Journal of the Korean
Data & Information Science Society
2001, Vol. 12, No.1, pp. 103 ~ 125

SAMPLE ENTROPY IN ESTIMATING THE BOX-COX TRANSFORMATION

Mezbahur Rahman [1] · Larry M. Pearson [2]

Abstract

The Box-Cox transformation is a well known family of power transformation that brings a set of data into agreement with the normality assumption of the residuals and hence the response variable of a postulated model in regression analysis. This paper proposes a new method for estimating the Box-Cox transformation using maximization of the Sample Entropy statistic which forces the data to get closer to normal as much as possible. A comparative study of the proposed procedure with the maximum likelihood procedure, the procedure via artificial regression estimation, and the recently introduced maximization of the Shapiro-Francia W' statistic procedure is given. In addition, we generate a table for the optimal spacings parameter in computing the Sample Entropy statistic.

Key Words and Phrases: Artificial regression model; Normal likelihood estimation; Regression analysis; Shapiro-Francia W' statistic; Shapiro-Wilk W statistic.

1. Introduction

1. INTRODUCTION

In regression analysis, often the key assumption regarding normality of the error variable and hence the response variable are violated. The commonly used remedy is the Box-Cox family of power transformations (Box and Cox (1964)). The process is to select a parameter in the Box-Cox transformation which maximizes the normal likelihood using the data at hand and then apply regression analysis on the transformed response variable. There is no role of the estimates of the location and the scale parameters which were derived in the process of estimating the power transformation parameter in regression analysis. In practice, the regression model parameters are usually estimated seperately after the necessary Box-Cox power transformation parameter is selected.

[1] Minnesota State University
[2] Mankato, MN 56001, USA

In literature, the estimation procedures of the Box-Cox power transformation parameter are considered by many authors. But the notable ones are the normal likelihood method of Box and Cox (1964), the robustified version of the normal likelihood method of Carroll (1980) and of Bickel and Doksum (1981), the transformation to symmetry method of Hinkley (1975), the quick estimate of Hinkley (1977) and of Taylor (1985). Lin and Vonesh (1989) constructed a nonlinear regression model which is used to estimate the transformation parameter such that the normal probability plot of the data on the transformed scale is as close to linearity as possible. Recently, following the footsteps of Box and Cox (1982) and Lin and Vonesh (1989), Halawa (1996) considered the power transformation parameter estimation procedure using an artificial regression model which gives the estimates with very small variabilities compared to the normal likelihood procedure. Halawa (1996) conducted an exhaustive comparative study with normal likelihood procedure. In that study, he also considered estimation procedures of the location and the scale parameters in the likelihood.

Most recently, Rahman (1999) introduced a method of estimating the Box-Cox power transformation parameter using maximization of Shapiro-Francia W' (Shapiro and Francia (1972)) statistic along with a comparison study of the normal likelihood method (Carroll (1980)), and of the artificial regression model method (Halawa (1996)). In this paper the estimation procedure for the Box-Cox power transformation parameter is considered, using maximization of the Sample Entropy statistic which is also independent of the normal location and scale parameters. This leads to a transformed distribution that is closer to a normal distribution in terms of the higher value of the Sample Entropy statistic and of the Shapiro-Francia W' statistic.

2. BOX-COX TRANSFORMATION

Let $Y_1, Y_2, \cdots, Y_n$ be a random sample of size n from a population whose functional form is unknown. Box and Cox (1964) suggested that if the transformation

$$X = \begin{cases} \frac{Y^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln(Y), & \lambda = 0 \end{cases} \tag{1}$$

is performed on the data then X will have an approximate normal distribution with mean μ and variance σ^2. In equation (1), λ is unknown and considered as the Box-Cox power transformation parameter and 'ln' represents the natural logarithm.

3. NORMAL LIKELIHOOD ESTIMATOR

After applying the transformation mentioned in equation (1), the density function of the data can be written as

$$f(y;\lambda,\mu,\sigma) \doteq \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}\left(\frac{y^\lambda-1}{\lambda}-\mu\right)^2} \cdot y^{\lambda-1}, & \lambda \neq 0, \\ \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(\ln(y)-\mu)^2} \cdot \frac{1}{y}, & \lambda = 0. \end{cases} \tag{2}$$

The log-likelihood function $\ell_L = \ell(\lambda,\mu,\sigma; y_1, y_2, \cdots, y_n)$, where the subscript 'L' stands for the normal-likelihood, can be written as

$$\ell_L = \begin{cases} -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^n \left(\frac{y_i^\lambda-1}{\lambda}-\mu\right)^2 + (\lambda-1)\sum_{i=1}^n \ln(y_i), & \lambda \neq 0, \\ -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^n (\ln(y_i)-\mu)^2 - \sum_{i=1}^n \ln(y_i), & \lambda = 0. \end{cases} \tag{3}$$

Equation (3) is maximized when the partial derivatives of (3) with respect to μ, σ and λ are equated to zero and the solution to the corresponding system is found. This leads to solving

$$\sum_{i=1}^n \left(\frac{y_i^{\hat{\lambda}_L}-1}{\hat{\lambda}_L} - \hat{\mu}_L\right)\left(\frac{y_i^{\hat{\lambda}_L}\ln(y_i)}{\hat{\lambda}_L} - \frac{y_i^{\hat{\lambda}_L}-1}{\hat{\lambda}_L^2}\right) = \hat{\sigma}_L^2 \sum_{i=1}^n \ln(y_i),$$

$$\hat{\mu}_L = \frac{1}{n}\sum_{i=1}^n \frac{y_i^{\hat{\lambda}_L}-1}{\hat{\lambda}_L} \quad \text{and}$$

$$\hat{\sigma}_L = \sqrt{\frac{1}{n}\sum_{i=1}^n \left(\frac{y_i^{\hat{\lambda}_L}-1}{\hat{\lambda}_L} - \hat{\mu}_L\right)^2} \tag{4}$$

iteratively for $\hat{\lambda}_L \neq 0$. For $\lambda = 0$, the system of equations (4) becomes

$$\hat{\mu}_L = \frac{1}{n}\sum_{i=1}^n \ln(y_i) \quad \text{and}$$

$$\hat{\sigma}_L = \sqrt{\frac{1}{n}\sum_{i=1}^n (\ln(y_i) - \hat{\mu}_L)^2},$$

where $\hat{\mu}_L$, $\hat{\sigma}_L$ and $\hat{\lambda}_L$ are the corresponding maximum likelihood estimates of μ, σ and λ , respectively. In the process of maximizing the log-likelihood ℓ_L, to incorporate $\lambda = 0$ in the parameter space, in each iteration, the expression (3) needs to be evaluated. Note that $\hat{\mu}_L$, $\hat{\sigma}_L$ and $\hat{\lambda}_L$ are asymptotically normal, consistent, and asymptotically unbiased estimates.

4. ARTIFICIAL REGRESSION MODEL ESTIMATOR

Linearizing the trend between the observations and the normal percentiles in a normal probability plot justifies the normality assumption. On the other hand, the problem of observing large variances of the point estimates, of the model parameters, is reduced through augmenting the model by adding some covariates that are related to the response variable. Putting these two concepts together, Halawa (1996) proposed an artificial augmented regression model as

$$X_{(i)} = \mu + \sigma Z_{(i)} + \epsilon_{(i)}, \quad i = 1, 2, \cdots, n,$$

where $X_{(i)}$ is the i^{th} smallest transformed observation, using the transformation (1) for a fixed λ, with $Z_{(i)} = \Phi^{-1}((i - .375)/(n + .25))$ where Φ is the cumulative standard normal distribution function. The $\epsilon_{(i)}$ are assumed to be normal variates with zero means, unit variances and zero covariances. It follows by definition that $\sum_{i=1}^{n} Z_{(i)} = 0$ and $\lim_{n\to\infty} \sum_{i=1}^{n} Z_{(i)}^2/n = 1$.

