Journal of the Korean
Statistical Society
Vol. 26, No. 4, 1997

Asymptotic Relative Efficiency of t-test Following Transformations

In-Kwon Yeo [1]

Abstract

The two-sample t-test is not expected to be optimal when the two samples are not drawn from normal populations. According to Box and Cox (1964), the transformation is estimated to enhance the normality of the transformed data. We investigate the asymptotic relative efficiency of the ordinary t-test versus t-test applied transformation introduced by Yeo and Johnson (1997) under Pitman local alternatives. The theoretical and simulation studies show that two-sample t-test using transformed data gives higher power than ordinary t-test for location-shift models.

Key Words : Asymptotic relative efficiency; Pitman local alternative, Power transformation.

[1]Department of Statistics, Unversity of Wisconsin, Madison, WI 53706, USA.

1. INTRODUCTION

Many parametric techniques in statistics are based on the assumption that the underlying distribution of the data is normal. The validity of the results obtained depends, sometimes critically, on the assumed conditions being satisfied, at least approximately. However, there are several situations where the normal assumptions are seriously violated and the techniques are sensitive to departures from normality. In these cases, a transformation of the data may permit the valid use of the powerful normal theory.

Chen and Loh (1992) discussed asymptotic power properties of the Box-Cox (1964) transformed t-test. They obtained the Pitman asymptotic relative efficiency and show that the Box-Cox transformed two-sample t-test is asymptotically more powerful than the ordinary t-test under location-shift alternatives. The Box-Cox transformations are, however, defined only for positive variables. Although a shift parameter can be introduced to handle the situations where the variables are negative but bounded below, the standard asymptotic results of maximum likelihood theory may not apply since the range of distribution is determined by the unknown shift parameter (see Atkinson (1985)). Chen (1995) extended this work to a general family of transformations $h(\lambda, x)$, $\lambda \in \Theta$, where h is a specific function. Here Θ is a subset of the real line. It turns out, however, that the extended transformations mentioned in Chen (1995) are not appropriate for reducing skewness (see Yeo and Johnson (1997)).

In this article, we determine the asymptotic relative efficiency of the two sample t-test under location alternatives, scale alternatives, or location and scale alternatives. We compare the t-test applied to the original data with the t-test using data transformed according to the transformation

$$X^{(\lambda)} = \begin{cases} \big((X+1)^{\lambda} - 1\big)/\lambda & \text{for } X \geq 0,\ \lambda \neq 0, \\ \log(X+1) & \text{for } X \geq 0,\ \lambda = 0, \\ -\big((-X+1)^{2-\lambda} - 1\big)/(2-\lambda) & \text{for } X < 0,\ \lambda \neq 2, \\ -\log(-X+1) & \text{for } X < 0,\ \lambda = 2, \end{cases} \tag{1.1}$$

which is introduced by Yeo and Johnson (1997). According to the definition of relative skewness introduced by van Zwet (1964), the transformation (1.1) is appropriate to improve the symmetry of skewed distributions. It has properties similar to those of the Box-Cox transformation for positive variables. In particular, there is a convexity (or concavity) property in parameters. We refer the reader to Yeo and Johnson (1997) for the details of transformation (1.1).

2. LOCATION AND SCALE SHIFT MODELS

Let $X_1, \ldots, X_{n_1}$ be independent and distributed as X and let $X_{n_1+1}, \ldots, X_{n_1+n_2}$ be independent and also distributed as X. Assume that $X_1, \ldots, X_{n_1}$ and $X_{n_1+1}, \ldots, X_{n_1+n_2}$ are also independent. Let $n = n_1+n_2$ be the combined sample size with $n_1/n \to \xi \in (0,1)$ as $n_1, n_2 \to \infty$. We consider a general location and scale alternative, where $Y_1, \ldots, Y_{n_2}$ are defined by

$$Y_i = (1+\tau_n)^r \left(X_{n_1+i} + u\tau_n\right), \quad i = 1, \ldots, n_2,$$

with $r \geq 0$, $u \geq 0$, $u + r > 0$ and $\tau_n = \tau/\sqrt{n}$. To simplify the notation, we suppress subscripts on the random variables and write $Y = (1+\tau_n)^r \left(X + u\tau_n\right)$.

