On the Distribution of Weighted Sum of Two Chi-squares with Applications to Shape Analysis

Kanti V. Mardia[1] and Xiangyu Wu[1]

[1]Department of Statistics, University of Leeds, Leeds, United Kingdom.

Contributing authors: K.V.Mardia@leeds.ac.uk; mmxw@leeds.ac.uk;

Abstract

In certain cutting-edge applications, it is found that a weighted sum of two χ^2-distributions plays an important role. It is well-known that the work on both central and non-central χ^2-distributions is classical and the next obvious step for extension is the weighted sum of two χ^2-distributions. Although there has been considerable theoretical work on the distribution of general linear combinations of χ^2, there have been no dedicated work on either getting deep insight into the distribution even in the particular case of weighted sum of two χ^2-distributions or its applications. We first derive the most general distribution of the weighted sum of two non-central χ^2 and give some properties. Particular cases are considered, and one important case arises when one of the χ^2 has 2 degrees of freedom, so that it has an exponential distribution. We refer to the resulted weighted sum as the exponentially modified χ^2-distribution. Another important case is when one of the χ^2 has large degrees of freedom, hence approximates a normal distribution. The resulted weighted sum is known as the exponentially modified Gaussian distribution in the literature. We give further insight into these skew distributions and we also consider some inference problems for these distributions. This work is motivated by new challenges in shape analysis on how to deal with asymmetry of bilateral shapes and we illustrate our methodology by applying it to a shape analysis problem involving a smile data on the cleft lip patients.

Keywords: bilateral symmetry, cleft lip patients, exponentially modified χ^2-distribution, exponentially modified Gaussian distribution, generalized χ^2-distribution, linear combination of χ^2-distributions.

1

1 Introduction

There have been several papers dealing with general linear combinations of χ^2-distributions (see for example, McKay [1], Press [2] and Bausch [3]); also known as generalized χ^2-distribution. However, these papers lack deeper insight even in the simple case of sum of two χ^2-distributions. The work for

[1]For the special issue for C. R. Rao, edited by Professor A.S.R. Srinivasa Rao

single non-central χ^2-distribution has been classical and well-studied and the next natural step is to move to the weighted sum of two χ^2-distributions, which we propose to do in this paper. We note that there is an exception for a particular case of the weighted sum, when one of the χ^2 is approximated by a normal (for large degrees of freedom) and the other one has 2 degrees of freedom. That is, we have the sum of a normal and an exponential distributions which is known in the literature as the exponentially modified Gaussian distribution; it has been applied in various practical applications including chemical analysis (Grushka [4]) and stochastic frontier analysis (Kumbhakar and Lovell [5]).

Let χ^2 with k degrees of freedom and non-centrality parameter λ be denoted by $\chi^2_k(\lambda)$. We define our weighted sum as

$$Z \sim a\chi^2_{k_1}(\lambda_1) + b\chi^2_{k_2}(\lambda_2),\ k_1 > 0, k_2 > 0, \lambda_1 > 0, \lambda_2 > 0, \tag{1.1}$$

where the two χ^2-distributions are independent and $(a,\ b)$ can take any real value, that is, we allow negative values of a and b, though in most of the practical cases, a and b are positive. Further, we will write Z to be distributed as $\text{WS}((k_1, k_2)^T, (\lambda_1, \lambda_2)^T, (a, b)^T)$, where WS stands for *Weighted Sum*.

As mentioned earlier, work has been done for various linear combinations of χ^2-distributions. The particular case of two weighted central χ^2-distributions with the same degrees of freedom has been given by Bausch [3] and we re-derive it as a particular case of our general weighted distribution.

When $k_2 = 2$ and $\lambda_2 = 0$ in Z, the χ^2 is equivalent to an exponential variable, so we call this distribution "exponentially modified χ^2-distribution" and denote it by ex-χ^2. Furthermore, when k_1 is large, $\chi^2_{k_1}(\lambda_1)$ approximates a normal distribution, so the resulting distribution is the ex-Gaussian distribution (denoted by ex-Gaussian in this paper) as indicated before.

This work is motivated by new challenges in shape analysis/size and shape analysis to deal with asymmetry of bilateral shapes (see for example Bock and Bowman [6], Patel et al. [7], Ajmera et al. [8] and Ajmera et al. [9]). In shape analysis, the object is invariant under translation, scaling and rotation (i.e., Euclidean similarity transformations), whereas in size-and-shape analysis, the object is invariant only under translation and rotation, see for example, Dryden and Mardia [10]. In this paper, we will work on an example using size-and-shape analysis. Our data are landmark-based, which means that a set of landmarks has been identified apriori for each dataset. Hence, each object in the data is represented by a configuration matrix X containing the coordinates of each landmark.

In some cases, the data is registered first in order to perform any statistical analysis. Figure 1 shows one way of registering the objects in size-and-shape analysis (the Bookstein registration). There are four landmarks for a rectangular shapes and their indices have been labeled $1, \ldots, 4$. The left of Figure 1 shows the original configuration; the second translated figure of Figure 1 displays the translated configuration where landmark 1 is at the origin; the right of Figure 1 shows the registered configuration after rotation where the line joining landmarks 1 and 2 becomes the x-axis. The codes for example are contained in the code chunk 1 of the `R` script `Sankhya Chi`. There are other methods for carrying out registration, see for example, Dryden and Mardia [10].

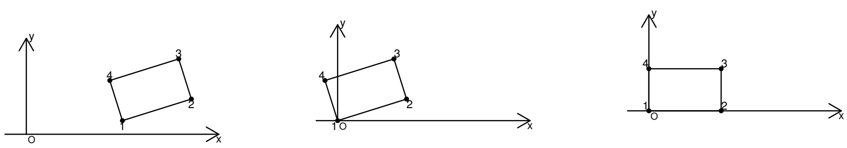

Fig. 1: From left to right: the original configuration; the translated configuration (landmark 1 is at the origin); the registered configuration after rotation of the translated configuration.

An object is said to be bilateral symmetry if its left and right parts are exactly the same with respect to some midline or midplane. For a bilateral shape, its landmarks can be divided into two categories: paired and solo. For the two landmarks which form a pair, they should lie on both sides of the midplane. The solo landmarks are unpaired and they should lie on the midplane in the symmetric case. Figure 2 shows an example of a bilaterally symmetric shape in two-dimensions. The midline is the y-axis. From the figure, for example, the landmarks 6 and 8 form a landmark pair, while landmarks 7 and 19 are solos.

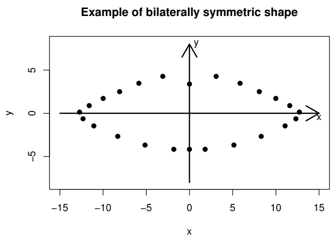

Fig. 2: An example of a bilaterally symmetric shape.

The key focus is on the asymmetry measure, ϕ_{L_2} (referred to in this paper as AS = ASymmetry measure for simplicity), given in Mardia et al. [11]. For practical applications such as the smile data (cleft lip patients) illustrated here, the distribution of the AS statistics is required, which under certain assumptions is distributed as the WS distribution. The smile data is used to illustrate how to estimate parameters and carry out one important test of hypothesis. The test is straightforward to carry out and the result is consistent with the medical opinion. Details are given in Section 7.

The paper is organized as following: Section 2 provides notations and several well-known important formulae. The main distribution given in equation (1.1) and corresponding special cases are discussed in Section 3. The ex-χ^2 is considered in Section 4 and in Section 5, ex-Gaussian distribution is discussed. In these sections, we give mainly their properties related to moments. In Section 6, we provide the parameter estimations of these distributions. The application to the cleft lip data in shape analysis is given in Section 7. We provide some discussions in Section 8.

2 Preliminaries

We provide here some preliminary background and well-known results with notations which are used subsequently.

We write $X \sim \chi^2_k$ which is a chi-square distribution with degrees of freedom (df) $k > 0$. The probability density function (pdf) of X is

$$f_X(x;k) = \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}},\ x > 0, \tag{2.1}$$

and the characteristic function of $X \sim \chi^2_k$ is given by

$$\phi_X(t;k) = (1-2it)^{-\frac{k}{2}}. \tag{2.2}$$

Let $X \sim \chi_k^2(\lambda)$ be a non-central chi-square distribution with df $k > 0$ and non-centrality parameter $\lambda > 0$, which has the pdf

$$f_X(x; k, \lambda) = \frac{1}{2} e^{-(x+\lambda)/2} \left(\frac{x}{\lambda}\right)^{\frac{k}{4}-\frac{1}{2}} I_{\frac{k}{2}-1}\left(\sqrt{\lambda x}\right), \ x > 0, \tag{2.3}$$

where $I_p(z)$ is the modified Bessel function of first kind:

$$I_p(z) = \sum_{m=0}^{\infty} \frac{\left(\frac{1}{4}z^2\right)^{m+\frac{p}{2}}}{m!\Gamma(p+m+1)}. \tag{2.4}$$

The characteristic function of $X \sim \chi_k^2(\lambda)$ is

$$\phi_X(t; k, \lambda) = \frac{\exp\left(\frac{i\lambda t}{1-2it}\right)}{(1-2it)^{\frac{k}{2}}}. \tag{2.5}$$

We write the lower incomplete gamma function defined as

$$\gamma(s, x) = \int_0^x t^{s-1} e^{-t} dt, \ x > 0 \tag{2.6}$$

and the beta function defined as

$$B(z_1, z_2) = \frac{\Gamma(z_1)\Gamma(z_2)}{\Gamma(z_1 + z_2)}, \ z_1, z_2 > 0. \tag{2.7}$$

The other function, we will need is the confluent hypergeometric function ${}_1F_1(a; b; z)$:

$${}_1F_1(a; b; z) = 1 + \frac{a}{b}\frac{z}{1!} + \frac{a(a+1)}{b(b+1)}\frac{z^2}{2!} + \cdots = \sum_{r=0}^{\infty} \frac{(a)_r z^r}{(b)_r r!}, \tag{2.8}$$

where $(a)_r = a(a+1)\cdots(a+r-1)$ and $(a)_0 = 1$. Further,

$${}_1F_1(a; b; z) = \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \int_0^1 e^{zt} t^{a-1} (1-t)^{b-a-1} dt \tag{2.9}$$

When $z \to 0$ and $b \neq -n$ for $n \in \mathbb{N}$, we have (see Section 13.5 on page 508 of Abramowitz and Stegun [12])

$${}_1F_1(a; b; z) \to 1. \tag{2.10}$$

Suppose X and Y are two independent random variables with pdf $f_X(\cdot)$ and $f_Y(\cdot)$ respectively. Define the sum

$$Z = X + Y.$$

Then the pdf of Z is given by the convolution formula

$$f_Z(z) = \int_{\mathcal{X}} f_X(x) f_Y(z-x) dx, \tag{2.11}$$

where $\mathcal{X}$ is the domain of X. Let $\phi_X(\cdot)$ and $\phi_Y(\cdot)$ denote the characteristic functions of X and Y respectively. Then the characteristic function of Z is given by

$$\phi_Z(t) = \phi_X(t)\phi_Y(t). \tag{2.12}$$

We note the following well-known results of sum Z (which can be derived using (2.12)):

- For $Z = \chi^2_{k_1}(\lambda_1) + \chi^2_{k_2}(\lambda_2)$, we have $Z \sim \chi^2_{k_1+k_2}(\lambda_1+\lambda_2)$.
- For $Z = \chi^2_{k_1} + \chi^2_{k_2}$, we have $Z \sim \chi^2_{k_1+k_2}$.
- For $Z = \chi^2_{k_1}(\lambda_1) + \chi^2_{k_2}$, we have $Z \sim \chi^2_{k_1+k_2}(\lambda_1)$.