It follows that the log-likelihood $\ell_R = \ell(\lambda, \mu, \sigma; y_1, y_2, \cdots, y_n)$, where the subscript 'R' denotes the artificial regression model procedure, can be written as

$$\ell_R = \begin{cases} -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_{(i)}^{\lambda}-1}{\lambda} - \mu - \sigma Z_{(i)}\right)^2 + (\lambda - 1)\sum_{i=1}^{n}\ln(y_i), & \lambda \neq 0, \\ -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\left(\ln(y_{(i)}) - \mu - \sigma Z_{(i)}\right)^2 - \sum_{i=1}^{n}\ln(y_i), & \lambda = 0. \end{cases} \tag{5}$$

Equation (5) is maximized when the partial derivatives of (5) with respect to μ, σ and λ are equated to zero and the resulting equations are solved simultaneously in which $\sum_{i=1}^{n} Z_{(i)} = 0$ and $\lim_{n\to\infty} \sum_{i=1}^{n} Z_{(i)}^2/n = 1$ are substituted. This leads to solving

$$\sum_{i=1}^{n}\ln(y_i) = \sum_{i=1}^{n}\left(\frac{y_{(i)}^{\hat\lambda_R} - 1}{\hat\lambda_R} - \hat\mu_R - \hat\sigma_R Z_{(i)}\right)\left(y_{(i)}^{\hat\lambda_R}\ln(y_{(i)}) - \frac{y_{(i)}^{\hat\lambda_R} - 1}{\hat\lambda_R}\right)\frac{1}{\hat\lambda_R},$$

$$\hat\mu_R = \frac{1}{n}\sum_{i=1}^{n}\frac{y_{(i)}^{\hat\lambda_R} - 1}{\hat\lambda_R} \quad \text{and}$$

$$\hat\sigma_R = \frac{1}{n}\sum_{i=1}^{n} Z_{(i)}\left(\frac{y_{(i)}^{\hat\lambda_R} - 1}{\hat\lambda_R}\right) \tag{6}$$

iteratively for $\hat\lambda_R \neq 0$. In the process of maximizing the log-likelihood ℓ_R, to incorporate $\lambda = 0$ in the parameter space, in each iteration, the expression (5) needs to be evaluated. For $\lambda = 0$, the system of equations (6) becomes

$$\hat\mu_R = \frac{1}{n}\sum_{i=1}^{n}\ln(y_i) \quad \text{and}$$

$$\hat{\sigma}_R = \frac{1}{n}\sum_{i=1}^{n} Z_{(i)}\ln(y_{(i)}).$$

The asymptotic properties of $\hat{\mu}_R$, $\hat{\sigma}_R$ and $\hat{\lambda}_R$ are given in Halawa (1996).

5. SHAPIRO-FRANCIA W' STATISTIC ESTIMATOR

The Shapiro-Francia W' test statistic (Shapiro and Francia (1972)) is obtained by dividing the square of an appropriate linear combination of the sample order statistics by the usual symmetric estimate of the variance.

Let $(X_1, X_2, \cdots, X_n)$ be a random sample to be tested for normality, ordered $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$, and let $(m_1, m_2, \cdots, m_n)'$ denote the vector of expected values of standard normal order statistics. Define

$$W' = \frac{\left(\sum_{i=1}^{n} m_i X_{(i)}\right)^2}{\sum_{i=1}^{n} m_i^2 \times \sum_{i=1}^{n}(X_i - \bar{X})^2}.$$

Note that W' equals the square of the standard product-moment correlation coefficient between the $X_{(i)}$ and m_i, and therefore measures the straightness of the normal probability plot of the $X_{(i)}$; small values of W' indicate non-normality. The ratio W' is both scale and origin invariant and hence the statistic is appropriate for a test of the composite hypothesis of normality. One useful feature of the Shapiro-Francia W' test is that several independent goodness-of-fit tests may be combined into one overall test of normality. W' can be written in another form, that is,

$$W' = \frac{\left(\sum_{i=1}^{n} a_i X_{(i)}\right)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

where

$$a_i = \frac{m_i}{\sqrt{\sum_{i=1}^{n} m_i^2}}.$$

The value of W' is closer to 1 means the data is closer to normality and the maximum value of W' is 1. By utilizing this fact, an estimation procedure for λ in the Box-Cox transformation (1) is proposed by Rahman (1999), which maximizes W'. That is, maximize

$$W' = \frac{\left(\sum_{i=1}^{n} a_i X_{(i)}\right)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

where

$$X_{(i)} = \begin{cases} \frac{Y_i^{\lambda} - 1}{\lambda}, & \lambda \neq 0, \\ \ln(Y_i), & \lambda = 0. \end{cases}$$

Note that the X_i's are ordered for a fixed λ. This maximization can be done either by using the Newton-Raphson root finding procedure after taking the first derivative of W' with respect to λ and equating to zero or by the Numerical grid search procedure. The explicit solutions are not presented because the expressions become tedious when it is taken into consideration that the relation between X and Y is not always monotone for all λ values. And, the statistic W' depends on the ordered X's. This estimate of λ is denoted by $\hat{\lambda}_W$.

The Shapiro and Wilk (1965) W statistic has been shown to yield a powerful test of normality for a variety of nonnormal distributions (Pearson, D'Agostino, and Bowman (1977) and Shapiro, Wilk, and Chen (1968)). The Shapiro-Francia W' statistic (referred to as the W' statistic) is an approximation of the W statistic.

The values of the a_i's are tabulated in Shapiro and Wilk (1965) for $n = 2(1)50$. For other sample sizes, the a_i's can be estimated using the following approximate results. By definition,

$$a = \frac{m'V^{-1}}{(m'V^{-1}V^{-1}m)^{1/2}} = \frac{m'V^{-1}}{C}$$

is such that $a'a = 1$, m is the vector of the expected values of standard normal order statistics and V is the covariance matrix of the standard normal order statistics. With $a^{*\prime} = m'V^{-1}$, it can be shown that $C^2 = a^{*\prime}a^*$. Suggested approximations are

$$\hat{a}_i^* = 2m_i, \quad i = 2, 3, \cdots, n-1, \quad \text{and}$$

$$\hat{a}_1^2 = \hat{a}_n^2 = \begin{cases} \frac{\Gamma(\frac{1}{2}n)}{\sqrt{2}\Gamma(\frac{1}{2}(n+1))}, & n \le 20, \\ \frac{\Gamma(\frac{1}{2}(n+1))}{\sqrt{2}\Gamma(\frac{1}{2}n+1)}, & n > 20. \end{cases}$$

The values of the m_i's are tabulated in Harter (1961) for $n = 2(1)100$, 125(25)250, 300, 350, and 400.

Using simulation, Rahman (1999) showed that the maximization of the Shapiro-Francia W' statistic procedure leads to the estimate of λ which forces the data to be closer to normal in the sense of a higher Shapiro-Francia W' statistic value than that of the normal likelihood procedure and the artificial regression model procedure. And, $\hat{\lambda}_W$ has smaller bias than that of the normal likelihood procedure and the artificial regression model procedure.

6. SAMPLE ENTROPY STATISTIC ESTIMATOR

A test of the composite hypothesis of normality is introduced by Vasicek (1976). The test is based on the property of the normal distribution that its entropy exceeds that of any other distribution with a density that has the same variance. The test statistic is based on a class of estimators of entropy. This is a consistent test of the null hypothesis for all alternatives without a singular continuous part. The test compared favourably with other tests of normality for the tests investigated by

Stephens (1974) which are Kolmogorov-Smirnov D, Cramér$-$*vonMises* W^2, Kuiper V, Watson U^2, Anderson-Darling A^2, Shapiro-Wilk W, and hence Shapiro-Francia W'.

A well-known theorem of information theory (Shannon, 1949, p.55) states that among all distributions that possess a density function f and have a given variance σ^2, the entropy

$$H(f) = -\int_{-\infty}^{\infty} f(x) \log f(x) dx$$

is maximized by the normal distribution. Vasicek (1976) used an estimate of $H(f)$,

$$H_{mn} = \frac{1}{n}\sum_{i=1}^{n} \log\left\{\frac{n}{2m}\left(X_{(i+m)} - X_{(i-m)}\right)\right\},$$

where $X_{(i)} = X_{(1)}$, $i < 1$, and $X_{(i)} = X_{(n)}$, $i > n$, for a suitably chosen m. Goodness properties of H_{mn} can be seen in Vasicek (1976). Using H_{mn}, a goodness of fit test for normality was developed by Vasicek (1976), the test statistic was given as,

$$K_{mn} = \frac{n}{2mS}\left\{\prod_{i=1}^{n}\left(X_{(i+m)} - X_{(i-m)}\right)\right\}^{\frac{1}{n}} \tag{7}$$

and

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2.$$

Due to intractability of the distribution of K_{mn}, Monte Carlo simulation was used to obtain the decision rules. Here we give a table of m values for $n = 4, 5, \cdots, 500$ which maximizes K_{mn}. For each sample size from the standard normal distribution 100,000 samples were taken to obtain m. In Table 1 the values of m that maximizes K_{mn} for a given n are presented. It is to be noted that according to Vasicek (1976, p.58), $0 \le K_{mn} < \sqrt{2\pi e} = 4.133\ldots$.

Here we introduce a new method of estimation of λ in (1) using maximization of K_{mn}. In equation (7), we consider

$$X_{(i)} = \begin{cases} \frac{Y_i^{\lambda}-1}{\lambda}, & \lambda \neq 0 \\ \ln(Y_i), & \lambda = 0 \end{cases}$$

and maximize K_{mn} with respect to λ using a numerical grid search procedure. Note that X_i's are ordered for a fixed λ.

