

Hypothesis Testing on the Location Parameter of a Skew-Normal Distribution (SND) with Application

Phontita Thiuthad¹ and Nabendu Pal^{2,*}

¹Department of Mathematics and Statistics, Faculty of Science, Prince of Songkla University, Hatyai, Songkhla 90110, Thailand

²Department of Mathematics, University of Louisiana at Lafayette, Lafayette, Louisiana 70504, USA

Abstract. This work deals with testing a hypothesis on the location parameter (μ) of a skew-normal distribution (SND) based on a random sample of size n . The details of this work can be summarized in four major components: (a) First we review some useful results on SND, including the approximate probability distribution of the sample average. (b) Next, we develop several tests to test a hypothesis on μ based on the sample mean when the scale (σ) and shape (λ) parameters are known. (c) The tests for the known scale and shape are then extended for unknown scale and shape. (d) Finally, the test methods have been used for a real-life data set.

Keywords: Fisher information, Likelihood ratio test, Neyman-Pearson Lemma, Power function, Size of a test.

MSC 2010 Subject Classifications: 62F03, 62F05, 62E17.

1 Introduction

1.1 Preliminaries

A random variable W is said to have a skew-normal distribution with location parameter μ , scale parameter σ , and shape parameter λ , henceforth denoted by ' $W \sim \text{SND}(\mu, \sigma, \lambda)$ ', provided its probability density function (*pdf*) is given as

$$f(w|\mu, \sigma, \lambda) = (2/\sigma)\phi((w - \mu)/\sigma)\Phi(\lambda(w - \mu)/\sigma), \quad -\infty < w < \infty, \quad (1.1)$$

where ϕ and Φ are the standard normal *pdf* and *cdf* respectively, for any $-\infty < \mu < \infty$, $-\infty < \lambda < \infty$, and $\sigma > 0$. The objective of this work is to develop test procedures for testing $H_0 : \mu = \mu_0$ vs. $H_A : \mu \neq \mu_0$. The two-sided alternative has been used only for convenience since an one-sided test can be developed easily with a suitable one-sided critical value.

*e-mail: nxp3695@louisiana.edu

Note that the above SND family includes the normal family as a special case (i.e., $N(\mu, \sigma^2)$) when the skew parameter $\lambda = 0$. Also, (1.1) includes positively skewed shape (when $\lambda > 0$) and negatively skewed shape (when $\lambda < 0$) which makes SND more appealing than the normal distribution.

The structure of SND was introduced in literature by several researchers independently (see Roberts (1966), O'Hagan and Leonard (1976), Aigner and Lovell (1977)), but it was given the present name by Azzalini (1985) who studied many characterization properties and applications of this distribution. For a nice review of many interesting properties of SND see Arnold and Lin (2004), and Gupta, Nguyen and Sanqui (2004) and other references therein. For applications of SND in reliability studies see Gupta and Brown (2001) which also used the SND to model IQ data taken from Roberts (1988). Figueiredo and Gomes (2013) used SND to model and monitor in statistical process control. Ngunkeng and Ning (2014) considered an information approach to detect changes in all three parameters in an SND, and then applied the proposed method to study Brazilian and Chilean stock market data (originally analyzed by Arellano-Valle et al. (2013)). For a multivariate generalization of the SND see Azzalini and Dalla Valle (1996), and Azzalini and Capitanio (1999).

1.2 Some Useful Properties of SND

The $SND(0, 1, \lambda)$ distribution is called the standard skew-normal distribution, and the following are some useful properties of the $SND(\mu, \sigma, \lambda)$:

- (i) $W \sim SND(\mu, \sigma, \lambda)$ if and only if $Z = (W - \mu)/\sigma \sim SND(0, 1, \lambda)$;
- (ii) $\lambda = 0$ corresponds to the normal distribution $N(\mu, \sigma^2)$;
- (iii) If $W \sim SND(\mu, \sigma, \lambda)$, then $(-W) \sim SND(-\mu, \sigma, -\lambda)$;
- (iv) As $\lambda \rightarrow \pm\infty$, $SND(0, 1, \lambda)$ tends to the half-normal distribution, i.e., the distribution of $\pm|Z_*|$ when $Z_* \sim N(0, 1)$;
- (v) If $Z \sim SND(0, 1, \lambda)$, then $Z^2 \sim \chi_1^2$ (Chi-square with 1 degree freedom);
- (vi) If $U, V \stackrel{iid}{\sim} N(0, 1)$, then $Z = (\lambda/\sqrt{1+\lambda^2})|U| + (1/\sqrt{1+\lambda^2})V \sim SND(0, 1, \lambda)$.
- (vii) The MGF of the standard SND is $M(t) = 2exp(t^2/2)\Phi(\delta t) = 2\psi(t)\Phi(\delta t)$, where $\delta = \lambda/\sqrt{1+\lambda^2}$ and $\psi(t)$ is the MGF of $Z_* \sim N(0, 1)$.
- (viii) If $Z_* \sim N(0, 1)$ is independent of $Z \sim SND(0, 1, \lambda)$, then for any $a, b \in \mathbb{R}$

$$(aZ_* + bZ)/\sqrt{a^2 + b^2} \sim SND(0, 1, (b\lambda)/\sqrt{a^2(1 + \lambda^2) + b^2}).$$

- (ix) The first three central moments of $W \sim SND(\mu, \sigma, \lambda)$ are $E(W) = \mu + \sigma\lambda\sqrt{2/\pi}/\sqrt{1+\lambda^2}$, $Var(W) = \sigma^2\{1 - (2/\pi)(\lambda^2/(1+\lambda^2))\}$, and $E(W - E(W))^3 = \sigma^3\sqrt{(2/\pi)((4/\pi) - 1)}\{\lambda/\sqrt{1+\lambda^2}\}^3$, respectively.

The above property (vi) is important as it is often used to generate SND observations in simulation studies. Property (ix) will be useful in method of moments estimation (MME). For more properties and further generalizations, see Azzalini (2005), Henze (1986), and the other references therein.

1.3 Approximate Distribution of an SND Sample Average

In a recent study Thiuthad and Pal (2018) have shown that the exact distribution of the average of a random sample from SND, though extremely complicated, can be approximated fairly well by a suitable SND as described below.

Assume W_1, W_2, \dots, W_n are iid $SND(\mu, \sigma, \lambda)$. Let $\bar{W} = \sum W_i/n$. It is proposed that $\bar{W} \sim SND(\bar{\mu}, \bar{\sigma}, \bar{\lambda})$ where $\bar{\mu}, \bar{\sigma}, \bar{\lambda}$ are functions of (μ, σ, λ) and n , which can be found (with the details omitted here) as follows.

(A) First, obtain $\bar{\lambda}$ as

$$\bar{\lambda} = \text{sign}(\lambda) \left\{ n^{1/3} [(1 + \lambda^2)/\lambda^2 - (2/\pi)] - [1 - (2/\pi)] \right\}^{-1/2} \tag{1.2}$$

where $\text{sign}(\lambda)$ is > 0 if $\lambda > 0$; < 0 if $\lambda < 0$; and $= 0$, if $\lambda = 0$.

(B) Next, obtain $\bar{\sigma}$ as

$$\bar{\sigma} = \sqrt{\frac{\sigma^2 \{1 - (2/\pi)(\lambda^2/(1 + \lambda^2))\}}{n \{1 - (2/\pi)(\bar{\lambda}^2/(1 + \bar{\lambda}^2))\}}} \tag{1.3}$$

(C) Finally, $\bar{\mu}$ is obtained as

$$\bar{\mu} = \mu + (\sigma\lambda) \sqrt{2/\pi} / \sqrt{1 + \lambda^2} - (\bar{\sigma}\bar{\lambda}) \sqrt{2/\pi} / \sqrt{1 + \bar{\lambda}^2} \tag{1.4}$$

Remark 1.1 Note that $\bar{\mu}, \bar{\sigma}, \bar{\lambda}$ are functions of $(\mu, \sigma, \lambda, n)$. Further, $(\bar{\sigma}, \bar{\lambda})$ is dependent only on (σ, λ, n) . Interestingly, it is easy to see that when $\lambda = 0$, then $\bar{\lambda} = 0$ and $\bar{\sigma} = \sigma/\sqrt{n}$ as well as $\bar{\mu} = \mu$, which perfectly matches with the result of a normal distribution. It has been shown through extensive simulation that the proposed approximation (of \bar{W} by $SND(\bar{\mu}, \bar{\sigma}, \bar{\lambda})$) is very close even for small n .

