Inference about the Population Kurtosis with Confidence: Parametric and Bootstrap Approaches

Guensley Jerome and B.M. Golam Kibria*

Department of Mathematics and Statistics, Florida International University, Miami, FL, 33196, USA

Abstract: This paper considers some classical and bootstrap methods in constructing confidence intervals for the kurtosis parameter of a distribution. The bootstrap techniques used are: Bias-Corrected Standard Bootstrap, Efron's Percentile Bootstrap, Hall's Percentile Bootstrap and Bias-Corrected Percentile Bootstrap. The performance of these estimators is compared through confidence intervals by determining the average width and probabilities of capturing the kurtosis parameter of a distribution. We observed that the parametric method works well in terms of coverage probability when data come from a normal distribution, while the bootstrap intervals struggled in constantly reaching a 95% confidence level. When sample data are from a distribution with negative kurtosis, both parametric and bootstrap confidence intervals performed well, although we noticed that bootstrap methods tend to have shorter intervals. When it comes to positive kurtosis, bootstrap methods perform slightly better than classical methods in the sense of high coverage probability. For illustration purposes, two real life health related data are analyzed.

Keywords: Beta Distribution, Bootstrap Techniques, Confidence Interval, Kurtosis Parameter, Simulation.

1. INTRODUCTION

In Statistics, kurtosis of a distribution is one of the more obscure population parameters and has not been discussed by many. To begin, we would first want to define what is kurtosis. The historical misconception is that the kurtosis is a characterization of the peakedness of a distribution. In various books, the kurtosis is described as the "flatness or peakedness of a distribution" [1], and the paper entitled: Kurtosis as Peakedness, 1908 - 2014, R.I.P. [2] strongly addressed said misconception. Westfall wrote: "Kurtosis tells you virtually nothing about the shape of the peak – its only unambiguous interpretation is in terms of tail extremity." His claims were supported with numerous examples of why you cannot relate the peakedness of distributions to kurtosis. We can now define the kurtosis of a distribution as the measurement of its tails. Distributions with positive kurtosis, also known as leptokurtic, have longer tails and tend to generate more outliers while distribution with negative kurtosis, or platykurtic, produce fewer to no outliers. Distributions with zero kurtosis are referred to as mesokurtic and the most prominent of such distributions is the Normal Distribution [1].

The kurtosis parameter of a probability distribution was first defined by Karl Pearson in 1905 [2] as:

$$\kappa(X) = \frac{\mu_4}{\sigma^2} = \frac{E(X-\mu)^4}{(E(X-\mu)^2)^2}$$

where μ_4 is the fourth moment about the mean and σ^2, the variance. The Normal distribution with a mean μ and variance σ^2 has a kurtosis of 3. Often statisticians adjust the kurtosis of the normal distribution to zero by simply subtracting 3 from the kurtosis. When this adjustment is made, it is usually referred to as the Excess Kurtosis. In this paper, excess kurtosis is defined as

$$Kurt(X) = \frac{\mu_4}{\sigma^4} - 3$$
$$= \frac{E(X-\mu)^4}{\sigma^4} - 3.$$

Since the kurtosis defined above is the parameter of a distribution, often estimators of the parameter are derived and the three most common kurtosis estimators proposed are g_2, G_2 and b_2. For more on kurtosis, we refer our readers to [3] and [4] among others.

2. KURTOSIS ESTIMATORS

2.1. Estimator g_2

For a sample of size n, the sample moment is defined as

$$m_r = \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^r. \quad (1)$$

In the definition of the excess kurtosis of a population, by replacing the population moments with sample moments, the first estimator, g_2, is derived as follows:

$$g_2 = \frac{m_4}{m_2^2} - 3 \quad (2)$$

with variance of g_2:

$$var(g_2) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)} . \quad (3)$$

Fisher showed that g_2 is a biased estimator because $E(g_2) = -\frac{6}{n+1}$ [5]. By simply applying a correction of $-\frac{n+1}{6}$, g_2 would be an unbiased estimator, but [6] suggests that instead of using this

*Address correspondence to this author at the Department of Mathematics and Statistics, Florida International University, Miami, FL, 33196, USA; Tel: (305) 348 1419; Fax: (305) 348 6895; E-mail: kibriag@fiu.edu, jjero003@fiu.edu

E-ISSN: 1929-6029/18

simple correction mentioned above, it is preferable to use ratios of unbiased cumulants to construct unbiased estimators of kurtosis.

2.2. Estimator G_2

First, we describe a cumulant generating function, $K(t)$. From moment generating functions, the cumulant generating function is defined as the natural log of an MGF as:

$$K_X(t) = \ln\,[M_X(t)]. \quad (4)$$

The cumulants are obtained from a power series expansion of the cumulant generating function:

$$K_X(t) = \ln\left[\sum_{r=0}^{\infty} \frac{\mu_r t^r}{r!}\right]. \quad (5)$$

Expanding equation (5) will result in creating a Maclaurin series. Therefore, *n-th* cumulant can be obtained by taking the derivative of the above expansion n times and evaluating the result at $t = 0$.

Based on the general formula above, we get the first four cumulants:

$$K_1 = E(X) = \mu_1' = \mu$$
$$K_2 = Var(X) = \mu_2' - (\mu_1')^2 = \mu_2$$
$$K_3 = 2{\mu_1'}^3 - 3\mu_1'\mu_2' + \mu_3' = \mu_3$$
$$K_4 = -6\mu_1'^4 + 12\mu_1'^2\mu_2' - 3\mu_2'^2 - 4\mu_1'\mu_3' + \mu_4'$$

where K_4 can be rewritten as $K_4 = \mu_4 - 3\mu^2$.

