The Exponential-Centred Skew-Normal Distribution

Guillermo Martínez-Flórez [1,2], Carlos Barrera-Causil [3] and Fernando Marmolejo-Ramos [4,*]

[1] Departamento de Matemáticas y Estadística, Facultad de Ciencias Básicas, Universidad de Córdoba, Córdoba 2300, Colombia; guillermomartinez@correo.unicordoba.edu.co

[2] Programa de Pós-Graduação em Modelagem e Métodos Quantitativos, Universidade Federal do Ceará, Fortaleza 60020-181, Brazil

[3] Davinci Research Group, Faculty of Applied and Exact Sciences, Metropolitan Technological Institute, Medellín 050034, Colombia; carlosbarrera@itm.edu.co

[4] Centre for Change and Complexity in Learning, The University of South Australia, Adelaide 5000, Australia

[*] Correspondence: fernando.marmolejo-ramos@unisa.edu.au

check for updates

Abstract: Data from some research fields tend to exhibit a positive skew. For example, in experimental psychology, reaction times (RTs) are characterised as being positively skewed. However, it is not unlikely that RTs can take a normal or, even, a negative shape. While the Ex-Gaussian distribution is suitable to model positively skewed data, it cannot cope with negatively skewed data. This manuscript proposes a distribution that can deal with both negative and positive skews: the exponential-centred skew-normal (ECSN) distribution. The mathematical properties of the proposed distribution are reported, and it is featured in two non-synthetic datasets.

Keywords: Ex-Gaussian distribution; skew-normal distribution; moments; maximum likelihood estimation; score function; observed information matrix; reaction times

1. Introduction

Hohle [1] used the convolution between two random variables, exponential and Gaussian, to model simple reaction times (RTs) elicited via psychological experiments. The distributions resulting from such convolutions are known as exponential-Gaussian distributions or, simply, Ex-Gaussian (EG) distributions. Other distributions, though, have been proposed to model RT data. For example, it has been shown that the shifted Wald, shifted Birnbaum–Saunders [2] and shifted Gamma distributions [3] provide good fits to RT data (the Birnbaum–Saunders distribution has also been used to model neural activity [4], and a flexible version of it was proposed recently [5]. The slashed quasi-Gamma distribution (an extension of the generalised Gamma distribution) has been proposed to model data with varying skewness and high kurtosis [6]. To our knowledge, neither the flexible Birnbaum–Saunders nor the slashed quasi-Gamma distributions have been tested in the modelling of RT data; hence, this is a task for future research). Nonetheless, distributions of RTs resulting from psychological experiments are commonly adjusted with the EG distribution model.

EG distributions were first introduced in an RT analysis carried out by McGill [7], who, when studying the RT cumulative distribution function as a result of different intensities of auditory stimuli, found that they had an exponential form with similar constant times. This led to the assumption that, at least, one component of the total RT had an exponential distribution, and since the time constants that the curves implied seemed to be almost independent of stimulus intensity, McGill [7] hypothesised that this component was the necessary time for the required motor response. Moreover, as the resulting distributions changed according to stimulus intensity, RTs are assumed to contain a non-exponential component that represents the decision time.

As part of a model designed to represent the disjunctive structure of RTs, Christie and Luce [8] assumed that the observed response time consisted of an exponentially distributed decision time plus a variable time or motor response time whose distribution was not specified. Hence, these two studies, by McGill [7] and Christie and Luce [8], offer opposite interpretations of the component with the exponential distribution.

Hohle [1] proposed a very simple hypothesis: the observed RT results from the sum or convolution of two independent random variables: (a) an exponentially distributed dominant variable with parameter τ, as proposed by McGill [7] and Christie and Luce [8], and (b) another variable with a normal distribution, $N(\mu, \sigma^2)$, which may be the sum of a number of component variables with similar variations.

Therefore, under this assumption, for $Y_1 \sim EXP(\tau)$ and $Y_2 \sim N(\mu, \sigma^2)$, where Y_1 and Y_2 are independent, it follows that the probability density function (PDF) of the random variable $X = Y_1 + Y_2$ resulting from the convolution of Y_1 and Y_2 is given by:

$$f_{EG}(x) = \frac{1}{\tau} e^{-\frac{x-\mu}{\tau} + \frac{\sigma^2}{2\tau}} \Phi\left(\frac{x-\mu}{\sigma} - \frac{\sigma}{\tau}\right), \quad \tau > 0, \quad -\infty < \mu < \infty, \quad \sigma > 0. \tag{1}$$

This will be denoted by $EG(\tau, \mu, \sigma^2)$. The cumulative distribution $F_{EG}(\cdot)$, survival $S_{EG}(\cdot)$, relative risk $r_{EG}(\cdot)$ and hazard functions $h_{EG}(\cdot)$ are given by:

$$F_{EG}(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) - \tau f_{EG}(x), \quad S_{EG}(t) = S_N(t) + \tau f_{EG}(t), \tag{2}$$

$$r_{EG}(t) = \frac{e^{-\frac{x-\mu}{\tau} + \frac{\sigma^2}{2\tau}} \Phi\left(\frac{x-\mu}{\sigma} - \frac{\sigma}{\tau}\right)}{\tau \left(\Phi\left(\frac{x-\mu}{\sigma}\right) + e^{-\frac{x-\mu}{\tau} + \frac{\sigma^2}{2\tau}} \Phi\left(\frac{x-\mu}{\sigma} - \frac{\sigma}{\tau}\right)\right)}, \quad h_{EG}(t) = \frac{e^{-\frac{x-\mu}{\tau} + \frac{\sigma^2}{2\tau}} \Phi\left(\frac{x-\mu}{\sigma} - \frac{\sigma}{\tau}\right)}{\tau \left(S_N(t) + e^{-\frac{x-\mu}{\tau} + \frac{\sigma^2}{2\tau}} \Phi\left(\frac{x-\mu}{\sigma} - \frac{\sigma}{\tau}\right)\right)}. \tag{3}$$

where $S_N(t)$ is the survival function of the normal distribution. Additionally,

$$\mathbb{E}(X) = \tau + \mu \quad \text{and} \quad Var(X) = \tau^2 + \sigma^2.$$

and the asymmetry and kurtosis coefficients are given by:

$$skew = \frac{2\left(\frac{\tau}{\sigma}\right)^3}{\left[1 + \left(\frac{\tau}{\sigma}\right)^2\right]^{3/2}} \quad \text{and} \quad kurt = \frac{3\left(1 + 2\left(\frac{\tau}{\sigma}\right)^2 + 3\left(\frac{\tau}{\sigma}\right)^4\right)}{\left[1 + \left(\frac{\tau}{\sigma}\right)^2\right]^2}.$$

It can be then concluded that the EG model has positive asymmetry. Furthermore, when $\tau \to 0$, then $\mathbb{E}(X) \to \mu$, $Var(X) \to \sigma^2$, $skew \to 0$, and $kurt \to 3$, which entails that:

$$EG(\tau, \mu, \sigma^2) \to N(\mu, \sigma^2).$$

Although the EG distribution can be useful in modelling positively skewed data (e.g., RTs), in practice, it is not unlikely that data take other shapes (e.g., normal or negatively skewed). Hence, having more flexible distributions is key to model data appropriately. This article proposes a modified EG distribution in which the Gaussian component is replaced by a skew-normal (SN) distribution. Doing so gives origin to the exponential skew-normal distribution; a distribution capable of adapting to various data shapes. First, the statistical properties of this distribution are presented. Given singularity issues associated with the skew-normal distribution, a centred version of this distribution is then described. Next, the performance of the proposed distribution is illustrated via two published datasets. Finally, this distribution is considered in the bigger picture of data analyses and modelling.