7. SIMULATION STUDY

Table 1: Choices of m for maximized K_{mn}

n	m	n	m	n	m	n	m	n	m
4-5	1	190-196	19	291-295	37	368-371	55	436-440	73
6-16	2	197-203	20	296-298	38	372-375	56	441-444	74
17-29	3	204-210	21	299-304	39	376-379	57	445-447	75
30-42	4	211-216	22	305-308	40	380-382	58	448-452	76
43-55	5	217-222	23	309-313	41	383-386	59	453-455	77
56-68	6	223-229	24	314-317	42	387-391	60	456-459	78
69-81	7	230-235	25	318-321	43	392-395	61	460-462	79
82-94	8	236-240	26	322-326	44	396-398	62	463-466	80
95-105	9	241-246	27	327-330	45	399-402	63	467-470	81
106-117	10	247-251	28	331-334	46	403-406	64	471-474	82
118-127	11	252-257	29	335-337	47	407-410	65	475-477	83
128-138	12	258-262	30	338-343	48	411-413	66	478-479	84
139-147	13	263-266	31	344-347	49	414-417	67	480-485	85
148-156	14	267-271	32	348-351	50	418-421	68	486-488	86
157-166	15	272-277	33	352-354	51	422-425	69	489-492	87
167-173	16	278-282	34	355-358	52	426-429	70	493-495	88
174-181	17	283-285	35	359-363	53	430-433	71	496-498	89
182-189	18	286-290	36	364-367	54	434-435	72	499-500	90

This section presents simulation results in which there exist values of the transformation parameter that satisfy model (2) under the transformation (1). These λ values, in combination with certain choices of the other model parameters are used to generate data as

$$Y = (1 + \lambda(\mu + \sigma\epsilon))^{\frac{1}{\lambda}}$$

where ϵ is a pseudo $N(0,1)$ random vector and $\lambda \neq 0$. When λ=0, $Y = \exp(\mu + \sigma\epsilon)$ is used to generate data. The choices of parameters are made in such a way that the Y vector is always positive, even in cases when $\lambda \neq 0$.

To see the effects of the two estimates under study, data from a skewed distribution is generated using Gamma and Weibull distributions under certain choices of their scale and shape parameters.

For a Gamma distribution

$$f(y; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} y^{\alpha-1} e^{-\frac{y}{\beta}}, \quad y > 0$$

is considered and for a Weibull distribution

$$f(y; \alpha, \beta) = \frac{\alpha}{\beta} \left(\frac{y}{\beta}\right)^{\alpha-1} e^{-(y/\alpha)^\alpha}$$

is considered in generating random samples.

The simulation study was performed using Fortran 77 program. The Fortran 77 codes are available from the authors. For each parameter configuration in Tables 2-11, 1000 samples were generated. Each sample was used to obtain the W' and K_{mn} statistics. The means (m) and the standard deviations (s) are displayed beside the labels Sample W' and Sample K_{mn}. The point estimates of λ are obtained using maximization of the likelihood as described in Section 3 (denoted by $\hat{\lambda}_L$), via artificial regression as in Section 4 (denoted by $\hat{\lambda}_R$), maximization of Shapiro-Francia W' statistic as in Section 5 (denoted by $\hat{\lambda}_W$), and maximization of entopy statistic K_{mn} as in Section 6 (denoted by $\hat{\lambda}_K$). In all computations, numerical grid search procedures are used. For all estimates of λ, the resulting W' statistic and K_{mn} statistic are computed. Finally, the means and standard deviations for all the estimates of λ and their corresponding W' and K_{mn} statistics are displayed in Tables 2-11.

In the case of normal samples, λ= -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, μ= -10, -5, 0, 5, 10 and σ= 1, 2 are considered. For a particular λ, the choices of μ and σ are restricted by the fact that $1 + \lambda(\mu + \sigma\epsilon) > 0$ where ϵ is a standard normal variate to avoid taking the power of a negative number while the powers are not necessarily an integer or taking logarithm of a negative number. The results for normal samples are displayed in Tables 2-8.

In Tables 2-8 we notice that the biases in estimation of λ are the highest for the case of $\hat{\lambda}_R$ in all situations considered. The other three estimates have similar biases. However, among these remaining three estimates, $\hat{\lambda}_W$ has the smallest bias throughout. Among $\hat{\lambda}_L$ and $\hat{\lambda}_K$, $\hat{\lambda}_K$ has smaller bias for $n = 20$ and $\hat{\lambda}_L$ has smaller bias for $n = 40$ and $n = 100$. In terms of standard errors, there is a consistent order from the smallest to the largest in the order of $\hat{\lambda}_R$, $\hat{\lambda}_L$, $\hat{\lambda}_K$, and $\hat{\lambda}_W$ for the sample sizes 20 and 40 but for the sample size of 100 the order from the smallest to the largest is $\hat{\lambda}_R$, $\hat{\lambda}_K$, $\hat{\lambda}_L$, and $\hat{\lambda}_W$, minor exceptions are noted for $\lambda = 0$. There is a consistent order in terms of mean squared errors among the estimates of λ for $n = 20$, which is from the smallest to the largest in the order of $\hat{\lambda}_R$, $\hat{\lambda}_L$, $\hat{\lambda}_K$, and $\hat{\lambda}_W$. For $n = 40$ there is no consistent order in the values of the mean squared errors. Again, for $n = 100$, there is a consistent rank order in terms of the values of mean squared errors, from the smallest to the largest as $\hat{\lambda}_K$, $\hat{\lambda}_L$, $\hat{\lambda}_W$, and $\hat{\lambda}_R$ throughout, minor exceptions are noticed for $\lambda = 0$. The resulting W' statistics are in the order from the largest to the smallest as $\hat{\lambda}_W$, $\hat{\lambda}_L$, $\hat{\lambda}_K$, and $\hat{\lambda}_R$. And, the resulting K_{mn} statistics are in the order from the largest to the smallest as $\hat{\lambda}_K$, $\hat{\lambda}_L$, $\hat{\lambda}_W$, and $\hat{\lambda}_R$.

For the Gamma and Weibull distributions, we also use sample sizes 20, 40, and 100. For Gamma samples, we use α=4, 9 and β=5, 10. For Weibull samples, we use α=2, 4 and β=5, 10. The results are displayed in Tables 9-11. Here we can not compare the biases as the values of λ are not known. The standard errors for all the estimates have similar patterns for Gamma samples as for normal samples. But in the case of Weibull samples there is a consistent pattern irrespective to the sample sizes which is from the smallest to the largest as $\hat{\lambda}_R$, $\hat{\lambda}_L$, $\hat{\lambda}_K$, and $\hat{\lambda}_W$. The order in values for W' and K_{mn} are exactly similar in both Gamma and Weibull samples as seen in samples derived from normal distributions.

Table 2: Normal Samples

	Likelihood		Regression		W'-Statistic		Entropy	
	m	s	m	s	m	s	m	s
	Normal: μ=-5, σ=1, λ=-2, n=20							
	Sample W': m=0.9258, s=0.0575				Sample K_{mn}: m=3.0576, s=0.2687			
$\hat{\lambda}$	-1.8212	1.8335	-3.2450	0.2381	-2.0057	1.9660	-1.8514	1.8652
W'	0.9686	0.0181	0.9537	0.0281	0.9688	0.0180	0.9685	0.0181
K_{mn}	3.2383	0.2003	3.1641	0.2199	3.2374	0.2003	3.2390	0.2003
	Normal: μ=-10, σ=2, λ=-2, n=20							
	Sample W': m=0.9226, s=0.0605				Sample K_{mn}: m=3.0444, s=0.2757			
$\hat{\lambda}$	-1.8121	1.7811	-2.5443	0.1715	-2.0017	1.9203	-1.8436	1.8160
W'	0.9687	0.0179	0.9557	0.0269	0.9690	0.0179	0.9687	0.0180
K_{mn}	3.2390	0.1999	3.1780	0.2178	3.2381	0.1999	3.2397	0.1999
	Normal: μ=-5, σ=1, λ=-1.5, n=20							
	Sample W': m=0.9216, s=0.0608				Sample K_{mn}: m=3.0381, s=0.2769			
$\hat{\lambda}$	-1.3391	1.5744	-2.5412	0.2029	-1.4948	1.7157	-1.3659	1.6105
W'	0.9694	0.0173	0.9532	0.0284	0.9697	0.0173	0.9693	0.0174
K_{mn}	3.2416	0.1987	3.1616	0.2201	3.2405	0.1987	3.2423	0.1986
	Normal: μ=-10, σ=2, λ=-1.5, n=20							
	Sample W': m=0.9168, s=0.0652				Sample K_{mn}: m=3.0180, s=0.2874			
$\hat{\lambda}$	-1.3316	1.5064	-1.9475	0.1422	-1.4914	1.6503	-1.3604	1.5470
W'	0.9695	0.0172	0.9557	0.0269	0.9698	0.0172	0.9695	0.0173
K_{mn}	3.2422	0.1984	3.1778	0.2177	3.2410	0.1985	3.2430	0.1983
	Normal: μ=-5, σ=1, λ=-1, n=20							
	Sample W': m=0.9133, s=0.0670				Sample K_{mn}: m=2.999, s=0.2928			
$\hat{\lambda}$	-0.8765	1.2197	-1.8219	0.1655	-0.9915	1.3662	-0.8994	1.2640
W'	0.9699	0.0169	0.9525	0.0287	0.9703	0.0168	0.9699	0.0169
K_{mn}	3.2439	0.1977	3.1575	0.2204	3.2427	0.1979	3.2447	0.1975
	Normal: μ=-10, σ=2, λ=-1, n=20							
	Sample W': m=0.9047, s=0.0743				Sample K_{mn}: m=2.9624, s=0.3118			
$\hat{\lambda}$	-0.8709	1.1261	-1.3441	0.1105	-0.9887	1.2760	-0.8952	1.1729
W'	0.9700	0.0169	0.9556	0.0268	0.9704	0.0167	0.9700	0.0169
K_{mn}	3.2441	0.1976	3.1773	0.2177	3.2428	0.1979	3.2450	0.1974
	Normal: μ=-5, σ=1, λ=-0.5, n=20							
	Sample W': m=0.8909, s=0.0817				Sample K_{mn}: m=2.8910, s=0.3343			
$\hat{\lambda}$	-0.4348	0.7386	-1.0761	0.1233	-0.4936	0.8535	-0.4464	0.7735
W'	0.9701	0.0168	0.9509	0.0296	0.9705	0.0166	0.9701	0.0168
K_{mn}	3.2444	0.1975	3.1488	0.2210	3.2431	0.1980	3.2452	0.1974
	Normal: μ=-10, σ=2, λ=-0.5, n=20							
	Sample W': m=0.8676, s=0.0978				Sample K_{mn}: m=2.7861, s=0.3831			
$\hat{\lambda}$	-0.4333	0.6292	-0.7285	0.0756	-0.4931	0.7329	-0.4453	0.6610
W'	0.9701	0.0168	0.9556	0.0267	0.9705	0.0166	0.9701	0.0168
K_{mn}	3.2444	0.1975	3.1764	0.2174	3.2430	0.1982	3.2453	0.1973

