

Analyzing Near-Normal Data Using a New Density Function within Azzalini's Class of Skew Distributions

Hassan Elsalloukh¹, Assistant Professor of Statistics, Department of Mathematics and Statistics, University of Arkansas at Little Rock, hxelsalloukh@ualr.edu

Jose H. Guardiola, Assistant Professor of Statistics, Department of Computing and Mathematical Sciences, Texas A&M University at Corpus Christi, jose.guardiola@tamucc.edu.

Abstract

A family of distributions, which was first introduced by O'Hagan and Leonard [9] for Bayesian analysis of normal means and was later investigated in detail by Azzalini [3, 4], is modified leading to a new class of asymmetric distributions. A new score test is derived for detecting non-normality within the new class of asymmetric distributions. The maximum likelihood estimates are derived for location, scale, shape, and skewness parameters. Then, the new score test is applied on an example of a real data set within the new class of asymmetric distributions to detect non-normality. Maximum likelihood estimators are used to fit the data with a skew distribution and compared to studies in which researchers used the normal distribution.

1. Introduction

During the last decade, interest has grown in the construction of flexible classes of distributions that, unlike the normal distribution, exhibit considerable skewness and kurtosis. Such distributions are useful and practical for modeling skewed data, including environmental and financial data modeling that often do not follow the normal law. O'Hagan and Leonard [9] proposed one such a family for Bayesian analysis of normal means. Azzalini [3] also investigated this family in detail where he defined his density function as the skew-normal (SN) and later provided an extension of the SN to a family of skew-exponential distributions [4]. Elsalloukh [7] developed a new flexible parametric asymmetric distribution that can be used on skewed data. Arellano-Valle [1] conducted in depth inference on a general class of asymmetric distributions. In this paper, the parameters of Azzalini's family of skew-exponential distributions are modified to provide a more applicable and tractable family.

A score test is also derived in this paper within the new family of skewed distributions. Score tests are equivalent to the Lagrange multiplier test statistics [13], and are widely used for testing hypotheses of the form $H_0 : \psi = 0$ against $H_1 : \psi \neq 0$, where ψ is a scalar parameter of interest. In many practical situations, the score test is an attractive competitor to the likelihood ratio (LR) and the Wald test, because it requires the model parameters to be estimated only under the null hypothesis. Furthermore, the LR, Wald, and score test statistics have the same local power and are all asymptotically equivalent [6]. However, the score test has a computational advantage over the LR and Wald tests.

The remainder of this paper is organized as follows: maximum likelihood estimators are derived for all the parameters for an example of real data and then used in the score test within the new family of skew distributions to detect non-normality and to be compared with the normal case.

¹ Principal author, email: <mailto:hxelsalloukh@ualr.edu>

2. Azzalini's Class of Skew Distributions

Many families of density functions approach the normal one as a certain parameter tends to an appropriate value. However, there are only a few parametric classes of distributions, which include the normal distribution as a proper member and not as a limiting distribution. Azzalini [3] introduced the skew-normal class of distributions, as a class or family able to reflect varying degrees of skewness, which is mathematically tractable and which includes the normal distributions as a special case. One such class of distributions was defined by Azzalini [3, 4]. Azzalini [3] defined a random variable Z to be a skew-normal random variable with a parameter λ ; that is, Z is $SN(\lambda)$ with a density function

$$\phi(Z; \lambda) = 2\phi(z)\Phi(\lambda z) \quad (-\infty < z < \infty), \quad (1)$$

where ϕ and Φ are the standard normal density and distribution functions, respectively.

One limitation of the family of (1) is that the parameter λ can produce only tails thinner than the normal distribution. However, we are often interested in analyzing data from heavy-tailed distributions. Azzalini [4] suggested a class of densities, which includes the normal family and allows thick tails, that is,

$$g(y; \omega) = C_\omega \exp\left\{-\frac{|y|^\omega}{\omega}\right\} \quad (-\infty < y < \infty), \quad (2)$$

where ω is a positive tail weight parameter and

$$C_\omega = \left\{2\omega^{\frac{1}{\omega-1}}\Gamma(1/\omega)\right\}^{-1};$$

see Box [5]. The density $g(y; 2)$ is the $N(0, 1)$ density and $g(y; 1)$ is the Laplace density. As $\omega \rightarrow \infty$, $g(y; \omega)$ converges to the uniform density on $(-1, 1)$. Azzalini [4] introduces skewness in (2) in the form of

$$2G(\lambda y)g(y; \omega), \quad (3)$$

where G satisfies the conditions of Lemma 1 and is

$$G(y) = \Phi\left\{\text{sgn}(y)\frac{|y|^\psi}{\sqrt{\psi}}\right\}, \quad (4)$$