2. Exponential Skew-Normal Distribution

The EG model exhibits positive asymmetry, because the exponential model is positive asymmetric since the second component of the convolution is symmetric around μ. In his proposal, Hohle [1] assumed that the second component of the convolution, from which the EG model results, follows a normal distribution that is located at μ and has a constant variation that represents the sum of a number of variables with similar variations that represents the necessary time for the required motor response, according to McGill's hypothesis [7], or the variable motor response time, according to Christie and Luce's [8].

It is assumed herein that the variable motor response time, or the necessary time for the motor response, follows a skew-normal distribution (see [9]) whose PDF is given by:

$$\varphi(z; \lambda) = 2\phi(z)\Phi(\lambda z), \; z \in \mathbb{R}, \tag{4}$$

where ϕ and Φ denote the standard normal density and cumulative distribution functions, respectively. The asymmetry is controlled by the parameter λ, and the notation $Z \sim SN(\lambda)$ is used herein; for $\lambda = 0$, it is the normal model. The asymmetry and kurtosis ranges of this distribution are $(-0.9953, 0.9953)$ and $[3, 3.8692)$, respectively (these values are extracted from [10]). The PDF of the location-scale model can be written as:

$$\varphi(y; \lambda) = \frac{2}{\eta}\phi\left(\frac{y - \xi}{\eta}\right)\Phi\left(\lambda\frac{y - \xi}{\eta}\right), \; y \in \mathbb{R}, \tag{5}$$

with $\xi \in \mathbb{R}$ and $\eta > 0$ as the location-scale parameters. The notation $SN(\xi, \eta, \lambda)$ is used herein.

Then, for $Y_1 \sim \exp(\tau)$ and $Y_2 \sim SN(\mu, \sigma, \alpha)$, where Y_1 and Y_2 are independent, the PDF of the random variable $X = Y_1 + Y_2$ resulting from the convolution of Y_1 and Y_2 is given by:

$$
\begin{aligned}
f_{ESN}(x) &= \int_{-\infty}^{\infty} f(z)f(z - x)dz, \quad z - x > 0 \\
&= \int_{-\infty}^{x} \frac{2}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{z - \mu}{\sigma}\right)^2}\Phi\left(\alpha\frac{z - \mu}{\sigma}\right)\frac{1}{\tau}e^{-\frac{z - x}{\tau}}dz \\
&= \frac{1}{\tau}e^{-\frac{x - \mu}{\tau} + \frac{\sigma^2}{2\tau^2}}\int_{-\infty}^{x}\frac{2}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{z - \mu}{\sigma} - \frac{\sigma}{\tau}\right)^2}\Phi\left(\alpha\frac{z - \mu}{\sigma}\right)dz \\
&= \frac{1}{\tau}e^{-\frac{x - \mu}{\tau} + \frac{\sigma^2}{2\tau^2}}2\int_{-\infty}^{\frac{x - \mu}{\sigma} - \frac{\sigma}{\tau}}\int_{-\infty}^{\alpha\frac{x - \mu}{\sigma}}\phi(u)\phi(y)dudy \\
&= \frac{2}{\tau}e^{-\frac{x - \mu}{\tau} + \frac{\sigma^2}{2\tau^2}}\int_{-\infty}^{\frac{x - \mu}{\sigma} - \frac{\sigma}{\tau}}\int_{-\infty}^{\frac{\sigma\delta}{\tau}}\frac{1}{\sqrt{1 - \delta^2}}\phi(t_1)\phi\left(\frac{t_2 + \delta t_1}{\sqrt{1 - \delta^2}}\right)dt_2 dt_1 \\
&= \frac{2}{\tau}e^{-\frac{x - \mu}{\tau} + \frac{\sigma^2}{2\tau^2}}\Phi_B\left(\frac{x - \mu}{\sigma} - \frac{\sigma}{\tau}, \frac{\sigma\delta}{\tau}; -\delta\right)
\end{aligned}
\tag{6}
$$

where $\delta = \frac{\alpha}{\sqrt{1 + \alpha^2}}$ and:

$$\Phi_B(h, k; \rho) = \frac{1}{2\pi\sqrt{1 - \delta^2}}\int_{-\infty}^{k}\int_{-\infty}^{h}e^{-\frac{1}{2(1 - \delta^2)}\left(u^2 + v^2 - 2\rho uv\right)}dudv.$$

This distribution is denoted as $ESN(\tau, \mu, \sigma, \alpha)$. It can be observed that, if $\alpha = 0$, then:

$$\Phi_B\left(\frac{x - \mu}{\sigma} - \frac{\sigma}{\tau}, \frac{\sigma\delta}{\tau}; -\delta\right) = \Phi\left(\frac{x - \mu}{\sigma} - \frac{\sigma}{\tau}\right),$$

that is, the EG distribution is obtained. Therefore, the exponential skew-normal (ESN) model contains the EG model as a special case, which entails that the ESN distribution is more flexible in terms of asymmetry than the EG model.

The ESN convolution model can be written in terms of the cumulative distribution function of the extended skew-normal model (here E; see [10], p. 40) and setting $\lambda_0 = \frac{\delta\sigma}{\tau}$:

$$f_{ESN}(x) = \frac{2}{\tau}e^{-\frac{x-\mu}{\tau}+\frac{\sigma^2}{2\tau^2}}\Phi(\lambda_0)\Phi_E(x;\lambda_0,\delta) \tag{7}$$

where $\Phi_E(x;\lambda_0,\delta)$ is given by

$$\Phi_E(x;\lambda_0,\delta) = \frac{\Phi_B\left(\frac{x-\mu}{\sigma}-\frac{\sigma}{\tau},\lambda_0;-\delta\right)}{\Phi(\lambda_0)}.$$

Figure 1 shows the ESN's density behaviour for certain parameter values; (a) and (b) present a comparison between the ESN and EG distributions (the EG ensues when ESN has $\alpha = 0$); (c) illustrates the case of negatively skewed ESN distributions.
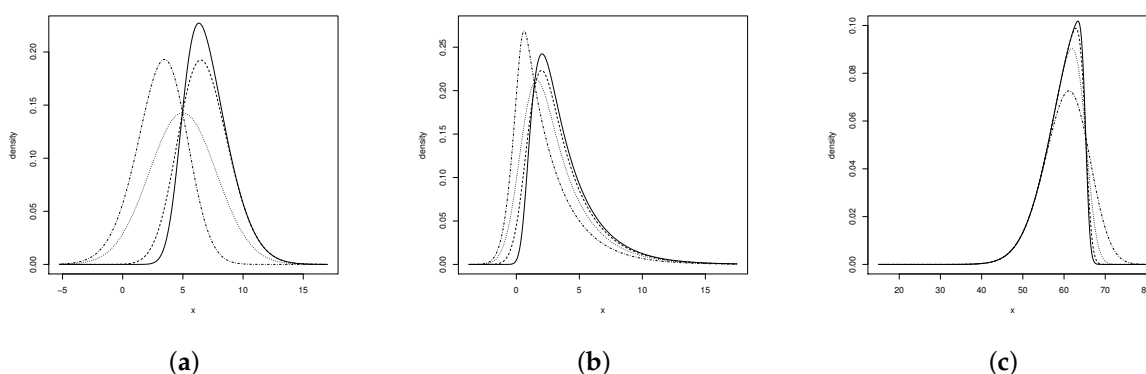