Table 3: Normal Samples (Continued)

	Likelihood		Regression		W'-Statistic		Entropy	
	m	s	m	s	m	s	m	s
	Normal: μ=0, σ=1, λ=0, n=20							
	Sample W': m=0.7402, s=0.1352				Sample K_{mn}: m=2.0698, s=0.5059			
$\hat{\lambda}$	-0.0022	0.2143	-0.0219	0.4991	-0.0035	0.2482	-0.0027	0.2237
W'	0.9702	0.0168	0.9299	0.0359	0.9705	0.0166	0.9702	0.0168
K_{mn}	3.2443	0.1977	3.0289	0.2765	3.2432	0.1985	3.2451	0.1976
	Normal: μ=0, σ=2, λ=0, n=20							
	Sample W': m=0.5229, s=0.1551				Sample K_{mn}: m=0.8680, s=0.4768			
$\hat{\lambda}$	-0.0009	0.1072	-0.0053	0.1395	-0.0018	0.1243	-0.0014	0.1118
W'	0.9702	0.0168	0.9640	0.0178	0.9705	0.0166	0.9702	0.0168
K_{mn}	3.2442	0.1977	3.2101	0.2045	3.2431	0.1985	3.2450	0.1976
	Normal: μ=5, σ=1, λ=0.5, n=20							
	Sample W': m=0.9493, s=0.0338				Sample K_{mn}: m=3.1446, s=0.2277			
$\hat{\lambda}$	0.4212	0.7382	1.0761	0.1207	0.4723	0.8604	0.4274	0.7747
W'	0.9701	0.0167	0.9505	0.0304	0.9705	0.0165	0.9701	0.0168
K_{mn}	3.2441	0.1980	3.1476	0.2195	3.2428	0.1989	3.2450	0.1979
	Normal: μ=10, σ=2, λ=0.5, n=20							
	Sample W': m=0.9472, s=0.0357				Sample K_{mn}: m=3.1313, s=0.2310			
$\hat{\lambda}$	0.4215	0.6289	0.7288	0.0748	0.4732	0.7369	0.4288	0.6617
W'	0.9700	0.0167	0.9554	0.0272	0.9705	0.0165	0.9701	0.0168
K_{mn}	3.2441	0.1981	3.1757	0.2174	3.2426	0.1991	3.2450	0.1979
	Normal: μ=5, σ=1, λ=1, n=20							
	Sample W': m=0.9554, s=0.0273				Sample K_{mn}: m=3.1806, s=0.2180			
$\hat{\lambda}$	0.8526	1.2179	1.8220	0.1624	0.9537	1.3618	0.8647	1.2610
W'	0.9699	0.0169	0.9522	0.0295	0.9703	0.0167	0.9699	0.0169
K_{mn}	3.2435	0.1984	3.1565	0.2192	3.2424	0.1987	3.2444	0.1983
	Normal: μ=10, σ=2, λ=1, n=20							
	Sample W': m=0.9554, s=0.0273				Sample K_{mn}: m=3.1806, s=0.2180			
$\hat{\lambda}$	0.8487	1.1247	1.3444	0.1099	0.9529	1.2737	0.8952	1.1729
W'	0.9699	0.0168	0.9555	0.0272	0.9704	0.0167	0.9699	0.0168
K_{mn}	3.2437	0.1983	3.1769	0.2176	3.2424	0.1988	3.2446	0.1982
	Normal: μ=5, σ=1, λ=1.5, n=20							
	Sample W': m=0.9538, s=0.0289				Sample K_{mn}: m=3.1761, s=0.2201			
$\hat{\lambda}$	1.3107	1.5755	2.5414	0.1999	1.4480	1.7157	1.3248	1.6131
W'	0.9693	0.0174	0.9530	0.0290	0.9696	0.0173	0.9693	0.0175
K_{mn}	3.2413	0.1995	3.1608	0.2191	3.2403	0.1997	3.2420	0.1975
	Normal: μ=10, σ=2, λ=1.5, n=20							
	Sample W': m=0.9535, s=0.0291				Sample K_{mn}: m=3.1755, s=0.2204			
$\hat{\lambda}$	1.3039	1.5056	1.9482	0.1411	1.4465	1.6513	1.3205	1.5472
W'	0.9694	0.0172	0.9556	0.0272	0.9698	0.0172	0.9694	0.0173
K_{mn}	3.2418	0.1992	3.1774	0.2178	3.2407	0.1995	3.2426	0.1992

Table 4: Normal Samples (Continued)

	Likelihood		Regression		W'-Statistic		Entropy	
	m	s	m	s	m	s	m	s
	Normal: μ=5, σ=1, λ=2, n=20							
	Sample W': m=0.9517, s=0.0310				Sample K_{mn}: m=3.1685, s=0.2231			
$\hat{\lambda}$	1.7148	2.1796	3.2453	0.2345	1.9155	0.2458	1.7396	2.2516
W'	0.9697	0.0169	0.9534	0.0288	0.9701	0.0168	0.9697	0.0170
K_{mn}	3.2431	0.1986	3.1633	0.2190	3.2419	0.1989	3.2439	0.1985
	Normal: μ=10, σ=2, λ=2, n=20							
	Sample W': m=0.9513, s=0.0314				Sample K_{mn}: m=3.1672, s=0.2237			
$\hat{\lambda}$	1.7099	2.0938	2.5448	0.1706	1.9140	2.3437	1.7367	2.1698
W'	0.9698	0.0169	0.9556	0.0272	0.9702	0.0168	0.9698	0.0169
K_{mn}	3.2433	0.1985	3.1777	0.2178	3.2420	0.1989	3.2441	0.1984
	Normal: μ=-5, σ=1, λ=-2, n=40							
	Sample W': m=0.9349, s=0.0513				Sample K_{mn}: m=3.3328, s=0.2207			
$\hat{\lambda}$	-1.8463	1.4791	-3.4259	0.2308	-1.9756	1.6045	-1.8256	1.4827
W'	0.9811	0.0095	0.9692	0.0172	0.9813	0.0095	0.9810	0.0096
K_{mn}	3.5388	0.1414	3.4710	0.1557	3.5379	0.1414	3.5392	0.1414
	Normal: μ=-10, σ=2, λ=-2, n=40							
	Sample W': m=0.9305, s=0.0554				Sample K_{mn}: m=3.3140, s=0.2315			
$\hat{\lambda}$	-1.8447	1.4089	-2.7208	0.1641	-1.9753	1.5309	-1.8242	1.4137
W'	0.9811	0.0095	0.9730	0.0145	0.9813	0.0095	0.9810	0.0096
K_{mn}	3.5388	0.1413	3.4950	0.1515	3.5378	0.1414	3.5392	0.1414
	Normal: μ=-5, σ=1, λ=-1.5, n=40							
	Sample W': m=0.9294, s=0.0552				Sample K_{mn}: m=3.3061, s=0.2318			
$\hat{\lambda}$	-1.3841	1.1499	-2.6851	0.1967	-1.4803	1.2536	-1.3682	1.1520
W'	0.9811	0.0095	0.9684	0.0175	0.9813	0.0095	0.9810	0.0096
K_{mn}	3.5388	0.1414	3.4665	0.1562	3.5379	0.1415	3.5392	0.1414
	Normal: μ=-10, σ=2, λ=-1.5, n=40							
	Sample W': m=0.9228, s=0.0610				Sample K_{mn}: m=3.2777, s=0.2478			
$\hat{\lambda}$	-1.3830	1.0779	-2.0905	0.1362	-1.4807	1.1768	-1.3676	1.0809
W'	0.9811	0.0095	0.9728	0.0146	0.9813	0.0095	0.9810	0.0096
K_{mn}	3.5388	0.1414	3.4940	0.1516	3.5378	0.1415	3.5392	0.1414
	Normal: μ=-5, σ=1, λ=-1, n=40							
	Sample W': m=0.9185, s=0.0623				Sample K_{mn}: m=3.2532, s=0.2529			
$\hat{\lambda}$	-0.9236	0.8142	-1.9260	0.1589	-0.9872	0.8871	-0.9127	0.8151
W'	0.9811	0.0095	0.9671	0.0182	0.9813	0.0095	0.9810	0.0096
K_{mn}	3.5388	0.1414	3.4590	0.1570	3.5379	0.1415	3.5392	0.1415
	Normal: μ=-10, σ=2, λ=-1, n=40							
	Sample W': m=0.9068, s=0.0718				Sample K_{mn}: m=3.2012, s=0.2806			
$\hat{\lambda}$	-0.9226	0.7425	-1.4508	0.1064	-0.9869	0.8106	-0.9120	0.7443
W'	0.9811	0.0095	0.9726	0.0147	0.9813	0.0095	0.9810	0.0096
K_{mn}	3.5388	0.1414	3.4922	0.1517	3.5378	0.1415	3.5392	0.1414