For the standard normal distribution $N(0,1)$, we will denote the pdf as $\varphi(\cdot)$ and cumulative distribution function (cdf) as $\Phi(\cdot)$.

3 Distribution of a Weighted Sum of Two Non-Central χ^2-Distributions

In this section, we first consider our general weighted sum of the random variable

$$Z \sim \mathrm{WS}((k_1,k_2)^T, (\lambda_1,\lambda_2)^T, (a,b)^T) = aX + bY, \tag{3.1}$$

where $X \sim \chi^2_{k_1}(\lambda_1)$ and $Y \sim \chi^2_{k_2}(\lambda_2)$ are independent, a, $b > 0$, k_1, $k_2 > 0$, λ_1, $\lambda_2 > 0$. In most of our works, df k_1 and k_2 are positive integers, but our theorems in general are applicable for any positive k_1 and k_2.

3.1 General Weighted Sum

We now prove the following general theorem for the distribution of the random variable Z given by (3.1).
Theorem 3.1. ***Non-central case.*** *The pdf of the weighted sum of two non-central χ^2-distributions with different df*

$$Z \sim WS((k_1,k_2)^T, (\lambda_1,\lambda_2)^T, (a,b)^T) = a\chi^2_{k_1}(\lambda_1) + b\chi^2_{k_2}(\lambda_2)$$

is given by

$$f_Z(z; a, b, k_1, k_2, \lambda_1, \lambda_2) = \frac{e^{-\frac{1}{2}(\frac{z}{b}+\lambda_1+\lambda_2)}}{4ab} \int_0^z e^{-\frac{x}{2}(\frac{1}{a}-\frac{1}{b})} \left(\frac{x}{a\lambda_1}\right)^{\frac{k_1}{4}-\frac{1}{2}} \left(\frac{z-x}{b\lambda_2}\right)^{\frac{k_2}{4}-\frac{1}{2}} I_{\frac{k_1}{2}-1}\left(\sqrt{\frac{\lambda_1 x}{a}}\right) I_{\frac{k_2}{2}-1}\left(\sqrt{\frac{\lambda_2(z-x)}{b}}\right) dx. \tag{3.2}$$

Further, we have the alternative form

$$f_Z(z; a, b, k_1, k_2, \lambda_1, \lambda_2) = \frac{e^{-\frac{1}{2}(\frac{z}{b}+\lambda_1+\lambda_2)}}{a^{\frac{k_1}{2}} b^{\frac{k_2}{2}} 2^{\frac{k_1+k_2}{2}}} \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{\left(\frac{\lambda_1}{4a}\right)^m \left(\frac{\lambda_2}{4b}\right)^n z^{\frac{k_1+k_2}{2}+m+n-1}}{m!n!\Gamma(\frac{k_1+k_2}{2}+m+n)} {}_1F_1\left(\frac{k_1}{2}+m, \frac{k_1+k_2}{2}+m+n, -\frac{b-a}{2ab}z\right), \tag{3.3}$$

where $z > 0$ and ${}_1F_1(\cdot)$ is the confluent hypergeometric function given at (2.8).

Proof. On substituting the pdf of non-central χ^2-distribution given in (2.3) after scaling X and Y (using aX and bY) into the convolution formula (equation (2.11)):

$$f_Z(z) = \int_0^z f_X(x) f_Y(z-x) dx, \tag{3.4}$$

(3.2) follows. On substituting the series expansion of Bessel function given in (2.4) and then changing the order of integral and summation, we find that the density becomes

$$\frac{e^{-\frac{1}{2}(\frac{z}{b}+\lambda_1+\lambda_2)}}{a^{\frac{k_1}{2}}b^{\frac{k_2}{2}}2^{\frac{k_1+k_2}{2}}}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{\left(\frac{\lambda_1}{4a}\right)^m\left(\frac{\lambda_2}{4b}\right)^n}{m!n!\Gamma(k_1/2+m)\Gamma(k_2/2+n)}\int_0^z e^{-\frac{b-a}{2ab}x}x^{\frac{k_1}{2}+m-1}(z-x)^{\frac{k_2}{2}+n-1}dx, \tag{3.5}$$

After changing the variable $x = zt$, we see that the integral becomes

$$z^{\frac{k_1+k_2}{2}+m+n-1}\int_0^1 e^{-\frac{b-a}{2ab}zt}t^{\frac{k_1}{2}+m-1}(1-t)^{\frac{k_2}{2}+n-1}dt \tag{3.6}$$

Noting that the integral is equal to a confluent hypergeometric function in (2.9), given by

$$\frac{\Gamma(\frac{k_1}{2}+m)\Gamma(\frac{k_2}{2}+n)}{\Gamma(\frac{k_1+k_2}{2}+m+n)}\,{}_1F_1\left(\frac{k_1}{2}+m,\frac{k_1+k_2}{2}+m+n,-\frac{b-a}{2ab}z\right) \tag{3.7}$$

and substituting (3.6) and (3.7) in (3.5), we obtain (3.3) and the proof of Theorem 3.1 follows. □

Several special cases of Theorem 3.1 are given in the following corollaries, which can be derived using equation (3.2).

Corollary 3.1.1. ***Central case.*** *The pdf of the weighted sum of central χ^2-distributions with equal df*

$$Z \sim WS((k,k)^T,(0,0)^T,(a,b)^T) = a\chi_k^2 + b\chi_k^2,$$

with k, a, $b > 0$ and $a \neq b$, is given by

$$f_Z(z;a,b,k) = \frac{\Gamma\left(\frac{1+k}{2}\right)}{(ab)^{\frac{1}{2}}2^{\frac{3}{2}-\frac{k}{2}}\Gamma(k)}\exp\left\{-\frac{a+b}{4ab}z\right\}z^{\frac{k-1}{2}}(|a-b|)^{\frac{1-k}{2}}I_{\frac{k-1}{2}}\left(\frac{|a-b|}{4ab}z\right), \tag{3.8}$$

where $z > 0$.

Proof. When $k_1 = k_2 = k$ and by using the pdf of central χ^2-distribution in the convolution formula in (2.11), (3.2) becomes

$$f_Z(z;a,b,k) = \frac{e^{-\frac{z}{2b}}}{(ab)^{\frac{k}{2}}2^k\Gamma(\frac{k}{2})^2}\int_0^z \exp\left\{-\frac{x}{2}\left(\frac{1}{a}-\frac{1}{b}\right)\right\}x^{\frac{k}{2}-1}(z-x)^{\frac{k}{2}-1}dx. \tag{3.9}$$

After changing the variable $x = zt$, the integral becomes

$$z^{k-1}\int_0^1 \exp\left\{-\frac{b-a}{2ab}zt\right\}t^{\frac{k}{2}-1}(1-t)^{\frac{k}{2}-1}dt. \tag{3.10}$$

It can be seen the above integral is equal to the confluent hypergeometric function in (2.9), so we have

$$\frac{\Gamma(\frac{k}{2})^2}{\Gamma(k)}\,{}_1F_1\left(\frac{k}{2};k;\frac{a-b}{2ab}z\right). \tag{3.11}$$

Substituting the Kummer's second transform (see Section 13.6 on page 509 of Abramowitz and Stegun [12]):

$${}_1F_1\left(\frac{k}{2};k;\frac{a-b}{2ab}z\right) = \exp\left\{\frac{a-b}{4ab}z\right\}\left(\frac{a-b}{8ab}z\right)^{\frac{1-k}{2}}\Gamma\left(\frac{1+k}{2}\right)I_{\frac{k-1}{2}}\left(\frac{a-b}{4ab}z\right) \tag{3.12}$$

into (3.9) and using (3.10) and (3.11), the result follows. □

Corollary 3.1.2. ***Central case.*** *The pdf of the weighted sum of central χ^2-distributions with different df*

$$Z \sim WS((k_1,k_2)^T,(0,0)^T,(a,b)^T) = a\chi^2_{k_1} + b\chi^2_{k_2},$$

with k_1, k_2, a, $b > 0$ and $a \neq b$, is given by

$$f_Z(z;a,b,k_1,k_2) = \frac{z^{\frac{k_1+k_2}{2}-1}e^{-\frac{z}{2b}}}{(2a)^{\frac{k_1}{2}}(2b)^{\frac{k_2}{2}}\Gamma(\frac{k_1+k_2}{2})}\,{}_1F_1\left(\frac{k_1}{2};\frac{k_1+k_2}{2};\frac{a-b}{2ab}z\right),\ \ z>0. \tag{3.13}$$

Proof. When $k_1 \neq k_2$ and by using the pdf of central χ^2-distribution in the convolution formula in (2.11), (3.2) becomes

$$f_Z(z;a,b,k_1,k_2) = \frac{e^{-\frac{z}{2b}}}{a^{\frac{k_1}{2}}b^{\frac{k_2}{2}}2^{\frac{k_1+k_2}{2}}\Gamma(\frac{k_1}{2})\Gamma(\frac{k_2}{2})}\int_0^z \exp\left\{-\frac{x}{2}\left(\frac{1}{a}-\frac{1}{b}\right)\right\}x^{\frac{k_1}{2}-1}(z-x)^{\frac{k_2}{2}-1}dx. \tag{3.14}$$

on changing the variable $x = zt$ in the integral and using the confluent hypergeometric function given in (2.9), the result follows. □

Corollary 3.1.3. ***Mixed case.*** *The pdf of the weighted sum of one central and one non-central χ^2-distributions with different df*

$$Z \sim WS((k_1,k_2)^T,(\lambda,0)^T,(a,b)^T) = a\chi^2_{k_1}(\lambda) + b\chi^2_{k_2},$$

with λ, k_1, k_2, a, $b > 0$ and $a \neq b$, is given by

$$f_Z(z;a,b,k_1,k_2,\lambda) = \frac{\exp\left\{-\frac{1}{2}(\frac{z}{b}+\lambda)\right\}}{2^{\frac{k_1+k_2}{2}}a^{\frac{k_1}{2}}b^{\frac{k_2}{2}}}\sum_{m=0}^{\infty}\frac{\left(\frac{\lambda}{4a}\right)^m z^{\frac{k_1+k_2}{2}+m-1}}{m!\Gamma(\frac{k_1+k_2}{2}+m)}\,{}_1F_1\left(\frac{k_1}{2}+m,\frac{k_1+k_2}{2}+m,-\frac{b-a}{2ab}z\right), \tag{3.15}$$

where $z > 0$.