We consider testing the hypotheses of equal populations

$$H_0 : \tau = 0 \quad \text{versus} \quad H_A : \tau > 0.$$

The alternative hypothesis H_A relates X and Y according to a location-scale-shift model in original scale where τ_n enters both location and scale parameters. The location-shift and scale-shift models of H_A correspond to the cases $(r = 0, u = 1)$ and $(r = 1, u = 0)$, respectively.

We assume that, for some λ, the transformed variables $X_1^{(\lambda)}, \ldots, X_{n_1}^{(\lambda)}$ and $Y_1^{(\lambda)}, \ldots, Y_{n_2}^{(\lambda)}$ can be treated as normally, as closely as possible, distributed with same variances. Let $\bar{X}^{(\lambda)} = \sum_{i=1}^{n_1} X_i^{(\lambda)}/n_1$, $\bar{Y}^{(\lambda)} = \sum_{i=1}^{n_2} Y_i^{(\lambda)}/n_2$, and let $S_n^2(\lambda)$ be the pooled sample variance for the two transformed samples,

$$S_n^2(\lambda) = \frac{1}{n-2}\left\{\sum_{i=1}^{n_1}\left(X_i^{(\lambda)} - \bar{X}^{(\lambda)}\right)^2 + \sum_{i=1}^{n_2}\left(Y_i^{(\lambda)} - \bar{Y}^{(\lambda)}\right)^2\right\}, \quad n = n_1 + n_2.$$

Then, the t-test statistic, based on the transformed samples, is

$$t_n(\lambda) = \sqrt{\frac{n_1 n_2}{n}}\left(\bar{Y}^{(\lambda)} - \bar{X}^{(\lambda)}\right)/S_n(\lambda). \tag{2.1}$$

Under the assumptions above, we obtain the maximum likelihood estimator (M.L.E.) of the transformation parameter, $\hat{\lambda}$, by minimizing the function

$$P_n(\lambda) = S_n^2(\lambda)/J^{2\lambda} \tag{2.2}$$

where

$$\log(J) = \frac{1}{n}\left(\sum_{i=1}^{n_1}\left(I_{(X_i\geq 0)}\log(X_i+1) - I_{(X_i<0)}\log(-X_i+1)\right)\right.$$
$$\left. + \sum_{i=1}^{n_2}\left(I_{(Y_i\geq 0)}\log(Y_i+1) - I_{(Y_i<0)}\log(-Y_i+1)\right)\right).$$

Suppose that the expected values $E\left[I_{(X<0)}(-X)^{2(2-\lambda)}\right]$ and $E\left[I_{(X\geq 0)}X^{2\lambda}\right]$ are finite. Then, for fixed λ, the strong law of large numbers ensures that under H_0,

$$P_n(\lambda) \xrightarrow{a.s} P_0(\lambda) = \sigma^2(\lambda)\exp(-2\lambda\eta) \tag{2.3}$$

where $\sigma^2(\lambda) = \mathrm{var}\left[X^{(\lambda)}\right]$ and $\eta = E\left[I_{(X\geq 0)}\log(X+1) - I_{(X<0)}\log(-X+1)\right]$.

3. ASYMPTOTIC RESULTS AND A.R.E.

Throughout this paper, we assume that the M.L.E. is calculated on a compact interval $\Lambda = [a,b]$ with $-\infty < a < b < \infty$ and the expected values $E\left[I_{(X<0)}(-X)^{2(2-a)}(\log(-X+1))^2\right]$ and $E\left[I_{(X\geq 0)}X^{2b}(\log(X+1))^2\right]$ are finite. When $a=2$, $E\left[I_{(X<0)}(-X)^{2(2-a)}(\log(-X+1))^2\right]$ is replaced by $E\left[I_{(X<0)}(\log(-X+1))^4\right]$. When $b=0$, $E\left[I_{(X\geq 0)}X^{2b}(\log(X+1))^2\right]$ is replaced by $E\left[I_{(X\geq 0)}(\log(X+1))^4\right]$. We write $\lambda_0 =$ arg min $P_0(\lambda)$ and assume that λ_0 is an interior point of Λ.