The rest of the paper is organized as follows. Section 2 gives a brief summary of two types of point estimators as they will be essential in constructing the test statistics. Section 3 deals with various tests on μ with known (σ, λ) , whereas Section 4 provides details in the most general case, that is with unknown (σ, λ) . Finally, a real-life data set has been used in Section 5 as a demonstration.

2 A Brief Review of Point Estimation of μ

Thiuthad and Pal (2018) have provided the details of point estimation of μ under two separate cases as follows.

2.1 σ and λ are known

Here, the method of moments estimator (MME) of μ , say $\hat{\mu}_{MM}$, is given as

$$\hat{\mu}_{MM} = \bar{W} - \sigma \sqrt{2/\pi} (\lambda / \sqrt{1 + \lambda^2}); \tag{2.1}$$

and the maximum likelihood estimator (MLE) of μ , say $\hat{\mu}_{ML}$, is found by solving

$$(\bar{W} - \mu)/\sigma = (\lambda/n) \sum_{i=1}^n \phi(\lambda(W_i - \mu)/\sigma)/\Phi(\lambda(W_i - \mu)/\sigma). \quad (2.2)$$

In this case, $\hat{\mu}_{MM}$ is unbiased for μ and $Var(\hat{\mu}_{MM}) = (\sigma^2/n)\{1 - (2/\pi)(\lambda^2/(1 + \lambda^2))\}$. By CLT, $\hat{\mu}_{MM} \xrightarrow{d} N(\mu, Var(\hat{\mu}_{MM}))$ as $n \rightarrow \infty$. By MLE's asymptotic property, $\hat{\mu}_{ML} \sim N(\mu, 1/(nI_0))$, where $I_0 =$ Fisher information per observation $= (1/\sigma^2)[1 + \lambda^2 E\{\{\phi(\lambda Z)/\Phi(\lambda Z)\}^2\}]$.

2.2 σ and λ are unknown

The MME of μ depends on those of λ and σ . Therefore, the following steps are followed in a sequential order.

Let $m_1 = \sum_{i=1}^n W_i/n$ and $m_j = \sum_{i=1}^n (W_i - m_1)^j/n, j = 2, 3$, be the first three sample moments. First obtain $\hat{\lambda}$ by solving (provided a solution exists)

$$\{((1 + \hat{\lambda}^2)/\hat{\lambda}^2) - (2/\pi)\}^3 - (2/\pi)((4/\pi) - 1)^2(m_2^3/m_3^2) = 0. \quad (2.3)$$

A solution of (2.3) exists provided $m_2^3/m_3^2 > (\pi - 2)^3/(2(4 - \pi)^2) = c_0$ (say). If a solution of (2.3) in terms of $\hat{\lambda}^2$ does not exist, then $\hat{\lambda}^2$ is deemed to have taken the value ∞ . Thus, the MME of λ , i.e., $\hat{\lambda}_{MM}$ is defined as

$$\hat{\lambda}_{MM} = \begin{cases} \text{sign}(m_3) \times \hat{\lambda}, & \text{if (2.3) has a solution.} \\ \text{sign}(m_3) \times \infty, & \text{otherwise.} \end{cases} \quad (2.4)$$

For practical purposes, in (2.4) replace ' ∞ ' by a sufficiently large value, say M which makes the estimated distribution almost half normal. After estimating λ as above, obtain $\hat{\sigma}_{MM}$ as

$$\hat{\sigma}_{MM} = \begin{cases} \sqrt{m_2/\{1 - (2/\pi)(\hat{\lambda}_{MM}^2/(1 + \hat{\lambda}_{MM}^2))\}}, & \text{if (2.3) has a solution.} \\ \sqrt{m_2/\{1 - (2/\pi)\}}, & \text{otherwise.} \end{cases} \quad (2.5)$$

Finally,

$$\hat{\mu}_{MM} = \begin{cases} m_1 - \hat{\sigma}_{MM} \sqrt{2/\pi}(\hat{\lambda}_{MM}/\sqrt{1 + \hat{\lambda}_{MM}^2}), & \text{if (2.3) has a solution.} \\ m_1 - \hat{\sigma}_{MM} \sqrt{2/\pi}, & \text{otherwise.} \end{cases} \quad (2.6)$$

Note that even if $\hat{\lambda}_{MM}$ takes the infinite value with a positive probability, it does not prevent the MME of σ and λ from taking finite values with probability 1.

When all parameters are unknown, the MLE of μ may not be a viable option as it may result into infinite value with a positive probability. Hence, a maximum penalized likelihood estimate (MPLE) looks more reasonable which maximizes

$$L_{**} = n \ln(2/\sigma) + \sum_{i=1}^n \ln\phi((W_i - \mu)/\sigma) + \sum_{i=1}^n \ln\Phi(\lambda(W_i - \mu)/\sigma) - c_1 \ln(1 + c_2 \lambda^2), \quad (2.7)$$

with $c_1 = 3\pi^2/32$ and $c_2 = 8/\pi^2$. The last term in (2.7) is the penalty function as suggested by Azzalini and Arellano-Valle (2012).

Thiuthad and Pal (2018) provided results of their extensive simulation study comparing $\hat{\mu}_{MM}$ and $\hat{\mu}_{MPL}$. Further, as noted by these authors, the Fisher information matrix $I = nI_0$ where I_0 is the 3×3 information matrix per observation) has at least three singularity points rendering I^{-1} useless as a benchmark dispersion matrix for the estimators of the parameter vector (μ, σ, λ) .

3 Tests on the location of an SND with known (σ, λ)

3.1 The Case of a Single Observation

To develop the idea of testing on μ with known (σ, λ) we first consider the case of a single observation $W \sim \text{SND}(\mu, \sigma, \lambda)$. (The method developed here can easily be transported to a test based on $\bar{W} \sim \text{SND}(\bar{\mu}, \bar{\sigma}, \bar{\lambda})$.) As mentioned earlier, we test $H_0 : \mu = \mu_0$ against a suitable alternative which, for convenience, has been taken as the two sided one, i.e., $H_A : \mu \neq \mu_0$. To develop the idea further, we consider a simple alternative as $H_A : \mu = \mu_A$, where $\mu_A \neq \mu_0$.

The likelihood ratio test (LRT) statistic for the above simple null against the simple alternative becomes

$$\Lambda = \frac{f(W|\mu_0, \sigma, \lambda)}{f(W|\mu_A, \sigma, \lambda)} = \frac{\phi((W - \mu_0)/\sigma)\Phi(\lambda(W - \mu_0)/\sigma)}{\phi((W - \mu_A)/\sigma)\Phi(\lambda(W - \mu_A)/\sigma)}. \tag{3.1}$$

It is more convenient to work with the natural logarithm of the LRT statistic, i.e.,

$$\Lambda_* \text{ (say)} = \ln \Lambda = \frac{-(W - \mu_0)^2}{2\sigma^2} + \frac{(W - \mu_A)^2}{2\sigma^2} + \ln \Phi(\lambda(W - \mu_0)/\sigma) - \ln \Phi(\lambda(W - \mu_A)/\sigma). \tag{3.2}$$

Write $g_0(W) = \ln \Phi(\lambda(W - \mu_0)/\sigma)$, which has the following Taylor series expression

$$g_0(W) = g(\mu_0) + (W - \mu_0)g'(\mu_0) + \frac{(W - \mu_0)^2}{2}g''(\mu_0) + \dots \approx -\ln 2 + \frac{\lambda}{\sigma} \sqrt{\frac{2}{\pi}}(W - \mu_0) - \frac{\lambda^2}{\pi\sigma^2}(W - \mu_0)^2, \tag{3.3}$$

by retaining up to the second order terms. Similarly,

$$g_A(W) = \ln \Phi(\lambda(W - \mu_A)/\sigma) \approx -\ln 2 + \frac{\lambda}{\sigma} \sqrt{\frac{2}{\pi}}(W - \mu_A) - \frac{\lambda^2}{\pi\sigma^2}(W - \mu_A)^2. \tag{3.4}$$

Using (3.3) and (3.4), we get a second order approximation of Λ_* , say $\Lambda_*^{(2)}(W)$, as

$$\Lambda_*(W) \approx \Lambda_*^{(2)}(W) := (\mu_0 - \mu_A) \left\{ (2W - \mu_0 - \mu_A) \left(\frac{1}{2\sigma^2} + \frac{\lambda^2}{\pi\sigma^2} \right) - \frac{\lambda}{\sigma} \sqrt{\frac{2}{\pi}} \right\}. \tag{3.5}$$

Let $\delta = \mu_0 - \mu_A$ and $W' = W - \mu_0$. Then, $W' \sim \text{SND}(0, \sigma, \lambda)$ under H_0 and,

$$\Lambda_*^{(2)}(W) = \Lambda_*^{(2)}(W') = \delta \left\{ (2W' + \delta) \left(\frac{1}{2\sigma^2} + \frac{\lambda^2}{\pi\sigma^2} \right) - \frac{\lambda}{\sigma} \sqrt{\frac{2}{\pi}} \right\}. \quad (3.6)$$

Using the Neyman-Pearson Lemma, retain H_0 if $\Lambda_* \geq k$, and reject H_0 otherwise. Therefore, using the second order approximation, retain H_0 if $\Lambda_*^{(2)} \geq k$, and reject H_0 if $\Lambda_*^{(2)} < k$, where k is found by the level condition $P(\text{Type-I error}) = \alpha$. This is considered in two cases:

Case (a) If $\mu_A > \mu_0$, then $\delta = \mu_0 - \mu_A < 0$. Hence, $\Lambda_*^{(2)} < k$ if and only if $W' > k'_1$.