Proof. By using the convolution formula in (2.11) with one central and another non-central pdf, we can write (3.2) directly as

$$f_Z(z;a,b,k_1,k_2,\lambda) = \frac{\exp\left\{-\frac{z}{2b}-\frac{\lambda}{2}\right\}}{2^{\frac{k_2}{2}+1}a^{\frac{k_1}{4}+\frac{1}{2}}b^{\frac{k_2}{2}}\lambda^{\frac{k_1}{4}-\frac{1}{2}}\Gamma(\frac{k_2}{2})}\int_0^z e^{-\frac{x}{2}(\frac{1}{a}-\frac{1}{b})}x^{\frac{k_1}{4}-\frac{1}{2}}I_{\frac{k_1}{2}-1}\left(\sqrt{\frac{\lambda x}{a}}\right)(z-x)^{\frac{k_2}{2}-1}dx. \tag{3.16}$$

By using the series expansion of Bessel function $I_p(z)$ given in (2.4) and changing the order of summation and integration, we obtain

$$\frac{\exp\left\{-\frac{1}{2}(\frac{z}{b}+\lambda)\right\}}{2^{\frac{k_1+k_2}{2}}\Gamma(\frac{k_2}{2})a^{\frac{k_1}{2}}b^{\frac{k_2}{2}}}\sum_{m=0}^{\infty}\frac{\left(\frac{\lambda}{4a}\right)^m}{m!\Gamma(\frac{k_1}{2}+m)}\int_0^z e^{-\frac{b-a}{2ab}x}x^{\frac{k_1}{2}+m-1}(z-x)^{\frac{k_2}{2}-1}dx. \tag{3.17}$$

Again, on changing the variable: $x = zt$ and using the confluent hypergeometric function given in (2.9), the result follows. □

The first two moments of random variable Z given by (3.1) can be shown to be the following:

$$\mathbb{E}\{Z\} = a(k_1+\lambda_1)+b(k_2+\lambda_2),\ \ \text{var}\{Z\} = 2a^2(k_1+2\lambda_1)+2b^2(k_2+2\lambda_2). \tag{3.18}$$

These moments are used later in the moment estimators for this distribution.

4 Ex-χ^2 Distribution

In this Section and the following Section 5, we will assume that $k_2 = 2$ for df and the non-centrality parameter $\lambda_2 = 0$ in Z, so that the second variable in this weighted sum is as an exponential distribution written as $Exp(c)$ with the pdf given by

$$f(x; c) = ce^{-cx}, \ x > 0, \ c > 0, \tag{4.1}$$

which in the case of χ_2^2, $c = \frac{1}{2}$, but we will use a general c. We will denote the pdf of $Exp(c)$ given by (4.1) as $\zeta(\cdot; c)$.

In this section, we consider the following ex-χ^2 random variable defined by:

$$Z \sim \text{ex-}\chi^2(k, \lambda, c, (a, b)^T) = a\chi_k^2(\lambda) + bExp(c), \tag{4.2}$$

where a, b, k and $c > 0$. We will use ex-$\chi^2(k, \lambda, c, (a, b)^T)$ as an abbreviation for ex-χ^2 distribution.

We start from the case where $\lambda = 0$ in Z.
Theorem 4.1. ***Central case.*** *The pdf of the weighted sum of the central case*

$$Z \sim ex\text{-}\chi^2(k, 0, c, (a, b)^T) = a\chi_k^2 + bExp(c)$$

is given by

1. *for* $c < \frac{b}{2a}$,

$$f_Z(z; a, b, k, c) = \frac{\left(\frac{b}{b-2ac}\right)^{\frac{k}{2}}}{\Gamma(\frac{k}{2})} \zeta\left(z; \frac{c}{b}\right) \gamma\left(\frac{k}{2}, \left(\frac{1}{2a} - \frac{c}{b}\right) z\right), \ z > 0; \tag{4.3}$$

2. *for* $c > \frac{b}{2a}$,

$$f_Z(z; a, b, k, c) = \frac{2\left(\frac{z}{2a}\right)^{\frac{k}{2}}}{k\Gamma(\frac{k}{2})} \zeta\left(z; \frac{c}{b}\right) {}_1F_1\left(\frac{k}{2}, \frac{k}{2} + 1, \left(\frac{c}{b} - \frac{1}{2a}\right) z\right), \ z > 0, \tag{4.4}$$

where $\zeta(\cdot; c)$ *is the pdf of* $Exp(c)$.

Proof. The convolution formula given in (2.11) becomes

$$f_Z(z; a, b, k, c) = \frac{ce^{-\frac{c}{b}z}}{(2a)^{\frac{k}{2}}\Gamma(\frac{k}{2})b} \int_0^z x^{\frac{k}{2}-1} e^{-\left(\frac{1}{2a} - \frac{c}{b}\right)x} dx \tag{4.5}$$

$c < \frac{b}{2a}$: by changing the variable: $u = \left(\frac{1}{2a} - \frac{c}{b}\right) x$, we see that the above integral is the same as the lower incomplete gamma function defined in (2.6) and is given as

$$\left(\frac{1}{2a} - \frac{c}{b}\right)^{-\frac{k}{2}} \gamma\left(\frac{k}{2}, \left(\frac{1}{2a} - \frac{c}{b}\right) z\right). \tag{4.6}$$

We substitute (4.6) into (4.5) and the result follows.

$c > \frac{b}{2a}$: we use the integral identity quoted in Section 3.38, page 318, of Gradshteyn and Ryzhik [13]:

$$\int_0^u x^{v-1}(u - x)^{\mu-1} e^{\beta x} dx = B(\mu, v) u^{\mu+v-1} {}_1F_1(v, \mu + v, \beta u), \tag{4.7}$$

where $B(\cdot)$ is the beta function given in (2.7). Thus, the integral becomes

$$B\left(1, \frac{k}{2}\right) {}_1F_1\left(\frac{k}{2}, \frac{k}{2}+1, \left(\frac{c}{b} - \frac{1}{2a}\right) z\right) \tag{4.8}$$

By substituting (4.8) into (4.5), the result follows. $\square$

Next, we consider the case where $\lambda > 0$ in Z.

Theorem 4.2. ***Non-central case.*** *The pdf for the weighted sum of the non-central case*

$$Z \sim ex\text{-}\chi^2(k, \lambda, c, (a,b)^T) = a\chi_k^2(\lambda) + bExp(c)$$

is given by

- *when* $a < \frac{b}{2c}$,

$$f_Z(z; a, b, k, \lambda) = e^{-\frac{\lambda}{2}} \left(\frac{b}{b-2ac}\right)^{\frac{k}{2}} \zeta\left(z; \frac{c}{b}\right) \sum_{m=0}^{\infty} \frac{\left(\frac{\lambda b}{2(b-2ac)}\right)^m}{m!\Gamma(\frac{k}{2}+m)} \gamma\left(\frac{k}{2}+m, \frac{b-2ac}{2ab} z\right); \tag{4.9}$$

- *when* $a > \frac{b}{2c}$,

$$f_Z(z; a, b, k, \lambda) = e^{-\frac{\lambda}{2}} \left(\frac{z}{2a}\right)^{\frac{k}{2}} \zeta\left(z; \frac{c}{b}\right) \sum_{m=0}^{\infty} \frac{\left(\frac{\lambda z}{4a}\right)^m {}_1F_1\left(\frac{k}{2}+m, \frac{k}{2}+m+1, \frac{2ac-b}{2ab} z\right)}{m!\Gamma(\frac{k}{2}+m+1)}, \tag{4.10}$$

where $z > 0$ *in the above two equations and* $\zeta(\cdot; c)$ *is the pdf of* $Exp(c)$.

Proof. The convolution integral given in (2.11) becomes

$$f_Z(z; a, b, k, \lambda) = \frac{c \exp\left\{-\frac{cz}{b} - \frac{\lambda}{2}\right\}}{2a^{\frac{k}{4}+\frac{1}{2}} b \lambda^{\frac{k}{4}-\frac{1}{2}}} \int_0^z e^{-x\left(\frac{1}{2a}-\frac{c}{b}\right)} x^{\frac{k}{4}-\frac{1}{2}} I_{\frac{k}{2}-1}\left(\sqrt{\frac{\lambda x}{a}}\right) dx. \tag{4.11}$$

By using the series expansion of $I_p(z)$ given in (2.4), we obtain

$$\left(\frac{1}{2}\sqrt{\frac{\lambda}{a}}\right)^{\frac{k_1}{2}-1} \sum_{m=0}^{\infty} \frac{\left(\frac{\lambda}{4a}\right)^m}{m!\Gamma(\frac{k}{2}+m)} \int_0^z e^{-x\left(\frac{1}{2a}-\frac{c}{b}\right)} x^{\frac{k}{2}+m-1} dx. \tag{4.12}$$

When $a < \frac{b}{2c}$, by using the following integral identity in Section 3.38, page 317, of Gradshteyn and Ryzhik [13] which is given as

$$\int_0^u x^{v-1} e^{-\mu x} dx = \mu^{-v} \gamma(v, \mu u), \tag{4.13}$$

where $\gamma(\cdot)$ is the lower incomplete gamma function given in (2.6). Hence, the integral given in (4.12) becomes

$$\left(\frac{b-2ac}{2ab}\right)^{-\left(\frac{k}{2}+m\right)} \gamma\left(\frac{k}{2}+m, \frac{b-2ac}{2ab} z\right), \tag{4.14}$$

Substituting equation (4.12) and (4.14) into (4.11), the result follows.

When $a > \frac{b}{2c}$, using (4.7), the integral given in (4.12) becomes

$$\left(\frac{k}{2}+m\right)^{-1} z^{\frac{k}{2}+m} {}_1F_1\left(\frac{k}{2}+m, \frac{k}{2}+m+1, \frac{2ac-b}{2ab} z\right). \tag{4.15}$$

Substituting equation (4.12) and (4.15) into (4.11), the result follows.

