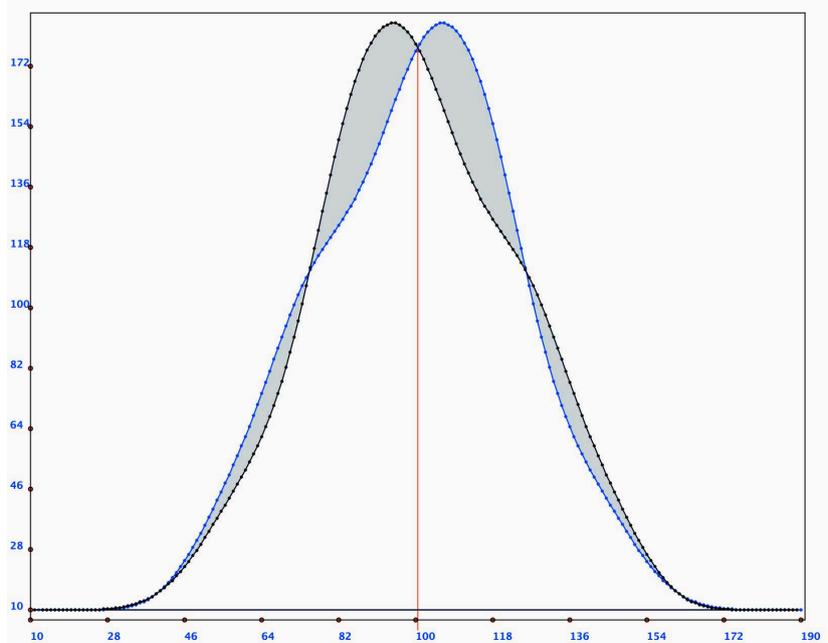

Figure 25. Input polygon for sample 1.

Figure 26. Input polygon for sample 2.

For sample 1, node count = 188, overlap area: 0.46, polygon area: 2.91, and the geometric measure of skweness = overlap area/polygon area = 0.1581.

For the second simulated example, **Figure 27** and **Figure 28** show the input polygon and the overlap area, respectively. For sample 2, node count = 40, overlap area: 0.41, polygon area: 2.93, and the geometric measure of skweness = overlap area/polygon area = 0.1387.
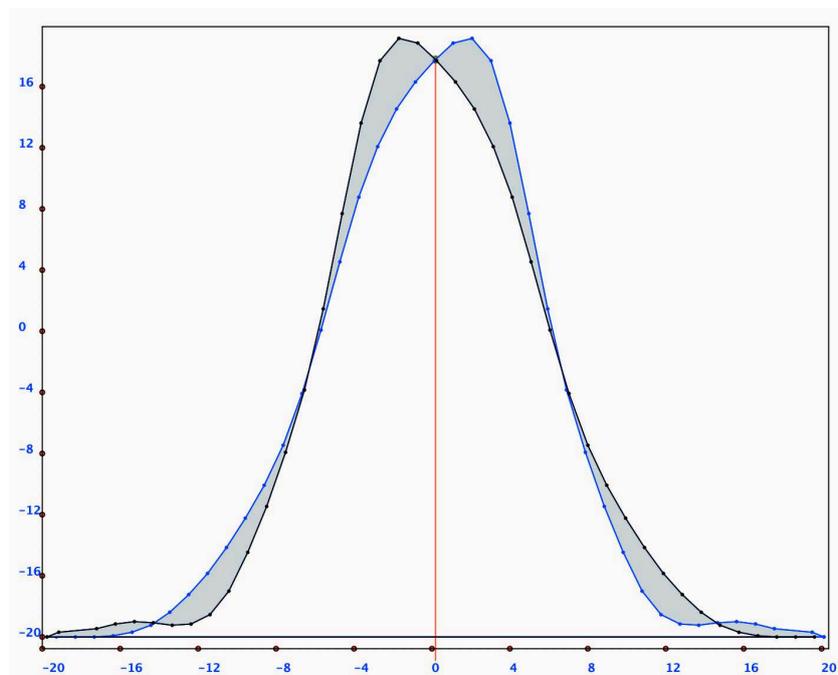

Figure 27. Input polygon for sample 2.

Figure 28. Overlap area for the second sample.

5. Discussion and Results

We have proposed two different skewness measures: area skewness and geometric skewness. The standard skewness measures suffer from one drawback: they do not have known lower and upper bounds. The absolute values of both of the proposed skewness estimates fall in the range (0, 1). We have used Monte Carlo simulations to compute confidence intervals from the area skewness estimate, and we intend to do the same for the geometric skewness estimate in the near future.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Devore, J.L. and Berk, K.N. (2011) Modern Mathematical Statistics with Applications. Springer Science & Business Media, Berlin, 302.

[2] Norman, G.R. and Streiner, D.L. (2008) Biostatistics: The Bare Essentials. PMPH USA, Raleigh, 80.

[3] LaMorte, W.W. (2016) Central Limit Theorem.
http://sphweb.bumc.bu.edu/otlt/MPH-modules/BS/BS704_Probability/BS704_Probability12.html

[4] Vieira Jr., E.T. (2017) Introduction to Real World Statistics: With Step-by-Step SPSS Instructions. Taylor & Francis, Abingdon-on-Thames, 67-68.
https://doi.org/10.4324/9781315233024-6

[5] Singh, A.K., Lucas, A.F., Dalpatadu, R.J. and Murphy, D.J. (2013) Casino Games and the Central Limit Theorem. *UNLV Gaming Research & Review Journal*, **17**, 45-61.