where $\psi = \frac{\omega}{2}$. The choice of G in (4) is the distribution function of $\text{sgn}(U)\left|\sqrt{\psi}U\right|^{\frac{1}{\psi}}$, where $U \sim N(0, 1)$. Reversing the sign of λ in (3) gives the density of $-Z$; therefore, no loss of generality occurs in assuming $\lambda \geq 0$. Therefore, the density (3) that was considered in [4] is

$$h(y) = 2C_{2\psi} \exp\left\{-\frac{|y|^{2\psi}}{2\psi}\right\} \Phi\left\{\operatorname{sgn}(\lambda y) \frac{|\lambda y|^\psi}{\sqrt{\psi}}\right\}.$$

3. A New Density Function within Azzalini's Class of Skew Distributions

Many choices of G and $g(y; \omega)$ are possible for (2) and (4). The choices that are considered in this paper are modified to produce a new density function of the form

$$h(y_i) = 2G(\lambda u_i)g(u_i; \alpha), \quad (5)$$

where,

$$u_i = \frac{y_i - \mu}{\sigma},$$

$$g(u_i | \alpha) = w(\alpha)\sigma^{-1} \exp\left\{-c(\alpha)|u_i|^{\frac{2}{1+\alpha}}\right\},$$

$$c(\alpha) = \Gamma\left[\frac{3(1+\alpha)}{2}\right]^{\frac{1}{1+\alpha}} \Gamma\left[\frac{(1+\alpha)}{2}\right]^{\frac{1}{1+\alpha}},$$

$$w(\alpha) = (1+\alpha)^{-1} \left\{\Gamma\left[\frac{3(1+\alpha)}{2}\right]\right\}^{\frac{1}{2}} \left\{\Gamma\left[\frac{(1+\alpha)}{2}\right]\right\}^{\frac{-3}{2}},$$

and

$$G(\lambda u_i | \alpha) = \Phi\left[\operatorname{sgn}(\lambda u_i)\sqrt{1+\alpha}|\lambda u_i|^{\frac{1}{1+\alpha}}\right],$$

for $\lambda \geq 0$. Note that when $\alpha = 0$ and $\lambda = 0$, $h(y_i)$, (5), reduces to a standard normal pdf. The density function, $g(u_i | \alpha)$, was also used by [10] to test for normality within the exponential family distributions.

4. A Score Test for Detecting Non-Normality within the New Density Function

The problem of testing hypotheses of univariate normality of a set of observations has been of interest to experimenters for many years due to the popularity of normal-based inference procedures. As a result, many test statistics have been suggested as possible solutions to the testing-normality problem. One such is the score test or Lagrange multiplier test within the exponential power family of distributions formulated by [10]. A score test of normality within the family of skew distributions in (5) will be developed now.

Since the score test testing procedure requires estimation only under the null hypothesis, an asymptotically uniformly most powerful unbiased test of the normality assumption $H_0: \alpha = 0, \lambda = 0$ versus $H_A: \alpha \neq 0$ or $\lambda \neq 0$ can be easily constructed.

Let y_1, \dots, y_n be random variables from a skew-normal distribution with a pdf (5) then the likelihood for $\varphi = (\lambda, \alpha, \mu, \sigma)$ is

$$l(\varphi) = 2^n \prod_{i=1}^n G(\lambda u_i | \alpha) \prod_{i=1}^n g(u_i | \alpha),$$

the corresponding log-likelihood is

$$L(\varphi) = n \ln 2 + \sum_{i=1}^n \ln G(\lambda u_i | \alpha) + \sum_{i=1}^n \ln g(u_i | \alpha).$$

To construct the score test, the following score with respect to α is required:

$$\frac{\partial L(\varphi)}{\partial \alpha} = \sum_{i=1}^n \frac{\frac{\partial}{\partial \alpha} G(\lambda u_i | \alpha)}{G(\lambda u_i | \alpha)} + \sum_{i=1}^n \frac{\partial}{\partial \alpha} \ln g(u_i | \alpha). \quad (6)$$