As it was previously defined, the excess kurtosis is

$$Kurt(X) = \frac{\mu_4}{\sigma^4} - 3 \quad (6)$$

Then, in terms of the population cumulant, excess kurtosis can be defined in [6] as

$$\gamma = \frac{K_4}{(K_2)^2} - 3. \quad (7)$$

Now, assume an unbiased cumulant estimator, c_j for which $E(c_j) = K_j$, then [3] shows that the unbiased sample cumulants c_j are:

$$c_2 = \frac{n}{n-1} m_2$$

$$c_3 = \frac{n^2}{(n-1)(n-2)} m_3$$

$$c_4 = \frac{n^2}{(n-1)(n-2)(n-3)} \{(n+1)m_4 - 3(n-1)\}m_2^2\} \quad (8)$$

where m_r, is the sample moment, which was first defined in equation (1). Then, the kurtosis estimator, G_2, solely using unbiased cumulants is defined in [6] as

$$G_2 = \frac{c_4}{c_2^2}$$

$$G_2 = \frac{n-1}{(n-2)(n-3)} \{(n+1)g_2 + 6\} \quad (9)$$

G_2, the estimator derived is the sample kurtosis adopted by statistical packages such as SAS and SPSS [7]. It is generally biased, but unbiased for the normal distribution.

Its variance is

$$var(G_2) = \left[\frac{(n+1)(n-1)}{(n-2)(n-3)}\right]^2 var(g_2). \quad (10)$$

The variance of G_2 can be approximated with the following:

$$var(G_2) = \left(1 + \frac{10}{n}\right) var(G_2) \text{ for all } n > 3 \quad (11)$$

2.3. Estimator b_2

Recall that g_2 is defined as $g_2 = \frac{m_4}{m_2^2}$, where m_2 is the second sample moment, a biased estimator of the sample standard deviation. Using the unbiased standard deviation of the sample instead would give us the third estimator. We refer to it as b_2 and this sample kurtosis estimator is used by computer software packages such as MINITAB and BMDP [6]. It is defined as

$$b_2 = \frac{m_4}{s^4} - 3 \quad (12)$$

with

$$s = \sqrt{\frac{\sum_i^n (x_i - \bar{x})^2}{n-1}}$$

Expanding the definition of b_2, we can rewrite b_2 as:

$$b_2 = \left(\frac{n-1}{n}\right)^2 \frac{m_4}{m_2^2} - 3 \quad (13)$$

To get the variance of b_2, let us first rewrite b_2 in terms of g_2 as

$$b_2 = \left(\frac{n-1}{n}\right)^2 g_2 + 3\left[\left(\frac{n-1}{n}\right)^2\right] - 1 \quad (14)$$

Then its variance is

$$var(b_2) = \left(\frac{n-1}{n}\right)^4 var(g_2) \quad (15)$$

which can be approximated as:

$$var(b_2) = \left(1 - \frac{4}{n}\right) var(g_2). \quad (16)$$

From the approximations mentioned in (11) and (16), the following inequality holds:

$$var(b_2) \le var(g_2) \le var(G_2). \quad (17)$$

The organization of this paper is as follows: we define both parametric and non-parametric confidence intervals in **Section 3**. A simulation study is conducted in **Section 4**. Two real life data sets are analyzed in

Section 5. Finally, some concluding remarks are given in **Section 6.**

3. CONFIDENCE INTERVALS

Let $X_1, X_2, \ldots, X_n$ independent and identically distributed random sample of size *n* from a population with mean *μ* and variance σ^2. Given a specific level of confidence, we can construct confidence intervals to estimate the parameter of the distribution of concern. Since we are studying kurtosis in this paper, then the excess kurtosis parameter, $Kurt(X) = \kappa(X) - 3$, is the value we want to estimate. We will rely on parametric and non-parametric approaches to construct confidence intervals with $(1 - \alpha)100\%$ confidence level.

3.1. Parametric Approaches

The general format of parametric confidence intervals is

estimator ± critical value × standard error of estimator

Given this general format, to construct confidence intervals for excess kurtosis parameter of a given population, we will use one of the three estimators g_2, G_2 and b_2 for a sample of size *n*, with their respective standard error and critical value $z_{\alpha/2}$ which is the upper *α/*2 percentile of the standard normal distribution [6].

For estimator g_2 with sample size *n*, the $(1-\alpha)100\%$ confidence interval is

$$g_2 \pm z_{\alpha/2} \cdot \sqrt{\frac{24n(n-2)(n-3)}{(n-1)^2(n+3)(n+5)}}. \quad (18)$$

For estimator G_2 with sample size *n*, the $(1 - \alpha)100\%$ confidence interval is:

$$G_2 \pm z_{\alpha/2} \cdot \sqrt{\frac{24n(n-1)^2}{(n-2)(n-3)(n+3)(n+5)}}. \quad (19)$$

For estimator b_2 with sample size *n*, the $(1 - \alpha)100\%$ confidence interval is:

$$b_2 \pm z_{\alpha/2} \cdot \sqrt{\frac{24n(n-1)^4(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}}$$

3.2. Bootstrap Approaches

[8] argued that parametric confidence intervals can be quite inaccurate in practice since they rely on asymptotic approximation; this means that the sample size n used to estimate parameter of a population is assumed to grow indefinitely [9]. Bootstrap process, on the other hand, does not need to worry about such assumption. The basic idea of the bootstrap is to resample from a sample of size n with replacement. The new samples obtained from resampling from the original sample are referred to as the Bootstrap Samples, and must be of size *n* as well. When a bootstrap sample vector $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ is obtained, a statistic is then computed. The statistic of concern in this paper is any of the three kurtosis estimators g_2, G_2 and b_2 previously defined. This process is repeated *B*-times, where *B* is expected to be at least 1000 to get reliable results [10]. The bootstrap method is a non-parametric tool where the need to know about the underlying distribution to make statistical inference, such as constructing confidence intervals to estimate the parameter of a population, is not needed. Bootstrapping process is best used through the aid of a computer since the number of bootstrap samples needed are required to be large to get reliable results. We will consider the following bootstrap confidence intervals.