Case (b) If $\mu_A < \mu_0$, then $\delta = \mu_0 - \mu_A > 0$. Hence, $\Lambda_*^{(2)} < k$ if and only if $W' < k'_2$.

Note that the cut-off points k'_1 and k'_2 depends only on the level condition, not the exact value of μ_A on either side of μ_0 . Therefore, for a two sided alternative $H_A : \mu \neq \mu_0$, we reject H_0 based on the second order LRT statistic if $W' < k'_2$ or $W' > k'_1$, and retain H_0 if $k'_2 \leq W' \leq k'_1$, where k'_1 and k'_2 can be chosen such that

$$G(k'_2) = \int_{-\infty}^{k'_2} (2/\sigma)\phi(u/\sigma)\Phi(\lambda(u/\sigma))du = (\alpha/2), \quad (3.7)$$

where, G represents the *cdf* of $\text{SND}(0, \sigma, \lambda)$; i.e., $k'_2 = G^{-1}(\alpha/2)$. Similarly,

$$1 - G(k'_1) = \int_{k'_1}^{\infty} (2/\sigma)\phi(u/\sigma)\Phi(\lambda(u/\sigma))du = (\alpha/2), \quad (3.8)$$

i.e., $k'_1 = G^{-1}(1 - (\alpha/2))$.

Further, note that the cut-off points k'_1 and k'_2 can be expressed as $k'_i = \sigma k'_{i*}$, $i = 1, 2$, where k'_{i*} ($i = 1, 2$) are found by the following equations.

$$\int_{-\infty}^{k'_{2*}} 2\phi(u_*)\Phi(\lambda u_*)du_* = (\alpha/2) \text{ and } \int_{-\infty}^{k'_{1*}} 2\phi(u_*)\Phi(\lambda u_*)du_* = 1 - (\alpha/2). \quad (3.9)$$

The following Table 3.1 tabulates the values of k'_{i*} ($i = 1, 2$) for various values of α . Note that when $\lambda = 0$, i.e., $\text{SND}(0, \sigma, 0)$, the cut-off points are exactly those of $N(0, \sigma^2)$. Also note that for a one sided alternative one should use only one of k'_{1*} and k'_{2*} with $(\alpha/2)$ replaced by α in one of the equations in (3.9).

Remark 3.1 The above Table 3.1 lists the two sided cut-off points only for $\lambda > 0$. Note that $(-W') = -(W - \mu_0) \sim \text{SND}(0, \sigma, -\lambda)$. Therefore, for $\lambda < 0$, one can get the cut-off points easily by interchanging k'_{1*} by k'_{2*} , and k'_{2*} by k'_{1*} with signs reversed.

3.2 The Case of a Sample of Size $n (\geq 1)$

We first look at the two point estimators, i.e., MME and MLE, given earlier as $\hat{\mu}_{MM}$ and $\hat{\mu}_{ML}$ respectively. The exact variance of $\hat{\mu}_{MM}$ is

$$V(\hat{\mu}_{MM}) = V(\bar{W}) = (\sigma^2/n)\{1 - (2/\pi)(\lambda^2/(1 + \lambda^2))\}. \quad (3.10)$$

Table 3.1: The two-sided cut-off points (k'_{2*}, k'_{1*}) satisfying (3.9) (i.e., each tail probability = $\alpha/2$) for various values of $\lambda > 0$.

λ	$\alpha = 0.01$		$\alpha = 0.02$		$\alpha = 0.05$		$\alpha = 0.10$		$\alpha = 0.20$	
	k'_{2*}	k'_{1*}	k'_{2*}	k'_{1*}	k'_{2*}	k'_{1*}	k'_{2*}	k'_{1*}	k'_{2*}	k'_{1*}
0.0	-2.576	2.576	-2.326	2.326	-1.960	1.960	-1.645	1.645	-1.282	1.282
0.5	-2.030	2.786	-1.801	2.548	-1.464	2.199	-1.173	1.900	-0.838	1.556
1.0	-1.471	2.807	-1.282	2.575	-1.002	2.239	-0.760	1.955	-0.478	1.632
1.5	-1.079	2.807	-0.925	2.576	-0.696	2.241	-0.496	1.960	-0.261	1.644
2.0	-0.823	2.807	-0.695	2.576	-0.503	2.241	-0.334	1.960	-0.134	1.645
2.5	-0.652	2.807	-0.543	2.576	-0.378	2.241	-0.232	1.960	-0.056	1.645
3.0	-0.532	2.807	-0.437	2.576	-0.292	2.241	-0.163	1.960	-0.006	1.645
3.5	-0.445	2.807	-0.360	2.576	-0.232	2.241	-0.115	1.960	0.028	1.645
4.0	-0.379	2.807	-0.303	2.576	-0.186	2.241	-0.080	1.960	0.051	1.645
4.5	-0.328	2.807	-0.259	2.576	-0.152	2.241	-0.054	1.960	0.068	1.645
5.0	-0.288	2.807	-0.224	2.576	-0.125	2.241	-0.035	1.960	0.081	1.645
5.5	-0.255	2.807	-0.196	2.576	-0.104	2.241	-0.019	1.960	0.090	1.645
6.0	-0.228	2.807	-0.173	2.576	-0.087	2.241	-0.007	1.960	0.097	1.645
6.5	-0.205	2.807	-0.154	2.576	-0.073	2.241	0.004	1.960	0.103	1.645
7.0	-0.186	2.807	-0.137	2.576	-0.061	2.241	0.012	1.960	0.107	1.645
7.5	-0.170	2.807	-0.123	2.576	-0.051	2.241	0.019	1.960	0.111	1.645
8.0	-0.155	2.807	-0.112	2.576	-0.042	2.241	0.024	1.960	0.113	1.645
8.5	-0.143	2.807	-0.101	2.576	-0.035	2.241	0.029	1.960	0.116	1.645
9.0	-0.132	2.807	-0.092	2.576	-0.029	2.241	0.033	1.960	0.117	1.645
9.5	-0.122	2.807	-0.084	2.576	-0.023	2.241	0.037	1.960	0.119	1.645
10.0	-0.114	2.807	-0.077	2.576	-0.018	2.241	0.040	1.960	0.120	1.645

But the exact variance expression of $\hat{\mu}_{ML}$ is not known. Hence we can use the asymptotic variance (AV) which is quite close to the exact one (see Thiuthad and Pal (2018)).

$$AV(\hat{\mu}_{ML}) = 1/(nI_0). \tag{3.11}$$

Based on the above expressions (3.10) and (3.11), and the asymptotic distributions presented in Subsection 2.1, we propose two tests, say T_1 and T_2 respectively, defined as follows:

$$T_1 : \text{Rej } H_0 \text{ if } |\hat{\mu}_{MM} - \mu_0| / \sqrt{V(\hat{\mu}_{MM})} > z_{\alpha/2} \tag{3.12}$$

$$T_2 : \text{Rej } H_0 \text{ if } |\hat{\mu}_{ML} - \mu_0| / \sqrt{AV(\hat{\mu}_{ML})} > z_{\alpha/2} \tag{3.13}$$

A third test, say T_3 , can be constructed as follows. Based on a random sample of size n from $SND(\mu, \sigma, \lambda)$, we use \bar{W} to test $H_0 : \mu = \mu_0$ versus $H_A : \mu \neq \mu_0$. From Subsection 1.3 it is easily seen that $\bar{W} \sim SND(\bar{\mu}, \bar{\sigma}, \bar{\lambda})$ where $\bar{\mu}$ is given in (1.4), and $\bar{\mu} = \mu +$ (an expression dependent solely on (σ, λ)). Therefore, testing the above H_0 vs. H_A is equivalent to testing $\bar{H}_0 : \bar{\mu} = \bar{\mu}_0$ where $\bar{\mu}_0$ is the same expression in (1.4) with μ replaced by μ_0 .