(a)   (b)   (c)

Figure 1. ESN distribution of (**a**) $\tau = 0.4$, $\mu = 4.5$, $\sigma = 2.75$, and $\alpha = -1.5$ (dash-dotted line); $\alpha = 0$ (dotted line); $\alpha = 1.5$ (dashed line); and $\alpha = 3.0$ (solid line); (**b**) $\tau = 2.5$, $\mu = 0.5$, $\sigma = 1.0$, and $\alpha = -3.5$ (dash-dotted line); $\alpha = 0$ (dotted line); $\alpha = 0.75$ (dashed line); and $\alpha = 2.5$ (solid line); (**c**) $\tau = 0.28$, $\mu = 65$, $\sigma = 7.5$, and $\alpha = -1.5$ (dash-dotted line); $\alpha = -3.5$ (dotted line); $\alpha = -6.5$ (dashed line); and $\alpha = -9.5$ (solid line).

The cumulative distribution function (CDF) of a random variable with an ESN distribution does not have a closed-form expression and should be calculated using a numerical approximation algorithm. It is given by:

$$F_{ESN}(x) = \frac{2}{\tau}e^{\frac{\sigma^2}{2\tau^2}}A(x) \tag{8}$$

where:

$$A(x) = \int_{-\infty}^{x}e^{\frac{t-\mu}{\sigma}}\Phi_B\left(\frac{t-\mu}{\sigma}-\frac{\sigma}{\tau},\frac{\sigma\delta}{\tau};-\delta\right)dt.$$

Moreover, the survival, hazard, and inverse risk functions are given by:

$$S_{ESN}(t) = 1 - \frac{2}{\tau}e^{\frac{\sigma^2}{2\tau^2}}A(t), \quad r(t) = \frac{e^{\frac{t-\mu}{\sigma}}\Phi_B\left(\frac{t-\mu}{\sigma}-\frac{\sigma}{\tau},\frac{\sigma\delta}{\tau};-\delta\right)}{\frac{\tau}{2}e^{-\frac{\sigma^2}{2\tau^2}}+A(t)}, \quad h(t) = \frac{e^{\frac{t-\mu}{\sigma}}\Phi_B\left(\frac{t-\mu}{\sigma}-\frac{\sigma}{\tau},\frac{\sigma\delta}{\tau};-\delta\right)}{A(t)}.$$

2.1. Moments

The moments of the ESN distribution can be obtained directly from the stochastic representation of the random variable $X = X_1 + X_2 \sim ESN(\tau, \mu, \sigma, \alpha)$, where $X_1 \sim Exp(\tau)$ and $X_2 \sim SN(\mu, \sigma, \alpha)$ are independent random variables. Thus, the r$^{\text{th}}$ moment of the random variable $X = X_1 + X_2$ is given by:

$$\mathbb{E}(X^r) = (X_1 + X_2)^r = \sum_{j=0}^{r} \binom{r}{j} \mathbb{E}(X_1^{r-j} X_2^j) = \sum_{j=0}^{r} \binom{r}{j} \mathbb{E}(X_1^{r-j}) \mathbb{E}(X_2^j). \tag{9}$$

Given that $\mathbb{E}(X_1^{r-j}) = (r-j)!$, the even moments of the SN distribution match those of the normal distribution and by expressing the odd moments of the SN distribution (see [11]), it results that:

$$\mathbb{E}(X^r) = \sum_{j=0}^{r} \sum_{l=0}^{j} \binom{r}{j} \binom{j}{l} (r-j)! \tau^{r-j} \mu^{j-l} \sigma^l \mathbb{E}(Z^l) \tag{10}$$

where:

$$\mathbb{E}(Z^l) = \begin{cases} \frac{1}{\sqrt{2\pi}} 2^{k+\frac{1}{2}} \Gamma\left(k+\frac{1}{2}\right) & \text{if } l = 2k, \quad \text{with } k \in \mathbb{Z} \cup \{0\}, \\ \sqrt{\frac{2}{\pi}} \alpha (1+\alpha^2)^{-(k+\frac{1}{2})} 2^{-k} (2k+1)! \sum_{t=0}^{k} \frac{t!(2\alpha)^{2t}}{(2t+1)!(k-t)!}, & \text{if } l = 2k+1, \quad \text{with } k \in \mathbb{Z} \cup \{0\}. \end{cases}$$

Hence, from the central moments given by $\mathbb{E}(X - \mathbb{E}(X))^r$, the expected value and variance of the ESN random variable are:

$$\mathbb{E}(X) = \tau + \mu + \sqrt{\frac{2}{\pi}} \sigma \delta, \quad Var(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = \tau^2 + \left(1 - \frac{2}{\pi} \delta^2\right) \sigma^2, \tag{11}$$

Grushka [12] found that the asymmetry and kurtosis coefficients of the EG random variable depended on the scale parameters of the convolution distributions, i.e., τ and σ. As for the ESN model, the asymmetry (*skew*) and kurtosis (*kurt*) coefficients are given by the following expressions:

$$skew = \frac{2\left(\frac{\tau}{\sigma}\right)^3 + \left(\frac{2}{\pi}\right)^{3/2} \left(2 - \frac{\pi}{2}\right) \delta^3}{\left(\left(\frac{\tau}{\sigma}\right)^2 + \left(1 - \frac{2}{\pi}\delta^2\right)\right)^{3/2}}$$

and:

$$kurt = \frac{3\left(3\left(\frac{\tau}{\sigma}\right)^4 + 2\left(\frac{\tau}{\sigma}\right)^2 \left(1 - \frac{2}{\pi}\delta^2\right)\right) + 2\left(\frac{2}{\pi}\right)^2 (\pi - 3)\delta^4}{\left(\left(\frac{\tau}{\sigma}\right)^2 + \left(1 - \frac{2}{\pi}\delta^2\right)\right)^2}.$$

where $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$ and $sgn(\cdot)$ denotes the sign function. Therefore, in line with Grushka's findings [12], (*skew*) and (*kurt*)depend on $\frac{\tau}{\sigma}$ and the asymmetry parameter α.