□

Corollary 4.2.1. ***Non-central case.*** *The pdf of the weighted sum of $\chi_2^2(\lambda)$ and $\chi_2^2 \equiv Exp(\frac{1}{2})$*

$$Z \sim ex\text{-}\chi^2\left(2, \lambda, \frac{1}{2}, (a,b)^T\right) = a\chi_2^2(\lambda) + bExp\left(\frac{1}{2}\right)$$

is given by

- *when $a < b$,*

$$f_Z(z; a, b, \lambda) = \frac{e^{-\frac{\lambda}{2}}\zeta(z; \frac{1}{2b})}{1 - \frac{a}{b}} \sum_{m=0}^{\infty} \frac{(\frac{\lambda b}{2(b-a)})^m}{(m!)^2} \gamma\left(m+1, \frac{b-a}{2ab}z\right), \ z > 0; \tag{4.16}$$

- *when $a > b$,*

$$f_Z(z; a, b, \lambda) = \frac{e^{-\frac{\lambda}{2}} z\zeta(z; \frac{1}{2b})}{2a} \sum_{m=0}^{\infty} \frac{(\frac{\lambda z}{4a})^m}{(m!)^2(m+1)} {}_1F_1\left(m+1, m+2, \frac{a-b}{2ab}z\right), \ z > 0, \tag{4.17}$$

where $\zeta(\cdot; c)$ is the pdf of $Exp(c)$.

Proof. By substituting $k = 2$ and $c = \frac{1}{2}$ in equation (4.9) and (4.10), the results follow. □

Remark 1. *Note that the proof depends on $a < \frac{b}{2c}$ or $a > \frac{b}{2c}$ for both central and non-central cases, which is due to the $Exp(c)$ and the weights.. Further, it is easier to derive the distribution of Z directly in the particular case of $\lambda = 0$ as we have done in Theorem 4.1 rather than obtaining it as a corollary of Theorem 4.2 with $\lambda > 0$.*

The first two moments of random variable Z given by (4.2) can be shown to be the following:

$$\mathbb{E}\{Z\} = a(k+\lambda) + \frac{b}{c}, \ \text{var}\{Z\} = 2a^2(k+2\lambda) + \frac{b^2}{c^2}. \tag{4.18}$$

These moments are used later in the moment estimators for this distribution.

5 Ex-Gaussian Distribution

In this section, we consider a particular case of the ex-χ^2 defined in equation (4.2) with df $k \to \infty$. It is well known that as $k \to \infty$, the $\chi_k^2(\lambda)$ can be approximated by

$$N(k+\lambda, 2(k+2\lambda)).$$

Hence, the random variable is given by

$$Z \sim \text{exG}(\mu, \sigma^2, c, (a,b)^T) = aN(\mu, \sigma^2) + bExp(c), \ a > 0, \ b > 0, \ c > 0, \tag{5.1}$$

which is the weighted sum of a normal and an exponential distribution and known as the ex-Gaussian distribution, which we have denoted by $\text{exG}(\mu, \sigma^2, c, (a,b)^T)$. Note that here we have got one normal variable with range $(-\infty, \infty)$, whereas for the exponential part, the variable has domain $(0, \infty)$. Hence, there is a choice we could use for the convolution formula obtained from (2.11) by interchanging X

and Y and substituting in the proper domain for $\mathcal{X}$ in two ways as follows, namely for the normal integration part (X with $Z = X + Y$) we have in the convolution formula is

$$f_Z(z) = \int_{-\infty}^{z} f_X(x) f_Y(z - x) dx, \tag{5.2}$$

whereas for the exponential integration part (Y) we have in the convolution formula is

$$f_Z(z) = \int_{0}^{\infty} f_X(z - y) f_Y(y) dy. \tag{5.3}$$

For (5.3), note that since the exponential random variable Y is always positive, so the lower limit of the convolution integral for this case is 0. Further, the normal random variable can be negative, hence it is possible that $Z < Y$, so Y is not bounded above (unlike the normal random variable X which is bounded above by Z). Consequently, the upper limit of the integral in (5.3) is ∞.

Theorem 5.1. *The pdf of the weighted sum of normal and exponential random variables*

$$Z \sim exG(\mu, \sigma^2, c, (a,b)^T) = aN(\mu, \sigma^2) + bExp(c)$$

is given by

$$f_Z(z; a, b, c, \mu, \sigma^2) = \exp\left\{\frac{c}{b}a\mu + \frac{c^2a^2\sigma^2}{2b^2}\right\} \zeta\left(z; \frac{c}{b}\right) \Phi\left(\frac{\sqrt{2}(z-\alpha)}{\beta}\right), \; z \in \mathbb{R}, \tag{5.4}$$

where $\zeta(\cdot; c)$ is the pdf of $Exp(c)$ and

$$\alpha = a\mu + \frac{ca^2\sigma^2}{b}, \; \beta = \sqrt{2a^2\sigma^2}.$$

Proof. Using the convolution formula given in (2.11), we have

$$\begin{aligned} f_Z(z; a, b, c, \mu, \sigma^2) &\propto e^{-\frac{cz}{b}} \int_{-\infty}^{z} \exp\left\{-\frac{(x - (a\mu + \frac{ca^2\sigma^2}{b}))^2}{2a^2\sigma^2}\right\} dx \\ &\propto e^{-\frac{cz}{b}} \Phi\left(\frac{\sqrt{2}(z-\alpha)}{\beta}\right), \end{aligned} \tag{5.5}$$

where we have changed the variable: $t = \frac{x-(a\mu+\frac{ca^2\sigma^2}{b})}{|a|\sigma}$ in the first expression and let

$$\alpha = a\mu + \frac{ca^2\sigma^2}{b}, \; \beta = \sqrt{2a^2\sigma^2} \tag{5.6}$$

in the second expression. By integrating over (5.5), the normalizing constant is

$$\frac{c}{b}\exp\left\{\frac{c}{b}a\mu + \frac{c^2a^2\sigma^2}{2b^2}\right\}$$

and the result hence follows from (5.5). □

The following corollary is a special case of Theorem 5.1 and can be derived easily from (5.4).

Corollary 5.1.1. *The pdf of the weighted sum of normal and χ^2_2 distributions*

$$Z \sim exG\left(\mu, \sigma^2, \frac{1}{2}, (a,b)^T\right) = aN(\mu, \sigma^2) + bExp\left(\frac{1}{2}\right) = aN(\mu, \sigma^2) + b\chi^2_2$$

is given by

$$f_Z(z;a,b,\mu,\sigma^2) = \exp\left\{\frac{\sigma^2 a^2}{8b^2} + \frac{a\mu}{2b}\right\}\zeta\left(z;\frac{1}{2b}\right)\Phi\left(\frac{\sqrt{2}(z-\alpha)}{\beta}\right),\ z\in\mathbb{R}, \tag{5.7}$$

where $\zeta(\cdot;c)$ *is the pdf of* $Exp(c)$ *and*

$$\alpha = a\mu + \frac{a^2\sigma^2}{2b},\ \beta = \sqrt{2\sigma^2 a^2}.$$

In the following theorem, we consider the *negative ex-Gaussian distribution*, which has $b<0$.

Theorem 5.2. *When* $b<0$*, the pdf of negative ex-Gaussian distribution*

$$Z \sim exG(\mu,\sigma^2,c,(a,-|b|)^T) = aN(\mu,\sigma^2) - |b|Exp(c)$$

is given by

$$f_Z(z;a,b,\mu,\sigma^2) = \frac{c}{|b|}\exp\left\{\frac{a^2\sigma^2c^2}{2b^2} + \frac{c(z-a\mu)}{|b|}\right\}\Phi\left(-\frac{\sqrt{2}(z-\alpha')}{\beta}\right) \tag{5.8}$$

where

$$\alpha' = a\mu - \frac{a^2\sigma^2}{2|b|},\ \beta = \sqrt{2\sigma^2 a^2},$$

Proof. Since $Y\sim Exp(c)$, thus for $b<0$, the pdf of bY is

$$f_Y(y;b) = \frac{c}{|b|}e^{\frac{c}{|b|}y}, y<0.$$

In this case, we use the convolution formula given in (2.11) so that

$$\begin{aligned} f_Z(z;a,b,\mu,\sigma^2) &\propto e^{\frac{c(z-a\mu)}{|b|}}\int_{-\infty}^{0}\exp\left\{-\frac{[y-(z-a\mu+\frac{a^2\sigma^2c}{|b|})]^2}{2a^2\sigma^2}\right\}dy \\ &\propto e^{\frac{c(z-a\mu)}{|b|}}\Phi\left(-\frac{z-(a\mu-\frac{a^2\sigma^2c}{|b|})}{|a|\sigma}\right) \\ &\propto e^{\frac{c(z-a\mu)}{|b|}}\Phi\left(-\frac{\sqrt{2}(z-\alpha')}{\beta}\right), \end{aligned} \tag{5.9}$$

where we have changed the variable: $t = \frac{y-(z-a\mu+\frac{a^2\sigma^2c}{|b|})}{|a|\sigma}$ in the first integral and have set $\alpha' = a\mu - \frac{a^2\sigma^2c}{|b|}$, $\beta=\sqrt{2\sigma^2a^2}$ in the third expression. By integrating over the last step of (5.9), the normalizing constant is

$$\frac{c}{|b|}e^{\frac{a^2\sigma^2c^2}{2b^2}},$$

and hence, the result follows from (5.9). □

We note that the negative ex-Gaussian has been derived in Carr et al. [14] using a different method. Carr et al. [14] uses this distribution for modeling the option prices. We have drawn several plots of the two densities: ex-χ^2 and ex-Gaussian and both are skewed distribution which have been already noted in previous papers.

The first three moments of random variable Z given by (5.1) can be shown to be the following.

$$\mathbb{E}\{Z\} = a\mu + \frac{b}{c}, \ \text{var}\{Z\} = a^2\sigma^2 + \frac{b^2}{c^2}, \ \beta_1 = \frac{2(\frac{b}{c})^3}{(a^2\sigma^2 + (\frac{b}{c})^2)^{\frac{3}{2}}}. \tag{5.10}$$

Note that the skewness coefficient β_1 in (5.10) is always positive for this distribution. These moments are used later for moment estimators for this distribution.

6 Estimation

6.1 Parameter Estimation for Weighted Sum of two χ^2-Distributions

Recall the random variable Z defined in equation (1.1):

$$Z \sim \text{WS}((k_1, k_2)^T, (\lambda_1, \lambda_2)^T, (a, b)^T) = a\chi^2_{k_1}(\lambda_1) + b\chi^2_{k_2}(\lambda_2),$$

with a and b known and inference should be provided on λ_1 and λ_2. We give below first the method of moment (MoM) estimates (which are easy to compute) and then give some methods on how to compute maximum likelihood estimate (MLE) numerically since the maximum likelihood equations are analytically intractable.

