How Large Does n Have to Be for Z and t Intervals?

Dennis D. BOOS and Jacqueline M. HUGHES-OLIVER *

Abstract

Students invariably ask the question “How large does n have to be for Z and t intervals to give appropriate coverage probabilities?” In this article we review the role of $\sqrt{\beta_1}(X)/\sqrt{n}$, where $\sqrt{\beta_1}(X)$ is the skewness coefficient of the random sample, in the answer to this question. We also comment on the opposite effect that $\sqrt{\beta_1}(X)$ has on the behavior of t intervals compared to Z intervals. Finally, we suggest a simple exercise for determining rules of thumb for n that result in appropriate confidence interval coverage.

KEY WORDS: Confidence interval; Convergence to normality; Central Limit Theorem; Edgeworth expansion; Kurtosis; Skewness; t statistic.

Institute of Statistics Mimeo Series #2506

February 1998

*Dennis D. Boos is Professor and Jacqueline M. Hughes-Oliver is Assistant Professor, Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203. Email addresses are: boos@unity.ncsu.edu, hughesol@unity.ncsu.edu

1. INTRODUCTION

In many courses we typically present the Central Limit Theorem (CLT) and related “Z” intervals

$$\left(\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right), \quad (1)$$

followed by the t statistic and related “t” intervals

$$\left(\bar{X} - t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2, n-1} \frac{S}{\sqrt{n}} \right). \quad (2)$$

These formulas assume a random sample $X_1, \dots, X_n$ with $E(X_i) = \mu$, $\text{var}(X_i) = \sigma^2 < \infty$, $\bar{X}$ is the sample mean, S^2 is the sample variance, and $z_{\alpha/2}$ and $t_{\alpha/2, n-1}$ are the $1-\alpha/2$ quantiles of the standard normal and t with $n-1$ degrees of freedom distributions, respectively.

As instructors, we make the point that the intervals (1) and (2) have exact $1-\alpha$ coverage for normally distributed data and approximate $1-\alpha$ coverage for non-normal data, where the approximation improves with increasing n . Students invariably then ask “How large does n have to be?”

The answer given might be “30.” But a proper answer is that it depends mostly on the skewness of the X density (and to a lesser degree on kurtosis and other aspects of non-normality). In introductory courses it usually suffices to mention skewness and to give a few histograms of $Z_n = \sqrt{n}(\bar{X} - \mu)/\sigma$ and $t_n = \sqrt{n}(\bar{X} - \mu)/S$ for, say, a symmetric density such as the uniform and a skewed density such as the exponential. But in later courses one may want to give more detail.

To demonstrate that the speed of convergence of Z_n to normality cannot be expressed solely in terms of n , consider the following simple example. Let $X_1, \dots, X_{10}$ be independent and identically distributed random variables and define $Y_1 = X_1 + X_2, \dots, Y_5 = X_9 + X_{10}$. It is clear that $Z_{X,10}$ based on the X ’s is exactly the same as $Z_{Y,5}$ based on the Y ’s. Consequently, $Z_{X,10}$ and $Z_{Y,5}$ have equal speed of convergence, even though the sample “sizes” are different.

The Pearson skewness coefficient $\sqrt{\beta_1}(X) = E\{X - E(X)\}^3 / \{\text{var}(X)\}^{3/2}$ offers additional information on the speed of convergence of Z_n . Since any normal distribution has $\sqrt{\beta_1}(X) = 0$, we can compute the skewness of Z_n and compare it to 0. To calculate $\sqrt{\beta_1}(Z_n)$, use the facts that $\sqrt{\beta_1}(a + bX) = \sqrt{\beta_1}(X)$ and $E\{\bar{X} - E(X)\}^3 = n^{-3} \sum E\{X_i - E(X)\}^3$ to obtain

$$\sqrt{\beta_1}(Z_n) = \sqrt{\beta_1}(\bar{X}) = \sqrt{\beta_1}(X) / \sqrt{n}. \quad (3)$$

Similarly, the skewness coefficient of the Y 's in the above example is equal to the skewness coefficient of the X 's divided by $\sqrt{2}$. As a result, we see that

$$\sqrt{\beta_1}(Z_{Y,5}) = \sqrt{\beta_1}(\bar{Y}) = (\sqrt{\beta_1}(X) / \sqrt{2}) / \sqrt{5} = \sqrt{\beta_1}(X) / \sqrt{10} = \sqrt{\beta_1}(\bar{X}) = \sqrt{\beta_1}(Z_{X,10}).$$

In other words, the skewness coefficient of $Z_{X,10}$ based on the X 's is the same as the skewness coefficient of $Z_{Y,5}$ based on the Y 's. Thus it appears that the quantity $\sqrt{\beta_1}(X) / \sqrt{n}$ in (3) is more important than n in assessing the convergence of Z_n to normality.