A $(0, \infty) \times (-1000, 1000)$ grid is constructed for $\frac{\tau}{\sigma} \in (0, \infty)$ and $\alpha \in (-1000, 1000)$ values and *skew* values within $(-0.995267, 1.99999)$; these values are obtained numerically using *R* statistical software. Doing so indicated that if $\frac{\tau}{\sigma} \to \infty$; $\alpha \to \pm\infty$, then *skew* $\to 2$; while if $\frac{\tau}{\sigma} \to 0$; $\alpha \to \infty$, then *skew* $\to 0.9953$. Moreover, if $\frac{\tau}{\sigma} \to 0$; $\alpha \to -\infty$, then *skew* $\to -0.9953$; and if $\frac{\tau}{\sigma} \to 0$; $\alpha \to 0$, then *skew* $\to 0$. Therefore, the range of asymmetry of the ESN model is $(-0.9953, 2]$, which makes it more flexible than the EG model and other asymmetric models such as the SN model [9] and the power-normal model [13,14].

Similarly, for the same range of $\frac{\tau}{\sigma}$ and α values, the kurtosis coefficient is within the $(0.006039, 8.99999)$ range (these values were obtained numerically using the *R*statistical software). Moreover, if $\frac{\tau}{\sigma} \to \infty$; $\alpha \to \pm\infty$, then *kurt* $\to 9$; while if $\frac{\tau}{\sigma} \to 0$; $\alpha \to \pm\infty$, then *kurt* $\to 0.8692$, and if $\frac{\tau}{\sigma} \to 0$; $\alpha \to 0$, then *kurt* $\to 0$. Therefore, the kurtosis range of the ESN model makes the ESN more

flexible than asymmetric models, including the SN model [9], the power-normal model [13,14] and the SN Alpha-power model [15].

The moment- and characteristic-generating functions of the ESN model are given by:

$$M_X(t) = 2e^{\mu t + \frac{1}{2}\sigma^2 t^2} \left(1 - \frac{t}{\tau}\right)^{-1} \Phi(\sigma\delta t)$$

and:

$$\phi_c(t) = 2e^{it\mu - \frac{1}{2}\sigma^2 t^2} \left(1 - \frac{it}{\tau}\right)^{-1} \Phi(\sigma\delta it).$$

Note that, for $a \in \mathbb{R}$ and $b > 0$,

$$M_{a+bX}(t) = 2e^{(\mu+a)t + \frac{1}{2}(b\sigma)^2 t^2} \left(1 - \frac{t}{\tau/b}\right)^{-1} \Phi((b\sigma)\delta t),$$

which means that $a + bX \sim ESN(\tau/b, a + \mu, b\sigma, \alpha)$. In other words, the ESN distribution is closed for the linear combination of the form $a + bX$.

2.2. Exponential-Centred Sn Distribution

The skew-normal model is a suitable alternative for datasets with a significant degree of asymmetry; however, its drawback is that it has a singular information matrix if its asymmetry parameter equals zero, i.e., $\alpha = 0$ [9]. Regarding this case (in which α is in the proximity of zero), some authors have proposed a reparameterization [16]. Another option is to use the so-called centred skew-normal distribution [9], which is obtained as shown below.

For $Z \sim SN(\alpha)$, with $\mathbb{E}(Z) = b\delta$ and $Var(Z) = 1 - (b\delta)^2$ where $b = \sqrt{\frac{2}{\pi}}$ and $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$, let the Y random variable be defined as:

$$Y = \xi + \eta \frac{Z - \mathbb{E}(Z)}{\sqrt{Var(Z)}}, \quad -\infty < \xi < \infty, \ \eta > 0.$$

For this random variable, it follows that:

$$\mathbb{E}(Y) = \xi \quad \text{and} \quad Var(Y) = \eta^2.$$

Given that the random variable Y is a linear combination of $Z \sim SN(\alpha)$, Y also follows an SN distribution, and after some calculations, it results that:

$$Y = \mu + \sigma Z \sim SN(\mu, \sigma, \alpha)$$

where:

$$\mu = \xi - c\eta\gamma_1^{1/3}, \quad \sigma = \eta\sqrt{1 + c^2\gamma_1^{2/3}}, \quad \alpha = \frac{c\gamma_1^{1/3}}{\sqrt{b^2 + c^2(b^2 - 1)\gamma_1^{2/3}}} \tag{12}$$

with $c = \{2/(4 - \pi)\}^{1/3}$ and $-0.9953 < \gamma_1 < 0.9953$ (these values are provided in [9]).

This is known as the centred SN (CSN) distribution and is denoted by $CSN(\xi, \eta, \gamma_1)$.

Ref. [9] showed that the information matrix of the CSN model can be written as:

$$I(\boldsymbol{\theta}) = D^T I(\boldsymbol{\theta}_1)D, \tag{13}$$

where $I(\boldsymbol{\theta}_1)$ is the information matrix of the SN model; D the matrix of the derivatives of the parameter vector $\boldsymbol{\theta}_1 = (\mu, \sigma, \gamma_1)$ with respect to the parameter vector $\boldsymbol{\theta} = (\xi, \eta, \lambda)$; [9] showed that $I(\boldsymbol{\theta})$ is a non-singular matrix.

Hence, based on this very reparameterization, the exponential-centred SN model is given by:

$$g(x) = \frac{2}{\tau} e^{-\frac{x - \xi + c\eta\gamma_1^{1/3}}{\tau} + \frac{\eta^2 \left(1 + c^2\gamma_1^{2/3}\right)}{2\tau^2}} \Phi_B \left(\frac{x - \xi + c\eta\gamma_1^{1/3}}{\eta\sqrt{1 + c^2\gamma_1^{2/3}}} - \frac{\eta\sqrt{1 + c^2\gamma_1^{2/3}}}{\tau}, \frac{\eta\delta_1\sqrt{1 + c^2\gamma_1^{2/3}}}{\tau}; -\delta_1 \right) \quad (14)$$

with $\delta_1 = \frac{\alpha}{\sqrt{1+\alpha^2}}$ and α as defined in Equation (12). This model will be denoted by $ECSN(\tau, \xi, \eta, \gamma_1)$. When $\gamma_1 = 0$, the $EG(\tau, \mu, \sigma^2)$ model is then obtained.

Note that this centred representation of the SN model reduces the space range of the asymmetry parameter since $\gamma_1 \in (-0.9953; 0.9953)$ (see [17]).

3. Estimation

For a random sample $X_1, X_2, \cdots, X_n$ where $X_i \sim ECSN(\tau, \xi, \eta, \gamma_1)$, it is possible to estimate the parameters of the $ECSN(\tau, \xi, \eta, \gamma_1)$ model by maximising the log-likelihood function:

$$\begin{aligned}
\ell(\theta) &= -n\log(\tau) + \frac{n\eta^2 \left(1 + c^2\gamma_1^{2/3}\right)}{2\tau^2} - \sum_{i=1}^{n} \frac{x_i - \xi + c\eta\gamma_1^{1/3}}{\tau} \\
&\quad + \sum_{i=1}^{n} \log \left(\Phi_B \left(\frac{x_i - \xi + c\eta\gamma_1^{1/3}}{\eta\sqrt{1 + c^2\gamma_1^{2/3}}} - \frac{\eta\sqrt{1 + c^2\gamma_1^{2/3}}}{\tau}, \frac{\eta\delta_1\sqrt{1 + c^2\gamma_1^{2/3}}}{\tau}; -\delta_1 \right) \right).
\end{aligned} \quad (15)$$

Another alternative could be to maximise the log-likelihood function of the parameter vector $\theta_1 = (\tau, \mu, \sigma, \alpha)$, which is given by:

$$\ell(\theta_1) = -n\log(\tau) + \frac{n\sigma^2}{2\tau^2} - \sum_{i=1}^{n} \frac{x_i - \mu}{\tau} + \sum_{i=1}^{n} \log \left(\Phi_B \left(\frac{x_i - \mu}{\sigma} - \frac{\sigma}{\tau}, \frac{\sigma\delta}{\tau}; -\delta \right) \right), \quad (16)$$

and then use the relationships defined in Equation (12) to find the estimates of vector $\theta_1 = (\tau, \xi, \eta, \gamma_1)$. The elements of the score function $U(\theta_1) = (U(\tau), U(\mu), U(\sigma), U(\alpha))$, i.e., the derivative of the log-likelihood function with respect to the parameters, are given in Appendix A.