The rest of the methodology is same as the above case (Subsection 3.1) of a single observation, i.e., retain \bar{H}_0 if $\bar{W}' = (\bar{W} - \bar{\mu}_0)$ is between \bar{k}'_2 and \bar{k}'_1 , and reject \bar{H}_0 otherwise, where $\bar{k}'_i = \bar{\sigma} \bar{k}'_{i*}$, $i = 1, 2$, with \bar{k}'_{i*} satisfying (3.9) with λ replaced by $\bar{\lambda}$, i.e.,

$$T_3 = \begin{cases} \text{retain } \bar{H}_0, & \text{if } \bar{k}'_2 < \bar{W}' < \bar{k}'_1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.14}$$

The following Tables 3.2 - 3.7 provide the simulated power (or size) of the above three tests, based on 2.5×10^4 replications as μ deviates from $\mu_0 = 0$ (w.l.g.) for various values of λ . All the tables have been constructed using $\alpha = 0.05$. Similar tables for other values of α have been provided in the Appendix A.1.

Table 3.2: Power (or size, when $\mu = \mu_0 = 0$) of three tests for various μ with fixed $\lambda = 0.0$

μ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.0512	0.0512	0.0512	0.0492	0.0492	0.0492	0.0485	0.0485	0.0485
0.2	0.0747	0.0747	0.0747	0.0966	0.0966	0.0966	0.1715	0.1715	0.1715
0.4	0.1490	0.1490	0.1490	0.2452	0.2452	0.2452	0.5120	0.5120	0.5120
0.6	0.2685	0.2685	0.2685	0.4756	0.4756	0.4756	0.8536	0.8536	0.8536
0.8	0.4314	0.4314	0.4314	0.7144	0.7144	0.7144	0.9780	0.9780	0.9780
1.0	0.6102	0.6102	0.6102	0.8829	0.8829	0.8829	0.9988	0.9988	0.9988
2.0	0.9940	0.9940	0.9940	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 3.3: Power (or size, when $\mu = \mu_0 = 0$) of three tests for various μ with fixed $\lambda = 1.0$

μ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.0487	0.0481	0.0484	0.0502	0.0502	0.0504	0.0494	0.0491	0.0495
0.2	0.0732	0.0758	0.0701	0.0923	0.0926	0.0887	0.1553	0.1575	0.1528
0.4	0.1468	0.1507	0.1414	0.2458	0.2489	0.2395	0.5110	0.5123	0.5050
0.6	0.2983	0.3042	0.2890	0.5117	0.5136	0.5034	0.8788	0.8778	0.8762
0.8	0.5051	0.5115	0.4944	0.7876	0.7887	0.7822	0.9918	0.9911	0.9916
1.0	0.7248	0.7288	0.7159	0.9455	0.9437	0.9427	0.9997	0.9997	0.9997
2.0	0.9999	0.9999	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Remark 3.2 The first three tables (Tables 3.2 - 3.4) provide the power/size as a function of μ for fixed λ , whereas the last three tables (Tables 3.5 - 3.7) provide the same as a function of λ for fixed μ . Therefore, the entire Table 3.5 provides size and the Tables 3.6 - 3.7 provide pure power of the three proposed tests. On the other hand, the first row of Tables 3.2 - 3.4 provide the size, and the remaining rows of these tables show power. All the three tests adhere to the level condition very closely, even for a sample size as small as 5. The patterns emerging from these tables are as follows.

Table 3.4: Power (or size, when $\mu = \mu_0 = 0$) of three tests for various μ with fixed $\lambda = 2.0$

μ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.0474	0.0506	0.0481	0.0497	0.0507	0.0496	0.0498	0.0495	0.0496
0.2	0.0650	0.0782	0.0584	0.0808	0.0889	0.0743	0.1202	0.1268	0.1122
0.4	0.1288	0.1476	0.1133	0.1909	0.2070	0.1747	0.3956	0.4020	0.3781
0.6	0.2548	0.2753	0.2280	0.4253	0.4349	0.4006	0.7831	0.7797	0.7700
0.8	0.4382	0.4549	0.4049	0.7014	0.7046	0.6785	0.9721	0.9685	0.9683
1.0	0.6560	0.6654	0.6217	0.9014	0.8967	0.8904	0.9989	0.9987	0.9987
2.0	0.9998	0.9996	0.9996	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 3.5: Size of three tests for various λ with fixed $\mu = 0.0$

λ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.0466	0.0466	0.0466	0.0492	0.0492	0.0492	0.0485	0.0485	0.0485
0.2	0.0504	0.0503	0.0503	0.0495	0.0494	0.0495	0.0491	0.0491	0.0491
0.4	0.0486	0.0488	0.0487	0.0508	0.0510	0.0509	0.0491	0.0488	0.0491
0.6	0.0480	0.0478	0.0480	0.0477	0.0477	0.0477	0.0483	0.0483	0.0483
0.8	0.0490	0.0492	0.0491	0.0508	0.0498	0.0504	0.0510	0.0509	0.0507
1.0	0.0504	0.0505	0.0498	0.0515	0.0513	0.0520	0.0517	0.0521	0.0521
2.0	0.0475	0.0519	0.0476	0.0496	0.0503	0.0488	0.0491	0.0494	0.0495
3.0	0.0478	0.0578	0.0486	0.0475	0.0528	0.0473	0.0505	0.0502	0.0494
4.0	0.0483	0.0697	0.0485	0.0489	0.0580	0.0496	0.0502	0.0532	0.0490
6.0	0.0453	0.0934	0.0484	0.0487	0.0678	0.0480	0.0489	0.0574	0.0479
8.0	0.0458	0.1173	0.0483	0.0479	0.0826	0.0490	0.0493	0.0622	0.0497

Table 3.6: Power of three tests for various λ with fixed $\mu = 1.0$

λ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.6070	0.6070	0.6070	0.8857	0.8857	0.8857	0.9986	0.9986	0.9986
0.2	0.6668	0.6667	0.6667	0.9176	0.9176	0.9176	0.9996	0.9996	0.9996
0.4	0.6998	0.7010	0.6988	0.9356	0.9358	0.9353	0.9999	0.9999	0.9999
0.6	0.7175	0.7195	0.7145	0.9448	0.9454	0.9436	1.0000	1.0000	1.0000
0.8	0.7216	0.7245	0.7163	0.9452	0.9459	0.9438	0.9998	0.9998	0.9998
1.0	0.7257	0.7290	0.7152	0.9455	0.9437	0.9427	0.9997	0.9997	0.9997
2.0	0.6575	0.6648	0.6222	0.9043	0.9016	0.8938	0.9991	0.9990	0.9990
3.0	0.5841	0.5978	0.5370	0.8397	0.8251	0.8166	0.9954	0.9923	0.9946
4.0	0.5284	0.5366	0.4751	0.7922	0.7595	0.7642	0.9890	0.9781	0.9864
6.0	0.4661	0.4890	0.4082	0.7061	0.6434	0.6684	0.9661	0.9207	0.9601
8.0	0.4273	0.4780	0.3684	0.6548	0.5727	0.6118	0.9447	0.8556	0.9357

(a) When $\lambda = 0$, i.e., the distribution is $N(\mu, \sigma^2)$, then all the three tests become the identical normal test (based on the sample mean), and hence they provide identical

Table 3.7: Power of three tests for various λ with fixed $\mu = 2.0$

λ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.9936	0.9936	0.9936	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.2	0.9989	0.9989	0.9989	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.4	0.9998	0.9998	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.6	0.9998	0.9998	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.8	0.9997	0.9997	0.9997	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1.0	0.9999	0.9999	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2.0	0.9994	0.9995	0.9992	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	0.9982	0.9974	0.9977	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	0.9961	0.9926	0.9948	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	0.9905	0.9710	0.9860	1.0000	0.9989	1.0000	1.0000	1.0000	1.0000
8.0	0.9862	0.9488	0.9810	0.9999	0.9962	0.9999	1.0000	1.0000	1.0000

performance.