In this paper we: (1) explain the ubiquitous role of $\sqrt{\beta_1}(X) / \sqrt{n}$ in the convergence of both Z_n and t_n , (2) explain the dramatic difference between “Z” and “t” confidence intervals for skewed data, and (3) give some rough rules of thumb for n that result in appropriate confidence interval coverage.

In Section 2 we will discuss the effect of skewness on the CLT and on the associated “Z” confidence intervals. In Section 3 we will show how this effect is reversed for the t statistic and for “t” intervals. Finally, in Section 4 we will use regression for several common skewed distributions to obtain rules of thumb for how large n needs to be.

There is a very large literature on the CLT and on the t_n statistic. We cannot cite all the relevant books and articles, but we would like to mention Cressie (1980) and three recent articles that are closely related to this one and provide a good list of references: Bartkowiak and Sen (1992), Chaffin and Rhiel (1993), and Rhiel and Chaffin (1996).

2. SKEWNESS AND THE CLT

The central limit theorem tells us that

$$P(Z_n \leq t) \rightarrow P(Z \leq t) \text{ as } n \rightarrow \infty, \text{ for all } t \in (-\infty, \infty), \quad (4)$$

where Z is a standard normal random variable. We have already mentioned that the Pearson skewness coefficient is one possible way to quantify the speed of this convergence to normality. A heuristic argument for this is that $\sqrt{\beta_1}(Z_n) = \sqrt{\beta_1}(X)/\sqrt{n}$ approaches 0, the value corresponding to the standard normal skewness coefficient, as either $n \rightarrow \infty$ or $\sqrt{\beta_1}(X) \rightarrow 0$. Thus, in terms of the skewness coefficient, Z_n inherits the direction of skewness of the X distribution and converges to Z at a rate $O(1/\sqrt{n})$.

It is worth mentioning that the Pearson kurtosis coefficient $\beta_2(X) = E\{X - E(X)\}^4 / \{\text{var}(X)\}^2$ is another way to quantify convergence of Z_n to Z . Kurtosis is a measure of tail length or peakedness (see, for example, Ruppert 1987, and Balanda and MacGillivray 1988). Since all normal random variables including Z have $\beta_2 = 3$, we can compare

$$\beta_2(Z_n) = \beta_2(\bar{X}) = 3 + (\beta_2 - 3)/n$$

to $\beta_2 = 3$ and note that Z_n converges to Z in terms of kurtosis at a rate $O(1/n)$. Comparing rates of convergence, we can see that skewness ($O(1/\sqrt{n})$) is more important than kurtosis ($O(1/n)$) in quantifying the convergence of Z_n to the distribution of Z .

For a more direct reasoning that the convergence in (4) depends on $\sqrt{\beta_1}(X)/\sqrt{n}$, we consider the one-term Edgeworth expansion of Z_n :

$$P(Z_n \leq t) \approx P(Z \leq t) - \frac{\sqrt{\beta_1}(X)}{\sqrt{n}} \frac{(t^2 - 1)}{6} \phi(t), \quad (5)$$

where $\phi(t)$ is the standard normal density function (see, for example, Feller 1966, p. 539). The closer $\sqrt{\beta_1}(X)/\sqrt{n}$ is to 0, the closer the distribution of Z_n is to Z .