Table 5: Normal Samples (Continued)

	Likelihood		Regression		W'-Statistic		Entropy	
	m	s	m	s	m	s	m	s
	Normal: μ=-5, σ=1, λ=-0.5, n=40							
	Sample W': m=0.8892, s=0.0787				Sample K_{mn}: m=3.1067, s=0.3049			
$\hat{\lambda}$	-0.4626	0.4782	-1.1364	0.1158	-0.4937	0.5198	-0.4567	0.4780
W'	0.9812	0.0095	0.9643	0.0193	0.9813	0.0095	0.9811	0.0096
K_{mn}	3.5388	0.1415	3.4438	0.1585	3.5380	0.1416	3.5392	0.1416
	Normal: μ=-10, σ=2, λ=-0.5, n=40							
	Sample W': m=0.8574, s=0.0988				Sample K_{mn}: m=2.9607, s=0.3695			
$\hat{\lambda}$	-0.4619	0.4071	-0.7946	0.0731	-0.4936	0.4434	-0.4562	0.4077
W'	0.9811	0.0095	0.9720	0.0149	0.9813	0.0095	0.9810	0.0096
K_{mn}	3.5388	0.1414	3.4883	0.1519	3.5379	0.1415	3.5392	0.1415
	Normal: μ=0, σ=1, λ=0, n=40							
	Sample W': m=0.6986, s=0.1331				Sample K_{mn}: m=2.0498, s=0.4734			
$\hat{\lambda}$	-0.0003	0.1384	0.0114	0.3489	0.0001	0.1498	-0.0004	0.1379
W'	0.9812	0.0096	0.9546	0.0262	0.9813	0.0095	0.9811	0.0097
K_{mn}	3.5387	0.1421	3.3885	0.2161	3.5382	0.1421	3.5391	0.1421
	Normal: μ=0, σ=2, λ=0, n=40							
	Sample W': m=0.4358, s=0.1468				Sample K_{mn}: m=0.6652, s=0.3707			
$\hat{\lambda}$	-0.0003	0.0692	0.0020	0.0922	0.0001	0.0749	-0.0002	0.0691
W'	0.9811	0.0096	0.9779	0.0104	0.9812	0.0095	0.9811	0.0096
K_{mn}	3.5387	0.1421	3.5196	0.1475	3.5381	0.1421	3.5391	0.1421
	Normal: μ=5, σ=1, λ=0.5, n=40							
	Sample W': m=0.9653, s=0.0224				Sample K_{mn}: m=3.4541, s=0.1611			
$\hat{\lambda}$	0.4605	0.4718	1.1337	0.1195	0.4950	0.5143	0.4545	0.4722
W'	0.9811	0.0096	0.9638	0.0197	0.9812	0.0095	0.9810	0.0097
K_{mn}	3.5386	0.1427	3.4420	0.1528	3.5377	0.1428	3.5390	0.1427
	Normal: μ=10, σ=2, λ=0.5, n=40							
	Sample W': m=0.9625, s=0.0245				Sample K_{mn}: m=3.4365, s=0.1651			
$\hat{\lambda}$	0.4605	0.4006	0.7930	0.0756	0.4948	0.4380	0.4547	0.4019
W'	0.9810	0.0096	0.9717	0.0151	0.9812	0.0095	0.9809	0.0097
K_{mn}	3.5385	0.1428	3.4873	0.1488	3.5376	0.1429	3.5390	0.1428
	Normal: μ=5, σ=1, λ=1, n=40							
	Sample W': m=0.9734, s=0.0144				Sample K_{mn}: m=3.5030, s=0.1504			
$\hat{\lambda}$	0.9206	0.8014	1.9222	0.1644	0.9899	0.8757	0.9091	0.8040
W'	0.9810	0.0096	0.9666	0.0185	0.9812	0.0095	0.9809	0.0097
K_{mn}	3.5385	0.1428	3.4573	0.1520	3.5376	0.1429	3.5390	0.1428
	Normal: μ=10, σ=2, λ=1, n=40							
	Sample W': m=0.9734, s=0.0144				Sample K_{mn}: m=3.5030, s=0.1504			
$\hat{\lambda}$	0.9201	0.7297	1.4485	0.1099	0.9897	0.7990	0.9092	0.7329
W'	0.9810	0.0096	0.9723	0.0148	0.9812	0.0095	0.9809	0.0097
K_{mn}	3.5385	0.1429	3.4913	0.1491	3.5374	0.1430	3.5390	0.1429

Table 6: Normal Samples (Continued)

	Likelihood		Regression		W'-Statistic		Entropy	
	m	s	m	s	m	s	m	s
	Normal: μ=5, σ=1, λ=1.5, n=40							
	Sample W': m=0.9716, s=0.0168				Sample K_{mn}: m=3.4976, s=0.1539			
$\hat{\lambda}$	1.3803	1.1306	2.6804	0.2029	1.4853	1.2347	1.3639	1.1347
W'	0.9810	0.0096	0.9680	0.0179	0.9812	0.0095	0.9809	0.0097
K_{mn}	3.5385	0.1428	3.4648	0.1516	3.5375	0.1430	3.5390	0.1429
	Normal: μ=10, σ=2, λ=1.5, n=40							
	Sample W': m=0.9713, s=0.0171				Sample K_{mn}: m=3.4968, s=0.1543			
$\hat{\lambda}$	1.3797	1.0587	2.0877	0.1404	1.4848	1.1594	1.3637	1.0641
W'	0.9810	0.0096	0.9726	0.0147	0.9812	0.0095	0.9809	0.0097
K_{mn}	3.5385	0.1429	3.4931	0.1492	3.5374	0.1430	3.5389	0.1429
	Normal: μ=5, σ=1, λ=2, n=40							
	Sample W': m=0.9691, s=0.0196				Sample K_{mn}: m=3.4878, s=0.1582			
$\hat{\lambda}$	1.8412	1.4557	3.4204	0.2385	1.9807	1.5851	1.8200	1.4618
W'	0.9810	0.0096	0.9688	0.0175	0.9812	0.0095	0.9809	0.0097
K_{mn}	3.5385	0.1429	3.4695	0.1514	3.5375	0.1430	3.5389	0.1429
	Normal: μ=10, σ=2, λ=2, n=40							
	Sample W': m=0.9687, s=0.0202				Sample K_{mn}: m=3.4861, s=0.1589			
$\hat{\lambda}$	1.8402	1.3856	2.7175	0.1689	1.9803	1.5118	1.8190	1.3923
W'	0.9810	0.0096	0.9727	0.0146	0.9812	0.0096	0.9809	0.0097
K_{mn}	3.5385	0.1429	3.4942	0.1493	3.5374	0.1431	3.5389	0.1430
	Normal: μ=-5, σ=1, λ=-2, n=100							
	Sample W': m=0.9429, s=0.0221				Sample K_{mn}: m=3.6085, s=0.1654			
$\hat{\lambda}$	-1.9057	0.8563	-3.5602	0.2134	-1.9727	0.8994	-1.8031	0.8287
W'	0.9906	0.0044	0.9795	0.0101	0.9906	0.0043	0.9904	0.0045
K_{mn}	3.8224	0.0954	3.7501	0.0999	3.8217	0.0952	3.8233	0.0956
	Normal: μ=-10, σ=2, λ=-2, n=100							
	Sample W': m=0.9374, s=0.0470				Sample K_{mn}: m=3.5851, s=0.1779			
$\hat{\lambda}$	-1.9064	0.8136	-2.8990	0.1665	-1.9738	0.8557	-1.8036	0.7877
W'	0.9906	0.0044	0.9851	0.0076	0.9906	0.0043	0.9903	0.0045
K_{mn}	3.8224	0.0954	3.7857	0.0978	3.8217	0.0951	3.8233	0.0957
	Normal: μ=-5, σ=1, λ=-1.5, n=100							
	Sample W': m=0.9362, s=0.0457				Sample K_{mn}: m=3.5756, s=0.1759			
$\hat{\lambda}$	-1.4291	0.6634	-2.7843	0.1768	-1.4789	0.6966	-1.3520	0.6418
W'	0.9906	0.0044	0.9783	0.0104	0.9906	0.0043	0.9904	0.0045
K_{mn}	3.8224	0.0954	3.7435	0.1002	3.8218	0.0952	3.8233	0.0956
	Normal: μ=-10, σ=2, λ=-1.5, n=100							
	Sample W': m=0.9278, s=0.0527				Sample K_{mn}: m=3.5399, s=0.1942			
$\hat{\lambda}$	-1.4297	0.6210	-2.2288	0.1373	-1.4803	0.6526	-1.3526	0.6012
W'	0.9906	0.0044	0.9847	0.0077	0.9906	0.0043	0.9903	0.0045
K_{mn}	3.8224	0.0954	3.7832	0.0978	3.8217	0.0951	3.8233	0.0957