The score equations are estimated by equating the scores to zero ($U(\theta_{1j}) = 0$). Then, the maximum likelihood estimators $(\hat{\tau}, \hat{\mu}, \hat{\sigma}, \hat{\alpha})$ of the parameter vector $(\tau, \mu, \sigma, \alpha)$ are calculated by solving the system of equations $U(\theta_j) = 0$ by means of Newton–Raphson or quasi-Newton numerical methods, among others.

The observed information matrix of the parameter vector $\theta_1 = (\tau, \mu, \sigma, \alpha)$ is defined as $\Lambda(\boldsymbol{\theta_1})) = -\frac{d}{d\theta_1} U(\theta_1) = (\Lambda_{\theta_{1j}\theta_{1j'}})$, i.e., minus the second derivative of the log-likelihood function with respect to the parameters, provided in Appendix B.

The elements of the Fisher information matrix of vector θ_1, $K(\boldsymbol{\theta_1})$ are given by $\kappa_{\theta_{1j}\theta_{1j'}} = \mathbb{E}\left(\Lambda_{\theta_{1j}\theta_{1j'}}\right)$ and should be calculated numerically.

The information matrix of the ECSN model can be obtained as the relationship in Equation (13) using an appropriate matrix $B = (\partial\theta_1/\partial\theta)_{4\times4}$.

Furthermore, the observed information matrix $(\Lambda(\boldsymbol{\theta}))$ and Fisher information matrix $(K(\boldsymbol{\theta}))$ of this model, for the parameter vector $\theta = (\tau, \xi, \eta, \gamma_1)$, can be written in terms of the observed and expected information matrices of the ESN model, of the parameter vector $\theta_1 = (\tau, \mu, \sigma, \alpha)$, in the form:

$$\Lambda(\boldsymbol{\theta}) = B'\Lambda(\boldsymbol{\theta_1})B \quad \text{and} \quad K(\boldsymbol{\theta}) = B'K(\boldsymbol{\theta_1})B \quad (17)$$

where:

$$B = \begin{pmatrix} 1 & \mathbf{0}_3^T \\ \mathbf{0}_3 & D_{3\times3} \end{pmatrix} \quad (18)$$

The elements of this matrix $D_{3\times3}$ can be found in their general form in Azzalini and Capitanio ([17], p. 68). Note that when $\gamma_1 \to 0$, the inferior sub-matrix of size 3×3 is nonsingular and converges to a diagonal matrix $diag(\sigma^2, \sigma^2/2, 6)$. This event entails that the matrix $K(\boldsymbol{\theta})$ is nonsingular for $0 < \tau < \infty$.

Therefore, in regular conditions and by applying the asymptotic theory to the maximum likelihood estimator vector of the parameter vector θ, $\Lambda(\theta) \to K(\theta)$. Then, using the same stochastic convergence argument, it follows that $\hat{\theta}$ converges to a normal distribution; that is:

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N_4(\mathbf{0}, \Delta^{-1}(\theta)).$$

Hence, for the ECSN convolution model, the elements in the diagonal of $\Delta^{-1}(\hat{\theta})$ represent the estimates of the variances of $\hat{\tau}$, $\hat{\xi}$, $\hat{\eta}$ and $\hat{\gamma}_1$. The $100 \times (1 - \beta)\%$ confidence intervals can be estimated in the conventional $\theta_j \mp Z_{1-\beta/2}\sigma_{jj}$ form, where σ_{jj} represents the standard error of the parameter θ_j.

3.1. Illustrations

3.1.1. Reaction Times

Ref. [18] investigated the sensorimotor processing of words by using a masked priming (MP) task. The MP is a task commonly used in experimental psychology and in which a prime word (e.g., below) is "sandwiched" for a short time (e.g., 35 ms) between a preceding masking pattern (or forward mask, e.g., LKPFH; shown for, say, 200 ms) and a proceeding target word (e.g., above; shown until a response is made). In some variations, there can be a masking pattern between the prime and the target word (called backward mask), and this was the case in Ansorge et al.'s study (such a pattern was shown for approximately 35 ms). Traditionally, there are congruent and incongruent trials; in the former case, there is a match between the response of the participant and a correct or expected response, while in the latter, there is a mismatch between those responses. The time the participant takes in responding is known as the reaction time (RT) and is measured in milliseconds.

Twelve females and 12 males participated in an MP task, and their RTs were collected for both congruent and incongruent trials (see Experiment 2a). The descriptive statistics for the distribution of mean RTs during incongruent trials were as follows: $\bar{X} = 890.1149$, $s = 98.6875$, $b_1 = -0.5326$, and $b_2 = 4.1798$. These statistics indicated that the distribution of RTs was negatively skewed; hence, an EG distribution would not provide an appropriate fit as its asymmetry is always positive. When the data were fitted with the ECSN, the following estimates were obtained: $\hat{\tau} = 0.2927(0.00002)$, $\hat{\mu} = 977.1546(0.0058)$, $\hat{\sigma} = 130.3170(0.0058)$ and $\hat{\alpha} = -1.8182(0.0041)$. Considering the centred representation of the ECSN model, the estimated parameters were as follows: $\hat{\tau} = 0.2927(0.00002)$, $\hat{\xi} = 886.0465(0.1350)$, $\hat{\nu} = 93.1774(0.1319)$ and $\hat{\gamma} = -0.4012(0.0034)$. Therefore, the ECSN gave an optimal fit, and it was further corroborated via the Kolmogorov–Smirnov (D) test; $D = 0.16667$, p-value $= 0.9024$ (in this work, the D test is used to assess the suitability of a given statistical distribution to a dataset. However, other tests exist; e.g., Cramér–von Mises, Anderson–Darling, Shapiro–Wilk, χ^2, Hosmer–Lemeshow, Kuiper, Kernelized Stein discrepancy, Moran, and Zhang's Z_K, Z_C and Z_A tests). Figure 2 displays the distribution of mean RTs and the ECSN distribution's fits.