3.2.1. Bias-Corrected Standard Bootstrap Approach

Let θ^* be one of the three point estimates for kurtosis previously defined. Then the bias-corrected standard bootstrap confidence intervals is

$$\theta^* - Bias(\theta^*) \pm z_{\frac{\alpha}{2}} \cdot \hat{\sigma}_B \quad (20)$$

where

$$\hat{\sigma}_B = \sqrt{\frac{1}{B-1} \sum_i^B (\theta^* - \overline{\theta^*})^2}$$

$$Bias(\theta^*) = \overline{\theta^*} - \theta^*$$

$\overline{\theta^*}$ is the mean of all kurtosis estimators derived from the bootstrap process.

3.2.2. Efron's Percentile Bootstrap Approach

Introduced by [11], Efron's Percentile Bootstrap approach is to construct a $100(1-\alpha)\%$ percentile confidence interval. Let $\theta^*_{L,\alpha/2}$ be the value for which $(\alpha/2)\%$ bootstrap estimates are less than and $\theta^*_{H,\alpha/2}$, the value for which $(\alpha/2)\%$ bootstrap estimates exceed. Then the confidence interval would have the following lower and upper bounds:

$$[\theta^*_{L,\alpha/2}, \theta^*_{H,\alpha/2}]. \quad (21)$$

3.2.3. Hall's Percentile Bootstrap Approach

Introduced by [12], the method uses the bootstrap on distribution of $\theta^* - \theta$. For any of the excess kurtosis estimators previously defined, we calculate the B bootstrap kurtosis estimates $\theta_1^*, \theta_2^*, \ldots, \theta_B^*$. For a given distribution with its known kurtosis parameter, θ, we calculate the following differences $\theta_1^* - \theta, \theta_2^* - \theta, \ldots, \theta_B^* - \theta$. For simplicity, we may label each $\theta_i^* - \theta$ as δ_i^*. Then the differences can be written as:

$$\delta_1^*, \delta_2^*, \ldots, \delta_B^*$$

Like Efron's method, for a value $\delta^*_{L,\alpha/2}$ which $(\alpha/2)\%$ of the differences, δ^*s, are less than and for a value $\delta^*_{H,\alpha/2}$ which $(\alpha/2)\%$ of the differences exceed. Then, the lower and upper bound of the confidence interval is given by the following confidence interval:

$L = \theta - \delta^*_{[1-\alpha/2]*B}$ and U= $\theta - \delta^*_{[\alpha/2]*B}$ (22)

3.2.4. Bias Corrected Percentile Bootstrap

[11] proposed the method when sample estimators consistently under or overestimate its parameter. Efron suggested that instead of using the usual 0.025 and 0.975 percentiles of the bootstraps, we should use and $b_{0.025}$ and $b_{0.975}$ instead. Where:

$b_{0.975} = \Phi\left(p^* + \frac{p^*+1.96}{1-a(p^*+1.96)}\right)$ and

$$b_{0.025} = \Phi\left(p^* + \frac{p^*+1.96}{1-a(p^*-1.96)}\right) \quad (23)$$

- $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF).
- p^* is the bias-correction that is calculated as $\Phi^{-1}\left(\frac{|\theta_i^*<\theta}{B}\right)$ which is the inverse normal CDF of the proportion of bootstrap statistics values that are less than the empirical sample statistics.
- a is the "acceleration factor." For normal bootstrap processes, *a = 0.000.*

We calculate the confidence intervals as:

$$L = \theta^*_{\Phi(2p^*-1.96)} \text{ and } U = \theta^*_{\Phi(2p^*+1.96)} \quad (24)$$

4. SIMULATION STUDY

Since a theoretical comparison among estimators is outside the scope of the paper, a simulation study is conducted to compare the performance of the interval estimators in this section.

4.1. Simulation Techniques

The main objective of this paper is to compare the performance of the estimators. The criteria in judging performance is derived from the coverage probability and average width of constructed confidence intervals. In order to get these intervals, we had to simulate our dataset and the simulation was done the following way:

For sample sizes n = 30, 50, 100 and 300, we generated data for the following distributions using the statistical software R:

- Standard normal distribution to capture zero kurtosis
- Beta(2,2) distribution to capture negative kurtosis
- Standard logistic distribution to capture positive kurtosis.

In constructing confidence intervals with 95% confidence level using the parametric method, for any of the given distributions mentioned, we generate samples of size n = 30, 50, 100 and 300. Confidence intervals are calculated for each of the estimators g_2, G_2 and b_2. For each of the three distributions, the process was simulated 3,000 times to generate 3,000 lower and upper bound values for each of the three estimators. We then take the average width of each estimator as well as calculate the percentage of times the true kurtosis parameter of a given distribution is within the 3000 constructed intervals.

For the construction of confidence intervals with 95% confidence level using the bootstrap method, from any of the three distributions mentioned above, given a sample size n and one of the three estimators, θ^*, we generate the bootstrap confidence intervals using 1,000 bootstrap statistics. We then simulate the process 3,000 times to construct the bootstrap intervals using the various bootstrap confidence interval techniques we discussed in **Section 3**. We then take the average width of each intervals as well as the percent coverage every time the true kurtosis parameter is within the 3,000 constructed bootstrap intervals. Refer to [13] for more on simulation techniques.

4.2. Results and Discussion

As mentioned, for a given estimator, we are to construct confidence intervals using both parametric and bootstrap methods. We would then calculate the coverage probability as well as the average width of these intervals as our criteria to compare the performance of these estimators. We constructed confidence intervals for the normal distribution, Beta(2,2) and standard logistic distribution in order to capture zero excess kurtosis, positive and negative excess kurtosis. R-Software was used to complete the simulation procedures.