Computation of the information matrix is straightforward but tedious and is confined to Appendix A. Starting with the computation of the first term in the right-hand side of equality (6), that is

$$\sum_{i=1}^n \frac{\frac{\partial}{\partial \alpha} G(\lambda u_i | \alpha)}{G(\lambda u_i | \alpha)} = \sum_{i=1}^n \frac{\phi \left[\text{sgn}(\lambda u_i) \sqrt{1+\alpha} |\lambda u_i|^{\frac{1}{1+\alpha}} \right] \frac{\partial}{\partial \alpha} \left[\text{sgn}(\lambda u_i) \sqrt{1+\alpha} |\lambda u_i|^{\frac{1}{1+\alpha}} \right]}{\Phi \left[\text{sgn}(\lambda u_i) \sqrt{1+\alpha} |\lambda u_i|^{\frac{1}{1+\alpha}} \right]}, \quad (7)$$

where

$$\begin{aligned} \frac{\partial}{\partial \alpha} \left[\text{sgn}(\lambda u_i) \sqrt{1+\alpha} |\lambda u_i|^{\frac{1}{1+\alpha}} \right] &= \frac{1}{2} (1+\alpha)^{-\frac{1}{2}} \text{sgn}(u_i) |u_i|^{\frac{1}{1+\alpha}} \lambda^{\frac{1}{1+\alpha}} \\ &\quad + (1+\alpha)^{\frac{1}{2}} \text{sgn}(u_i) |u_i|^{\frac{1}{1+\alpha}} (-1)(1+\alpha)^{-2} \ln |u_i| \lambda^{\frac{1}{1+\alpha}} \\ &\quad + (1+\alpha)^{\frac{1}{2}} \text{sgn}(u_i) |u_i|^{\frac{1}{1+\alpha}} (-1)(1+\alpha)^{-2} \lambda^{\frac{1}{1+\alpha}} \ln \lambda. \end{aligned}$$

Under $H_0 : \alpha = 0, \lambda = 0$,

$$\frac{\partial}{\partial \alpha} \left[\text{sgn}(\lambda u_i) \sqrt{1+\alpha} |\lambda u_i|^{\frac{1}{1+\alpha}} \right]_{\substack{\alpha=0 \\ \lambda=0}} = 0,$$

providing that $\lim_{\lambda \rightarrow 0} \lambda \ln \lambda = 0$ and $\lim_{\lambda \rightarrow 0} u \ln u = 0$, and hence (7) reduces to 0. Computing the second term in the right-hand side of equality (6), that is

$$\sum_{i=1}^n \frac{\partial}{\partial \alpha} \ln g(u_i | \alpha) = n \frac{\frac{\partial}{\partial \alpha} w(\alpha)}{w(\alpha)} - \frac{\partial}{\partial \alpha} c(\alpha) \sum_{i=1}^n |u_i|^{\frac{2}{1+\alpha}} + 2(1+\alpha)^{-2} c(\alpha) \ln |u_i| \sum_{i=1}^n |u_i|^{\frac{2}{1+\alpha}}, \quad (8)$$

where

$$\begin{aligned} \frac{\frac{\partial}{\partial \alpha} w(\alpha)}{w(\alpha)} &= \frac{3}{4} \left\{ \psi \left[\frac{3(1+\alpha)}{2} \right] - \psi \left[\frac{(1+\alpha)}{2} \right] \right\} - (1+\alpha)^{-1}, \\ \frac{\partial}{\partial \alpha} c(\alpha) &= c(\alpha) \left\{ -(1+\alpha)^{-1} \ln [c(\alpha)] + \frac{3}{2(1+\alpha)} \psi \left[\frac{3(1+\alpha)}{2} \right] - \frac{1}{2} \psi \left[\frac{(1+\alpha)}{2} \right] \right\}, \end{aligned}$$

and $\psi(t) \equiv \frac{d}{dt}[\ln \Gamma(t)]$ is known as the digamma function. Under $H_0 : \alpha = 0, \lambda = 0$,

$$\frac{\frac{\partial w(0)}{\partial \alpha}}{w(0)} = c(0) = \frac{1}{2}, \frac{\partial c(0)}{\partial \alpha} = .8648186,$$

therefore

$$\sum_{i=1}^n \frac{\partial}{\partial \alpha} \ln g(u_i | \alpha) \Big|_{\substack{\alpha=0 \\ \lambda=0}} = \frac{n}{2} - .8648186 \sum_{i=1}^n u_i^2 + \sum_{i=1}^n u_i^2 \ln |u_i|,$$

and

$$\frac{\partial L(\hat{\varphi})}{\partial \alpha} = \frac{\partial L(\varphi)}{\partial \alpha} \Big|_{\substack{\alpha=0 \\ \lambda=0}} = \frac{n}{2} - .8648186 \sum_{i=1}^n \hat{u}_i^2 + \sum_{i=1}^n \hat{u}_i^2 \ln |\hat{u}_i|.$$