(b) For λ close to 0, the test T_2 seems to have the best power performance. But note that T_2 becomes very liberal (with size exceeding α substantially) as λ gets large (≥ 3).

(c) The test T_3 has good local power for small λ (≤ 2), but its performance deteriorates as λ gets larger. This is not surprising because \bar{W} is not a sufficient statistic for μ , and as shown in Thiuthad and Pal (2018), the amount of information about μ carried by \bar{W} diminishes rapidly as λ deviates away from 0.

(d) The mechanisms of the subsection 3.2 can be followed easily even for unknown (σ, λ) provided one has a large sample which can give reliable estimates of σ and λ . In such a case follow the methods of this section by replacing (σ, λ) by $(\hat{\sigma}, \hat{\lambda})$, for a large n , and then pretend that the nuisance parameters were known.

4 Hypothesis Testing on μ with unknown (σ, λ)

In order to develop tests on μ for unknown σ and λ , we can think of replacing the nuisance parameters by their suitable estimates (MME and/or MPL) and call the resultant tests (i.e., variants of T_1, T_2, T_3 as developed in Subsection 3.2) as T'_1, T'_2, T'_3 respectively. But the main difficulty we face here is that the asymptotic distributions of $\hat{\mu}_{MM}$ and $\hat{\mu}_{MPL}$ are far from clear, and the normal cut-off point as used in Section 4 produces bad results.

We can use a parametric bootstrap version of T_1 and T_2 , i.e., based on $\hat{\mu}_{MM}$ and $\hat{\mu}_{MPL}$, to create two tests, say T^*_1 and T^*_2 , respectively. The basic idea of a parametric bootstrap (PB) test based on a point estimate is described in the following.

Let $\hat{\mu}$ be an estimator based on the given data (W_1, W_2, \dots, W_n) ($\hat{\mu}$ can be either $\hat{\mu}_{MM}$ (to create T_1^*) or $\hat{\mu}_{MPL}$ (to create T_2^*)). Our objective is to test $H_0 : \mu = \mu_0$ versus $H_A : \mu \neq \mu_0$. The general steps of the PB method are as follows:

- Step-1: For a given set of n iid observations, obtain $\hat{\mu}$ and compute a deviant metric $D = (\hat{\mu} - \mu_0)$, which measures the deviation of the point estimator $\hat{\mu}$ from the null value of μ .
- Step-2: Assume that H_0 is true, i.e., $\mu = \mu_0$. Here, we obtain the estimates of σ and λ under H_0 which are called the restricted estimates (REs). For this step, we first derive the restricted method of moments (RMM) and the restricted maximum penalized likelihood (RMPL) estimators of $(\hat{\sigma}, \hat{\lambda})$ (for the details see in Appendix A.2), whichever is needed for the estimation method of μ . [If we use $D_{MM} = (\hat{\mu}_{MM} - \mu_0)$ as our deviant measure D , then we need to use corresponding $\hat{\sigma}_{RMM}$ and $\hat{\lambda}_{RMM}$ under H_0 for the PB test to work. Similarly, if we use $D_{MPL} = (\hat{\mu}_{MPL} - \mu_0)$ as our metric D , then we should use $\hat{\sigma}_{RMPL}$ and $\hat{\lambda}_{RMPL}$ accordingly.]
- Step-3: Generate iid observations under H_0 , i.e., $W_1^*, W_2^*, \dots, W_n^* \sim \text{SND}(\mu_0, \hat{\sigma}_{RE}, \hat{\lambda}_{RE})$. Obtain $\hat{\mu}$ based on $(W_1^*, W_2^*, \dots, W_n^*)$, and call it $\hat{\mu}^*$. Recalculate D as $D^* = (\hat{\mu}^* - \mu_0)$.
- Step-4: Repeat the above Step-3 a large number of times, say N times. This gives us N copies of D^* , i.e., $D^{*1}, D^{*2}, \dots, D^{*N}$. Order them as $D^{*(1)} \leq \dots \leq D^{*(N)}$. Then the cut off points for D are to be found by the (approximate) probability distribution of D under H_0 , i.e., the left cut-off point $D_L = D^{*(\alpha/2)N}$ and the right cut-off point $D_U = D^{*(1-\alpha/2)N}$ such that $\alpha \approx P(\text{Rej } H_0 | H_0 \text{ is true})$. If $D \in (D_L, D_U)$, then we retain H_0 . If $D < D_L$ or $D > D_U$, then we reject H_0 . Alternatively, we can compute the p-value as $p_{PB} = 2 \min\{p_L, p_U\}$ where $p_L = \{\text{number of } D^{*j} < D\}/N$ and $p_U = \{\text{number of } D^{*j} > D\}/N$.

Apart from T_1^* and T_2^* (based on D_{MM} and D_{MPL} , respectively), we propose a few other tests as follows. Though not all of them may not show good performance, the idea behind them seems reasonable, and hence worth trying.

It is tempting to modify the third test T_3 in (3.14) and propose T'_3 as follows:

$$T'_3 = \begin{cases} \text{retain } \bar{H}_0, & \text{if } \bar{k}''_2 < \bar{W}' < \bar{k}''_1, \\ 0, & \text{otherwise,} \end{cases} \quad (4.1)$$

where $\bar{W}' = (\bar{W} - \mu_0)$, $\bar{k}''_i = \hat{\sigma}_{RE} \bar{k}''_{i*}$, $i = 1, 2$, with \bar{k}''_{i*} satisfying (3.9) with λ replaced by $\hat{\lambda}_{RE}$ where $\hat{\sigma}_{RE}$ and $\hat{\lambda}_{RE}$ satisfy (1.2) and (1.3) with λ replaced by $\hat{\lambda}_{RE}$ and σ replaced by $\hat{\sigma}_{RE}$, respectively. (The subscript ‘RE’ indicates the restricted estimate of the nuisance parameters under the null hypothesis.) Here we can create two versions of T'_3 by using either RMM or RMPL estimators. The two versions of T'_3 are henceforth referred to as $T'_{3(RMM)}$ and $T'_{3(RMPL)}$, respectively. However, both the versions of this test show a very bad and strange Type-I probability performance (either 0 or 1). Therefore we opt for a parametric bootstrap version of T_3 , call it T_3^* , as follows.

Similar to T_1^* and T_2^* , we propose T_3^* with the deviant metric $D = (\bar{W} - \mu_0) = \bar{W}'$ whose cut-off points are to be found through a parametric bootstrap method. The steps to implement T_3^* are identical to those of T_1^* and T_2^* as mentioned above. However, in step 3, the *iid* bootstrap observations W_i^* are to be generated from $\text{SND}(\mu_0, \hat{\sigma}_{RE}, \hat{\lambda}_{RE})$, where $\hat{\sigma}_{RE}$ and $\hat{\lambda}_{RE}$ are the restricted estimates (RE) of σ and λ under H_0 . This gives rise to two subversions of T_3^* , called here as $T_{3(RMM)}^*$ and $T_{3(RMPL)}^*$, depending on the type of RE used. If we use the method of moments estimates of σ and λ under the restriction of H_0 then it is called $T_{3(RMM)}^*$ and similarly $T_{3(RMPL)}^*$.

We propose two other tests based on the likelihood ratio (LR) statistic as detailed below. In order to test $H_0 : \mu = \mu_0$ vs. $H_A : \mu \neq \mu_0$, the likelihood ratio statistic

$$\Lambda = \frac{\sup_{H_0} L(\mu, \sigma, \lambda|W)}{\sup L(\mu, \sigma, \lambda|W)} \approx \frac{L(\mu_0, \hat{\sigma}_{RMPL}, \hat{\lambda}_{RMPL})}{L(\hat{\mu}_{MPL}, \hat{\sigma}_{MPL}, \hat{\lambda}_{MPL})} \tag{4.2}$$

Note that the above Λ is supposed to be a likelihood based deviant metric which measures the compatibility of H_0 with the data.

Asymptotically, i.e., as $n \rightarrow \infty$, under H_0 ,

$$\Lambda_* = (-2 \ln \Lambda) \sim \chi_1^2.$$

So we use $\Lambda_* = 2 \sum_{i=1}^n [\ln f(W_i | \hat{\mu}_{MPL}, \hat{\sigma}_{MPL}, \hat{\lambda}_{MPL}) - \ln f(W_i | \mu_0, \hat{\sigma}_{RMPL}, \hat{\lambda}_{RMPL})]$ as our deviant measure. This leads us to the fourth test T'_4 as follows:

$$T'_4 := \begin{cases} \text{retain } H_0, & \text{if } \Lambda_* \leq \chi_{1,\alpha}^2, \\ \text{reject } H_0, & \text{if } \Lambda_* > \chi_{1,\alpha}^2, \end{cases} \tag{4.3}$$

which is the asymptotic LRT. This T'_4 is expected to work well for a large sample size.