Table 7: Normal Samples (Continued)

	Likelihood		Regression		W'-Statistic		Entropy	
	m	s	m	s	m	s	m	s
	Normal: μ=-5, σ=1, λ=-1, n=100							
	Sample W': m=0.9229, s=0.0525				Sample K_{mn}: m=3.5102, s=0.1957			
$\hat{\lambda}$	-0.9523	0.4705	-1.9901	0.1376	-0.9850	0.4931	-0.9011	0.4546
W'	0.9906	0.0044	0.9765	0.0109	0.9906	0.0043	0.9904	0.0045
K_{mn}	3.8224	0.0954	3.7331	0.1008	3.8218	0.0952	3.8232	0.0956
	Normal: μ=-10, σ=2, λ=-1, n=100							
	Sample W': m=0.9077, s=0.0637				Sample K_{mn}: m=3.4440, s=0.2273			
$\hat{\lambda}$	-0.9529	0.4282	-1.5475	0.1054	-0.9863	0.4497	-0.9016	0.4142
W'	0.9906	0.0044	0.9841	0.0079	0.9906	0.0043	0.9904	0.0045
K_{mn}	3.8224	0.0954	3.7794	0.0980	3.8217	0.0951	3.8233	0.0956
	Normal: μ=-5, σ=1, λ=-0.5, n=100							
	Sample W': m=0.8868, s=0.0683				Sample K_{mn}: m=3.3279, s=0.2447			
$\hat{\lambda}$	-0.4751	0.2769	-1.1655	0.0940	-0.4914	0.2896	-0.4497	0.2675
W'	0.9906	0.0044	0.9729	0.0119	0.9906	0.0043	0.9904	0.0045
K_{mn}	3.8224	0.0954	3.7132	0.1020	3.8219	0.0952	3.8231	0.0956
	Normal: μ=-10, σ=2, λ=-0.5, n=100							
	Sample W': m=0.8446, s=0.0921				Sample Kv: m=3.1386, s=0.3175			
$\hat{\lambda}$	-0.4763	0.2351	-0.8468	0.0691	-0.4926	0.2463	-0.4506	0.2274
W'	0.9906	0.0044	0.9827	0.0083	0.9906	0.0043	0.9904	0.0045
K_{mn}	3.8224	0.0954	3.7714	0.0982	3.8218	0.0952	3.8232	0.0956
	Normal: μ=0, σ=1, λ=0, n=100							
	Sample W': m=0.6532, s=0.1193				Sample K_{mn}: m=2.0416, s=0.3849			
$\hat{\lambda}$	0.0022	0.0810	-0.0030	0.2063	0.0025	0.0841	0.0022	0.0781
W'	0.9906	0.0044	0.9772	0.0245	0.9906	0.0043	0.9905	0.0044
K_{mn}	3.8223	0.0956	3.7452	0.1361	3.8220	0.0953	3.8227	0.0956
	Normal: μ=0, σ=2, λ=0, n=100							
	Sample W': m=0.3350, s=0.1250				Sample K_{mn}: m=0.5016, s=0.2513			
$\hat{\lambda}$	0.0012	0.0506	0.0003	0.0505	0.0011	0.0424	0.0010	0.0392
W'	0.9906	0.0044	0.9894	0.0045	0.9906	0.0043	0.9905	0.0045
K_{mn}	3.8222	0.0955	3.8149	0.0972	3.8219	0.0953	3.8227	0.0956
	Normal: μ=5, σ=1, λ=0.5, n=100							
	Sample W': m=0.9786, s=0.0134				Sample K_{mn}: m=3.7493, s=0.1073			
$\hat{\lambda}$	0.4905	0.2758	1.1699	0.0939	0.5084	0.2881	0.4644	0.2653
W'	0.9906	0.0044	0.9734	0.0117	0.9906	0.0043	0.9904	0.0045
K_{mn}	3.8220	0.0955	3.7153	0.1019	3.8215	0.0952	3.8227	0.0957
	Normal: μ=10, σ=2, λ=0.5, n=100							
	Sample W': m=0.9754, s=0.0151				Sample K_{mn}: m=3.7274, s=0.1101			
$\hat{\lambda}$	0.4891	0.2338	0.8500	0.0696	0.5071	0.2449	0.4629	0.2250
W'	0.9906	0.0044	0.9831	0.0082	0.9906	0.0043	0.9904	0.0045
K_{mn}	3.8219	0.0954	3.7727	0.0983	3.8213	0.0952	3.8227	0.0958

Table 8: Normal Samples (Continued)

	Likelihood		Regression		W'-Statistic		Entropy	
	m	s	m	s	m	s	m	s
	Normal: μ=5, σ=1, λ=1, n=100							
	Sample W': m=0.9875, s=0.0066				Sample K_{mn}: m=3.8079, s=0.0978			
$\hat{\lambda}$	0.9781	0.4675	1.9966	0.1376	1.0144	0.4895	0.9258	0.4506
W'	0.9906	0.0044	0.9770	0.0108	0.9906	0.0043	0.9904	0.0045
K_{mn}	3.8219	0.0954	3.7350	0.1008	3.8213	0.0952	3.8227	0.0958
	Normal: μ=10, σ=2, λ=1, n=100							
	Sample W': m=0.9875, s=0.0066				Sample K_{mn}: m=3.8079, s=0.0978			
$\hat{\lambda}$	0.9765	0.4255	1.5530	0.1064	1.0129	0.4460	0.9241	0.4099
W'	0.9906	0.0044	0.9844	0.0078	0.9906	0.0043	0.9903	0.0045
K_{mn}	3.8219	0.0954	3.7805	0.0981	3.8217	0.0952	3.8228	0.0958
	Normal: μ=5, σ=1, λ=1.5, n=100							
	Sample W': m=0.9851, s=0.0091				Sample K_{mn}: m=3.8011, s=0.0992			
$\hat{\lambda}$	1.4652	0.6590	2.7926	0.1771	1.5204	0.6901	1.3870	0.6351
W'	0.9906	0.0044	0.9788	0.0103	0.9906	0.0043	0.9904	0.0045
K_{mn}	3.8219	0.0954	3.7453	0.1003	3.8212	0.0952	3.8228	0.0958
	Normal: μ=10, σ=2, λ=1.5, n=100							
	Sample W': m=0.9848, s=0.0095				Sample K_{mn}: m=3.8002, s=0.0995			
$\hat{\lambda}$	1.4635	0.6169	2.2359	0.1391	1.5188	0.6471	1.3853	0.5945
W'	0.9906	0.0044	0.9850	0.0076	0.9906	0.0043	0.9903	0.0045
K_{mn}	3.8219	0.0954	3.7844	0.0981	3.8211	0.0952	3.8228	0.0958
	Normal: μ=5, σ=1, λ=2, n=100							
	Sample W': m=0.9822, s=0.0120				Sample K_{mn}: m=3.7891, s=0.1019			
$\hat{\lambda}$	1.9525	0.8508	3.5706	0.2138	2.0258	0.8924	1.8484	0.8198
W'	0.9906	0.0044	0.9799	0.0100	0.9906	0.0043	0.9903	0.0045
K_{mn}	3.8219	0.0954	3.7518	0.1000	3.8212	0.0952	3.8228	0.0958
	Normal: μ=10, σ=2, λ=2, n=100							
	Sample W': m=0.9816, s=0.0126				Sample K_{mn}: m=3.7870, s=0.1024			
$\hat{\lambda}$	1.9508	0.8086	2.9075	0.1693	2.0246	0.8484	1.8465	0.7789
W'	0.9906	0.0044	0.9853	0.0075	0.9906	0.0043	0.9903	0.0045
K_{mn}	3.8219	0.0954	3.7868	0.0981	3.8211	0.0952	3.8228	0.0958