Thus,

$$E \left[\frac{\partial L(\hat{\varphi})}{\partial \alpha} \right] = \frac{n}{2} - .8648186 \sum_{i=1}^n E(\hat{u}_i^2) + \sum_{i=1}^n E(\hat{u}_i^2 \ln |\hat{u}_i|),$$

$$= 0,$$

and

$$E \left[\frac{\partial L(\hat{\varphi})}{\partial \alpha} \right]^2 = \frac{n^2}{4} + \left(\frac{\partial c(0)}{\partial \alpha} \right)^2 \sum_{i=1}^n E(\hat{u}_i^4) + \sum_{i=1}^n E \left[\hat{u}_i^4 (\ln |\hat{u}_i|)^2 \right] - \frac{\partial c(0)}{\partial \alpha} \sum_{i=1}^n E(\hat{u}_i^2)$$

$$+ \sum_{i=1}^n E \left[\hat{u}_i^2 \ln |\hat{u}_i| \right] - 2 \frac{\partial c(0)}{\partial \alpha} \sum_{i=1}^n E \left[\hat{u}_i^4 \ln |\hat{u}_i| \right]$$

$$= .2011014n, \quad (9)$$

where $u_i \sim N(\mu, \sigma)$. Completion of constructing the score with respect to λ is also required, that is

$$\frac{\partial L(\varphi)}{\partial \lambda} = \sum_{i=1}^n \frac{\phi \left[\text{sgn}(\lambda u_i) \sqrt{1+\alpha} |\lambda u_i|^{\frac{1}{1+\alpha}} \right]}{\Phi \left[\text{sgn}(\lambda u_i) \sqrt{1+\alpha} |\lambda u_i|^{\frac{1}{1+\alpha}} \right]} \frac{\partial}{\partial \lambda} \left[\text{sgn}(\lambda u_i) \sqrt{1+\alpha} |\lambda u_i|^{\frac{1}{1+\alpha}} \right]$$

$$= \sum_{i=1}^n \frac{\phi \left[\text{sgn}(\lambda u_i) \sqrt{1+\alpha} |\lambda u_i|^{\frac{1}{1+\alpha}} \right]}{\Phi \left[\text{sgn}(\lambda u_i) \sqrt{1+\alpha} |\lambda u_i|^{\frac{1}{1+\alpha}} \right]} (1+\alpha)^{-\frac{1}{2}} \text{sgn}(u_i) |u_i|^{\frac{1}{1+\alpha}} (\lambda)^{\frac{-\alpha}{1+\alpha}}. \quad (10)$$

Evaluating (10) at $\alpha = 0$ and $\lambda = 0$, we have

$$\frac{\partial L(\hat{\varphi})}{\partial \lambda} = \sum_{i=1}^n \left[\frac{\phi(0)}{\Phi(0)} \hat{u}_i \right] = \frac{2}{\sqrt{2\pi}} \sum_{i=1}^n \hat{u}_i.$$

Also,

$$E\left[\frac{\partial L(\hat{\phi})}{\partial \lambda}\right] = \frac{2}{\sqrt{2\pi}} \sum_{i=1}^n E(\hat{u}_i) = 0,$$

and

$$E\left[\frac{\partial L(\hat{\phi})}{\partial \lambda}\right]^2 = \frac{4}{2\pi} E\left[\sum_{i=1}^n E(\hat{u}_i)\right]^2 = \frac{2n}{\pi},$$

where $u_i \sim N(\mu, \sigma)$. For the off diagonal elements for the information matrix

$$E\left[\frac{\partial L(\hat{\phi})}{\partial \alpha} \frac{\partial L(\hat{\phi})}{\partial \lambda}\right] = E\left[\left(\frac{n}{2} - .8648186 \sum_{i=1}^n \hat{u}_i^2 + \sum_{i=1}^n \hat{u}_i^2 \ln|\hat{u}_i|\right) \left(\frac{2}{\sqrt{2\pi}} \sum_{i=1}^n \hat{u}_i\right)\right] = 0.$$