4.2.1. Standard Normal Distribution: Zero Kurtosis

The average width and coverage probability for all confidence intervals using both parametric and non-parametric methods when data are generated from N(0,1) are reported in Table **1**. As expected, the larger the sample sizes, the smaller the average width of the intervals, regardless of confidence interval methods. When the width of all three estimators are compared, in every case, the average width of b_2 is always less than or equal to that of g_2. Furthermore, the average width of g_2 is also always less than or equal to G_2, regardless of sample sizes. This observation agrees with equation (17) since

$$var(b_2) \leq var(g_2) \leq var(G_2)$$

As for the coverage intervals, the classical method achieved 95% coverage for sample sizes n = 30 or higher for any of the three estimators. We noticed that the classical method does show higher coverage probability when compared to all non-parametric intervals for sample sizes 50 or higher. Last, we see that G_2 tends to also have the highest percent coverage when compared to the other two estimators regardless of sample sizes as well as confidence interval construction method. In all cases, G_2 achieved

Table 1: Average Width and Coverage Probability of The Intervals when The Data Are Generated from N(0,1)

Method	Coverage Probability	Width	Estimator	Sample Size	Distribution
Bias Corrected Standard Bootstrap	0.797	2.376	g2	30	Normal(0,1)
Bias Corrected Standard Bootstrap	0.869	2.827	G_2	30	Normal(0,1)
Bias Corrected Standard Bootstrap	0.720	2.224	b2	30	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.882	2.763	g2	30	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.917	3.286	G_2	30	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.834	2.592	b2	30	Normal(0,1)
Classical	0.963	2.745	g2	30	Normal(0,1)
Classical	0.953	3.264	G_2	30	Normal(0,1)
Classical	0.961	2.565	b2	30	Normal(0,1)
Efron's Percentile Bootstrap	0.901	2.366	g2	30	Normal(0,1)
Efron's Percentile Bootstrap	0.959	2.813	G_2	30	Normal(0,1)
Efron's Percentile Bootstrap	0.819	2.208	b2	30	Normal(0,1)
Hall's Percentile Bootstrap	0.740	2.349	g2	30	Normal(0,1)
Hall's Percentile Bootstrap	0.820	2.793	G_2	30	Normal(0,1)
Hall's Percentile Bootstrap	0.634	2.195	b2	30	Normal(0,1)
Bias Corrected Standard Bootstrap	0.830	1.940	g2	50	Normal(0,1)
Bias Corrected Standard Bootstrap	0.872	2.148	G_2	50	Normal(0,1)
Bias Corrected Standard Bootstrap	0.780	1.862	b2	50	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.881	2.116	g2	50	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.909	2.339	G_2	50	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.845	2.027	b2	50	Normal(0,1)
Classical	0.959	2.342	g2	50	Normal(0,1)
Classical	0.951	2.595	G_2	50	Normal(0,1)
Classical	0.963	2.250	b2	50	Normal(0,1)
Efron's Percentile Bootstrap	0.879	1.898	g2	50	Normal(0,1)
Efron's Percentile Bootstrap	0.921	2.103	G_2	50	Normal(0,1)
Efron's Percentile Bootstrap	0.818	1.824	b2	50	Normal(0,1)
Hall's Percentile Bootstrap	0.763	1.846	g2	50	Normal(0,1)
Hall's Percentile Bootstrap	0.815	2.042	G_2	50	Normal(0,1)
Hall's Percentile Bootstrap	0.701	1.770	b2	50	Normal(0,1)
Bias Corrected Standard Bootstrap	0.837	1.464	g2	100	Normal(0,1)
Bias Corrected Standard Bootstrap	0.872	1.540	G_2	100	Normal(0,1)
Bias Corrected Standard Bootstrap	0.807	1.434	b2	100	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.890	1.520	g2	100	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.905	1.599	G_2	100	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.867	1.492	b2	100	Normal(0,1)
Classical	0.962	1.783	g2	100	Normal(0,1)
Classical	0.958	1.875	G_2	100	Normal(0,1)
Classical	0.961	1.747	b2	100	Normal(0,1)
Efron's Percentile Bootstrap	0.874	1.434	g2	100	Normal(0,1)
Efron's Percentile Bootstrap	0.908	1.509	G_2	100	Normal(0,1)
Efron's Percentile Bootstrap	0.845	1.406	b2	100	Normal(0,1)
Hall's Percentile Bootstrap	0.819	1.434	g2	100	Normal(0,1)
Hall's Percentile Bootstrap	0.851	1.509	G_2	100	Normal(0,1)

(Table 1). Continued.

Method	Coverage Probability	Width	Estimator	Sample Size	Distribution
Hall's Percentile Bootstrap	0.784	1.406	b2	100	Normal(0,1)
Bias Corrected Standard Bootstrap	0.884	0.961	g2	300	Normal(0,1)
Bias Corrected Standard Bootstrap	0.892	0.977	G_2	300	Normal(0,1)
Bias Corrected Standard Bootstrap	0.874	0.955	b2	300	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.906	0.974	g2	300	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.913	0.990	G_2	300	Normal(0,1)
Bias Corrected Percentile Bootstrap	0.900	0.967	b2	300	Normal(0,1)
Classical	0.951	1.081	g2	300	Normal(0,1)
Classical	0.949	1.100	G_2	300	Normal(0,1)
Classical	0.955	1.074	b2	300	Normal(0,1)
Efron's Percentile Bootstrap	0.889	0.930	g2	300	Normal(0,1)
Efron's Percentile Bootstrap	0.901	0.946	G_2	300	Normal(0,1)
Efron's Percentile Bootstrap	0.874	0.924	b2	300	Normal(0,1)
Hall's Percentile Bootstrap	0.864	0.945	g2	300	Normal(0,1)
Hall's Percentile Bootstrap	0.877	0.961	G_2	300	Normal(0,1)
Hall's Percentile Bootstrap	0.852	0.939	b2	300	Normal(0,1)