In order to explore the limiting distribution of $t_n(\hat\lambda)$, we first show that $t_n(\hat\lambda) - t_n(\lambda_0) = o_p(1)$. Yeo and Johnson (1997) show that the M.L.E., $\hat\lambda$, is a strongly consistent estimator of λ_0 and $\sqrt{n}(\hat\lambda-\lambda_0)$ has a normal limiting distribution under H_0. We write $X^{(\lambda)} = \psi(\lambda, X)$ and $Y^{(\lambda)} = \psi(\tau_n,\lambda,X) = \psi\left(\lambda,(1+\tau_n)^r(X+u\tau_n)\right)$ to parallel Chen and Loh (1992). Then, letting r and u be finite fixed numbers, and setting $\tau_n\in[0,1]$, we can also obtain the strong consistency and limiting normality of $\hat\lambda$ under H_A (see Chen and Loh (1992)).

We next switch from their argument and expand $\sqrt{n}\left(\bar Y^{(\hat\lambda)} - \bar X^{(\hat\lambda)}\right)$ about λ_0 as in Doksum and Wong (1983). We have

$$\sqrt{n}\left(\bar Y^{(\hat\lambda)} - \bar X^{(\hat\lambda)}\right) = \sqrt{n}\left(\bar Y^{(\lambda_0)} - \bar X^{(\lambda_0)}\right) + \sqrt{n}(\hat\lambda-\lambda_0)\left(\nabla\bar Y^{(\lambda_*)} - \nabla\bar X^{(\lambda_*)}\right),$$

where $|\lambda_* - \lambda_0| \le |\hat{\lambda} - \lambda_0|$. Applying Rubin (1956) to $\nabla\bar{X}^{(\lambda_*)}$ and $\nabla\bar{Y}^{(\lambda_*)}$, we conclude that $\nabla\bar{X}^{(\lambda)} \xrightarrow{a.s} E\left[\nabla X^{(\lambda)}\right]$ and $\nabla\bar{Y}^{(\lambda)} \xrightarrow{a.s} E\left[\nabla Y^{(\lambda)}\right]$ uniformly in λ for $\lambda \in \Lambda$. Since $E\left[\nabla Y^{(\lambda)}\right]$ is uniformly continuous in $(\lambda, \tau_n) \in \Lambda \times [0,1]$, $E\left[\nabla Y^{(\lambda)}\right]$ converges to $E\left[\nabla X^{(\lambda)}\right]$ uniformly in λ as $\tau_n = \tau/\sqrt{n} \to 0$. Consequently, $\nabla\bar{Y}^{(\lambda_*)} \xrightarrow{a.s} E\left[\nabla X^{(\lambda_*)}\right]$ under H_0 as well as H_A. Since $\sqrt{n}\left(\hat{\lambda} - \lambda_0\right) = O_p(1)$,

$$\sqrt{n}\left(\bar{Y}^{(\hat{\lambda})} - \bar{X}^{(\hat{\lambda})}\right) = \sqrt{n}\left(\bar{Y}^{(\lambda_0)} - \bar{X}^{(\lambda_0)}\right) + o_p(1). \tag{3.1}$$