Therefore, the test statistic is

$$\begin{aligned} \Lambda &= \begin{bmatrix} \frac{\partial L(\hat{\phi})}{\partial \alpha} & \frac{\partial L(\hat{\phi})}{\partial \lambda} \end{bmatrix} \begin{pmatrix} E\left[\frac{\partial L(\hat{\phi})}{\partial \alpha}\right]^2 & E\left[\frac{\partial L(\hat{\phi})}{\partial \alpha} \frac{\partial L(\hat{\phi})}{\partial \lambda}\right] \\ E\left[\frac{\partial L(\hat{\phi})}{\partial \alpha} \frac{\partial L(\hat{\phi})}{\partial \lambda}\right] & E\left[\frac{\partial L(\hat{\phi})}{\partial \lambda}\right]^2 \end{pmatrix}^{-1} \begin{bmatrix} \frac{\partial L(\hat{\phi})}{\partial \alpha} \\ \frac{\partial L(\hat{\phi})}{\partial \lambda} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial L(\hat{\phi})}{\partial \alpha} & \frac{\partial L(\hat{\phi})}{\partial \lambda} \end{bmatrix} \begin{pmatrix} .2011014n & 0 \\ 0 & \frac{2n}{\pi} \end{pmatrix}^{-1} \begin{bmatrix} \frac{\partial L(\hat{\phi})}{\partial \alpha} \\ \frac{\partial L(\hat{\phi})}{\partial \lambda} \end{bmatrix} \\ &= \frac{\left[\frac{n}{2} - .8648186 \sum_{i=1}^n \hat{u}_i^2 + \sum_{i=1}^n \hat{u}_i^2 \ln|\hat{u}_i|\right]^2}{.2011014n} + \frac{\left[\sum_{i=1}^n \hat{u}_i\right]^2}{n} \\ &= \hat{\xi}_1 + \hat{\xi}_2, \end{aligned} \tag{11}$$

where $u_i \sim N(\mu, \sigma)$. Note that as $n \rightarrow \infty$, the asymptotic distribution of Λ is chi-square with two degrees of freedom [14], thus, the null hypothesis is rejected if $\Lambda < \chi_{(2,1-\alpha/2)}^2$ or $\Lambda > \chi_{(2,\alpha/2)}^2$. The first part of the test statistic, $\hat{\xi}_1$, measures kurtosis and the second part, $\hat{\xi}_2$, measures the skewness of the distribution of interest.

5. Example 1

The score test (11) computations involving (5) are used on the heights of 219 of the world's volcanoes (Source: National Geographic Society and the World Alamac 1966, pp. 282-283), [8, 14]. Figure 1 shows an exploratory data analysis in the form of a stem-and-leaf plot. The basic descriptive statistics for the volcano heights Y are: the mean $\bar{Y} = 70.246$, the standard deviation $S = 43.018$, the

median=65.000, and the coefficient of skewness $\sqrt{b_1} = 0.840$. This coefficient indicates that Y is asymmetric.

Therefore, the score test Λ , (11), was calculated for the volcano heights using SAS IML, $\Lambda = 0.0635$. Since Λ falls in the rejection region, at the 5% significant level, we conclude that the data does not come from a symmetric normal distribution; indeed it can be modeled using the asymmetric distribution in (5).

The maximum likelihood estimators for the mean, standard deviation, and skewness parameters were computed using SAS IML; they are: $\hat{\mu} = 41.134$, $\hat{\sigma} = 40.350$, and $\hat{\lambda} = 0.7$. These estimators were used to provide a better fit for the data as shown in Figure 2.

As seen in Figure 2, the red solid graph, which is the skew distribution, provides a satisfactory model for the volcano heights, especially as compared with the normal family represented in the blue dashed graph.

Stem Leaf	#
19 03379	5
18 5	1
17 29	2
16 25	2
15 667	3
14 00	2
13 03478	5
12 11244456	8
11 0112334669	10
10 0112233445689	13
9 000123344556779	15
8 122223335679	12
7 00001112334555678889	20
6 001144556666777889	18
5 00112223445566666677799	23
4 01112333334444678899999	22
3 011224455556667899	18
2 0011222444556667788999	22
1 0001366799	10
0 25666799	8

-----+-----+-----+-----+-----
 Multiply Stem.Leaf by 10**+1

Figure 1. Heights of 219 of the world's volcanoes

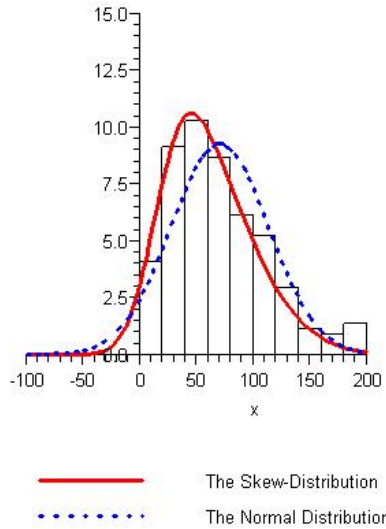


Figure 2. The normal and skew-normal for heights the world's volcanoes

6. Example 2

Daily rainfall in millimeters was recorded over a 47 year period in Turrumurra, Sydney, Australia [11]. For each year, the day with greatest rainfall was identified. The most rainfall in a day in each year and an exploratory data analysis are shown in Figure 3 in the form of a stem-and-leaf plot. The basic descriptive statistics for the daily rainfall Y are: the mean $\bar{Y} = 1369.106$, the standard deviation $S = 693.670$, the median = 1331, and the coefficient of skewness $\sqrt{b_1} = 1.295$. This coefficient indicates that Y is asymmetric.