Recall that the skewness coefficient of Z_n , $\sqrt{\beta_1}(X)/\sqrt{n}$, shows that Z_n inherits the

skewness of the X 's (but diluted by $\sqrt{n}$): if the X 's are positively skewed, then so is Z_n ; if the X 's are negatively skewed, then so is Z_n . This fact is also illustrated by the Edgeworth expansion. Suppose the X 's are positively skewed. For standard normal quantiles $z_{\alpha/2} > 1$, the Edgeworth expansion gives $P(Z_n \leq z_{\alpha/2}) < 1 - \alpha/2$, which implies that the true $1 - \alpha/2$ quantile $\xi_{1-\alpha/2}$ of Z_n is larger than $z_{\alpha/2}$. Similarly, the true lower quantiles of Z_n will be larger than the associated standard normal quantiles.

For example, if the X 's have a standard exponential density, where $\sqrt{\beta_1}(X) = 2$, and $n = 10$, then the .975 quantile of Z_n is 2.24 instead of 1.96, and the .025 quantile is -1.65 instead of -1.96 . Figure 1 illustrates this intuition more simply by overlaying the approximate density of Z_n (obtained as the derivative of the Edgeworth expansion given in (5)) with the standard normal density. To demonstrate the adequacy of the Edgeworth expansion, Figure 1 also displays the true density of Z_n , which is easily obtained in this case. The effect of sample size is demonstrated since Figure 1 shows results for $n = 10$ and $n = 20$. The inherited positive skewness of Z_n is obvious.

Figure 1: *The convergence of Z_n from a standard exponential random sample, for (a) $n = 10$ and (b) $n = 20$. The curves are as follows: exact density of Z_n (—), derivative of one-term Edgeworth expansion given in (5) (---), standard normal density (···).*

A $(1 - \alpha)100\%$ confidence interval for μ based on the true distribution of Z_n is

$$\left(\bar{X} - \xi_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X} - \xi_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right). \quad (6)$$

Because $\xi_{1-\alpha/2} > z_{\alpha/2}$ and $\xi_{\alpha/2} > -z_{\alpha/2}$ for positively skewed X 's, the quantity subtracted from $\bar{X}$ to form the left endpoint of (1) is actually too small, and the quantity added to $\bar{X}$ to form the right endpoint is too large. Hence, the “Z” confidence interval found in equation (1) is to the right of where it should be.

Consider again a random sample of size $n = 10$ from the standard exponential distribution. The 95% interval obtained from (6) is

$$\left(\bar{X} - 2.24 \frac{\sigma}{\sqrt{10}}, \bar{X} + 1.65 \frac{\sigma}{\sqrt{10}} \right),$$

whereas the interval obtained from (1) is

$$\left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{10}}, \bar{X} + 1.96 \frac{\sigma}{\sqrt{10}} \right)$$

and thus located to the right of where it should be. Moreover, exact calculations yield that the interval obtained from (1) is totally to the right of μ on average .039 of the time (instead of .025) and totally to the left of μ on average .006 of the time. The overall coverage of $1 - (.039 + .006) = .955$ is, however, just fine and is even conservative.

So the Z_n distribution inherits the skewness direction of the X 's, and the corresponding Z_n intervals are displaced in the same direction (to the right if $\sqrt{\beta_1}(X)$ is positive and to the left if it is negative).

Unfortunately, we rarely know σ in real problems. Worse yet, the intuition gained from Z_n and skewness is opposite of the correct intuition for t_n .

3. SKEWNESS AND THE t_n STATISTIC

Although the t statistic $t_n = \sqrt{n}(\bar{X} - \mu)/S$ is well-known to be remarkably robust to non-normality, it is still sensitive to highly skewed distributions (Johnson 1978).

First let us consider the skewness coefficient of t_n , $\sqrt{\beta_1}(t_n)$. Since the second and third moments of t_n are not easy to calculate, we simulated 10,000 samples of size $n = 10$ and $n = 20$ from a standard exponential density and obtained $\widehat{\sqrt{\beta_1}}(t_n) = -1.69$ and -1.11 , respectively. Compared to the skewness coefficient of Z_n for these situations, $2/\sqrt{10} = 0.63$ and $2/\sqrt{20} = 0.45$, we see that the skewness coefficient of t_n is larger in magnitude and has the *opposite sign*.