The uniform convergence of the pooled sample variance and the strong consistency of $\hat{\lambda}$ ensure that $S_n(\lambda_0) = \sigma(\lambda_0) + o_p(1)$ and $S_n^2(\hat{\lambda}) = \sigma^2(\hat{\lambda}) + o_p(1) = \sigma^2(\lambda_0) + o_p(1)$ under H_0 as well as H_A. Since $\sqrt{n}\left(\bar{Y}^{(\lambda_0)} - \bar{X}^{(\lambda_0)}\right) = O_p(1)$ and $S_n(\hat{\lambda}) - S_n(\lambda_0) = o_p(1)$, using (3.1), we obtain that

$$\begin{aligned} t_n(\hat{\lambda}) - t_n(\lambda_0) &= \sqrt{\frac{n_1 n_2}{n}}\left(\frac{\left(\bar{Y}^{(\hat{\lambda})} - \bar{X}^{(\hat{\lambda})}\right)}{S_n(\hat{\lambda})} - \frac{\left(\bar{Y}^{(\lambda_0)} - \bar{X}^{(\lambda_0)}\right)}{S_n(\lambda_0)}\right) \\ &= \sqrt{\frac{n_1 n_2}{n}}\left(\bar{Y}^{(\lambda_0)} - \bar{X}^{(\lambda_0)}\right)\left(\frac{S_n(\lambda_0) - S_n(\hat{\lambda})}{S_n(\hat{\lambda})S_n(\lambda_0)} + o_p(1)\right) = o_p(1). \end{aligned}$$

Hence, we obtain that $t_n(\hat{\lambda}) - t_n(\lambda_0) = o_p(1)$ under H_0 as well as H_A.

Next, we show that $t_n(\hat{\lambda}) \xrightarrow{d} N(\mu_{\lambda_0}, 1)$, where the limiting mean μ_{λ_0} is

$$\begin{aligned} \mu_{\lambda_0} = \tau\sqrt{\xi(1-\xi)}E\Big[(rX + u)\Big(&I_{(X \ge 0)}(X+1)^{\lambda_0 - 1} \\ &+ I_{(X<0)}(-X+1)^{1-\lambda_0}\Big)\Big]/\sigma(\lambda). \end{aligned} \tag{3.2}$$

In particular, for $\lambda_0 = 1$, $\mu_1 = \tau\sqrt{\xi(1-\xi)}\left(rE[X] + u\right)/\sigma$. Define

$$\begin{aligned} Z_n(\lambda) &= \sqrt{n_1 n_2/n}\left(\bar{Y}^{(\lambda)} - \bar{X}^{(\lambda)} - \left(E\left[Y^{(\lambda)}\right] - E\left[X^{(\lambda)}\right]\right)\right)/S_n(\lambda), \\ \mu_n(\lambda) &= \sqrt{n_1 n_2/n}\left(E\left[Y^{(\lambda)}\right] - E\left[X^{(\lambda)}\right]\right)/S_n(\lambda). \end{aligned}$$

Then, from (2.1), $t_n(\lambda) = Z_n(\lambda) + \mu_n(\lambda)$. By the central limit theorem and the almost sure convergence of $S_n(\lambda)$ to $\sigma(\lambda)$, we have $Z_n(\lambda) \xrightarrow{d} N(0,1)$.

Expanding $Y^{(\lambda)} = \psi(\tau_n, \lambda, X)$ about $\tau_n = 0$, we see that $Y^{(\lambda)} = \psi(0, \lambda, X) + \tau_n \left.\frac{\partial}{\partial \tau_n}\psi(\tau_n, \lambda, X)\right|_{\tau_n = \tau_n^*}$, where $0 \le \tau_n^* \le \tau_n$. Since $\psi(0, \lambda, X) = X^{(\lambda)}$ and

$$\lim_{n\to\infty} E\left[\left.\frac{\partial}{\partial \tau_n}\psi(\tau_n, \lambda, X)\right|_{\tau_n=\tau_n^*}\right]$$
$$= E\Big[(rX+u)\Big(I_{(X\ge 0)}(X+1)^{\lambda-1} + I_{(X<0)}(-X+1)^{1-\lambda}\Big)\Big],$$

$\mu_n(\lambda)$ converges to μ_λ as $n \to \infty$. Therefore, $t_n(\hat\lambda) \xrightarrow{d} N(\mu_{\lambda_0}, 1)$ since $t_n(\hat\lambda) - t_n(\lambda_0) = o_p(1)$.