[6] Veluchamy, S.K. (2005) A Graphical Approach for Verification of the Central Limit Theorem. M.S. Thesis, Department of Mathematical Sciences, University of Nevada, Las Vegas.

[7] Johnson, S., Sen, A. and Balyeat, B. (2012) A Skewness-Adjusted Binomial Model for Pricing Futures Options—The Importance of the Mean and Carrying-Cost Parameters. *Journal of Mathematical Finance*, **2**, 105-120.
https://doi.org/10.4236/jmf.2012.21013

[8] Kun, P.K. (2017) Importance of Skewness in Investor Utility: Evidence from the Chinese Stock Markets. *Journal of Mathematical Finance*, **7**, Article ID: 80137.
https://doi.org/10.4236/jmf.2017.74047

[9] Chateau, J.-P.D. (2014) Valuing European Put Options under Skewness and Increasing [Excess] Kurtosis. *Journal of Mathematical Finance*, **4**, 160-177.
https://doi.org/10.4236/jmf.2014.43015

[10] Pavlos, A. and Sickles, R. (2011) The Skewness Issue in Stochastic Frontiers Models: Fact or Fiction? In: Van Keilegom, I. and Wilson, P.W., Eds., *Exploring Research Frontiers in Contemporary Statistics and Econometrics: A Festschrift for Léopold Simar*, Springer Science & Business Media, Berlin, 201-228.

[11] Brys, G., Hubert, M. and Struyf, A. (2003) A Comparison of Some New Measures of Skewness. In: Dutter, R., Filzmoser, P., Gather, U. and Rousseeuw, P.J., Eds., *De-

velopments in Robust Statistics, *International Conference on Robust Statistics* 2001, Physica-Verlag, Heidelberg, 98-113. https://doi.org/10.1007/978-3-642-57338-5_8

[12] Brys, G., Hubert, M. and Struyf, A. (2004) A Robust Measure of Skewness. *Journal of Computational and Graphical Statistics*, **13**, 996-1017. https://doi.org/10.1198/106186004X12632

[13] Joanes, D.N. and Gill, C.A. (1998) Comparing Measures of Sample Skewness and Kurtosis. *Journal of the Royal Statistical Society. Series D* (*The Statistician*), **47**, 183-189. https://doi.org/10.1111/1467-9884.00122

[14] Lim, W.K. and Lim, A.W. (2016) A Comparison of Usual t-Test Statistic and Modified t-Test Statistics on Skewed Distribution Functions. *Journal of Modern Applied Statistical Methods*, **15**, Article 8. https://doi.org/10.22237/jmasm/1478001960 http://digitalcommons.wayne.edu/jmasm/vol15/iss2/8

[15] Sharma, K.K., Kumar, A. and Chaudhary, A. (2009) Statistics in Management Studies. Krishna Media, Meerut, India, 213-214.

[16] Doane, D.P. and Seward, L.E. (2011) Measuring Skewness: A Forgotten Statistic? *Journal of Statistics Education*, **19**, 18. https://doi.org/10.1080/10691898.2011.11889611

[17] R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org

[18] Hotelling, H. and Solomon, L.M. (1932) The Limits of a Measure of Skewness. *The Annals of Mathematical Statistics*, **3**, 141-142. https://doi.org/10.1214/aoms/1177732911

[19] Nesselroade Jr., P.K. and Grimm, L.G. (2018) Statistical Applications for the Behavioral and Social Sciences. John Wiley & Sons, Berlin, 83. https://doi.org/10.1002/9781119531708

[20] Triola, M.F. and Franklin, L.A. (1994) Business Statistics: Understanding Populations and Processes. Addison-Wesley, Reading.

[21] Jaisingh, L.R. (2006) Statistics for the Utterly Confused. 2nd Edition, McGraw Hill Professional, New York, 38.

[22] von Hippel, P.T. (2005) Mean, Median, and Skew: Correcting a Textbook Rule. *Journal of Statistics Education*, **13**, 2. https://doi.org/10.1080/10691898.2005.11910556

[23] Hothorn, T. and Everitt, B.S. (2014) A Handbook of Statistical Analyses Using R. Chapman and Hall/CRC, London.

[24] Deng, H. and Wickham, H. (2014) Density Estimation in R. https://www.semanticscholar.org/paper/Density-estimation-in-R-Deng-Wickham/74a6589c40a55b75e36ebcc1fb279472b00feb2b

[25] Singh, A.K., Singh, A. and Engelhardt, M. (1997) The Lognormal Distribution in Environmental Applications. EPA Technology Support Center Issue Paper, Las Vegas, EPA/600/R-97/006.

[26] Balakrishnan, N. and Nevzorov, V.B. (2004) A Primer on Statistical Distributions. John Wiley & Sons, Hoboken, 123. https://doi.org/10.1002/0471722227

[27] Westfall, P. and Henning, K.S.S. (2013) Understanding Advanced Statistical Methods. CRC Press, London, 68-69. https://doi.org/10.1201/b14398

[28] De Berg, M., van Kreveld, M., Overmars, M. and Schwarzkopf, O.C. (2000) Computational Geometry: Algorithms and Applications. Springer, Berlin. https://doi.org/10.1007/978-3-662-04245-8