Another way to see this reversal is from the one-term Edgeworth expansion for t_n (see, for example, Hall 1987)

$$P(t_n \leq t) \approx P(Z \leq t) + \frac{\sqrt{\beta_1}(X)(2t^2 + 1)}{\sqrt{n}} \frac{\phi(t)}{6}. \quad (7)$$

First note the reappearance of $\sqrt{\beta_1}(X)/\sqrt{n}$. Compared to the related expansion (5) for Z_n , we see that the correction is of the opposite sign so that we should expect the t_n distribution for positively skewed X 's to be negatively skewed. In Figure 2 we plot the derivative of the above expansion for a random sample from a standard exponential distribution where $\sqrt{\beta_1}(X) = 2$, for $n = 10$ and $n = 20$. The t_9 and t_{19} densities are also shown. The negative skewness of t_n is obvious.

Why does the distribution of t_n have a skewness that is opposite to the skewness of the X 's? The asymptotic correlation between $\bar{X}$ and S^2 is (as the reader might now expect) $\sqrt{\beta_1}(X)/\sqrt{n}$. Thus, when the X 's are positively skewed, then $\bar{X}$ and S^2 are positively correlated, and a large random occurrence for $\bar{X}$ tends to be counteracted by a large S^2 resulting in a lower value for t_n than for Z_n . On the other hand, when $\bar{X} - \mu$ is small and negative, then S^2 tends to be small (it is bounded by 0), and inflates the negative value of t_n , stretching the tail of the distribution to the left.

Figure 2: *The convergence of t_n from a standard exponential random sample, for (a) $n = 10$ and (b) $n = 20$. The curves are as follows: derivative of one-term Edgeworth expansion given in (6) (——), density of the t -distribution having $n - 1$ degrees of freedom (---).*

A $(1 - \alpha)100\%$ confidence interval for μ based on the true distribution of t_n is

$$\left(\bar{X} - \tau_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X} - \tau_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right),$$

where τ_p is the $p100^{th}$ percentile of t_n . Because $\tau_{1-\alpha/2} < t_{\alpha/2, n-1}$ and $\tau_{\alpha/2} < -t_{\alpha/2, n-1}$ for positively skewed X 's, the t confidence interval given by (2) is to the left of where it should be. For example, for standard exponential data with $\alpha = .05$ and $n = 10$, a simulation of 10,000 data sets results in .004 missing on the right of μ (interval totally to the right of μ) and 0.093 missing on the left of μ (interval totally to the left of μ). (Recall that the exact values for Z_n were .039 and .006, respectively.)

Thus, skewness has the opposite effect on t_n as it has on Z_n . If the X 's are positively skewed, then the distribution of t_n is negatively skewed, and confidence intervals are to the left of where they should be in order to have symmetric coverage; if the X 's are negatively skewed, then t_n is positively skewed and intervals are to the right of where they should be. This property of “ t ” intervals is known (see, for example, Cressie 1980 and Johnson 1978) but

not widely taught. Although the reverse property of “Z” intervals is more straightforward, it is also not widely taught. We suggest that these properties be discussed simultaneously in order to emphasize the opposite effect that skewness has on intervals (1) and (2).

Skewness also has a larger effect on t_n than on Z_n , since the adjustment term in (7) is larger (in absolute value) than the adjustment term in (5). Hence the convergence of t_n is slower than the convergence of Z_n .

4. HOW LARGE SHOULD n BE?

Our concern here is to be able to suggest sample sizes that will ensure good coverage properties for the t_n -based confidence interval (2). It would be simple to work with the Edgeworth expansion (7) for t_n , but unfortunately it is not very accurate, as seen in Figure 2. One reason is that the expansion is taken around the limiting standard normal distribution rather than around the approximating t distribution with $n - 1$ degrees of freedom.

Instead we take a very simple empirical approach that can easily be given as a class assignment. We simulate samples of sizes $n = 10$ and $n = 30$ from some skewed distributions and record estimated coverage probabilities for one-sided and two-sided intervals based on 10,000 simulation replicates. The distributions used are: the gamma distribution with shape parameter $\gamma = 1, 1.78, 4,$ and 16 having respective skewness coefficients $\sqrt{\beta_1}(X) = 2/\sqrt{\gamma}$ of $2, 1.5, 1.0,$ and 0.5 ; the Weibull distribution with shape parameter $c = 1.2, 1.6,$ and 2.2 having respective skewness coefficients $1.52, 0.96, 0.51$ (see p. 633 of Johnson, Kotz, and Balakrishnan, 1994); and the Tukey-Lambda distributions 7, 8, and 12 of Table 1 of Randles, *et al.* (1980) having respective skewness coefficients $0.5, 1.5,$ and 2.0 . The results for one-sided $\alpha = .05$ and two-sided $\alpha = .10$ intervals are displayed in Table 1 and plotted in Figure 3.