Theorem 1. Let $t(\lambda_0)$ be the test based on $t_n(\hat\lambda)$ and let t be the ordinary t-test, $t = t_n(1)$, based on the untransformed data. Then, the A.R.E. of $t(\lambda_0)$ to t is

$$ARE(t(\lambda_0), t) = \Big(pA_{\lambda_0} + qB_{\lambda_0}\Big)^2 \tag{3.3}$$

where $p = u/|rE[X]+u|$, $q = rE[X]/|rE[X]+u|$ and

$$A_{\lambda_0} = \sigma E\Big[I_{(X\ge 0)}(X+1)^{\lambda_0-1} + I_{(X<0)}(-X+1)^{1-\lambda_0}\Big]/\sigma(\lambda), \tag{3.4}$$
$$B_{\lambda_0} = \sigma\Big\{E\Big[X\Big(I_{(X\ge 0)}(X+1)^{\lambda_0-1} + I_{(X<0)}(-X+1)^{1-\lambda_0}\Big)\Big]/E[X]\Big\}/\sigma(\lambda). \tag{3.5}$$

Furthermore, the A.R.E. of location-shift model, $(r = 0, u = 1)$, is

$$ARE(t(\lambda_0), t) = A_{\lambda_0}^2 \ge 1, \tag{3.6}$$

and the A.R.E. of scale-shift model, $(r = 1, u = 0)$, is

$$ARE(t(\lambda_0), t) = B_{\lambda_0}^2 \ge 1, \tag{3.7}$$

when $\Big(\lambda_0 > 1, E[X] > 0\Big)$ or $\Big(\lambda_0 < 1, E[X] < 0\Big)$.

Proof. Let $\beta(\tau_n|t^*, n)$ be the power of test t^* at the local alternatives based on sample size n. Since $t_n(\hat\lambda) \xrightarrow{d} N(\mu_{\lambda_0}, 1)$ where μ_{λ_0} is defined in (3.2), we conclude that the limiting power $1 - \Phi(z_\alpha - \mu_{\lambda_0}) = \lim_{n\to\infty}\beta(\tau_n|t(\lambda_0), n)$, where Φ denotes the standard normal distribution function. Modifying the sample size to n' for $t = t(1)$, and matching limiting power, we determine

$$\lim_{n\to\infty}\beta(\tau_n|t, n') = \lim_{n\to\infty}\beta\Big((n'/n)^{1/2}\tau_{n'}|t, n'\Big) = 1 - \Phi(z_\alpha - \gamma^{1/2}\mu_1).$$

Therefore, we have

$$\Phi(z_\alpha - \mu_{\lambda_0}) = \Phi(z_\alpha - \gamma^{1/2}\mu_1).$$

According to (3.2),

$$\begin{aligned}\mu_{\lambda_0} &= \tau\sqrt{\xi(1-\xi)}\Big\{uE\Big[I_{(X\geq 0)}(X+1)^{\lambda_0-1} + I_{(X<0)}(-X+1)^{1-\lambda_0}\Big] \\ &\quad + rE\Big[X\left(I_{(X\geq 0)}(X+1)^{\lambda_0-1} + I_{(X<0)}(-X+1)^{1-\lambda_0}\right)\Big]\Big\}/\sigma(\lambda_0).\end{aligned}$$

Thus, the limiting power are equal when $\gamma = \mu_{\lambda_0}^2/\mu_1^2 = \left(pA_{\lambda_0} + qB_{\lambda_0}\right)^2$.