For small samples, we propose a parametric bootstrap version for our fourth test to create the PB LRT, say T_4^* , where the cut-off point is found by simulation. The method comes with the following steps:

Step-1: Given the original data W_1, W_2, \dots, W_n , compute Λ_* .

Step-2: Under the null hypothesis H_0 , we get $\mu = \mu_0$, $\sigma = \hat{\sigma}_{RMPL}$, and $\lambda = \hat{\lambda}_{RMPL}$.

Step-3: Generate *iid* observations under H_0 , i.e., $W_1^*, W_2^*, \dots, W_n^* \sim \text{SND}(\mu_0, \hat{\sigma}_{RMPL}, \hat{\lambda}_{RMPL})$. Based on $\underline{W}^* = (W_1^*, \dots, W_n^*)$, obtain all the unrestricted and restricted estimates of the parameters marked with the superscript ‘*’ for easy identification, i.e., $\hat{\mu}_{MPL}^*, \hat{\sigma}_{MPL}^*, \hat{\lambda}_{MPL}^*, \hat{\sigma}_{RMPL}^*$, and $\hat{\lambda}_{RMPL}^*$, and get $\Lambda_*^* = 2 \sum_{i=1}^n [\ln f(W_i^* | \hat{\mu}_{MPL}^*, \hat{\sigma}_{MPL}^*, \hat{\lambda}_{MPL}^*) - \ln f(W_i^* | \mu_0, \hat{\sigma}_{RMPL}^*, \hat{\lambda}_{RMPL}^*)]$.

Step-4: Repeat the above Step-3 a large number of times, (N times). This gives us N copies of Λ_*^* , i.e., $\Lambda_*^{*1}, \Lambda_*^{*2}, \dots, \Lambda_*^{*N}$. Order them as $\Lambda_*^{*(1)} \leq \dots \leq \Lambda_*^{*(N)}$. Then the cut-off points for Λ_* are to be found by the (approximate) probability distribution of Λ_* under H_0 , i.e., $\Lambda_{*U} = \Lambda_*^{*((1-\alpha)N)}$ such that $\alpha \approx P(\text{Rej } H_0 | H_0 \text{ is true})$. If $\Lambda_* > \Lambda_{*U}$, then we will reject H_0 . Alternatively, the p-value = $(1/N)(\text{Number of } \Lambda_*^j > \Lambda_*)$.

Now, to study the power/size of a test, we have to generate the original data (W_1, W_2, \dots, W_n) multiple times, say Q times, and then the power/size = proportion (out of Q times) of times H_0 is rejected. This simulation is quite time consuming since all the PB tests are subjected to double loops, and hence we had to restrict the number of replications to $N = Q = 1000$. The results of our simulation study for $n = 10$ have been provided in the following Tables 4.1 - 4.2.

Table 4.1: Size of six tests using $n = 10$ and $\alpha = 0.05$

λ	$\mu = 0$ (size)					
	T_1^*	T_2^*	$T_{3(RMM)}^*$	$T_{3(RMPL)}^*$	T_4'	T_4^*
0	0	0.0220	0	0	0.0010	0.0210
1	0.0140	0.0250	0	0	0.0230	0.0380
2	0.0240	0.0200	0	0	0.0660	0.0700

Table 4.2: Power of six tests using $n = 10$ and $\alpha = 0.05$

λ	$\mu = 1.0$ (power)					
	T_1^*	T_2^*	$T_{3(RMM)}^*$	$T_{3(RMPL)}^*$	T_4'	T_4^*
0	0.0260	0.0140	0	0	0.0920	0.1050
1	0.0310	0.0090	0	0	0.6240	0.6020
2	0.0050	0.0050	0	0	0.9280	0.9060

Remark 4.1 Strangely, the above simulation (Tables 4.1 - 4.2) shows the strange behavior of $T_{3(RMM)}^*$ and $T_{3(RMPL)}^*$, and hence they should not be considered as reliable test methods. In terms of size T_1^* appears to be very conservative. The tests T_4' and T_4^* appear to be a bit liberal when λ gets large (i.e., SND becomes very skewed). On the other hand, the test T_2^* seems to be more reliable, and appears to maintain the level consistently, though a bit conservative. Therefore, we recommend using T_2^* for hypothesis testing on μ . However, a more comprehensive simulation with higher values of N and Q will be needed along with other values of n .

5 An application with a Real-life Data Set

In this section our objective is the usage of the tests provided in Section 4 for a realistic data set. For this purpose we use Krung Thai Bank (KTB)'s quarterly stock price data set taken from Thiuthad and Pal (2018) where the quarterly price change (QPC) in % has been observed over 16 quarters. Note that μ indicates the location

Table 5.1: Quarterly stock price change of KTB from Thai stock market (Thiuthad and Pal (2018))

Period	1	2	3	4	5	6	7	8
QPC(%)	20.65	11.17	9.72	-5.42	3.17	-25.33	2.40	-0.60
Period	9	10	11	12	13	14	15	16
QPC(%)	16.88	-11.89	4.76	-0.56	13.33	-8.29	0.00	1.05

shift of the distribution, but not the mean of the distribution. Also, SND was found to have a good fit for this KTB data.

The objective is to test $H_0 : \mu = 0$ vs. $H_A : \mu \neq 0$. The following are the p-value of all the six tests as discussed in the previous section.

Test	T_1^*	T_2^*	$T_{3(RMM)}^*$	$T_{3(RMPL)}^*$	T_4'	T_4^*
p-value	0.2480	0.7732	0.9908	0.6831	0.9081	0.3984

As mentioned earlier, T_2^* seems to be the most reliable test, and for the given data its p-value = 0.7732 which suggest that we retain H_0 .

References

[1] Aigner, D.J. and Lovell, C.A.K.. Formulation and estimation of stochastic frontier production function model. *Econometrics* **12**: 21-37 (1977).

[2] Arellano-Valle, R. B., Castro, L. M. and Loschi, R. H.. Change Point Detection in the Skew-Normal Model Parameters. *Communications in Statistics-Theory and Methods* **42**: 603-618 (2013).

[3] Arnold, B.C. and Lin, G.D.. Characterizations of the Skew-normal and Generalized Chi Distributions. *Sankhyā: The Indian Journal of Stat.* **66**: 593-606 (2004).

[4] Azzalini, A.. A class of distribution which includes the normal ones. *Scan. J. Statist.* **12**: 171-178 (1985).

[5] Azzalini, A.. The Skew-normal Distribution and Related Multivariate Families. *Scan. J. Statist.* **32**: 159-188 (2005).

[6] Azzalini, A. and Capitanio, A.. Statistical applications of the multivariate skew-normal distribution. *J. Roy. Statist. Soc. B* **61**: 579-602 (1999).

[7] Azzalini, A. and Dalla Valle, A.. The multivariate skew-normal distribution. *Biometrika* **83**: 715-726 (1996).

- [8] Figueiredo, F. and Gomes, M.I.. The Skew-Normal Distribution in SPC. *REVSTAT – Statistical Journal* **11**: 83-104 (2013).
- [9] Gupta, A.K., Nguyen, T.T. and Sanqui, J.T.. Characterization of the Skew-Normal Distribution. *Ann. Inst. Statist. Math.* **56**: 351-360 (2004).
- [10] Gupta, R.C. and Brown, N.. Reliability Studies of the Skew Normal Distribution and Its Application to a Strength-Stress Model. *Commun. Statist.-Theory Meth.* **11**: 2427-2445 (2001).
- [11] Henze, N.. A probabilistic representation of the skew-normal distribution. *Scand. J. Statist.* **13**: 271-275 (1986).
- [12] Ngunkeng, G. and Ning, W.. Information Approach for the Change-Point Detection in the Skew Normal Distribution and Its Applications. *Sequential Analysis*, **33**: 475-490 (2014).
- [13] O'Hagan, A. and Leonard, T.. Bayes estimation subject to uncertainty about parameter constraints, *Biometrika*, **63**: 201-202 (1976).
- [14] Roberts, C.. A correlation model useful in the study of twins. *Journal of the American Statistical Association.* **61**: 1184-1190 (1966).
- [15] Roberts, H.V.. *Data Analysis for Managers with Minitab*. (Scientific Press: Redwood City, CA, 1988).
- [16] Thiuthad, P. and Pal, N.. On Point Estimation of the Location Parameter of a Skew-Normal Distribution. *Submitted* (2018).