Table 2: Average Width and Coverage Probability of The Intervals When The Data Are Generated from Beta(2,2)

Method	Coverage Probability	Width	Estimator	Sample Size	Distribution
Bias Corrected Standard Bootstrap	0.960	1.581	g2	30	Beta(2,2)
Bias Corrected Standard Bootstrap	0.975	1.879	G_2	30	Beta(2,2)
Bias Corrected Standard Bootstrap	0.905	1.476	b2	30	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.959	1.616	g2	30	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.954	1.919	G_2	30	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.936	1.508	b2	30	Beta(2,2)
Classical	0.997	2.745	g2	30	Beta(2,2)
Classical	0.997	3.264	G_2	30	Beta(2,2)
Classical	0.999	2.565	b2	30	Beta(2,2)
Efron's Percentile Bootstrap	0.996	1.562	g2	30	Beta(2,2)
Efron's Percentile Bootstrap	0.994	1.856	G_2	30	Beta(2,2)
Efron's Percentile Bootstrap	0.990	1.460	b2	30	Beta(2,2)
Hall's Percentile Bootstrap	0.897	1.568	g2	30	Beta(2,2)
Hall's Percentile Bootstrap	0.927	1.863	G_2	30	Beta(2,2)
Hall's Percentile Bootstrap	0.793	1.464	b2	30	Beta(2,2)
Bias Corrected Standard Bootstrap	0.958	1.113	g2	50	Beta(2,2)
Bias Corrected Standard Bootstrap	0.966	1.234	G_2	50	Beta(2,2)
Bias Corrected Standard Bootstrap	0.917	1.070	b2	50	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.958	1.106	g2	50	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.951	1.225	G_2	50	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.946	1.062	b2	50	Beta(2,2)
Classical	0.999	2.342	g2	50	Beta(2,2)
Classical	0.999	2.595	G_2	50	Beta(2,2)
Classical	1.000	2.250	b2	50	Beta(2,2)

(Table 2). Continued.

Method	Coverage Probability	Width	Estimator	Sample Size	Distribution
Efron's Percentile Bootstrap	0.987	1.099	g2	50	Beta(2,2)
Efron's Percentile Bootstrap	0.986	1.218	G_2	50	Beta(2,2)
Efron's Percentile Bootstrap	0.978	1.055	b2	50	Beta(2,2)
Hall's Percentile Bootstrap	0.918	1.095	g2	50	Beta(2,2)
Hall's Percentile Bootstrap	0.936	1.211	G_2	50	Beta(2,2)
Hall's Percentile Bootstrap	0.848	1.050	b2	50	Beta(2,2)
Bias Corrected Standard Bootstrap	0.952	0.714	g2	100	Beta(2,2)
Bias Corrected Standard Bootstrap	0.957	0.751	G_2	100	Beta(2,2)
Bias Corrected Standard Bootstrap	0.927	0.700	b2	100	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.952	0.710	g2	100	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.952	0.747	G_2	100	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.947	0.696	b2	100	Beta(2,2)
Classical	1.000	1.783	g2	100	Beta(2,2)
Classical	1.000	1.875	G_2	100	Beta(2,2)
Classical	1.000	1.747	b2	100	Beta(2,2)
Efron's Percentile Bootstrap	0.968	0.706	g2	100	Beta(2,2)
Efron's Percentile Bootstrap	0.967	0.743	G_2	100	Beta(2,2)
Efron's Percentile Bootstrap	0.962	0.693	b2	100	Beta(2,2)
Hall's Percentile Bootstrap	0.937	0.712	g2	100	Beta(2,2)
Hall's Percentile Bootstrap	0.945	0.748	G_2	100	Beta(2,2)
Hall's Percentile Bootstrap	0.903	0.697	b2	100	Beta(2,2)
Bias Corrected Standard Bootstrap	0.947	0.385	g2	300	Beta(2,2)
Bias Corrected Standard Bootstrap	0.950	0.391	G_2	300	Beta(2,2)
Bias Corrected Standard Bootstrap	0.939	0.382	b2	300	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.952	0.384	g2	300	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.952	0.390	G_2	300	Beta(2,2)
Bias Corrected Percentile Bootstrap	0.949	0.381	b2	300	Beta(2,2)
Classical	1.000	1.081	g2	300	Beta(2,2)
Classical	1.000	1.100	G_2	300	Beta(2,2)
Classical	1.000	1.074	b2	300	Beta(2,2)
Efron's Percentile Bootstrap	0.952	0.383	g2	300	Beta(2,2)
Efron's Percentile Bootstrap	0.952	0.390	G_2	300	Beta(2,2)
Efron's Percentile Bootstrap	0.951	0.381	b2	300	Beta(2,2)
Hall's Percentile Bootstrap	0.940	0.384	g2	300	Beta(2,2)
Hall's Percentile Bootstrap	0.944	0.390	G_2	300	Beta(2,2)
Hall's Percentile Bootstrap	0.927	0.381	b2	300	Beta(2,2)

at least 94% coverage, regardless of confidence interval method used. On the other hand, out of all three estimators, g_2 performs the worst in almost every case. From these observations, we can say that for the normal distribution, the best method in estimating the true kurtosis parameter is to use the classical method with G_2 estimator.

4.2.2. Negative Kurtosis

To assess performance of estimators on distributions with negative kurtosis, we simulated data from *Beta(2,2)* with excess kurtosis Kurt(X) = -0.8571429 and results are reported in Table **2**. As for interval average width, the higher the sample size, the shorter the intervals, as expected. Also, regardless of methods, the average width inequality $b_2 \leq g_2 \leq G_2$ for all three estimator holds since this was guaranteed by equation (17). For large sample sizes ($n > 30$), the parametric method has higher average width when compared to non-parametric methods.