By (2.3) and the definition of λ_0, $P_0(\lambda_0) \leq P_0(1)$ so

$$\frac{\sigma^2(\lambda_0)}{\sigma^2} \leq \exp\Big(2(\lambda_0 - 1)\eta\Big). \tag{3.8}$$

By the definition of η above (2.3), Jensen's inequality yields

$$\exp\Big(2(\lambda_0-1)\eta\Big) \leq E^2\Big[I_{(X\geq 0)}(X+1)^{\lambda_0-1} + I_{(X<0)}(-X+1)^{1-\lambda_0}\Big]. \tag{3.9}$$

The right hand side equals $A_{\lambda_0}^2\sigma^2(\lambda_0)/\sigma^2$, according to (3.4), which proves (3.6).

Suppose $\lambda_0 > 1$ and $E[X] > 0$. Then, $I_{(X\geq 0)}(X+1)^{\lambda-1} + I_{(X<0)}(-X+1)^{1-\lambda}$ is clearly increasing in X, so

$$cov\Big[X, I_{(X\geq 0)}(X+1)^{\lambda_0-1} + I_{(X<0)}(-X+1)^{1-\lambda_0}\Big] \geq 0$$

or, by definition of covariance,

$$\begin{aligned}&E\Big[I_{(X\geq 0)}(X+1)^{\lambda_0-1} + I_{(X<0)}(-X+1)^{1-\lambda_0}\Big] \\ &\quad \leq E\Big[X\left(I_{(X\geq 0)}(X+1)^{\lambda_0-1} + I_{(X<0)}(-X+1)^{1-\lambda_0}\right)\Big]/E[X].\end{aligned}$$

Thus, by (3.8) and (3.9),

$$\frac{\sigma^2(\lambda_0)}{\sigma^2} \leq \left(E\Big[X\left(I_{(X\geq 0)}(X+1)^{\lambda_0-1} + I_{(X<0)}(-X+1)^{1-\lambda_0}\right)\Big]/E[X]\right)^2. \tag{3.10}$$

Similarly, for $\lambda_0 < 1$ and $E[X] < 0$, $I_{(X\geq 0)}(X+1)^{\lambda-1} + I_{(X<0)}(-X+1)^{1-\lambda}$ is decreasing in X so

$$cov\Big[X, I_{(X\geq 0)}(X+1)^{\lambda_0-1} + I_{(X<0)}(-X+1)^{1-\lambda_0}\Big] \leq 0$$

and then,

$$E\Big[I_{(X\geq 0)}(X+1)^{\lambda_0-1}+I_{(X<0)}(-X+1)^{1-\lambda_0}\Big]$$
$$\leq E\Big[X\left(I_{(X\geq 0)}(X+1)^{\lambda_0-1}+I_{(X<0)}(-X+1)^{1-\lambda_0}\right)\Big]/E[X],$$

which is the same inequality as (3.10). Therefore, when $(\lambda_0>1, E[X]>0)$ or $(\lambda_0<1, E[X]<0)$, $\mathrm{ARE}(t(\lambda_0),t)\geq 1$.

Remark 1. When E(X) is negative, the tests might have a left-sided rejection region, $t<-t_\alpha$, under a scale shift.

4. NUMERICAL EXAMPLES

Example 3.1 : Scale Shift Model. To compare the performance of the tests based on the original data and the transformed data, we suppose that a random variable X have a mixture distribution: $f_X(x)=0.7\,f_1(x)+0.3\,f_2(x)$ where $f_1(x)$ is gamma distribution with $(\alpha,\beta,\gamma)=(4,1,3.5)$, that is,

$$f_1(x)=\frac{(x+3.5)^3\exp(-(x+3.5))}{\Gamma(4)},\qquad x>-3.5$$

and $f_2(x)$ is normal$(-1.5,1)$. Then the distribution of X is skewed to the right and $E(X)=-0.1$. A numerical calculation gives $\lambda_0=0.5913$ and the $ARE(t(\lambda_0),t)=38.360$ for a scale-shift within this model. Table 1 gives the powers of the ordinary t-test and the transformed t-test for $n_1=20$ and $n_2=25$, at nominal level 0.05. Each result is based on 40,000 Monte Carlo trials. The results of the simulation are consistent with the fact that the transformed t-test is asymptotically more efficient than the ordinary t-test under the scale shift in this example.