Table 9: Skewed Samples

	Likelihood		Regression		W'-Statistic		Entropy	
	m	s	m	s	m	s	m	s
	Gamma: α=4, β=5, n=20							
	Sample W': m=0.9216, s=0.0568				Sample K_{mn}: m=2.9892, s=0.2781			
$\hat{\lambda}$	0.2944	0.4057	0.7078	0.1094	0.3349	0.4822	0.3001	0.4221
W'	0.9705	0.0158	0.9504	0.0294	0.9710	0.0155	0.9705	0.0159
K_{mn}	3.2422	0.1943	3.1386	0.2135	3.2404	0.1968	3.2431	0.1940
	Gamma: α=4, β=10, n=20							
	Sample W': m=0.9216, s=0.0568				Sample K_{mn}: m=2.9892, s=0.2781			
$\hat{\lambda}$	0.2944	0.4056	0.5915	0.0836	0.3347	0.4825	0.3001	0.4220
W'	0.9705	0.0158	0.9546	0.0270	0.9710	0.0155	0.9705	0.0159
K_{mn}	3.2422	0.1943	3.1638	0.2121	3.2403	0.1968	3.2431	0.1940
	Gamma: α=9, β=5, n=20							
	Sample W': m=0.9216, s=0.0568				Sample K_{mn}: m=2.9892, s=0.2781			
$\hat{\lambda}$	0.2944	0.4057	0.7078	0.1094	0.3349	0.4822	0.3001	0.4221
W'	0.9705	0.0158	0.9504	0.0294	0.9710	0.0155	0.9705	0.0159
K_{mn}	3.2422	0.1943	3.1386	0.2135	3.2404	0.1968	3.2431	0.1940
	Gamma: α=9, β=10, n=20							
	Sample W': m=0.9216, s=0.0568				Sample K_{mn}: m=2.9892, s=0.2781			
$\hat{\lambda}$	0.2944	0.4056	0.5915	0.0836	0.3347	0.4825	0.3001	0.4220
W'	0.9705	0.0158	0.9546	0.0270	0.9710	0.0155	0.9705	0.0159
K_{mn}	3.2422	0.1943	3.1638	0.2121	3.2403	0.1968	3.2431	0.1940
	Gamma: α=4, β=5, n=40							
	Sample W': m=0.9306, s=0.0441				Sample K_{mn}: m=3.2540, s=0.2200			
$\hat{\lambda}$	0.2955	0.2691	0.7409	0.0982	0.3171	0.2923	0.2928	0.2679
W'	0.9817	0.0097	0.9611	0.0202	0.9818	0.0097	0.9816	0.0098
K_{mn}	3.5411	0.1395	3.4219	0.1528	3.5402	0.1397	3.5416	0.1395
	Gamma: α=4, β=10, n=40							
	Sample W': m=0.9306, s=0.0441				Sample K_{mn}: m=3.2540, s=0.2200			
$\hat{\lambda}$	0.2956	0.2691	0.6355	0.0805	0.3171	0.2924	0.2928	0.2679
W'	0.9817	0.0097	0.9677	0.0178	0.9818	0.0097	0.9816	0.0098
K_{mn}	3.5411	0.1395	3.4609	0.1502	3.5402	0.1397	3.5416	0.1395
	Gamma: α=9, β=5, n=40							
	Sample W': m=0.9306, s=0.0441				Sample K_{mn}: m=3.2540, s=0.2200			
$\hat{\lambda}$	0.2955	0.2691	0.7409	0.0982	0.3171	0.2923	0.2928	0.2679
W'	0.9817	0.0097	0.9611	0.0202	0.9818	0.0097	0.9816	0.0098
K_{mn}	3.5411	0.1395	3.4219	0.1528	3.5402	0.1397	3.5416	0.1395
	Gamma: α=9, β=10, n=40							
	Sample W': m=0.9306, s=0.0441				Sample K_{mn}: m=3.2540, s=0.2200			
$\hat{\lambda}$	0.2956	0.2691	0.6355	0.0805	0.3171	0.2924	0.2928	0.2679
W'	0.9817	0.0097	0.9677	0.0178	0.9818	0.0097	0.9816	0.0098
K_{mn}	3.5411	0.1395	3.4609	0.1502	3.5402	0.1397	3.5416	0.1395

Table 10: Skewed Samples

	Likelihood		Regression		W'-Statistic		Entropy	
	m	s	m	s	m	s	m	s
	Gamma: α=4, β=5, n=100							
	Sample W': m=0.9368, s=0.0310				Sample K_{mn}: m=2.4929, s=0.1515			
$\hat{\lambda}$	0.3085	0.1575	0.7507	0.0719	0.3195	0.1642	0.2931	0.1530
W'	0.9907	0.0044	0.9686	0.0120	0.9907	0.0044	0.9905	0.0045
K_{mn}	3.8171	0.0950	3.6778	0.1020	3.8165	0.0949	3.8177	0.0953
	Gamma: α=4, β=10, n=100							
	Sample W': m=0.9368, s=0.0310				Sample K_{mn}: m=3.4929, s=0.1515			
$\hat{\lambda}$	0.3085	0.1575	0.6570	0.0631	0.3195	0.1643	0.2931	0.1531
W'	0.9907	0.0044	0.9763	0.0102	0.9907	0.0044	0.9905	0.0045
K_{mn}	3.8171	0.0950	3.7251	0.0998	3.8165	0.0949	3.8178	0.0953
	Gamma: α=9, β=5, n=100							
	Sample W': m=0.9368, s=0.0310				Sample K_{mn}: m=3.4929, s=0.1515			
$\hat{\lambda}$	0.3085	0.1575	0.7509	0.0719	0.3195	0.1642	0.2931	0.1530
W'	0.9907	0.0044	0.9686	0.0120	0.9907	0.0044	0.9905	0.0045
K_{mn}	3.8171	0.0950	3.6778	0.1020	3.8165	0.0949	3.8178	0.0953
	Gamma: α=9, β=10, n=100							
	Sample W': m=0.9368, s=0.0310				Sample K_{mn}: m=3.4929, s=0.1515			
$\hat{\lambda}$	0.3085	0.1575	0.6570	0.0631	0.3195	0.1643	0.2931	0.1531
W'	0.9907	0.0044	0.9763	0.0102	0.9907	0.0044	0.9905	0.0045
K_{mn}	3.8171	0.0950	3.7251	0.0998	3.8165	0.0949	3.8178	0.0953
	Weibull: α=2, β=5, n=20							
	Sample W': m=0.9413, s=0.0368				Sample K_{mn}: m=3.0559, s=0.2409			
$\hat{\lambda}$	0.4727	0.3294	1.1881	0.2019	0.5424	0.4023	0.4881	0.3544
W'	0.9709	0.0142	0.9342	0.0281	0.9718	0.0139	0.9712	0.0141
K_{mn}	3.2297	0.1996	3.0015	0.2076	3.2274	0.2009	3.2311	0.1993
	Weibull: α=2, β=10, n=20							
	Sample W': m=0.9413, s=0.0368				Sample K_{mn}: m=3.0559, s=0.2409			
$\hat{\lambda}$	0.4727	0.3294	0.9214	0.1353	0.5424	0.4024	0.4881	0.3544
W'	0.9709	0.0142	0.9531	0.0230	0.9718	0.0139	0.9712	0.0141
K_{mn}	3.2297	0.1996	3.1190	0.2060	3.2274	0.2009	3.2311	0.1993
	Weibull: α=4, β=5, n=20							
	Sample W': m=0.9587, s=0.0233				Sample K_{mn}: m=3.1764, s=0.2127			
$\hat{\lambda}$	0.9453	0.6589	1.7775	0.1917	1.0794	0.7895	0.9752	0.7047
W'	0.9709	0.0142	0.9534	0.0243	0.9718	0.0139	0.9711	0.0141
K_{mn}	3.2297	0.1996	3.1234	0.2100	3.2276	0.2008	3.2310	0.1993
	Weibull: α=4, β=10, n=20							
	Sample W': m=0.9587, s=0.0233				Sample K_{mn}: m=3.1764, s=0.2127			
$\hat{\lambda}$	0.9453	0.6591	1.2790	0.1141	1.0793	0.7895	0.9751	0.7046
W'	0.9709	0.0142	0.9604	0.0215	0.9718	0.0139	0.9711	0.0141
K_{mn}	3.2297	0.1996	3.1756	0.2104	3.2276	0.2008	3.2310	0.1993