In terms of coverage probability, Efron's Percentile Bootstrap performs well and sometimes better than the

Table 3: Average Width and Coverage Probability of The Intervals When The Data Are Generated from logistic(0,1)

Method	Coverage Probability	Width	Estimator	Sample Size	Distribution
Bias Corrected Standard Bootstrap	0.5773	3.3279	g2	30	Logistic(0,1)
Bias Corrected Standard Bootstrap	0.6850	3.9585	G_2	30	Logistic(0,1)
Bias Corrected Standard Bootstrap	0.5137	3.1085	b2	30	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.7530	4.0430	g2	30	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.8180	4.8215	G_2	30	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.6897	3.7799	b2	30	Logistic(0,1)
Classical	0.5900	2.7451	g2	30	Logistic(0,1)
Classical	0.7573	3.2643	G_2	30	Logistic(0,1)
Classical	0.4560	2.5651	b2	30	Logistic(0,1)
Efron's Percentile Bootstrap	0.6220	3.2284	g2	30	Logistic(0,1)
Efron's Percentile Bootstrap	0.7427	3.8376	G_2	30	Logistic(0,1)
Efron's Percentile Bootstrap	0.5400	3.0187	b2	30	Logistic(0,1)
Hall's Percentile Bootstrap	0.5090	3.2410	g2	30	Logistic(0,1)
Hall's Percentile Bootstrap	0.6100	3.8516	G_2	30	Logistic(0,1)
Hall's Percentile Bootstrap	0.4370	3.0290	b2	30	Logistic(0,1)
Bias Corrected Standard Bootstrap	0.6190	3.0854	g2	50	Logistic(0,1)
Bias Corrected Standard Bootstrap	0.6893	3.4202	G_2	50	Logistic(0,1)
Bias Corrected Standard Bootstrap	0.5737	2.9659	b2	50	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.7293	3.3831	g2	50	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.7793	3.7516	G_2	50	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.6827	3.2458	b2	50	Logistic(0,1)
Classical	0.5800	2.3423	g2	50	Logistic(0,1)
Classical	0.6887	2.5946	G_2	50	Logistic(0,1)
Classical	0.5010	2.2496	b2	50	Logistic(0,1)
Efron's Percentile Bootstrap	0.6377	2.9113	g2	50	Logistic(0,1)
Efron's Percentile Bootstrap	0.7200	3.2268	G_2	50	Logistic(0,1)
Efron's Percentile Bootstrap	0.5800	2.7992	b2	50	Logistic(0,1)
Hall's Percentile Bootstrap	0.5643	2.9532	g2	50	Logistic(0,1)
Hall's Percentile Bootstrap	0.6270	3.2727	G_2	50	Logistic(0,1)
Hall's Percentile Bootstrap	0.5150	2.8361	b2	50	Logistic(0,1)
Bias Corrected Standard Bootstrap	0.6557	2.6514	g2	100	Logistic(0,1)
Bias Corrected Standard Bootstrap	0.6907	2.7890	G_2	100	Logistic(0,1)
Bias Corrected Standard Bootstrap	0.6307	2.5984	b2	100	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.7407	2.7079	g2	100	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.7670	2.8499	G_2	100	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.7140	2.6534	b2	100	Logistic(0,1)
Classical	0.5693	1.7826	g2	100	Logistic(0,1)
Classical	0.6117	1.8750	G_2	100	Logistic(0,1)
Classical	0.5310	1.7471	b2	100	Logistic(0,1)
Efron's Percentile Bootstrap	0.6793	2.5542	g2	100	Logistic(0,1)
Efron's Percentile Bootstrap	0.7247	2.6854	G_2	100	Logistic(0,1)
Efron's Percentile Bootstrap	0.6463	2.5043	b2	100	Logistic(0,1)
Hall's Percentile Bootstrap	0.6203	2.5317	g2	100	Logistic(0,1)
Hall's Percentile Bootstrap	0.6613	2.6629	G_2	100	Logistic(0,1)

(Table 3). Continued.

Method	Coverage Probability	Width	Estimator	Sample Size	Distribution
Hall's Percentile Bootstrap	0.5937	2.4781	b2	100	Logistic(0,1)
Bias Corrected Standard Bootstrap	0.7567	2.1501	g2	300	Logistic(0,1)
Bias Corrected Standard Bootstrap	0.7777	2.1868	G_2	300	Logistic(0,1)
Bias Corrected Standard Bootstrap	0.7510	2.1366	b2	300	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.7933	2.1595	g2	300	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.8067	2.1926	G_2	300	Logistic(0,1)
Bias Corrected Percentile Bootstrap	0.7830	2.1429	b2	300	Logistic(0,1)
Classical	0.5060	1.0814	g2	300	Logistic(0,1)
Classical	0.5277	1.0997	G_2	300	Logistic(0,1)
Classical	0.4893	1.0742	b2	300	Logistic(0,1)
Efron's Percentile Bootstrap	0.7460	2.0395	g2	300	Logistic(0,1)
Efron's Percentile Bootstrap	0.7650	2.0750	G_2	300	Logistic(0,1)
Efron's Percentile Bootstrap	0.7357	2.0274	b2	300	Logistic(0,1)
Hall's Percentile Bootstrap	0.7230	2.0705	g2	300	Logistic(0,1)
Hall's Percentile Bootstrap	0.7383	2.1050	G_2	300	Logistic(0,1)
Hall's Percentile Bootstrap	0.7160	2.0559	b2	300	Logistic(0,1)

classical method. The advantage of Efron's Percentile Bootstrap is that its average interval is always less than that of the classical method or any other non-parametric methods regardless of sample size or estimators. Thus, we can say that Efron's Percentile Bootstrap process is the best method when it comes to constructing confidence intervals for $Beta(2,2)$.