Table 1. Simulated powers of the ordinary t-test and the transformed t-test for scale shift in a mixture model.

shift	0.0	0.2	0.4	0.6	0.8	1.0
t	0.049	0.051	0.055	0.056	0.059	0.061
$t(\hat{\lambda})$	0.049	0.066	0.087	0.098	0.110	0.119

Example 3.2 : Location-and-Scale Shift Model. Suppose $Y=-X+0.5$, where X is defined as in Example 3.1. Then the distribution of Y is skewed

to the left and $E(Y) = 0.6 > 0$ so we expect the asymptotic power of the $t(\lambda)$-test to dominate that of the t-test over all shifts. In fact, we calculate that $\lambda_0 = 1.4004$ for this model and

$$ARE(t(\lambda_0), t) = \begin{cases} 1.371 & \text{for a location-shift} \\ 5.577 & \text{for a scale-shift} \\ 2.616 & \text{for a location-and-scale shift} \end{cases}$$

Table 2. Simulated powers of the ordinary t-test and the transformed t-test for location-and-scale shift.

Scale shift		0.0	0.1	0.2	0.3	0.4	0.5
Location shift = 0.0	t	0.050	0.066	0.084	0.095	0.117	0.137
	$t(\hat{\lambda})$	0.050	0.078	0.111	0.142	0.182	0.226
Location shift = 0.2	t	0.124	0.155	0.191	0.222	0.266	0.292
	$t(\hat{\lambda})$	0.137	0.190	0.246	0.316	0.374	0.426
Location shift = 0.4	t	0.249	0.309	0.369	0.416	0.462	0.506
	$t(\hat{\lambda})$	0.281	0.370	0.461	0.535	0.597	0.650
Location shift = 0.6	t	0.434	0.510	0.569	0.634	0.672	0.713
	$t(\hat{\lambda})$	0.481	0.589	0.672	0.749	0.789	0.834
Location shift = 0.8	t	0.621	0.707	0.759	0.807	0.841	0.866
	$t(\hat{\lambda})$	0.684	0.781	0.845	0.890	0.917	0.932
Location shift = 1.0	t	0.791	0.850	0.891	0.920	0.935	0.947
	$t(\hat{\lambda})$	0.846	0.902	0.942	0.964	0.972	0.979

Table 2 gives the powers of the ordinary t-test and the transformed t-test for $n_1 = 25$ and $n_2 = 30$, at nominal level 0.05 according to 40,000 Monte Carlo trials. It is shown that the transformed t-tests have a higher power than the ordinary t-test under all shifts.

Two simulations above re-confirm Chen and Loh's (1992) result that performing a two-sample t-test with transforming the data is preferable.

REFERENCES

(1) Atkinson, A. C. (1985). *Plots, Transformations and Regression,* Oxford University Press, Oxford .

(2) Chen, H (1995). Tests Following Transformations. *Annals of Statistics,* **23**, 1587-1593.

(3) Chen, H and Loh, W. Y. (1992). Bounds on AREs of Tests Following Box-Cox Transformations *Annals of Statistics,* **20**, 1485-1500.

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

(5) Doksum, K. A. and Wong, C. W. (1983). Statistical Tests Based on Transformed Data. *Journal of the American Statistical Association,* **78**, 411-415.

(6) Rubin, H. (1956). Uniform Convergence of Random Functions with Applications to Statistics. *Annals of Mathematical Statistics*, **27**, 200-203.

(7) VAN ZWET, W. R. (1964). *Convex Transformations of Random Variables.* Mathematisch Centrum, Amsterdam.

(8) Yeo, I. K. and Johnson, R. A. (1997). A New Family of Power Transformations to Improve Normality or Symmetry. (summit to *Annals of Statistics*, ms 857).