The score test (11) is now used to check symmetry and normality of the daily rainfall data set. Therefore, the score test Λ , (11), was calculated using SAS IML, $\Lambda = 0.0365$. Since Λ falls in the rejection region, at the 5% significant level, we conclude that the data does not come from a symmetric normal distribution; indeed it can be modeled using the asymmetric distribution in (5).

The maximum likelihood estimators for the mean, standard deviation, and skewness parameters were computed using SAS IML; they are: $\hat{\mu} = 1100.356$, $\hat{\sigma} = 580.230$, and $\hat{\lambda} = 0.8$. These estimators were used to provide a better fit for the data as shown in Figure 4.

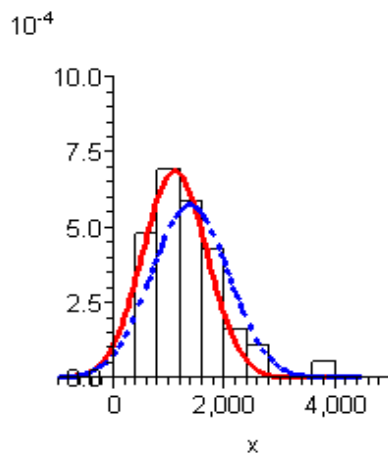
Once again, as seen in Figure 4, the red solid graph, which is the skew distribution, provides a satisfactory model for the daily rainfall data set, especially as compared with the normal family represented in the blue dashed graph.

Stem Leaf	#
38 3	1
36	
34	
32	
30	
28	
26 582	3
24 4	1
22	
20 0	1
18 0567	4
16 2248	4
14 07466	5
12 333679	6
10 149	3
8 4558116689	10
6 8025	4
4 58688	5

-----+-----+-----+-----+

Multiply Stem.Leaf by 10^{**+2}

Figure 3. Daily Rainfall in Millimeters



— The Skew-Distribution
 The Normal Distribution

Figure 4. The normal and skew-normal for rainfall data

7. Remark

This research can also be generalized to derive a score test for testing near-Laplace data using the density function (5). Such a score test can be used to detect asymmetry in data that follow the Laplace distribution and can be applied to fit and model data such as financial and environmental. The proofs shown in Appendix A are simple enough and sufficiently short making them useful for educational purposes and research.

8. Acknowledgment

The authors wish to thank Dr. Dean M. Young for useful comments and inspiration and the anonymous reviewers for their suggestions and comments that strengthened the paper.

9. Appendix A: Computation of the Information Matrix

Lemma 1. Poirier [10]. Suppose that V has a standard gamma distribution with a pdf

$$g(v) = \frac{v^{\delta-1} e^{-v}}{\Gamma(\delta)} \text{ for } v \geq 0 \text{ and } \delta > 0. \text{ Then,}$$

$$(a) \ E(\ln V) = \psi(\delta)$$

$$(b) \ E[(\ln V)^2] = [\psi(\delta)]^2 + \psi'(\delta),$$

where $\psi(\bullet)$ and $\psi'(\bullet)$ are the digamma and trigamma functions, respectively [2].

Using (8), we have

$$\begin{aligned} E \left[\frac{\partial L(\varphi)}{\partial \alpha} \right]^2 &= E \left[\left(\sum_i^n \frac{\partial \ln f(y_i; \varphi)}{\partial \alpha} \right)^2 \right] = E \left[\left(\sum_{i=1}^n \frac{1}{2} - \frac{\partial c(0)}{\partial \alpha} u_i^2 + u_i^2 \ln |u_i| \right)^2 \right] \\ &= \frac{n}{4} + \left(\frac{\partial c(0)}{\partial \alpha} \right)^2 \sum_{i=1}^n E(u_i^4) + \sum_{i=1}^n E \left[u_i^4 (\ln |u_i|)^2 \right] - \frac{\partial c(0)}{\partial \alpha} \sum_{i=1}^n E(u_i^2) \\ &\quad + \sum_{i=1}^n E \left[u_i^2 \ln |u_i| \right] - 2 \frac{\partial c(0)}{\partial \alpha} \sum_{i=1}^n E \left[u_i^4 \ln |u_i| \right]. \end{aligned} \quad (12)$$

Under H_0 , $E(u_i^2) = \sigma^2 = 1$ and $E(u_i^4) = 3\sigma^4 = 1$. We present evaluation of the remaining expectations in (12) in the following lemmas, which we draw, each from Lemma 1.