The following example is provided to illustrate that negatively skewed data can emerge in situations other than RTs. For example, in psychometrics, some rating scales can show a negative skew.
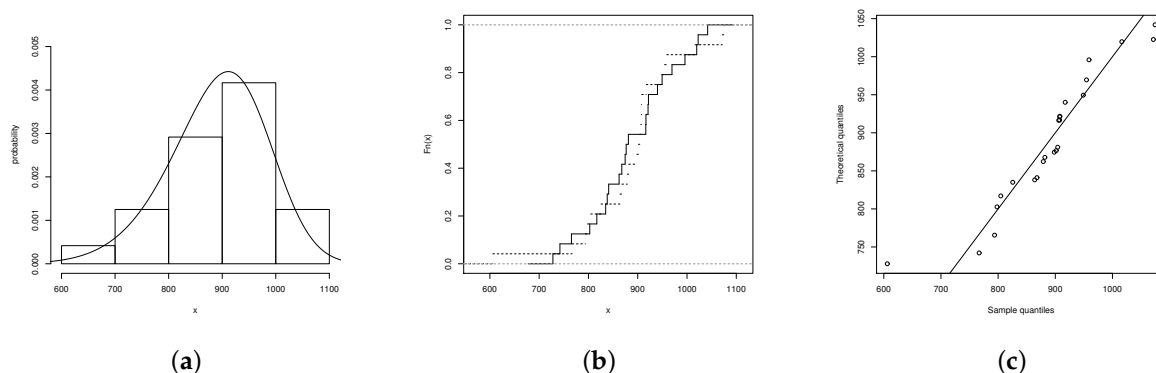
(a)

(b)

(c)

Figure 2. Histogram and overlaid PDF of the distribution of mean RTs (**a**); RTs' ECDF(solid line) and ECSN distribution's CDF (dashed line) (**b**); and QQ-plot representing the ECSN distribution's goodness of fit (**c**).

3.1.2. The Flourishing Scale

Ref. [19] validated the Flourishing Scale (FS) in a sample of adults ($n = 275$) with low and moderate levels of well-being. The FS rating scale is an eight item measure of social-psychological functioning that tackles aspects such as relationships, self-esteem, purpose and optimism.

The authors aimed to understand the distribution of the total Flourishing scores ($X = \text{TFS}$, $X \in \mathbb{Z}^{+}$); a global measure of the core aspects of individuals' optimal social and psychological functioning. The descriptive statistics of the TFS showed an average of $\bar{X} = 41.44$, a standard deviation of $s = 6.48$, a skewness of $skew = -1.4394$ and a kurtosis of $kurt = 5.8725$. The negative skewness suggested that the EG model was not the most suitable option because its range of asymmetry was positive. In addition, the high kurtosis left little room for models such as the SN, whose ranges of kurtosis and asymmetry were [3, 3.86) and (−0.9953, 0.9953), respectively.

The EG, the skew-normal Alpha-power (SNAP) and the ECSN models were fitted to the data. The exponential representation $\tau \exp(-\tau x)$ was used for those models. The four parameter model $SNAP(\mu, \sigma, \lambda, \alpha)$, studied by [15], is more flexible than its SN and power-normal (PN) counterparts because when $\alpha = 1$, the SN distribution was obtained, and when $\lambda = 0$, the PN distribution was obtained; a normal distribution ensues when $\lambda = 0$ and $\alpha = 1$. The SNAP's ranges of asymmetry and kurtosis were $[-1.4676, 0.9953)$ and $[1.4672, 5.4386]$, respectively [15], which made it a good fit to the TFS data.

These models were compared using the Akaike information criterion (AIC) [20] and the Bayesian information criterion (BIC) of Schwarz [21], defined as:

$$AIC = -2\ell(\theta) + 2p \quad \text{and} \quad BIC = -2\ell(\theta) + p\log(n)$$

where p is the number of parameters of the vector θ of the specific model. Based on the results of the AIC and the BIC, the most suitable model for the TFS data was the ECSN model (see Table 1). Figure 3 further shows the good fit of the ECSN model. Figure 4 shows the low degree of fit of the EG (a) and the SNAP (b) models; but it is evident that the EG gave a better fit than SNAP (see also the AIC and BIC values in Table 1).

A null hypothesis test of no difference between the ECSN model and the EG model can be expressed as,

$$H_{01} : \alpha = 0 \quad \text{versus} \quad H_{11} : \alpha \neq 0.$$

By using the likelihood-ratio test (models are nested),

$$\Lambda_1 = \frac{L_{EG}(\hat{\boldsymbol{\theta}})}{L_{ECSN}(\hat{\boldsymbol{\theta}})}.$$

it is found that:

$$-2\log(\Lambda_1) = -2(-903.665 + 865.955) = 75.4198,$$

This was a value larger than the critical 5% chi-squared value, i.e., $\chi^2_{1,95\%} = 3.8414$. Therefore, the null hypothesis was rejected, and it could be concluded that the ECSN model—which involved one extra parameter, making it more flexible in terms of asymmetry—fit the data better than the EG model (compare Figure 3b and Figure 4a).

Table 1. MLE parameter estimates (with standard errors in parentheses) of the EG, SNAP, and ECSN models. The estimates of the ECSN model correspond to the centred model.

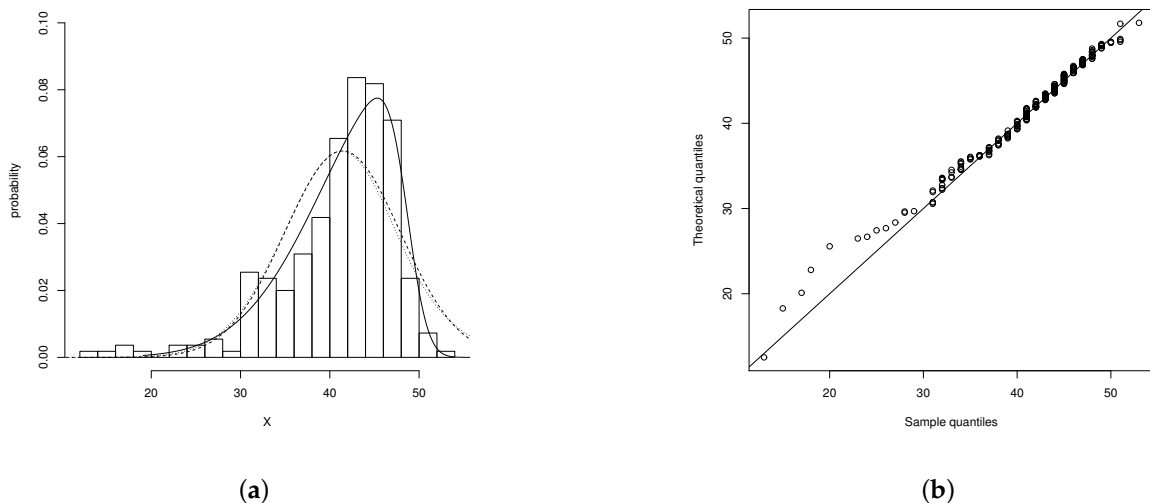| Parameter	EG Model	SNAP Model	ECSN Model
	Estimate	Estimate	Estimate
$\hat{\zeta}$	41.0257(0.7945)	46.1732(1.3908)	41.0952(0.4073)
$\hat{\eta}$	6.4577(0.2783)	6.8396(2.0609)	5.8089(0.5621)
$\hat{\tau}$	2.4131(4.0354)	2.4579(0.1903)	3.7056(0.0037)
$\hat{\gamma}$		0.1226(0.0350)	−0.8642(0.4082)
AIC	1813.33	1817.23	1739.91
BIC	1824.18	1831.69	1754.37