A Appendix

A.1 Power and size tables for Section 3 with other α values

The following Tables A.1 - A.6 provide the simulated power (or size) of the above three tests, based on 2.5×10^4 replications as μ deviates from $\mu_0 = 0$ (w.l.g.) for various values of λ . All the tables have been constructed using $\alpha = 0.01$.

Table A.1: Power (or size, when $\mu = \mu_0 = 0$) of three tests for various μ with fixed $\lambda = 0.0$

μ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.0093	0.0093	0.0093	0.0092	0.0092	0.0092	0.0102	0.0102	0.0102
0.2	0.0168	0.0168	0.0168	0.0274	0.0274	0.0274	0.0577	0.0577	0.0577
0.4	0.0478	0.0478	0.0478	0.0978	0.0978	0.0978	0.2806	0.2806	0.2806
0.6	0.1120	0.1120	0.1120	0.2481	0.2481	0.2481	0.6631	0.6631	0.6631
0.8	0.2146	0.2146	0.2146	0.4845	0.4845	0.4845	0.9233	0.9233	0.9233
1.0	0.3634	0.3634	0.3634	0.7174	0.7174	0.7174	0.9919	0.9919	0.9919
2.0	0.9714	0.9714	0.9714	0.9999	0.9999	0.9999	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table A.2: Power (or size, when $\mu = \mu_0 = 0$) of three tests for various μ with fixed $\lambda = 1.0$

μ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.0104	0.0103	0.0104	0.0091	0.0089	0.0087	0.0103	0.0107	0.0106
0.2	0.0191	0.0200	0.0169	0.0265	0.0276	0.0239	0.0516	0.0523	0.0493
0.4	0.0518	0.0539	0.0464	0.0990	0.1001	0.0910	0.2812	0.2829	0.2711
0.6	0.1372	0.1395	0.1239	0.2902	0.2913	0.2749	0.7168	0.7137	0.7069
0.8	0.2955	0.2986	0.2760	0.5963	0.5951	0.5784	0.9629	0.9622	0.9610
1.0	0.5181	0.5201	0.4944	0.8464	0.8463	0.8370	0.9986	0.9984	0.9986
2.0	0.9988	0.9986	0.9986	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table A.3: Power (or size, when $\mu = \mu_0 = 0$) of three tests for various μ with fixed $\lambda = 2.0$

μ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.0100	0.0119	0.0085	0.0109	0.0112	0.0100	0.0091	0.0106	0.0090
0.2	0.0191	0.0234	0.0134	0.0238	0.0278	0.0182	0.0378	0.0412	0.0311
0.4	0.0466	0.0554	0.0330	0.0770	0.0846	0.0589	0.1927	0.1973	0.1708
0.6	0.1167	0.1294	0.0884	0.2226	0.2336	0.1868	0.5737	0.5680	0.5394
0.8	0.2469	0.2623	0.1987	0.4918	0.4889	0.4400	0.9119	0.9012	0.8971
1.0	0.4553	0.4612	0.3914	0.7782	0.7650	0.7384	0.9944	0.9922	0.9927
2.0	0.9976	0.9967	0.9958	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table A.4: Size of three tests for various λ with fixed $\mu = 0.0$

λ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.0095	0.0095	0.0095	0.0094	0.0094	0.0094	0.0101	0.0101	0.0101
0.2	0.0099	0.0099	0.0099	0.0100	0.0100	0.0100	0.0098	0.0098	0.0097
0.4	0.0111	0.0112	0.0111	0.0102	0.0101	0.0101	0.0094	0.0094	0.0094
0.6	0.0104	0.0104	0.0108	0.0095	0.0096	0.0096	0.0112	0.0110	0.0112
0.8	0.0100	0.0103	0.0099	0.0100	0.0098	0.0100	0.0099	0.0092	0.0097
1.0	0.0111	0.0106	0.0106	0.0113	0.0110	0.0110	0.0104	0.0101	0.0104
2.0	0.0100	0.0121	0.0093	0.0112	0.0118	0.0107	0.0107	0.0113	0.0107
3.0	0.0112	0.0164	0.0095	0.0102	0.0128	0.0101	0.0103	0.0112	0.0098
4.0	0.0106	0.0241	0.0083	0.0105	0.0177	0.0088	0.0091	0.0113	0.0087
6.0	0.0103	0.0426	0.0070	0.0096	0.0227	0.0079	0.0106	0.0171	0.0093
8.0	0.0122	0.0631	0.0086	0.0111	0.0352	0.0093	0.0099	0.0192	0.0093

Table A.5: Power of three tests for various λ with fixed $\mu = 1.0$

λ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.3654	0.3654	0.3654	0.7201	0.7201	0.7201	0.9926	0.9926	0.9926
0.2	0.4276	0.4278	0.4274	0.7846	0.7846	0.7844	0.9969	0.9969	0.9969
0.4	0.4656	0.4667	0.4642	0.8277	0.8275	0.8262	0.9982	0.9982	0.9982
0.6	0.4992	0.5022	0.4922	0.8434	0.8432	0.8406	0.9991	0.9990	0.9991
0.8	0.5110	0.5130	0.4970	0.8518	0.8511	0.8456	0.9991	0.9992	0.9991
1.0	0.5168	0.5189	0.4924	0.8464	0.8464	0.8370	0.9986	0.9983	0.9986
2.0	0.4616	0.4673	0.3957	0.7798	0.7694	0.7438	0.9945	0.9928	0.9932
3.0	0.3914	0.4020	0.3092	0.6929	0.6625	0.6298	0.9802	0.9693	0.9737
4.0	0.3493	0.3648	0.2654	0.6204	0.5741	0.5455	0.9572	0.9276	0.9429
6.0	0.2956	0.3358	0.2124	0.5243	0.4621	0.4425	0.9048	0.8085	0.8782
8.0	0.2637	0.3430	0.1852	0.4720	0.4063	0.3878	0.8662	0.7015	0.8313

Table A.6: Power of three tests for various λ with fixed $\mu = 2.0$

λ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.9697	0.9697	0.9697	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.2	0.9909	0.9909	0.9909	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.4	0.9966	0.9966	0.9965	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.6	0.9985	0.9986	0.9985	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.8	0.9989	0.9989	0.9989	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1.0	0.9987	0.9989	0.9984	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2.0	0.9977	0.9970	0.9962	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	0.9933	0.9889	0.9886	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	0.9867	0.9729	0.9768	1.0000	0.9997	0.9998	1.0000	1.0000	1.0000
6.0	0.9752	0.9322	0.9552	0.9998	0.9955	0.9996	1.0000	1.0000	1.0000
8.0	0.9611	0.8964	0.9341	0.9995	0.9845	0.9989	1.0000	1.0000	1.0000

The following Tables A.7 - A.12 provide the simulated power (or size) of the above three tests, based on 2.5×10^4 replications as μ deviates from $\mu_0 = 0$ (w.l.g.) for various values of λ . All the tables have been constructed using $\alpha = 0.10$.

Table A.7: Power (or size, when $\mu = \mu_0 = 0$) of three tests for various μ with fixed $\lambda = 0.0$

μ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.1008	0.1008	0.1008	0.0993	0.0993	0.0993	0.0972	0.0972	0.0972
0.2	0.1306	0.1306	0.1306	0.1662	0.1662	0.1662	0.2642	0.2642	0.2642
0.4	0.2324	0.2324	0.2324	0.3570	0.3570	0.3570	0.6383	0.6383	0.6383
0.6	0.3856	0.3856	0.3856	0.5987	0.5987	0.5987	0.9103	0.9103	0.9103
0.8	0.5613	0.5613	0.5613	0.8107	0.8107	0.8107	0.9906	0.9906	0.9906
1.0	0.7219	0.7219	0.7219	0.9327	0.9327	0.9327	0.9997	0.9997	0.9997
2.0	0.9979	0.9979	0.9979	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table A.8: Power (or size, when $\mu = \mu_0 = 0$) of three tests for various μ with fixed $\lambda = 1.0$