Lemma 2. Under H_0 , $E[U_i^2 \ln |U_i|] = .3648186$.

Proof.

$$\begin{aligned} E[U_i^2 \ln |U_i|] &= \int_{-\infty}^{\infty} t^2 \ln |t| \phi(t) dt = 2 \int_0^{\infty} t^2 \ln t \phi(t) dt = \frac{1}{2} \left[\ln 2 + \psi \left(\frac{3}{2} \right) \right] \\ &= .3648186, \end{aligned}$$

where we used the fact that $\psi(\eta+1) = \psi(\eta) + \eta^{-1}$ and $\psi\left(\frac{1}{2}\right) = -1.963510$.

Lemma 3. Under H_0 , $E\left[U_i^4 \ln|U_i|\right] = 2.094458$.

Proof. We proceed as in Lemma 3 and we straightforwardly show that

$$E\left[U_i^4 \ln|U_i|\right] = 2 \int_0^\infty t^4 \ln t \phi(t) dt = \frac{3}{2} \left[\ln 2 + \psi\left(\frac{5}{2}\right) \right] = 2.094458.$$

Lemma 4. Under H_0 , $E\left[U_i^4 (\ln|U_i|)^2\right] = 1.830017$.

Proof. We again proceed as in Lemma 3 and we straightforwardly show that

$$E\left[U_i^4 (\ln|U_i|)^2\right] = 2 \int_0^\infty t^4 (\ln t)^2 \phi(t) dt = \frac{3}{4} \left\{ \left[\psi\left(\frac{5}{2}\right) \right]^2 + \psi'\left(\frac{5}{2}\right) \right\} = 1.830017,$$

where we use the fact that $\psi'(\eta+1) = \psi'(\eta) + \eta^{-2}$ and $\psi'\left(\frac{1}{2}\right) = 4.934802$.

Finally we substitute the results of Lemma 2 through Lemma 5 into (9), and we find that

$$E\left[\left.\frac{\partial L(\varphi)}{\partial \alpha}\right|_{\substack{\alpha=0 \\ \lambda=0}}\right]^2 = .2011014n.$$

Lemma 5. Under H_0 , $E\left[|U_i| \ln|U_i|\right] = \sqrt{2}\sigma \left[\frac{\ln \sigma}{2} - \frac{\ln 2}{4} + \frac{\psi(2)}{2} \right]$.

Proof.

$$\begin{aligned} E\left[|U_i| \ln|U_i|\right] &= \int_{-\infty}^\infty \frac{\sqrt{2}}{2\sigma} |u| \ln|u| \exp\left(\frac{-\sqrt{2}|u|}{\sigma}\right) du \\ &= \int_{-\infty}^0 \frac{\sqrt{2}}{2\sigma} \sigma t \ln(\sigma t) \exp(-\sqrt{2}t) (-\sigma dt) + \int_0^\infty \frac{\sqrt{2}}{2\sigma} \sigma t \ln(\sigma t) \exp(-\sqrt{2}t) \\ &= \sqrt{2}\sigma \left[(\ln \sigma / 2) \int_0^\infty z \exp(-z) dz + (1/2) \int_0^\infty z (\ln z - \ln \sqrt{2}) \exp(-z) dz \right] \\ &= \sqrt{2}\sigma \left[\frac{\ln \sigma}{2} - \frac{\ln 2}{4} + \frac{\psi(2)}{2} \right]. \end{aligned}$$

Lemma 6. Under H_0 , $E[U_i^2 \ln|U_i|] = \sigma^2 \left[\ln \sigma - \frac{\ln 2}{2} + \psi(3) \right]$.