6.1.1 Moment Estimates

Let $z_1, \cdots, z_n$ denote a sample of $Z \sim \text{WS}((k_1, k_2)^T, (\lambda_1, \lambda_2)^T, (a, b)^T)$. Let $\bar{z}$ denote the sample mean and s^2 denote the sample variance. By using the moments given in (3.18), the MoM equations are

$$\bar{z} = a(k_1 + \lambda_1) + b(k_2 + \lambda_2), \ s^2 = 2a^2(k_1 + 2\lambda_1) + 2b^2(k_2 + 2\lambda_2).$$

By solving the above two equations, we have

$$\hat{\lambda}_1 = \frac{4b\bar{z} - s^2 - (4ab - 2a^2)k_1 - 2b^2k_2}{4ab - 4a^2}, \ \hat{\lambda}_2 = \frac{\bar{z} - a(k_1 + \hat{\lambda}_1) - bk_2}{b}. \tag{6.1}$$

Now, let $u_1, \cdots, u_{k_1}$ be independently normally distributed such that $u_i \sim N(\mu_i, \sigma_1^2)$ for $i = 1, \ldots, k_1$, and let $u_{k_1+1}, \cdots, u_{k_1+k_2}$ be independently normally distributed such that $u_i \sim N(\mu_i, \sigma_2^2)$ for $i = k_1 + 1, \ldots, k_1 + k_2$. Thus, we can give that

$$\sum_{i=1}^{k_1} \frac{u_i^2}{\sigma_1^2} \sim \chi^2_{k_1}(\lambda_1) \Rightarrow \sum_{i=1}^{k_1} u_i^2 \sim \sigma_1^2\chi^2_{k_1}(\lambda_1); \ \sum_{i=k_1+1}^{k_1+k_2} \frac{u_i^2}{\sigma_2^2} \sim \chi^2_{k_2}(\lambda_2) \Rightarrow \sum_{i=k_1+1}^{k_1+k_2} u_i^2 \sim \sigma_2^2\chi^2_{k_2}(\lambda_2),$$

with

$$\lambda_1 = \sum_{i=1}^{k_1} \frac{\mu_i^2}{\sigma_1^2}, \ \lambda_2 = \sum_{i=k_1+1}^{k_1+k_2} \frac{\mu_i^2}{\sigma_2^2}. \tag{6.2}$$

Hence,

$$\sum_{i=1}^{k_1+k_2} u_i^2 \sim \sigma_1^2\chi^2_{k_1}(\lambda_1) + \sigma_2^2\chi^2_{k_2}(\lambda_2), \tag{6.3}$$

where $a = \sigma_1^2$ and $b = \sigma_2^2$. The estimates of λ_1 and λ_2 can be derived by plugging in the estimates of σ_1^2, σ_2^2 and μ_i, for $i = 1, \dots, k_1 + k_2$:

$$\hat{\lambda}_1 = \sum_{i=1}^{k_1} \frac{\hat{\mu}_i^2}{\hat{\sigma}_1^2}, \ \hat{\lambda}_2 = \sum_{i=k_1+1}^{k_1+k_2} \frac{\hat{\mu}_i^2}{\hat{\sigma}_2^2}. \tag{6.4}$$

The advantage of this method over the MoM given in (6.1) is that estimates of λ_1 and λ_2 cannot be negative. Further, the estimates of σ_1^2 and σ_2^2 can be obtained as the following:

1. Subtract the sample means $\hat{\mu}_i$ from realizations of u_i, for $i = 1, \dots, k_1$ and $i = k_1 + 1, \dots, k_1 + k_2$ separately. Then stack the observations into a long vector.
2. Compute sample variance on this stacked vector.

This method is applicable since the random variable Z has been derived originally from the matrix normal distribution, so the estimates of non-centrality parameters can be obtained by estimating the parameters of the normal distribution. Since this method of moment estimation uses different types of moments comparing with (6.1), so we will call these estimates *hybrid moment estimates*.

Similarly, for the ex-χ^2 random variable given in (4.2)

$$Z \sim \text{ex-}\chi^2(k, \lambda, c, (a, b)^T) = a\chi_k^2(\lambda) + bExp(c),$$

where a and b are known, the MoM equations by using (4.18) are

$$\bar{z} = a(k + \lambda) + \frac{b}{c}, \ s^2 = 2a^2(k + 2\lambda) + \frac{b^2}{c^2}.$$

By solving these two equations, we have

$$\hat{c} = \frac{4ab \pm \sqrt{16a^2b^2 - 4b^2(4a\bar{z} - s^2 - 2a^2k)}}{2(4a\bar{z} - s^2 - 2a^2k)}, \ \hat{\lambda} = \frac{\hat{c}(\bar{z} - ak) - b}{a\hat{c}}. \tag{6.5}$$

Note that there are two roots and depending on the data, an appropriate root can be selected in general. We do not use the moment estimates in our illustration in Section 7.2.2 since we could compute the MLEs.

6.1.2 MLE

We can write the MLE equations for the parameters, but the solutions are not in closed form. However, we can estimate these parameters in practical examples using simulated annealing (SA) algorithm as described in Section 7.2.2.

6.2 Parameter Estimation for ex-Gaussian

We will consider below two parameter estimation approaches – MoM and MLE – for the ex-Gaussian distribution. Let $z_1, \dots, z_n$ be a sample of random variable Z defined in (5.1)

$$Z \sim \text{exG}(\mu, \sigma^2, c, (a, b)^T) = aN(\mu, \sigma^2) + bExp(c)$$

with sample mean $\bar{z}$ and sample variance s^2, where a and b are given.

6.2.1 MoM

Using (5.10), we find that the MoM equations to estimate the parameters (μ, σ^2, c) are

$$\bar{z} = a\mu + \frac{b}{c}, \; s^2 = a^2\sigma^2 + \frac{b^2}{c^2}, \; b_1 = \frac{2(\frac{b}{c})^3}{(a^2\sigma^2 + (\frac{b}{c})^2)^{\frac{3}{2}}}, \tag{6.6}$$

where $\bar{z}$ is the sample mean, s^2 is the sample variance and b_1 is the sample skewness coefficient

$$b_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(z_i - \bar{z})^3}{(s^2)^{\frac{3}{2}}}. \tag{6.7}$$

For given a and b, it can be shown that the moment estimates $\hat{\mu}$, $\hat{\sigma}^2$ and $\hat{c}$ of μ, σ^2 and c are given by (see, for example, Dyson [15])

$$\hat{\mu} = \frac{1}{a}\left(\bar{z} - s\left(\frac{b_1}{2}\right)^{\frac{1}{3}}\right), \; \hat{\sigma}^2 = \frac{s^2}{a^2}\left(1 - \left(\frac{b_1}{2}\right)^{\frac{2}{3}}\right), \; \hat{c} = \frac{b}{s}\left(\frac{b_1}{2}\right)^{-\frac{1}{3}}. \tag{6.8}$$

However, in our illustration in Section 7.2.2, we will be using only MLE.

6.2.2 MLE

We rewrite the pdf of the ex-Gaussian distribution as follows:

$$f_Z(z; a, b, c, \mu, \sigma^2) = \frac{c}{b}\exp\left\{\frac{c}{b}a\mu + \frac{c^2a^2\sigma^2}{2b^2} - \frac{cz}{b}\right\}\Phi\left(\frac{\sqrt{2}(z-\alpha)}{\beta}\right), \; z \in \mathbb{R}, \tag{6.9}$$

where $\Phi(\cdot)$ is the cdf of $N(0,1)$ and

$$\alpha = a\mu + \frac{ca^2\sigma^2}{b}, \; \beta = \sqrt{2a^2\sigma^2}.$$

Suppose a and b are known as before and the aim is to estimate MLE of μ, σ^2 and c. It can be seen that the log-likelihood is

$$\ell(\mu, \sigma^2, c) = n\log\frac{c}{b} + \sum_{i=1}^{n}\left(\frac{c}{b}a\mu + \frac{c^2a^2\sigma^2}{2b^2} - \frac{cz_i}{b}\right) + \sum_{i=1}^{n}\log\Phi\left(\frac{\sqrt{2}(z_i-\alpha)}{\beta}\right) \tag{6.10}$$

Recall that

$$\Phi(x) = \int_{-\infty}^{x}\varphi(t)dt, \; \frac{d}{dx}\Phi(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},$$

where $\varphi(\cdot)$ is the pdf of $N(0,1)$. Hence, the MLE for $(\mu,\ \sigma^2,\ c)$ are the solutions of the following equations

$$
\begin{aligned}
\frac{\partial \ell(\mu,\sigma^2,c)}{\partial \mu} &= \frac{nac}{b} - \sum_{i=1}^{n} \frac{\varphi\left(\frac{\sqrt{2}(z_i-\alpha)}{\beta}\right)\frac{1}{\sqrt{\sigma^2}}}{\Phi\left(\frac{\sqrt{2}(z_i-\alpha)}{\beta}\right)} = 0 \\
\frac{\partial \ell(\mu,\sigma^2,c)}{\partial \sigma^2} &= \frac{na^2c^2}{2b^2} - \sum_{i=1}^{n} \frac{\varphi\left(\frac{\sqrt{2}(z_i-\alpha)}{\beta}\right)}{\Phi\left(\frac{\sqrt{2}(z_i-\alpha)}{\beta}\right)}\left[\frac{ac}{2b}(\sigma^2)^{-\frac{1}{2}} + \frac{z_i - a\mu}{2a}(\sigma^2)^{-\frac{3}{2}}\right] = 0 \qquad (6.11) \\
\frac{\partial \ell(\mu,\sigma^2,c)}{\partial c} &= \frac{n}{c} + \frac{na\mu}{b} + \frac{nca^2\sigma^2}{b^2} - \sum_{i=1}^{n}\frac{z_i}{b} - \sum_{i=1}^{n} \frac{\varphi\left(\frac{\sqrt{2}(z_i-\alpha)}{\beta}\right)\frac{\sqrt{a^2\sigma^2}}{b}}{\Phi\left(\frac{\sqrt{2}(z_i-\alpha)}{\beta}\right)} = 0
\end{aligned}
$$

It can be seen that there is no closed solution and these equations can only be solved numerically. We will use SA in our practical application since it is a reliable modern approach and details are in Section 7.2.2. Alternatively, one can use the older numerical methods such as quantile maximum likelihood method as done in Brown and Heathcote [16] to compute the MLE.

7 Application to Shape Analysis

7.1 Bilateral Symmetry

Recall from Section 1 that shape analysis deals with shapes of objects when the effects of translation, scaling and rotation are filtered out (see, for example, Dryden and Mardia [10]). An object is defined by a set of landmarks on which the samples are taken. Here, we focus on bilateral symmetry of objects (Mardia et al. [17]). An object in two or three dimensions is said to be bilaterally symmetric if its mirror image about some line, in two dimensions, or some plane, in three dimensions, is the same as the original form after relabeling paired (see below) landmarks. This mirroring locus will, in general, be called the mid-plane (Mardia et al. [17]). The same discussion applies for size-and-shape analysis, which is our focus in the application.