Regressing one-sided error rates of nominal level $\alpha = .05$ on $\sqrt{\beta_1}(X)/\sqrt{n}$ we obtain

Figure 3: *Estimated error rates of t intervals as a function of $\sqrt{\beta_1}(X)/\sqrt{n}$. The error rates for two-sided 90% intervals are shown in (a) and the error rates for one-sided 95% intervals are shown in (b). In (b), $\circ$ means the interval is on the left of the true μ and $\bullet$ means the interval is on the right of μ . The solid lines are linear least squares regression lines constrained to cross the vertical axis at the nominal values.*

(using least squares for $y = \text{error rate} - .05$ with no intercept)

$$\text{miss right} = .050 - .067 \frac{\sqrt{\beta_1}}{\sqrt{n}}$$

$$\text{miss left} = .050 + .116 \frac{\sqrt{\beta_1}}{\sqrt{n}}$$

and for the two-sided $\alpha = .10$

$$\text{miss right or left} = .100 + .050 \frac{\sqrt{\beta_1}}{\sqrt{n}},$$

where the standard errors of the estimated slopes are, respectively, .003, .006, and .005. A similar exercise for two-sided error rates of nominal level $\alpha = .05$ leads to the regression model

$$\text{miss right or left} = .05 + .051 \frac{\sqrt{\beta_1}}{\sqrt{n}},$$

with a standard error of .005 for the estimated slope.

Table 1: *Estimated Error Rates for t Intervals*

Distribution	n	$\sqrt{\beta_1}(X)$	$\beta_2(X)$	$\sqrt{\beta_1}(t_n)$	$\beta_2(t_n)$	$\sqrt{\beta_1}(X)/\sqrt{n}$	Miss	Miss	Miss
							Right $\alpha = .05$	Left $\alpha = .05$	Total $\alpha = .10$
Gamma ($\gamma = 1.0$)	10	2.00	9.00	-1.65	8.77	0.63	0.014	0.131	0.145
	30	2.00	9.00	-0.85	4.58	0.37	0.020	0.098	0.119
Gamma ($\gamma = 1.78$)	10	1.50	6.38	-1.25	7.52	0.47	0.020	0.107	0.127
	30	1.50	6.38	-0.56	3.93	0.27	0.026	0.082	0.109
Gamma ($\gamma = 4.0$)	10	1.00	4.50	-0.70	4.80	0.32	0.027	0.086	0.112
	30	1.00	4.50	-0.44	3.59	0.18	0.032	0.068	0.100
Gamma ($\gamma = 16.0$)	10	0.50	3.38	-0.52	4.65	0.16	0.033	0.069	0.102
	30	0.50	3.38	-0.20	3.38	0.09	0.041	0.057	0.098
Weibull (c=1.2)	10	1.52	6.24	-1.47	7.95	0.48	0.018	0.115	0.133
	30	1.52	6.24	-0.68	4.33	0.28	0.028	0.089	0.117
Weibull (c=1.6)	10	0.96	4.04	-0.76	5.10	0.30	0.028	0.087	0.115
	30	0.96	4.04	-0.44	3.64	0.18	0.030	0.073	0.102
Weibull (c=2.2)	10	0.51	3.04	-0.51	4.87	0.16	0.037	0.073	0.110
	30	0.51	3.04	-0.26	3.43	0.09	0.040	0.065	0.105
Tukey-Lambda (#7)	10	0.50	2.20	-0.97	7.75	0.16	0.035	0.070	0.105
	30	0.50	2.20	-0.16	3.34	0.09	0.045	0.057	0.102
Tukey-Lambda (#8)	10	1.50	5.80	-1.58	8.85	0.47	0.016	0.115	0.131
	30	1.50	5.80	-0.69	4.14	0.27	0.026	0.090	0.116
Tukey-Lambda (#12)	10	2.00	21.20	-0.71	4.35	0.63	0.020	0.092	0.112
	30	2.00	21.20	-0.39	3.23	0.37	0.030	0.086	0.116