(**a**)

(**b**)

Figure 3. (**a**) Histogram of the total Flourishing score (TFS) data and estimated densities (EG model (dashed line), SNAP model (dotted line) and ECSN model (solid line)); (**b**) QQ-plot of the ECSN model.
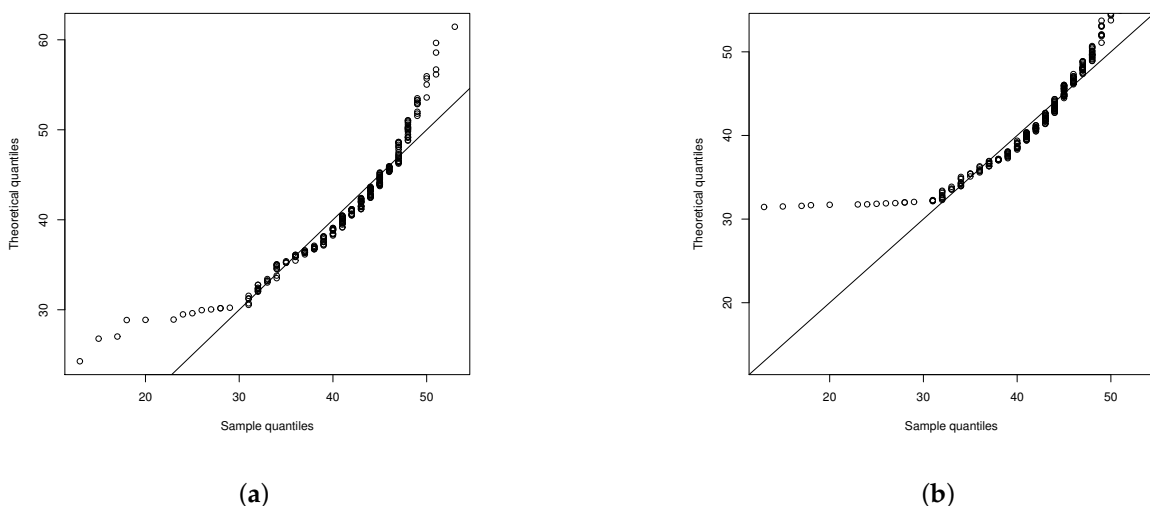

(**a**)

(**b**)

Figure 4. QQ-plots and TFS data. (**a**) QQ-plot of the EG model; (**b**) QQ-plot of the SNAP model.

Furthermore, the *D* test was used to determine the best fit for the TFS data. The results were as follows: for the EG model, $D = 0.200$ (*p*-value = 0.000); for the SNAP model, $D = 0.2290$ (*p*-value = 0.000); and for the ECSN model, $D = 0.0981$ (*p*-value = 0.1411). The results thus corroborated that the ECSN model was suitable for the TFS data.

4. Discussion

Data can take different distributional shapes, but some shapes are more common in certain research fields (for more information about other distributions and their applications, see the Special Issues in Volume 11 Issue 11 (year 2019) and Volume 12 Issue 5 (year 2020) in this journal). RTs, for example, are a ubiquitous metric in experimental psychology, and it is well known that their distribution is positively skewed. However, that is not necessarily always the case, and normally or negatively shaped RTs can ensue. In those cases, the EG distribution, a distribution traditionally used to model RT data, is not able to fit the data well. In this article, a distribution capable of fitting traditional RT data and RT data that deviate from a positive skew was proposed.

It is important, however, to frame the proposed distribution within the larger field of data analysis and modelling. While most research across scientific fields tends to focus on average differences, not much attention is paid to investigate group differences in variation. Indeed, tests designed to account for both location and scale (e.g., the Cucconi test) are rarely used. Adopting a distributional modelling approach enables tackling data differently by considering, at a minimum, the data's location (e.g., mean) and scale (e.g., SD). To our knowledge, the only existing statistical approach able to model data in a proper distributional manner is the generalised additive models for location, scale and shape (GAMLSS).

To conclude, the ECSN distribution was proposed as a new statistical distribution suitable to adjust the shape of RT data. As the ECSN was a four parameter distribution, it could be used to model RT data's location, scale and shape. It is our hope to see this distribution implemented in the GAMLSS R package so that it is put to the test in research contexts. The necessary Rcodes and datasets to reproduce the estimations and plots reported here can be found at https://figshare.com/projects/ Two_new_distributions_for_the_modelling_of_RT_data/81743.

Abbreviations

The following abbreviations are used in this manuscript:

CDF	Cumulative distribution function
CSN	Centred skew-normal distribution
ECSN	Exponential-centred skew-normal distribution
EG	Ex-Gaussian distribution
ESN	Exponential skew-normal distribution
EXP	Exponential distribution
FS	Flourishing scale
GAMLSS	Generalised additive models for location, scale and shape
MP	Masked priming
PDF	Probability distribution function

PN   Power-normal distribution
RTs   Reaction times
SN   Skew-normal distribution
SNAP   Skew-normal Alpha-power distribution

Appendix A. Score Function

The elements of the score function, obtained from (16), are given by:

$$U(\tau) = \frac{\partial \ell(\theta_1)}{\partial \tau} = -\frac{n}{\tau} + \sum_{i=1}^{n} \frac{x_i - \mu}{\tau^2} + \sum_{i=1}^{n} P_\tau(x_i; \theta_1), \tag{A1}$$

$$U(\mu) = \frac{\partial \ell(\theta_1)}{\partial \mu} = \frac{n}{\tau^2} + \sum_{i=1}^{n} P_\mu(x_i; \theta_1), \tag{A2}$$

$$U(\sigma) = \frac{\partial \ell(\theta_1)}{\partial \sigma} = \frac{n\sigma}{\tau^2} + \sum_{i=1}^{n} P_\sigma(x_i; \theta_1), \tag{A3}$$

$$U(\alpha) = \frac{\partial \ell(\theta_1)}{\partial \alpha} = \sum_{i=1}^{n} P_\alpha(x_i; \theta_1), \tag{A4}$$

where:

$$P_{\theta_j}(x_i; \theta_1) = \frac{\partial}{\partial \theta_j} \log\left(\Phi_B\left(z_{ci}, \frac{\sigma\delta}{\tau}; -\delta\right)\right) = \frac{\frac{\partial}{\partial \theta_j}\Phi_B\left(z_c, \frac{\sigma\delta}{\tau}; -\delta\right)}{\Phi_B\left(z_c, \frac{\sigma\delta}{\tau}; -\delta\right)},$$

with $z_{ci} = \frac{x_i - \mu}{\sigma} - \frac{\sigma}{\tau}$. Therefore, for $\theta_j = \tau, \mu, \sigma, \alpha$, the following is obtained:

$$P_\tau(x_i; \theta_1) = \frac{-\frac{\alpha}{\tau^2}\frac{1}{\sqrt{1+\alpha^2}}\phi\left(\frac{\alpha\sigma}{\tau}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci} + \frac{\sigma\delta}{\tau}\right) + \frac{1}{\tau^2}\phi(z_{ci})\Phi\left(\alpha z_{ci} + \frac{\alpha\sigma}{\tau}\right)}{\Phi_B\left(z_{ci}, \frac{\sigma\delta}{\tau}; -\delta\right)}$$