Table 11: Skewed Samples

	Likelihood		Regression		W'-Statistic		Entropy	
	m	s	m	s	m	s	m	s
	Weibull: α=2, β=5, n=40							
	Sample W': m=0.9542, s=0.0271				Sample K_{mn}: m=3.3381, s=0.1776			
$\hat{\lambda}$	0.5073	0.2058	1.1866	0.1503	0.5544	0.2354	0.5043	0.2126
W'	0.9816	0.0083	0.9413	0.0187	0.9820	0.0081	0.9816	0.0083
K_{mn}	3.5225	0.1435	3.2428	0.1465	3.5202	0.1445	3.5230	0.1436
	Weibull: α=2, β=10, n=40							
	Sample W': m=0.9542, s=0.0771				Sample K_{mn}: m=3.3381, s=0.1776			
$\hat{\lambda}$	0.5073	0.2058	0.9516	0.1100	0.5544	0.2353	0.5043	0.2126
W'	0.9816	0.0083	0.9628	0.0146	0.9820	0.0081	0.9816	0.0083
K_{mn}	3.5225	0.1435	3.3872	0.1426	3.5202	0.1445	3.5230	0.1436
	Weibull: α=4, β=5, n=40							
	Sample W': m=0.9754, s=0.0124				Sample K_{mn}: m=3.4932, s=0.1489			
$\hat{\lambda}$	1.0142	0.4116	1.8438	0.1667	1.1085	0.4711	1.0081	0.4253
W'	0.9816	0.0083	0.9642	0.0157	0.9820	0.0081	0.9816	0.0083
K_{mn}	3.5225	0.1435	3.3983	0.1453	3.5202	0.1445	3.5230	0.1436
	Weibull: α=4, β=10, n=40							
	Sample W': m=0.9754, s=0.0124				Sample K_{mn}: m=3.4932, s=0.1489			
$\hat{\lambda}$	1.0142	0.4116	1.3867	0.1120	1.1085	0.4711	1.0081	0.4254
W'	0.9816	0.0093	0.9756	0.0115	0.9820	0.0081	0.9816	0.0083
K_{mn}	3.5225	0.1435	3.4800	0.1442	3.5202	0.1445	3.5230	0.1436
	Weibull: α=2, β=5, n=100							
	Sample W': m=0.9645, s=0.0156				Sample K_{mn}: m=3.5848, s=0.1111			
$\hat{\lambda}$	0.5222	0.1235	1.1824	0.0974	0.5508	0.1357	0.5002	0.1248
W'	0.9909	0.0041	0.9469	0.0109	0.9910	0.0040	0.9906	0.0041
K_{mn}	3.7826	0.0898	3.4568	0.0923	3.7808	0.0900	3.7836	0.0901
	Weibull: α=2, β=10, n=100							
	Sample W': m=0.9645, s=0.0156				Sample K_{mn}: m=3.5848, s=0.1111			
$\hat{\lambda}$	0.5223	0.1235	0.9668	0.0746	0.5507	0.1357	0.5002	0.1248
W'	0.9909	0.0041	0.9697	0.0081	0.9910	0.0040	0.9906	0.0041
K_{mn}	3.7826	0.0898	3.6174	0.0883	3.7808	0.0900	3.7836	0.0901
	Weibull: α=4, β=5, n=100							
	Sample W': m=0.9881, s=0.0058				Sample K_{mn}: m=3.7710, s=0.0921			
$\hat{\lambda}$	1.0444	0.2474	1.8815	0.1174	1.1015	0.2713	1.0001	0.2497
W'	0.9909	0.0040	0.9717	0.0087	0.9910	0.0040	0.9906	0.0041
K_{mn}	3.7826	0.0898	3.6323	0.0893	3.7808	0.0900	3.7836	0.0901
	Weibull: α=4, β=10, n=100							
	Sample W': m=0.9881, s=0.0058				Sample K_{mn}: m=3.7710, s=0.0921			
$\hat{\lambda}$	1.0444	0.2475	1.4724	0.0926	1.1015	0.2714	1.0001	0.2498
W'	0.9909	0.0041	0.9854	0.0056	0.9910	0.0040	0.9906	0.0041
K_{mn}	3.7826	0.0898	3.7330	0.0867	3.7808	0.0900	3.7836	0.0901

8. APPLICATION

To demonstrate the proposed procedure of normalizing the response variable, the **BOQ Data** of Example 5.2 of Myers (1990) is used. A multiple regression line is fitted and the R^2 (coefficient of determination) value is 0.961. The values of W' and the K_{mn} for the residuals are computed as 0.8521 and 2.2878, respectively. Note that since the 5% significance level values of the W' and the K_{mn} are 0.918 and 2.93 (Vasicek (1976)), the departure from normality is detected and hence the Box-Cox transformation (1) is applied. The four methods described in this paper are applied and $\hat{\lambda}_L = 0.26$, $\hat{\lambda}_R = 0.20$, $\hat{\lambda}_W = 0.40$, and $\hat{\lambda}_K = 0.29$ are computed. Then, multiple regression lines are fitted for each transformed response. The R^2 values are reported as 0.896, 0.884, 0.919 and 0.901 with respect to $\hat{\lambda}_L$, $\hat{\lambda}_R$, $\hat{\lambda}_W$, and $\hat{\lambda}_K$, respectively. The corresponding W' statistics are computed as 0.9488, 0.9427, 0.9546, and 0.9510. And the corresponding K_{mn} statistics are computed as 2.9690, 2.9568, 2.9511, and 2.9706.

9. GENERAL REMARKS

Rahman (1999) showed that even though the Artificial Regression Model provides estimates with small variabilities, the estimates may often lead to nonnormal transformed data. Normal likelihood procedure estimates are asymptotically unbiased and consistent and will rarely lead to nonnormal transformed data. For small samples, the Shapiro-Francia W' statistic procedure is reliable.

We have demonstrated that maximization of the Entropy Estimate K_{mn} leads to the estimate of λ which has smaller mean squared errors for larger samples. To apply the method of maximizing the K_{mn} statistic, the dependence of statistical tables can be minimized using simulation as done in Table 1, the empirical distribution for K_{mn} and the values of m are results of simulation. The table for means of order statistics for the standard normal variates not needed, which is essential for the W' maximization method.

10. ACKNOWLEDGEMENTS

We thank the referees for constructive comments and suggestions which improved the earlier version of the paper.

References

1. Bickel, P. J. and K. A. Doksum (1981). "An Analysis of Transformations Revisited." *Journal of the American Statistical Association,* 76, 296-311.

2. Box, G. E. P. and D. R. Cox (1964). "An Analysis of Transformations." *Journal of the Royal Statistical Society, Series B.,* 26, 211-252.

3. Box, G. E. P. and D. R. Cox (1982). "An Analysis of Transformations Revisited (Rebutted)." *Journal of the American Statistical Association,* 77, 209-210.

4. Carroll, R. J. (1980). "A Robust Method for Testing Transformations to Achieve Approximate Normality." *Journal of the Royal Statistical Society, Series B.,* 42, 71-78.

5. Halawa, Adel M. (1996). "Estimating the Box-Cox Transformation via an Artificial Regression Model." *Communications in Statistics — Simulation and Computation,* 25(2), 331-350.

6. Harter, H. Leon (1961). "Expected Values of Normal Order Statistics." *Biometrika,* 48, 1 and 2, 151-165.

7. Hinkley, D. V. (1975). "On Power Transformation to Symmetry." *Biometrika,* 62, 101-111.

8. Hinkley, D. V. (1977). "On Quick Choice of Power Transformation." *Applied Statistics,* 26, 67-68.

9. Lin, L. I. and E. F. Vonesh (1989). "An Empirical Nonlinear Data-Fitting Approach for Transforming Data to Normality." *American Statistician,* 43, 237-243.

10. Myers, Raymond H. (1990). *Classical and Modern Regression with Applications,* Second Edition, Boston, The Duxbury Advanced Series in Statistics and Decision Sciences.

11. Pearson, E. S., R. B. D'Agostino, and K. O. Bowman (1977). "Tests for Departure from Normality: Comparison of Powers." *Biometrika,* 64, 231-246.

12. Rahman, Mezbahur (1999). "Estimating the Box-Cox Transformation via Shapiro-Wilk W Statistic." *Communications in Statistics - Simulation and Computation,* 28(1), 223-241.

13. Shannon, C. E. (1949). *The Mathematical Theory of Communication.* Urbana: University of Illinois Press.

14. Shapiro, S. S. and R. S. Francia (1972). "An Approximate Analysis of Variance Test for Normality." *Journal of the American Statistical Association,* 67(337), 215-216.

15. Shapiro, S. S. and M. B. Wilk (1965). "An Analysis of Variance Test for Normality." *Biometrika,* 52, 3 and 4, 591-611.

16. Shapiro, S. S., M. B. Wilk, and H. J. Chen (1968). "A Comparative Study of Various Tests of Normality." *Journal of the American Statistical Association,* 63, 1343-1372.

17. Stephens, M. A. (1974). "Asymptotic Results for Goodness-of-fit Statistics with Unknown Parameters." *The Annals of Statistics,* 4(2), 357-369.

18. Taylor, J. M. G. (1985). "Power Transformations to Symmetry." *Annals of Mathematical Statistics,* 33, 1-67.

19. Vasicek, Oldrich (1976). "A Test for Normality Based on Sample Entropy." *Journal of the Royal Statistical Society," Series B, 38(1), 54-59.*