μ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.0987	0.0984	0.0991	0.0996	0.1001	0.0999	0.1024	0.1021	0.1028
0.2	0.1262	0.1309	0.1242	0.1570	0.1606	0.1547	0.2413	0.2445	0.2390
0.4	0.2318	0.2394	0.2267	0.3523	0.3572	0.3473	0.6226	0.6257	0.6201
0.6	0.4123	0.4210	0.4059	0.6336	0.6359	0.6289	0.9285	0.9285	0.9278
0.8	0.6276	0.6324	0.6222	0.8660	0.8662	0.8633	0.9969	0.9963	0.9969
1.0	0.8176	0.8206	0.8140	0.9713	0.9716	0.9704	0.9999	0.9999	0.9999
2.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table A.9: Power (or size, when $\mu = \mu_0 = 0$) of three tests for various μ with fixed $\lambda = 2.0$

μ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.0992	0.0994	0.0997	0.1002	0.1023	0.1015	0.1018	0.1026	0.1024
0.2	0.1204	0.1336	0.1144	0.1411	0.1522	0.1360	0.1882	0.2028	0.1827
0.4	0.1992	0.2212	0.1862	0.2826	0.3009	0.2703	0.5070	0.5199	0.4972
0.6	0.3487	0.3746	0.3297	0.5390	0.5510	0.5231	0.8524	0.8529	0.8466
0.8	0.5444	0.5662	0.5240	0.7883	0.7904	0.7775	0.9852	0.9847	0.9844
1.0	0.7488	0.7597	0.7322	0.9430	0.9398	0.9382	0.9996	0.9994	0.9995
2.0	0.9998	0.9999	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table A.10: Size of three tests for various λ with fixed $\mu = 0.0$

λ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.1008	0.1008	0.1008	0.0993	0.0993	0.0993	0.0972	0.0972	0.0972
0.2	0.0953	0.0953	0.0954	0.1002	0.1002	0.1002	0.0983	0.0984	0.0983
0.4	0.1040	0.1042	0.1040	0.1002	0.0996	0.1003	0.0995	0.0996	0.0993
0.6	0.1018	0.1011	0.1016	0.1049	0.1042	0.1051	0.1036	0.1028	0.1039
0.8	0.0996	0.0985	0.0992	0.0998	0.0984	0.0994	0.0989	0.0988	0.0988
1.0	0.1009	0.1024	0.1014	0.1030	0.1022	0.1035	0.1028	0.1016	0.1024
2.0	0.1002	0.1011	0.1002	0.0999	0.1024	0.1016	0.1017	0.1009	0.1022
3.0	0.0985	0.1086	0.1008	0.0979	0.1004	0.0986	0.1008	0.1017	0.1006
4.0	0.0959	0.1176	0.1014	0.0971	0.1095	0.0992	0.0971	0.0990	0.0980
6.0	0.0932	0.1426	0.0989	0.0950	0.1176	0.0968	0.0996	0.1043	0.1004
8.0	0.0944	0.1702	0.1013	0.0992	0.1302	0.1016	0.0999	0.1104	0.1006

Table A.11: Power of three tests for various λ with fixed $\mu = 1.0$

λ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.7226	0.7226	0.7226	0.9339	0.9339	0.9339	0.9996	0.9996	0.9996
0.2	0.7706	0.7709	0.7705	0.9567	0.9568	0.9567	0.9998	0.9998	0.9998
0.4	0.7953	0.7962	0.7949	0.9682	0.9685	0.9681	1.0000	1.0000	1.0000
0.6	0.8121	0.8137	0.8112	0.9728	0.9729	0.9727	1.0000	1.0000	1.0000
0.8	0.8164	0.8189	0.8136	0.9736	0.9735	0.9731	1.0000	1.0000	1.0000
1.0	0.8176	0.8206	0.8140	0.9713	0.9716	0.9704	0.9999	0.9999	0.9999
2.0	0.7527	0.7610	0.7363	0.9453	0.9423	0.9411	0.9997	0.9996	0.9996
3.0	0.6825	0.6921	0.6559	0.8994	0.8913	0.8915	0.9981	0.9975	0.9978
4.0	0.6252	0.6369	0.5958	0.8544	0.8352	0.8417	0.9949	0.9902	0.9944
6.0	0.5603	0.5764	0.5271	0.7856	0.7378	0.7672	0.9831	0.9566	0.9808
8.0	0.5184	0.5538	0.4809	0.7447	0.6764	0.7222	0.9683	0.9106	0.9650

Table A.12: Power of three tests for various λ with fixed $\mu = 2.0$

λ	$n = 5$			$n = 10$			$n = 25$		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.0	0.9980	0.9980	0.9980	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.2	0.9996	0.9996	0.9996	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.4	0.9998	0.9998	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.8	0.9999	1.0000	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2.0	0.9998	0.9999	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	0.9993	0.9990	0.9992	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	0.9981	0.9954	0.9977	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	0.9960	0.9859	0.9947	1.0000	0.9995	1.0000	1.0000	1.0000	1.0000
8.0	0.9922	0.9700	0.9906	1.0000	0.9985	1.0000	1.0000	1.0000	1.0000

A.2 The restricted estimates (REs)

A.2.1 The restricted method of moments estimators (RMME) of σ and λ under

$$H_0 : \mu = \mu_0:$$

The estimators of σ and λ are found by solving

$$\mu_0 + \sigma\lambda \sqrt{2/\pi} / \sqrt{1 + \lambda^2} = \bar{W} = m_1 \text{ (say),} \tag{A.1}$$

$$\sigma^2 \{1 - (2/\pi)(\lambda^2/(1 + \lambda^2))\} = (1/n) \sum_{i=1}^n (W_i - \bar{W})^2 = m_2 \text{ (say),} \tag{A.2}$$

Note that, in equation (A.1), the estimated parameter $\hat{\lambda}_{RMM}$ carries the sign of $(m_1 - \mu_0)$. Simplification of (A.1) - (A.2) yields

$$\hat{\sigma}_{RMM} = \sqrt{m_2 + (m_1 - \mu_0)^2} \tag{A.3}$$

and solve

$$(1 + \hat{\lambda}^2)/\hat{\lambda}^2 = (2/\pi)\{1 + m_2/(m_1 - \mu_0)^2\} \tag{A.4}$$

to obtain $\hat{\lambda}$ which is the value of $\hat{\lambda}_{RMM}$. Since the LHS in (A.4) is > 1 , a finite solution of (A.4) exists only if

$$(2/\pi)\{1 + m_2/(m_1 - \mu_0)^2\} > 1, \\ \text{i.e., } m_2/(m_1 - \mu_0)^2 > (\pi/2) - 1 = 0.570796... \tag{A.5}$$

Here, we suggest an estimate of λ based on the RMME approach as

$$\hat{\lambda}_{RMM} = \begin{cases} \text{sign}(m_1 - \mu_0) \times \hat{\lambda}, & \text{if (A.5) holds.} \\ \text{sign}(m_1 - \mu_0) \times \infty, & \text{if (A.5) does not hold.} \end{cases} \tag{A.6}$$

The inequality (A.5) can be treated as an acceptance region for testing $H_0 : |\lambda| < \infty$ against $H_A : |\lambda| = \infty$. If H_0 is retained, then λ gets estimated by a finite $\hat{\lambda}$ value. Otherwise, λ is estimated by $\pm\infty$. However, for practical considerations, one should replace ∞ by a large M in (A.6) so that the SND with $\hat{\lambda} = \pm M$ becomes very close to the intended half normal distribution.

A.2.2 The restricted maximum penalized likelihood estimator (RMPLE) of σ and

$$\lambda \text{ under } H_0 : \mu = \mu_0:$$

The penalized log-likelihood function is

$$L_{**}(\mu_0) = L_*(\mu_0) - c_1 \ln(1 + c_2 \lambda^2), \tag{A.7}$$

where $L_*(\mu_0)$ is the log-likelihood function with $\mu = \mu_0$ and c_1, c_2 are constants mentioned in (2.7). So, the RMPLE should be found by solving

$$0 = \frac{\partial L_{**}(\mu_0)}{\partial \sigma} = -((\lambda/\sigma^2) \sum_{i=1}^n (W_i - \mu_0)\phi(\lambda(W_i - \mu_0)/\sigma)/\Phi(\lambda(W_i - \mu_0)/\sigma)) - (n/\sigma) + \sum_{i=1}^n (W_i - \mu_0)^2/\sigma^3, \quad (\text{A.8})$$

$$0 = \frac{\partial L_{**}(\mu_0)}{\partial \lambda} = \sum_{i=1}^n ((W_i - \mu_0)/\sigma)\phi(\lambda(W_i - \mu_0)/\sigma)/\Phi(\lambda(W_i - \mu_0)/\sigma) - 2c_1c_2\lambda/(1 + c_2\lambda^2). \quad (\text{A.9})$$