Further, in a population with perfect bilateral symmetry, we have two types of landmarks. Some are paired and they do not lie on the mid-plane, but appear separately on left and right sides. Additional landmarks must lie exactly on the mid-plane; they are unpaired/solos. Let there be K_P landmark pairs (so there are in total $2K_P$ landmarks) and K_S solos, so the total landmarks are $K = 2K_P + K_S$. Let $X \in \mathbb{R}^{K\times M}$ denote the random "configuration matrix" for an object with K landmarks in M dimensions, with $X[k,m]$ for the coordinate m of landmark k, where $k = 1,\ldots,K$ and $m = 1,\ldots,M$. We write the coordinate vector of landmark k as $X[k,\ ]$. Let (k_L, k_R) denote the indices for a typical landmark pair and k_S the index for a solo, where $k_L \in \{k : \text{landmark } k \text{ is on the left side of mid-plane}\}$ and similar for k_R, $k_S \in \{k : \text{landmark } k \text{ is solo}\}$.

Let $\text{vec}(\cdot)$ denote the matrix vectorization operator: $\text{vec}(X) \in \mathbb{R}^{KM}$ is the vector with length KM obtained by stacking columns of X together. We assume that

$$\text{vec}(X) \sim N(\text{vec}(\boldsymbol{\mu}), \Sigma), \qquad (7.1)$$

where $\boldsymbol{\mu} = \mathbb{E}\{X\} \in \mathbb{R}^{K\times M}$ is the mean configuration. $\Sigma \in \mathbb{R}^{KM\times KM}$ is the covariance matrix. For simplicity, we assume $\Sigma = \sigma^2 I_{KM}$, where I_{KM} is the $KM \times KM$ identity matrix. In other words, we assume an isotropic Gaussian distribution over the landmarks. Note that we are using the same σ^2 notation for different underlying distribution, but it will be clear from the context which σ^2 is meant.

Mardia et al. [11] have introduced a *measure of asymmetry* (AS) of bilateral objects defined by

$$\mathrm{AS} = \sum_{(k_L,k_R)} \sum_{m=1}^{M} d[(k_L,k_R),m]^2 + \sum_{k_S} d[(k_S)]^2, \tag{7.2}$$

where in the paper AS is written as $\phi_{L_2}(\boldsymbol{d})$ and $\boldsymbol{d} \in \mathbb{R}^{MK_P+K_S}$ is based on X as follows. Let $d[(k_L,k_R),m]$ and $d[(k_S)]$ denote the components in $\boldsymbol{d}$ corresponding to paired landmarks and solos respectively then d's are given by the following in terms of X's

$$\begin{aligned} d[(k_L,k_R),1] &= X[k_L,1] + X[k_R,1] \\ d[(k_L,k_R),m] &= X[k_L,m] - X[k_R,m],\ m = 2,\dots,M \\ d[(k_S)] &= X[k_S,1]. \end{aligned} \tag{7.3}$$

We now give two simple examples defining X's and d's in $M = 2$ dimensions with $K_P = 2$ landmark pairs and $K_S = 0$ solo. The configuration matrices for the two examples are taken as

$$X_1 = \begin{pmatrix} X_1[1,\ ]^T \\ X_1[2,\ ]^T \\ X_1[3,\ ]^T \\ X_1[4,\ ]^T \end{pmatrix} = \begin{pmatrix} -4.33 & 0 \\ 4.33 & 0 \\ 4.33 & 3.84 \\ -4.33 & 3.84 \end{pmatrix},\quad X_2 = \begin{pmatrix} X_2[1,\ ]^T \\ X_2[2,\ ]^T \\ X_2[3,\ ]^T \\ X_2[4,\ ]^T \end{pmatrix} = \begin{pmatrix} -4.66 & 0 \\ 4.66 & 0 \\ 6.55 & 3.01 \\ -4.79 & 2.24 \end{pmatrix}. \tag{7.4}$$

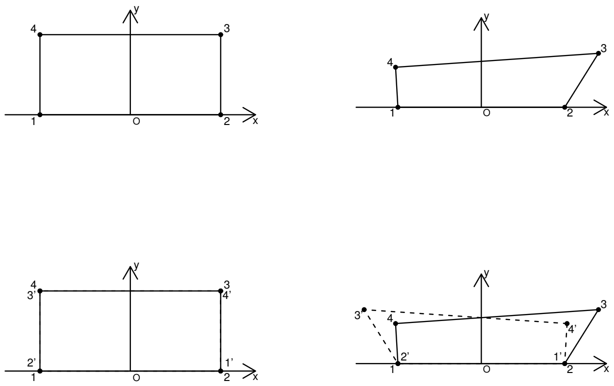

Fig. 3: The left and right figures in the top row show the original configurations of the symmetric rectangle X_1 and the asymmetric quadrilateral X_2 respectively. The left and right figures in the bottom row show the original objects (in solid lines) together with their respectively reflections $X_1^{(\mathrm{refl})}$ and $X_2^{(\mathrm{refl})}$ (in dotted lines) respectively, in order to illustrate the coordinate-wise elementary feature vector.

These configurations are displayed in the Figure 3 with their reflection; the landmark pairs are $(1,2)$ and $(3,4)$. The midline is the y-axis. The configurations X_1 and X_2 are shown in the left and right figures in the top row of Figure 3 respectively, whereas the bottom row of Figure 3 show their corresponding reflections $X_1^{(\mathrm{refl})}$ and $X_2^{(\mathrm{refl})}$, where the left is for X_1 and $X_1^{(\mathrm{refl})}$ and right is for X_2 and $X_2^{(\mathrm{refl})}$. Note that X_1 is bilaterally symmetric, with $\boldsymbol{d} = \boldsymbol{0}$. Table 1 gives the $MK_P + K_S = 4$ elements of the vector $\boldsymbol{d}$ for X_2. The two elements for the landmark pair (1, 2) are listed first, followed by the landmark pair (3, 4). Notice that the elements of $\boldsymbol{d}$ corresponding to landmarks 1 and 2 are 0, since these two landmarks are equally spaced on both sides of the midline y-axis. However, for the second configuration, the values

Table 1: Elements of the absolute elementary feature vector $\boldsymbol{d} \in \mathbb{R}^4$ for X_2 with $K = 4$ landmarks in $M = 2$ dimensions.

Landmark Indices	Coordinate Axis	Feature of $\boldsymbol{a}$	value of $\boldsymbol{d}$
Pair (1,2)	1	$d[(1,2),1] = X_2[1,1] + X_2[2,1]$	0
Pair (1,2)	2	$d[(1,2),2] = X_2[1,2] - X_2[2,2]$	0
Pair (3,4)	1	$d[(3,4),1] = X_2[3,1] + X_2[4,1]$	1.76
Pair (3,4)	2	$d[(3,4),2] = X_2[3,2] - X_2[4,2]$	0.77

of the d's are non-zero. Using the values in the tables, we can compute the value of AS defined in (7.2) as

$$\text{AS} = 1.76^2 + 0.77^2 = 3.69.$$

This value of AS quantifies the departure from symmetry. The codes for the symmetric and asymmetric examples are contained in the code chunks 2 and 3 respectively of the `R` script `Sankhya Chi`.

Since the d's given by (7.3) are linear functions of X's and the X's are normally distributed, hence, under the assumption (7.1), d's are also normally distributed. Therefore, (7.2) is a sum of two *quadratic functions* in normal variables, namely

$$AS_1 = \sum_{(k_L,k_R)} \sum_{m=1}^{M} d[(k_L,k_R),m]^2 \text{ and } AS_2 = \sum_{k_S} d[(k_S)]^2, \tag{7.5}$$

so we can write (7.2) as $\text{AS} = AS_1 + AS_2$, that is, we have

$$\text{AS} \sim \text{WS}((k_1,k_2)^T, (\lambda_1,\lambda_2)^T, (2\sigma^2,\sigma^2)^T) = 2\sigma^2\chi^2_{k_1}(\lambda_1) + \sigma^2\chi^2_{k_2}(\lambda_2), \tag{7.6}$$

where λ_1 and λ_2 are the two non-centrality parameters corresponding to landmark pairs and solos respectively:

$$\begin{aligned}\lambda_1 &= \sum_{(k_L,k_R)} \frac{(\boldsymbol{\mu}[k_L,1] + \boldsymbol{\mu}[k_R,1])^2 + (\boldsymbol{\mu}[k_L,2] - \boldsymbol{\mu}[k_R,2])^2 + \cdots + (\boldsymbol{\mu}[k_L,M] - \boldsymbol{\mu}[k_R,M])^2}{2\sigma^2} \\ \lambda_2 &= \sum_{k_S} \frac{(\boldsymbol{\mu}[k_S,1])^2}{\sigma^2}.\end{aligned} \tag{7.7}$$

Hence, it can be seen that (7.6) has the structure of our main distributional assumption given in (1.1) with $a = 2\sigma^2$, $b = \sigma^2$. Note that df for the first χ^2 distribution is $k_1 = MK_P$ and for the second one is $k_2 = K_S$. Further, we introduce a direct measure of asymmetry $\text{AS}_{\boldsymbol{\mu}}$ for the mean shape $\boldsymbol{\mu}$ as

$$\text{AS}_{\boldsymbol{\mu}} = 2\lambda_1\sigma^2 + \lambda_2\sigma^2. \tag{7.8}$$

For completeness, we note the following moments which can be deduced from (3.18):

$$\mathbb{E}\{\text{AS}\} = 2\sigma^2(k_1+\lambda_1) + \sigma^2(k_2+\lambda_2),\ \text{var}\{\text{AS}\} = 8\sigma^4(k_1+2\lambda_1) + 2\sigma^4(k_2+2\lambda_2), \tag{7.9}$$

The methods of moments for estimating λ_1 and λ_2 have already been given in Section 6.1, equation (6.1) (MoM estimates) and (6.4) (hybrid estimates). For hybrid estimates, we have $k_1 = MK_P$ and $k_2 = K_S$. The u's in (6.3) are the d's which are given in (7.3), and $\sigma_1^2 = 2\sigma^2$ and $\sigma_2^2 = \sigma^2$. The $\hat{\mu}$'s in (6.4) are obtained by plugging estimate $\hat{\boldsymbol{\mu}}$ of mean shape $\boldsymbol{\mu}$ in (7.7). The procedures for estimating σ^2 are given at the end of Section 6.1.

Suppose we have two groups (g_1 and g_2) of data and would like to test whether there are differences in asymmetry between the two groups. We note that the asymmetries are reflected through the λ_1 and λ_2, as can be seen in $\text{AS}_{\boldsymbol{\mu}}$ given by (7.8), so the test can be based on statistics corresponding to λ_1 and λ_2.

Table 2: Indices of landmark pairs and solos in Figure 4 on the lip.