One way to use these regressions is as follows. Suppose that we would be happy if the coverage probability for a one-sided $\alpha = .05$ interval (upper bound for μ) was at least .94. Setting $.06 = .050 + .116\sqrt{\beta_1}/\sqrt{n}$ and solving for $\sqrt{\beta_1}/\sqrt{n}$ yields $\sqrt{\beta_1}/\sqrt{n} = .086$. Thus if we have exponential data ($\sqrt{\beta_1} = 2$), then it would require a sample size of $n = (2/.086)^2 = 538$. That is pretty unrealistic but illustrates the slow convergence for the worst case.

Instead let us consider the two-sided interval (2) that is most-used. If .88 coverage of nominal .90 intervals is acceptable and our distribution has $\sqrt{\beta_1} = 2$, then we find that $n = 25$ is required. For more moderate skewness, say $\sqrt{\beta_1} = 1.0$, .88 coverage requires only $n = 7$, and .89 coverage requires only $n = 25$. For milder skewness, say $\sqrt{\beta_1} = .5$, .89 coverage requires only $n = 7$. Similarly, if .94 coverage of nominal .95 intervals is acceptable, then $\sqrt{\beta_1} = 2$ requires $n = 104$, $\sqrt{\beta_1} = 1$ requires $n = 26$, $\sqrt{\beta_1} = .5$ requires $n = 7$, and, in general, we require $n \geq 26.01\beta_1$. Cochran's (1977, p. 42) recommendation of $n > 25\beta_1$ is specific to .94 coverage of nominal .95 intervals and is in general agreement to the empirical

findings presented here.

So a rule of thumb for two-sided intervals might be to require $n \geq 30$ for very skewed data, $n \geq 20$ for moderately skewed data, and $n \geq 10$ for mildly skewed data; the required sample size increases for minimum acceptable error less than .02. These are rough generalizations and we encourage readers to find their own rules.

REFERENCES

- Balanda, K. P., and MacGillivray, H. L. (1988), "Kurtosis: A Critical Review," *The American Statistician*, 42, 111-119.
- Bartkowiak, A., and Sen, A. R. (1992), "Minimum Sample Size Ensuring Validity of Classical Confidence Intervals for Means of Skewed and Platykurtic Distributions," *Biometrical Journal*, 34, 367-382.
- Chaffin, W. W., and Rhiel, G. S. (1993), "The Effect of Skewness and Kurtosis on the One-Sample T test and the Impact of Knowledge of the Population Standard Deviation," *Journal of Statistical Computation and Simulation*, 46, 79-90.
- Cochran, W. G. (1977), *Sampling Techniques*, (third edition), Wiley: New York.
- Cressie, N. (1980), "Relaxing Assumptions in the One Sample t -test," *Australian Journal of Statistics*, 22, 143-153.
- Feller, W. (1966), *An Introduction to Probability Theory and Its Applications*, Vol. II, (2nd Edition), Wiley: New York.
- Hall, P. (1987), "Edgeworth Expansion for Student's t statistic under minimal moment conditions," *Annals of Probability*, 15, 920-931.
- Johnson, J. J. (1978), "Modified t Tests and Confidence Intervals for Asymmetrical Populations," *Journal of the American Statistical Association*, 73, 536-544.
- Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994), *Continuous Univariate Distributions*, Vol. 1, (2nd Edition), Wiley.
- Randles, R. H., Fligner, M. A., Policello II, G. E., and Wolfe, D. A. (1980), "An Asymptotically Distribution-Free Test for Symmetry Versus Asymmetry," *Journal of The American Statistical Association*, 75, 168-172.

- Rhiel, G. S., and Chaffin, W. W. (1996), "An Investigation of the Large-Sample/Small-Sample Approach to the One-Sample Test for a Mean (Sigma Unknown)," *Journal of Statistics Education*, 4, No. 3.
- Ruppert, D. (1987), "What is Kurtosis? An Influence Function Approach," *The American Statistician*, 41, 1-5.