Proof. We proceed as in Lemma 6 and we straightforwardly show that

$$\begin{aligned}
E[U_i^2 \ln|U_i|] &= \int_{-\infty}^{\infty} \frac{\sqrt{2}}{2\sigma} u^2 \ln|u| \exp\left(\frac{-\sqrt{2}|u|}{\sigma}\right) du \\
&= \int_0^{\infty} \frac{\sqrt{2}}{2\sigma} \sigma^2 t^2 \ln(\sigma t) \exp(-\sqrt{2}t) (-\sigma dt) + \int_0^{\infty} \frac{\sqrt{2}}{2\sigma} \sigma^2 t^2 \ln(\sigma t) \exp(-\sqrt{2}t) (\sigma dt) \\
&= \sqrt{2}\sigma^2 \left[(\ln \sigma / 2\sqrt{2}) \int_0^{\infty} z^2 \exp(-z) dz + (1/2\sqrt{2}) \int_0^{\infty} z^2 (\ln z - \ln \sqrt{2}) \exp(-z) dz \right] \\
&= \sigma^2 \left[\ln \sigma - \frac{\ln 2}{2} + \psi(3) \right].
\end{aligned}$$

Lemma 7. Under H_0 ,

$$E[U_i^2 (\ln|U_i|)^2] = \sigma^2 \left[(\ln \sigma)^2 + \frac{[\psi(3)]^2 + \psi'(3)}{2} + \frac{(\ln 2)^2}{4} - \ln 2\psi(3) + 2 \ln \sigma \psi(3) - \ln 2 \ln \sigma \right].$$

Proof. We again proceed as in Lemma 6 and we straightforwardly show that

$$\begin{aligned}
E[U_i^2 (\ln|U_i|)^2] &= \int_{-\infty}^{\infty} \frac{\sqrt{2}}{2\sigma} u^2 (\ln|u|)^2 \exp\left(\frac{-\sqrt{2}|u|}{\sigma}\right) du \\
&= \sqrt{2}\sigma^2 \int_0^{\infty} t^2 (\ln^2 \sigma + \ln^2 t + 2 \ln \sigma \ln t) \exp(-\sqrt{2}t) dt \\
&= \sqrt{2}\sigma^2 \left[(\ln^2 \sigma / 2\sqrt{2}) \int_0^{\infty} z^2 \exp(-z) dz \right. \\
&\quad \left. + (1/2\sqrt{2}) \int_0^{\infty} z^2 (\ln^2 z + \ln^2 \sqrt{2} - 2 \ln z \ln \sqrt{2}) \exp(-z) dz \right. \\
&\quad \left. + (\ln \sigma / \sqrt{2}) \int_0^{\infty} z^2 (\ln z - \ln \sqrt{2}) \exp(-z) dz \right] \\
&= \sigma^2 \left[(\ln \sigma)^2 + \frac{[\psi(3)]^2 + \psi'(3)}{2} + \frac{(\ln 2)^2}{4} - \ln 2\psi(3) + 2 \ln \sigma \psi(3) - \ln 2 \ln \sigma \right].
\end{aligned}$$

References

- [1] R. B. Arellano-Valle, H. W. Gómez, and F. A. Quintana, *Statistical inference for a general class of asymmetric distributions*. Journal of Statistical Planning and Inference 128 (2005), no. 2, 427-443.
- [2] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover Publications, 1965.
- [3] A. Azzalini, *A Class of Distributions Which Includes the Normal Ones*. Scand. J. Statist., 12(1985), 171-178.
- [4] A. Azzalini, *Further Results on a Class of Distributions Which Includes the Normal Ones*. Statistica, 46(1986), 199-208.
- [5] G. E. P. Box, *A Note on Regionns of Kurtosis*. Biometrika, 40(1953), 465-468.
- [6] D. R. Cox and D. V. Hinkley, *Theoretical Statistics*. London: Chapman and Hall, 1974
- [7] H. Elsalloukh, J. H. Guardiola, and D. M. Young, *The Epsilon-Skew Exponential Power Distribution Family*. Far East Journal of Theoretical Statistics 17(1) (2005), 97-112.
- [8] G. S. Mudholkar and A. D. Hutson, *The Epsilon-Skew-Normal Distribution for Analyzing Near-Normal Data*. Journal of Statistical Planning and Inference, 83(2000), 291- 309.
- [9] A. O'Hagan, and T. Leonard, *Bayes Estimation Subject to Uncertainty about Parameter Constraints*. Biometrika, 63(1976), 201-202.
- [10] D. J. Poirier, D. Tello, and E. Zin, *A diagnostic Test for Normality Within the Power Exponential Family*, Journal of Business and Economic Statistics, 4(1986), 359-373.
- [11] J. C. W. Rayner and D. J. Best, *Smooth Tests of Goodness of Fit*, Oxford: Oxford University Press, Reprinted in *Small Data Sets*, P. 123.
- [12] J. R. Serfling, *Approximation Theorems of Mathematical Statistics*, New York: Wiley, 1980
- [13] S. D. Silvey, *The Lagrangian Multiplier Test*. Ann. Math. Statist., 30(1954), 389-407.
- [14] J. W. Tukey, *Exploratory Data Analysis*. Addison-Wesley, Reading, MA, 1977.