Landmark notation	Indices
pair (k_L, k_R)	(1,13), (2,12), (3, 11), (4,10), (5,9), (6,8), (20,18), (21,17), (22,16), (23,15), (24,14)
k_S	7, 19

Let $\sigma^2_{g_i}$, $\lambda_1^{g_i}$ and $\lambda_2^{g_i}$ denote the variance of isotropic Gaussian distribution and two non-centrality parameters for group g_i, $i = 1, 2$. We assume that $\sigma^2_{g_i}$ can be estimated separately from $\lambda_1^{g_i}$ and $\lambda_2^{g_i}$ and can take $\sigma^2_{g_i}$ as known by plugging in its estimates. It is interested in testing whether significant differences exist among the non-centrality parameters of the two groups:

$$H_0 : \lambda_i^{g_1} = \lambda_i^{g_2} \text{ vs } H_1 : \lambda_i^{g_1} \neq \lambda_i^{g_2},\ i = 1, 2,$$

where $i = 1, 2$. We employ a simple test as the following: the test statistic is chosen simply as difference between λ_i for the two groups:

$$d_{\lambda_i} = |\hat{\lambda}_i^{g_1} - \hat{\lambda}_i^{g_2}|,\ i = 1, 2, \tag{7.10}$$

where $\hat{\lambda}_i^{g_1}$ is the estimate of $\lambda_i^{g_1}$ using equation (6.1) or (6.4), similarly for $\hat{\lambda}_i^{g_2}$.

In practice, permutation test is used to estimate the p-value corresponding to the test statistic given by (7.10). In each iteration, we resample from the pooled data of configuration matrices for the two groups, then we compute moment estimates $\hat{\lambda}_{i,b}^{g_1}$ and $\hat{\lambda}_{i,b}^{g_2}$ using (6.1) or (6.4), where $b = 1, \ldots, B$ is the index for iteration. We use $B = 10000$. The p-value is estimated by comparing the observed test statistic with the quantile of test statistic computed on resampled data. We will see below that for this case, there is only one non-centrality parameter of interest. Under this assumption, it is simpler to use this particular test statistics rather than using the likelihood-ratio test.

7.2 Application to the Smile Data

Mardia et al. [11] have introduced the smile data arising from the problem of measuring asymmetry after the cleft lip surgery. The data consists of three time frames and for illustrative purpose, we use here the first frame of the data. The details of the data are given in the paper but the point to note is the data has been pre-registered so that the effect of translation, scale and rotation has been removed, so the configuration matrix is an Euclidean random matrix.

Our main selected data is the cleft lip data which is 3-dimensional ($M = 3$) with sample size = 13 and there are 24 landmarks ($K = 24$). For simplicity, we refer to it as Data1. We have also used similar data for control subjects with sample size = 12 and refer to it as control data or Data2. Figure 4 shows the location of the landmarks in 2-dimension over the lip. Among these 24 landmarks, there are $K_P = 11$ landmark pairs and $K_S = 2$ solo landmarks. Table 2 shows the indices for landmark pairs and solos from Figure 4. The scale of the data is in mm.

Hence, the df for the two χ^2-distributions in weighted sum AS given in (7.6) are $k_1 = 33$ and $k_2 = 2$ respectively and here

$$\text{AS} \sim \text{WS}((33, 2)^T, (\lambda_1, \lambda_2)^T, (2\sigma^2, \sigma^2)^T) = 2\sigma^2\chi^2_{33}(\lambda_1) + \sigma^2\chi^2_2(\lambda_2). \tag{7.11}$$

In Section 7.2.1 below we first use the inference under normality, including estimation and testing, and in Section 7.2.2 the ex-χ^2 and ex-Gaussian models are fitted. We then give the comparison.

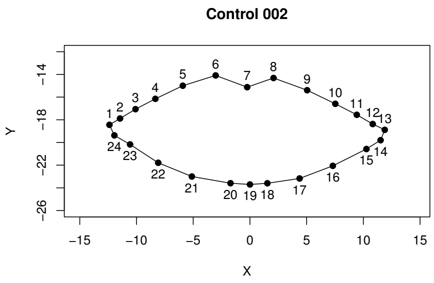

Fig. 4: Lip configuration with 24 landmarks used in the smile data.

Table 3: Sample mean and variance of AS of subjects from the smile data, together with the moment estimates of λ_1 and λ_2 by using (6.1). The first row is for Data1 and the second row is for Data2.

	$\bar{z}$	s^2	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Cleft	45.47	483.83	-0.29	-65.64
Control	19.89	211.19	-0.44	-63.26

7.2.1 Inference under Matrix Normality

In this section, we will use the random matrix normality of the configurations when needed, though our main focus will be directly on the measure of asymmetry AS.

Estimation: Let $z_1, \cdots, z_n$ be the observed AS realizations of cleft or control subjects, where AS is defined in (7.2) and $n = 13$ for cleft and $n = 12$ for control. Two different estimators given in (6.1) and (6.4) respectively, where (6.1) gives the MoM estimators and (6.4) gives the hybrid estimators, are used to estimate λ_1 and λ_2. However, σ^2 is estimated using the same method given at the end of Section 6.1.1.

We first use the MoM equations given in (6.1). The values of the moment estimates $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are shown in Table 3, together with sample mean $\bar{z}$ and sample variance s^2. The MoM estimates of λ_1 and λ_2 given in Table 3 are both negative, though a priori non-centrality parameters are non-negative. Thus, these estimates are inadequate for this data.

Then, we use the hybrid moment estimators given in (6.4) to estimate λ_1 and λ_2. Further, the estimates $\hat{\boldsymbol{\mu}}$ is obtained simply by taking the arithmetic mean of sample configuration matrices from each group, as our data has been pre-registered. We find that the hybrid moment estimates $\hat{\lambda}_2$ are negligible in both cases relative to $\hat{\lambda}_1$, thus we set $\lambda_2 = 0$. Hence, the random variable AS given in (7.6) is now simply

$$\text{AS} \sim \text{WS}((33, 2)^T, (\lambda_1, 0)^T, (2\sigma^2, \sigma^2)^T) = 2\sigma^2\chi^2_{33}(\lambda_1) + \sigma^2\chi^2_2.$$

In this case, the moments from (3.18) become:

$$\mathbb{E}\{\text{AS}\} = 2\sigma^2(k_1 + \lambda_1) + \sigma^2 k_2, \ \text{var}\{\text{AS}\} = 8\sigma^4(k_1 + 2\lambda_1) + 2\sigma^4 k_2. \tag{7.12}$$

The estimates of λ_1 and σ^2 are given in Table 4. We do not report the likelihood values since the likelihood function is not involved in these hybrid estimates. The estimate $\hat{\sigma}^2$ for cleft lip data is larger than control data, which indicates that there is more variation in the cleft lip data. Further, recall that AS is a measure of asymmetry and $\mathbb{E}\{\text{AS}\}$ and var$\{\text{AS}\}$ are larger for cleft group from Table 4, indicating that cleft lip has more asymmetry in mean as well as there is more variability, which matches the medical opinion.

Table 4: Estimations of $\hat{\lambda}_1$ and $\hat{\sigma}^2$ for the smile data (Data1: cleft lip, Data2: control) using hybrid estimators given in (6.4), MLEs for ex-χ^2 model and ex-Gaussian model, $k_1 = 33$ and $k_2 = 2$ in (7.6).

	$\hat{\lambda}_1$	$\hat{\sigma}^2$	$\mathbb{E}\{\text{AS}\}$	var{AS}	log-likelihood	$\widehat{\text{AS}}_{\mu}$
Hybrid Moment Estimator:						
Cleft	0.034	28.03	1907.95	210990	NA	1.91
Control	0.94	5.15	359.88	7506.93	NA	9.68
ex-χ^2:						
Cleft	0.58	0.66	45.65	120.78	-72.21	0.77
Control	0.23	0.28	19.17	21.30	-74.58	0.13
ex-Gaussian:						
Cleft	0.48	0.75	51.72	155.07	-74.03	0.72
Control	0.51	0.35	24.16	33.83	-75.58	0.36

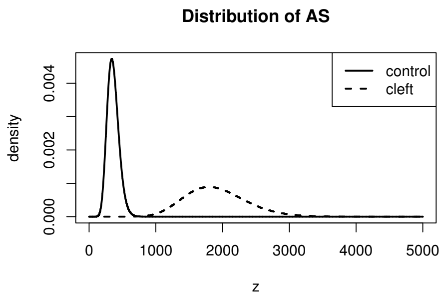

Fig. 5: Plot of pdf for AS for cleft and control subjects from the cleft lip data. The solid line shows the pdf for control data while the dashed line is for cleft data.

By plugging in the estimates of λ_1 and σ^2 in equation (7.12), the mean and variance of AS are computed and reported in Table 4. We also give the density plots of AS in both cases. To plot these, we have used the `convpow` function in `R` to compute the convoluted pdf of AS given by (7.6). Figure 5 shows the plot for pdf of AS for both cleft and control subjects. From the figure, it can be seen visually that their means are well-separated and there is more variation for cleft lip data. It also indicates that both distributions are approximately normal.

Two-sample test for asymmetry: We now test for differences between the asymmetry measures for the cleft lip data (Data1) and control data (Data2) using the method described in Section 7.1. Since the parameter λ_2 is taken to be 0 as described above, the parameter λ_1 is of interested so our test statistics is simply a single statistics given by

$$d_{\lambda_1} = |\hat{\lambda}_1^{g_1} - \hat{\lambda}_1^{g_2}| \tag{7.13}$$

where the superscript g_1 is for Data1 and g_2 is for Data2. Our $\hat{\lambda}_1^{g_1}$ and $\hat{\lambda}_1^{g_2}$ are estimated using (6.4) by first estimating σ^2 for each data separately from λ_1 and then plugging in the estimates $\hat{\sigma}^2$ into (6.4) for σ^2 for each case to compute estimates $\hat{\lambda}_1$ of λ_1. It is found that in our case $d_{\lambda_1} = 0.90$. Using the permutation test, the p-value is found to be 0.001. Hence, we conclude that there are significant differences between cleft lip data and control data based on this statistics, which matches medical opinion.

The codes for computing MoM estimators given in (6.1) and the hybrid estimators given in (6.4) are contained in the code chunk 3 in the `R` script `Sankhya Chi Inf`, together with carrying out the permutation test. The codes for computing the AS measures for cleft and control subjects are in code chunk 2 of the script.

7.2.2 Inference under Ex-χ^2 and Ex-Gaussian Models

We now give the MLE for ex-χ^2 and ex-Gaussian models using only observed values of AS. In the following, again, we will take $\lambda_2 = 0$ so there are only two parameters to be estimated: λ_1 and σ^2.