4.2.3. Positive Kurtosis

To assess performance of estimators with positive kurtosis, we simulated data from the standard logistic distribution and results are presented in Table **3**. Average width of the classical method is longer when compared to all other bootstrap confidence interval methods for sample sizes less than or equal to 50 and its coverage probability is also slightly higher than all other bootstrap methods. In constructing bootstrap intervals, estimator G_2 always generates coverage probability that is greater than or equal to coverage probability of the other two estimators. As for average width, it was observed that, for samples $n \geq 50$, the classical method performs a lot worst when it is compared to any of the bootstrap methods simulated. And in every case, we see that the coverage probability never reached its 95% threshold. Comparing the classical method and bootstrap methods, we observed that the bootstrap methods do have higher coverage probability, but none of these confidence interval methods consistently meet their 95% threshold. When it comes to positive kurtosis estimators, there are no clear winners since all methods failed to meet the 95% threshold. But Efron's method as well as Bias Corrected Percentile bootstrap do get closer than most. Last, in choosing an estimator, it is recommended to always use G_2 since it consistently will have higher coverage intervals guaranteed by equation (17).

5. APPLICATIONS

We will consider two real life health related data to illustrate the findings of the paper in this section.

5.1. Healthy Bones and PTH Data:

In this example, we consider the values of parathyroid hormone (PTH) measured on a sample of 30 boys and girls aged between 12 to 15 years. (Source, Moore, McCabe and Craig (2012), page 19). There was one missing value, so we have 29 observations to construct confidence intervals for the true Kurtosis parameter of the PTH data.

39, 59, 30, 48, 71, 31, 25, 31, 71, 50, 38, 63, 49, 45, 31, 33, 28, 40, 127, 49, 59, 50, 64, 28, 46, 35, 28, 19, 29

The histogram of the PTH data is given in Figure **1**, and it showed that the PTH data is right skewed. Thus, the PTH data was fitted to a log normal distribution with mean 3.731 and variance 0.161 and ks test give p-value=0.83. Therefore, there is no evidence against the log-normal distribution assumption. The excess kurtosis parameter for lognormal with variance 0.16, is

$$Kurt(X) = e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 6$$

Kurt(X) =3.26

The R code for estimating parameters is, **fitdistr(x,"log-normal")$estimate** and testing for log-normal distribution is, **ks.test(x, "plnorm", 2.992, 0.041)**. The confidence intervals and widths of the PTH data are provided in Table **4**. We observed that all the intervals but classical do not capture the true excess kurtosis, which is 3.26. Among those captured the true excess kurtosis, b2Hall has the shortest width followed by b2Efron, g2Hall.

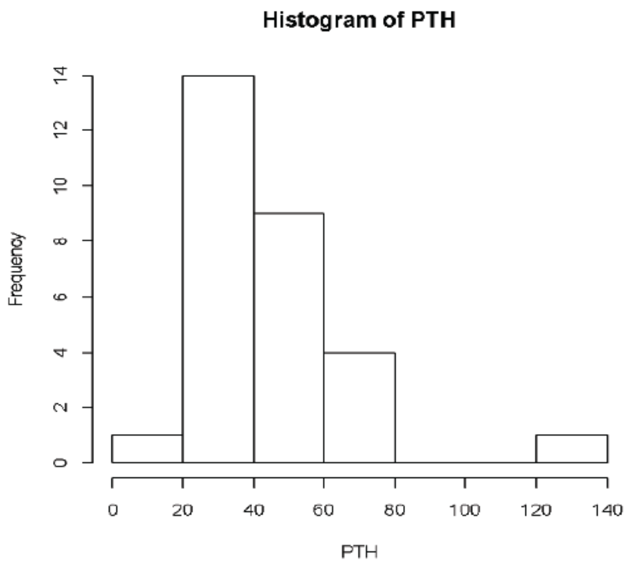

Figure 1: Histogram of the PTH data.

Table 4: The Upper and Lower Bounds and Width of the Intervals

	Lower Bound	Upper Bound	Width
g2Classic	4.14719	6.91614	2.76895
G2Classic	5.20176	8.51504	3.31328
b2Classic	3.8661	6.4phht4739	2.58128
g2Hall	2.02021	12.3694	10.3492
G2Hall	3.0743	15.0534	11.9791
b2Hall	2.18435	11.5375	9.35317
g2Efron	-1.3036	9.08888	10.3925
G2Efron	-1.3122	11.3518	12.664
b2Efron	-1.2188	8.20284	9.42167
g2BCorrPerc	0.624	14.2955	13.6715
G2BCorrPerc	1.26003	15.6546	14.3946
b2BCorrPerc	-0.2520	11.1544	11.4069
g2BCorr	1.99452	14.4326	12.4381
G2BCorr	2.32695	17.7747	15.4477
b2BCorr	1.95643	13.3926	11.4362

5.2. Diabetes and Glucose Data:

Here we consider the fasting plasma glucose levels (mg/dl) for 18 diabetics enrolled in a diabetes control class, five months after the end of class. (Source: Debora L. Arsenau "Comparison of diet management instruction for patients with non-insulin dependent diabetes: Learning activity packages vs. group instruction. [14, page 25].

141, 158, 112, 153, 134, 95, 96, 78, 148, 172, 200, 271, 103, 172, 359, 145, 147, 255

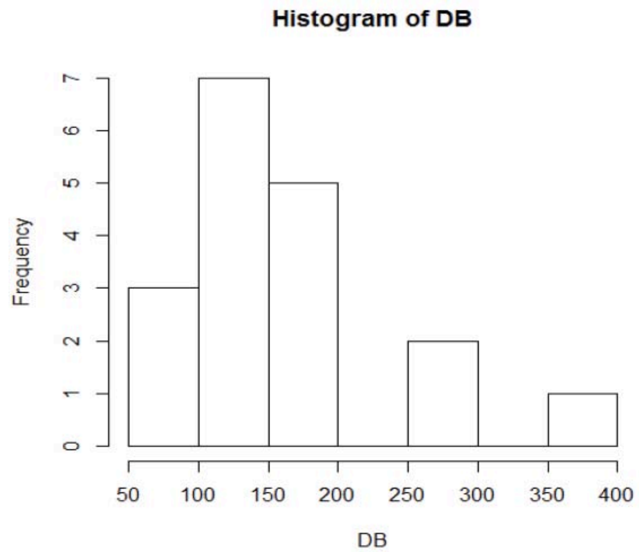

Figure 2: Histogram of the Diabetes and Glucose.

The histogram of the diabetes and glucose data is given in Figure **2**, and it showed that the data is right skewed.

The data was fitted to a gamma distribution with shape parameter 6.75 and rate 0.041 and KS-test give p-value=0.59. Thus, we may assume that the data follow a gamma distribution. Then the population kurtosis will be 6/6.751=0.89. We have computed the lower and upper bound and the width of the intervals and provided them in Table **5**. It appears that all the proposed estimators captured the true excess kurtosis 0.89. However, we obtain the shortest interval by b_2 classical interval followed by $g_2$classical, G_2 classical, b_2 Hall, g_2 Hall and b_2 Efron's method.

Table 5: The Upper and Lower Bounds and Width of the Intervals

	Lower Bound	Upper Bound	Width
g2Classic	-0.0352	2.98769	3.02272
G2Classic	0.3781	4.44593	4.06808
b2Classic	-0.0313	2.66494	2.69619
g2Hall	-2.5db304	4.14826	6.67869
G2Hall	-3.5907	5.90345	9.49411
b2Hall	-2.1887	3.6919	5.88058
g2Efron	-1.2206	5.68061	6.90121
G2Efron	-1.1382	8.3246	9.46283
b2Efron	-1.0823	4.91494	5.99726
g2BCorrPerc	-0.311	8.48216	8.79303
G2BCorrPerc	-0.0272	12.2041	12.2313
b2BCorrPerc	-0.37	6.9432	7.31283
g2BCorr	-1.4399	5.35257	6.79249
G2BCorr	-1.4034	7.59479	8.99823
b2BCorr	-1.2378	4.779	6.01678

6. SOME CONCLUDING REMARKS

This paper considers some classical confidence intervals based on g_2, G_2 and b_2 kurtosis estimators and bootstrap methods in constructing confidence intervals for the kurtosis parameter of a distribution. The bootstrap techniques used are: Bias-Corrected Standard Bootstrap, Efron's Percentile Bootstrap, Hall's Percentile Bootstrap and Bias-Corrected Percentile Bootstrap. Since a theoretical comparison is not possible, a simulation study was conducted to compare the performance of the interval estimators. The criteria of performance were judged on which had shorter intervals with coverage probability of at least meeting the 95% confidence threshold. We saw that when dealing with a normal distribution with kurtosis 0, the classical method with parameter b_2 performs best since it generates the smallest interval while meeting its 95% threshold. When dealing with beta(2,2), a distribution with negative kurtosis, the classical method did perform well. We observed that Efron's Percentile Method performed equally as well, but with shorter intervals. For the normal logistic distribution with positive kurtosis, it is best to use Efron's Percentile Method with sample estimator G_2 even though all methods had a hard time reaching that 95% coverage threshold. To illustrate the finding of this paper, two real life health related data are analyzed and the results supported the simulation study to some extent. Hope the findings of this paper will be useful for the researchers in various fields, not limited to health and medical sciences.

ACKNOWLEDGEMENTS

We thank the anonymous referee for the useful suggestions which have improved the quality of this paper.

REFERENCES

[1] Van Belle G, *et al*. Biostatistics: a methodology for the health sciences. John Wiley & Sons, Watkins, 2004; Vol. 519: p. 51. https://doi.org/10.1002/0471602396

[2] Westfall PH. Kurtosis as peakedness, 1905–2014. RIP. The American Statistician 2014; 68(3): 191-195. https://doi.org/10.1080/00031305.2014.917055

[3] Cramér H. Mathematical methods of statistics. Princeton University Press. Princeton 1991.

[4] DeCarlo LT. On the meaning and use of kurtosis. Psychological Methods 1997; 2(3): 292. https://doi.org/10.1037/1082-989X.2.3.292

[5] Fisher RA. The moments of the distribution for normal samples of measures of departure from normality. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. The Royal Society 1930; Vol. 130. 812: pp. 16-28. https://doi.org/10.1098/rspa.1930.0185

[6] Joanes DN, Gill CA. Comparing measures of sample skewness and kurtosis. Journal of the Royal Statistical Society: Series D (The Statistician) 1998; 47(1): 183-189. https://doi.org/10.1111/1467-9884.00122

[7] Bruin J. newtest: command to compute new test @ONLINE 2011. URL: http://stats.idre. ucla.edu/stata/ado/analysis/.

[8] DiCiccio TJ, Efron B. Bootstrap confidence intervals. Statistical Sciences 1996; 11(3): 189-228. https://doi.org/10.1214/ss/1032280214

[9] Efron B. Better bootstrap confidence intervals. Journal of the American statistical Association 1987; 82(397): pp. 171-185. BIBLIOGRAPHY 82

[10] Efron B. Bootstrap methods: another look at the jackknife. The annals of Statistics 1979; 1-26. https://doi.org/10.1214/aos/1176344552

[11] Efron B. Nonparametric standard errors and confidence intervals. Canadian Journal of Statistics 1981; 9(2): 139-158. https://doi.org/10.2307/3314608

[12] Hall P. The bootstrap and Edgeworth expansion. Springer Science & Business Media 2013.

[13] Banik S, Kibria BMG. A simulation study on some confidence intervals for the population standard deviation. SORT 2011; 35(2): 83-102.

[14] Moore D, McCabe GP, Craig BA. Introduction to the Practice of Statistics. W. H. Freeman and Company, New York 2012.

Received on 18-04-2018 Accepted on 23-04-2018 Published on 25-06-2018

https://doi.org/10.6000/1929-6029.2018.07.03.3