Ex-χ^2 Model: We assume the following model from equation (4.2)

$$AS \sim \text{ex-}\chi^2\left(k_1, \lambda_1, \frac{1}{2}, (2\sigma^2, \sigma^2)^T\right) = 2\sigma^2\chi^2_{k_1}(\lambda_1) + \sigma^2 Exp\left(\frac{1}{2}\right) = 2\sigma^2\chi^2_{k_1}(\lambda_1) + \sigma^2\chi^2_2,$$

where $k_1 = 33$, and obtain the MLE of parameters λ_1 and σ^2. As mentioned in Section 6.1.2, the MLE solutions can be obtained via SA using the pdf of AS given in Theorem 4.2. The function `convpow` in `R` is used to evaluate the likelihood. We use the proposal densities for cleft as $U(0, 2)$ for λ_1 and $U(0, 10)$ for σ^2 respectively, where $U(a, b)$ is the uniform distribution on interval (a, b). On the other hand, for control, we use $U(0, 1)$ for λ_1 and $U(0, 2)$ for σ^2 respectively. A chain of length 100000 is run for both cleft and control. The MLEs of λ_1 and σ^2 so obtained are given in Table 4, together with the values of the log-likelihood. The conclusion is similar as the hybrid estimates: the mean for cleft subjects is larger than control and there is more variation for cleft lip data. Moreover, for the direct asymmetry measure of the mean shape $\boldsymbol{\mu}$ given in (7.8), its estimate is larger for cleft which matches the medical opinion.

Ex-Gaussian Model: We assume the following model from equation (5.1)

$$\text{AS} \sim \text{exG}\left(k_1 + \lambda_1, 2(k_1 + 2\lambda_1), \frac{1}{2}, (2\sigma^2, \sigma^2)^T\right) = 2\sigma^2 N(k_1 + \lambda_1, 2(k_1 + 2\lambda_1)) + \sigma^2 Exp\left(\frac{1}{2}\right), \quad (7.14)$$

where $k_1 = 33$, and obtain the MLE for parameters λ_1 and σ^2. As mentioned in Section 6.1.2, the MLE solutions can be obtained via the SA algorithm using the pdf of AS given in (5.7). We use the proposals for cleft and control subjects as $U(0, 1)$ for λ_1 and $\Gamma(5, 1)$ for σ^2 respectively, where $\Gamma(\alpha, \beta)$ is the gamma distribution and $\alpha > 0$ is the shape parameter, $\beta > 0$ is the rate parameter. A chain of length 100000 is run for both cleft and control. The MLEs of λ_1 and σ^2 so obtained are given in Table 4, together with the values of the log-likelihood. The conclusion is similar as the hybrid estimates and MLE for ex-χ^2 model: the mean for cleft subjects is larger than control and there is more variation for cleft. The direct asymmetry measure of the mean shape $\boldsymbol{\mu}$ given in (7.8) has larger value for cleft subjects as well.

Summary: The parameter estimates obtained from all three approaches indicate that the distributions of AS are well-separated for cleft and control subjects. Further, the distribution of AS for cleft has larger mean and variation.

The MLEs of variance of isotropic Gaussian distribution, σ^2, obtained from ex-χ^2 and ex-Gaussian models are smaller than that of hybrid estimators. Note that σ^2 estimated directly from the original configurations in the hybrid estimator requires the original landmark data which seems to lead to higher variance in the $\hat{\sigma}^2$, which may be as we have a small dataset here. In Table 4, we have also given the estimates $\widehat{\text{AS}}_{\boldsymbol{\mu}} = 2\hat{\lambda}_1\hat{\sigma}^2$ of the mean shape $\boldsymbol{\mu}$ given by (7.8), using apriori $\lambda_2 = 0$. In the case of hybrid estimates, it indicates that the control is more asymmetric which is contrary to the medical opinion. However, under the assumption of ex-χ^2 or ex-Gaussian, there is a clear evidence from $\widehat{\text{AS}}_{\boldsymbol{\mu}}$ that the cleft is more asymmetric than the control which is consistent with the medical opinion. Further work is required to study the relative performance of these three estimation procedures. Here, these are all used simply for illustrative purpose.

The codes for computing the density function of ex-χ^2 distribution and performing the simulated annealing method to estimate the MLEs for ex-χ^2 and ex-Gaussian distributions are given in the code chunks 4 and 5 respectively in the `R` script `Sankhya Chi Inf`.

8 Discussion

We have focused here for the weighted sum of two χ^2-distributions. Our aim has been to give deeper insight into the distribution of the weighted sum and consider particular cases and provide the procedure to carry out inference in practice. Our approach can be extended to any general linear combination of χ^2-distributions; however, that will need some further research.

It is interesting to note that the density function given by (5.4) of ex-Gaussian distribution has a similar structure as Azzalini's skew distribution (Azzalini [18]), which has the pdf given by

$$f_Z(z;\alpha) = 2\varphi(z)\Phi(\alpha z), \tag{8.1}$$

where again $\varphi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of standard normal distribution respectively. In our (5.4), the density component in the product is proportional to the exponential distribution, but the second component in both case is normal. Future work will be involved in comparing the three skew distributions: Azzalini's skew distribution, ex-χ^2 and ex-Gaussian. Azzalini's distribution is well-established in contrast and therefore it would give a deeper insight into the other two distributions.

We have given the general form of ex-χ^2 and ex-Gaussian distributions in (4.2) and (5.1) respectively. In ex-χ^2, there are four parameters, a, b, λ and c, and in ex-Gaussian, there are five parameters, a, b, μ, σ^2 and c. For application, one should be aware there could be non-identifiability problems in the parameter space.

In our shape application, the underlying random variable is in fact a random matrix, on which samples are drawn. We have indicated how one can use it for inference by using a hybrid moment estimator approach. Future work is required to assess its performance in comparison to the MLE method.

We have assumed isotropic normal distribution for the application but future work will involve the non-isotropic normal distribution as a starting point, so that we will have the weighted sum of two general quadratic forms in random variable Z. Further, we have only considered the sum of two weighted χ^2 and there is a potential to work in the similar way for the sum of two general Wishart distributions (Mardia et al. [19]).

We have mentioned the first two moments of the WS distribution at (3.18) and have used third moments for ex-Gaussian. Future work will involve writing down the higher moments of WS and studying their behavior for general method of moments for weighted sums of χ^2-distributions. We have focused on the linear function of a and b as required by our illustrative example but, in particular, this future work will explore the moment estimators when both a and b are also unknown in the weighted sum.

The computer program used in this work has been deposited at GitHub with link: https://github.com/XW-2025-hub/Sankhya-paper, with read-me files.

Acknowledgments

We are grateful to all other members from the smile project team: Colin R. Goodall, John T. Kent, Balvinder S. Khambay and Sofya Titarenko for their helpful comments.

References

[1] McKay, A.T.: A Bessel function distribution. *Biometrika* **24**(1/2), 39–44 (1932) https://doi.org/10.2307/2333794

[2] Press, S.J.: Linear combinations of non-central chi-square variates. *The Annals of Mathematical Statistics* **37**(2), 480–487 (1966)

[3] Bausch, J.: On the efficient calculation of a linear combination of chi-square random variables with an application in counting string vacua. *Journal of Physics A: Mathematical and Theoretical* **46**(50) (2013) https://doi.org/10.1088/1751-8113/46/50/505202

[4] Grushka, E.: Characterization of exponentially modified Gaussian peaks in chromatography. *Analytical Chemistry* **44**(11), 1733–1738 (1972) https://doi.org/10.1021/ac60319a011

[5] Kumbhakar, S.C., Lovell, C.A.K.: *Stochastic Frontier Analysis*. Cambridge University Press, Cambridge (2000). https://doi.org/10.1017/CBO9781139174411

[6] Bock, M.T., Bowman, A.W.: On the measurement and analysis of asymmetry with applications to facial modelling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **55**(1), 77–91 (2006) https://doi.org/10.1111/j.1467-9876.2005.00532.x

[7] Patel, Y., Sharp, I., Enocson, L., Khambay, B.S.: An innovative analysis of nasolabial dynamics of surgically managed adult patients with unilateral cleft lip and palate using 3D facial motion capture. *Journal of Plastic, Reconstructive and Aesthetic surgery* **85**, 287–298 (2023) https://doi.org/10.1016/j.bjps.2023.07.007

[8] Ajmera, D.H., Hsung, R.T., Singh, P., Wong, N.S.M., Yeung, A.W.K., Lam, W.Y.H., Khambay, B.S., Leung, Y.Y., Gu, M.: Three-dimensional assessment of facial asymmetry in class III subjects. Part 1: a retrospective study evaluating postsurgical outcomes. *Clinical Oral Investigations* **26**, 4947–4966 (2022) https://doi.org/10.1007/s00784-022-04463-4

[9] Ajmera, D.H., Zhang, C., Ng, J.H.H., Hsung, R.T., Lam, W.Y.H., Wang, W., Leung, Y.Y., Khambay, B.S., Gu, M.: Three-dimensional assessment of facial asymmetry in class III subjects. Part 2: evaluating asymmetry index and asymmetry scores. *Clinical Oral Investigations* **27**, 5813–5826 (2023) https://doi.org/10.1007/s00784-023-05193-x

[10] Dryden, I.L., Mardia, K.V.: *Statistical Shape Analysis: With Applications in R*. Wiley, Chichester (2016). https://doi.org/10.1002/9781119072492

[11] Mardia, K.V., Wu, X., Kent, J.T., Goodall, C.R., Khambay, B.S.: Asymmetry analysis of bilateral shapes (2024) https://doi.org/10.48550/arXiv.2407.17225

[12] Abramowitz, M., Stegun, I.: *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. United States Department of Commerce, National Bureau of Standards (NBS), New York (1964). https://doi.org/10.1119/1.15378

[13] Gradshteyn, I.S., Ryzhik, I.M.: *Table of Integrals, Series, and Products*. Elsevier Academic Press, Boston (1980). https://doi.org/10.1016/C2010-0-64839-5

[14] Carr, P.P., Madan, D., Smith, R.: Saddle point methods for option pricing. *The Journal of Computational Finance* **13**, 49–61 (2009) https://doi.org/10.21314/JCF.2009.198

[15] Dyson, N.: *Chromatographic Integration Methods*. Royal Society of Chemistry, London (1998). https://doi.org/10.1039/9781847550514

[16] Brown, S., Heathcote, A.: QMLE: Fast, robust, and efficient estimation of distribution functions based on quantiles. *Behavior research methods, Instruments & Computers : a journal of the Psychonomic Society, Inc* **35**, 485–92 (2003) https://doi.org/10.3758/bf03195527

[17] Mardia, K.V., Bookstein, F.L., Moreton, I.J.: Statistical assessment of bilateral symmetry of shapes. Biometrika **87**(2), 285–300 (2000) https://doi.org/10.1093/biomet/87.2.285

[18] Azzalini, A.: *The Skew-Normal and Related Families*. Institute of Mathematical Statistics Monographs. Cambridge University Press, Cambridge (2013). https://doi.org/10.1017/CBO9781139248891

[19] Mardia, K.V., Kent, J.T., Taylor, C.C.: *Multivariate Analysis, Second Edition*. John Wiley, Chichester (2024)