$$P_\mu(x_i; \theta_1) = -\frac{\frac{1}{\sigma}\phi(z_{ci})\Phi\left(\alpha z_{ci} + \frac{\alpha\sigma}{\tau}\right)}{\Phi_B\left(z_{ci}, \frac{\sigma\delta}{\tau}; -\delta\right)}$$

$$P_\sigma(x_i; \theta_1) = -\frac{1}{\sigma} + \frac{\frac{\alpha}{\tau\sigma\sqrt{1+\alpha^2}}\phi\left(\frac{\alpha\sigma}{\tau}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci} + \frac{\sigma\delta}{\tau}\right) - \frac{1}{\sigma}\left(z_{ci} + \frac{2\sigma}{\tau}\right)\phi(z_{ci})\Phi\left(\alpha z_{ci} + \frac{\alpha\sigma}{\tau}\right)}{\Phi_B\left(z_{ci}, \frac{\sigma\delta}{\tau}; -\delta\right)}$$

$$
\begin{aligned}
P_\alpha(x_i; \theta_1) = {} & \frac{\frac{1}{\tau}\frac{1}{\sqrt{1+\alpha^2}}\phi\left(\frac{\alpha\sigma}{\tau}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci} + \frac{\sigma\delta}{\tau}\right) - \frac{1}{\sigma\sqrt{1+\alpha^2}}\phi\left(\frac{\sigma\delta}{\tau}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci} + \frac{\alpha\sigma\delta}{\tau}\right)}{\Phi_B\left(z_{ci}, \frac{\sigma\delta}{\tau}; -\delta\right)} \\
& -\frac{\frac{\delta^3\sigma}{\alpha\tau}\phi\left(\frac{\alpha\delta}{\tau}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci} + \frac{\alpha\sigma\delta}{\tau}\right)}{\Phi_B\left(z_{ci}, \frac{\sigma\delta}{\tau}; -\delta\right)}
\end{aligned}
$$

Appendix B. Observed Information Matrix

Knowing that $W_c = \frac{\phi_{SN}(z_{ci};\alpha)}{\Phi_{SN}(z_{ci};\alpha)}$ and $W_\alpha = \frac{\phi(\sqrt{1+\alpha^2}z_{ci};\alpha)}{\Phi_{SN}(z_{ci};\alpha)}$, the elements of the observed information matrix are given by:

$$\Lambda_{\tau\tau} = \frac{1}{\tau^4}\left[n\tau^2 + 2n\sigma\tau\bar{z} - 3n\sigma^2\right] + \sum_{i=1}^{n}\left[\frac{\frac{1}{\tau^2}\phi(z_{ci})\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)\left[\frac{2}{\tau}+\frac{\sigma}{\tau^2}z_{ci}\right]}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} + P_\tau^2(x_i;\theta_1)\right] -$$

$$\sum_{i=1}^{n}\frac{\frac{\delta}{\tau^2}\phi\left(\frac{\alpha\sigma}{\tau}\right)\left[\left(\frac{2}{\tau^2}-\frac{\alpha^2\sigma^2}{\tau^3}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right)+\frac{\sigma}{\tau^2}\left(\sqrt{1+\alpha^2}-\delta\right)\phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right)\right]}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} \quad (A5)$$

$$\Lambda_{\tau\mu} = \frac{n}{\tau^2} - \frac{1}{\tau^2}\sum_{i=1}^{n}\frac{z_{ci}\phi(z_{ci})\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} +$$

$$\frac{1}{\tau^2\sigma}\sum_{i=1}^{n}\frac{\phi(Z_c)\Phi\left(\alpha Z_c+\frac{\alpha\sigma}{\tau}\right)}{\Phi_B^2\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)}\left[\delta\phi\left(\frac{\alpha\sigma}{\tau}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\alpha\delta}{\tau}\right) -\right.$$

$$\left. \phi(z_{ci})\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)\right] \quad (A6)$$

$$\Lambda_{\tau\sigma} = \frac{2n\sigma}{\tau^3} + \sum_{i=1}^{n}\left[-\frac{1}{\sigma\tau^2}\phi(z_{ci})\frac{\left[z_{ci}\left(z_{ci}+\frac{2\sigma}{\tau}\right)\phi\left(\alpha Z_c+\frac{\alpha\sigma}{\tau}\right)-\left(\alpha Z_c+\frac{\alpha\sigma}{\tau}\right)\phi\left(\alpha Z_c+\frac{\alpha\sigma}{\tau}\right)\right]}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} + P_\tau(x_i;\theta_1)P_\sigma(x_i;\theta_1)\right]$$

$$-\frac{\delta}{\tau^2}\phi\left(\frac{\alpha\sigma}{\tau}\right)\sum_{i=1}^{n}\frac{\left[\frac{\alpha^2\sigma}{\tau^2}\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right)+\frac{1}{\sigma}\sqrt{1+\alpha^2}\left(z_{ci}+\frac{2\sigma}{\tau}-\frac{\sigma\delta}{\tau}\right)\phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\alpha\delta}{\tau}\right)\right]}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} \quad (A7)$$

$$\Lambda_{\tau\alpha} = -\frac{1}{\tau^2}\sum_{i=1}^{n}\frac{\left(z_{ci}+\frac{\sigma}{\tau}\right)\phi(z_{ci})\phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} + \sum_{i=1}^{n}P_\alpha(x_i;\theta_1)P_\tau(x_i;\theta_1)$$

$$+\frac{\delta}{\tau^2}\phi\left(\frac{\alpha\sigma}{\tau}\right)\sum_{i=1}^{n}\frac{\left[\left(\frac{\delta^2}{\alpha^3}-\frac{\alpha\sigma^2}{\tau^2}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right)+\left(2\delta z_{ci}+\frac{\sigma\delta^3}{\alpha^3}\right)\phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right)\right]}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} \quad (A8)$$

$$\Lambda_{\mu\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left[\frac{\phi(z_{ci})\left(z_{ci}\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)-\alpha\phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)\right)}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} + \left(\frac{\phi(z_{ci})\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)}\right)^2\right] \quad (A9)$$

$$\Lambda_{\mu\sigma} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{z_{ci}\left(z_{ci}+\frac{2\sigma}{\tau}\right)\phi(z_{ci})\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)-\alpha z_i\phi(z_{ci})\phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)}{\Phi_B^2\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)}$$

$$+\frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{-\frac{\delta}{\tau}\phi\left(\frac{\alpha\delta}{\tau}\right)\phi(z_{ci})\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right)+\left(z_{ci}+\frac{2\sigma}{\tau}\right)\left(\phi(z_{ci})\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)\right)^2}{\Phi_B^2\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} \quad (A10)$$

$$\Lambda_{\mu\alpha} = \frac{1}{\sigma}\sum_{i=1}^{n}\frac{z_i\phi(z_{ci})\phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} - \frac{\delta}{\sigma\tau\alpha}\sum_{i=1}^{n}\frac{\phi(z_{ci})\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)}{\Phi_B^2\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)}\left[\phi\left(\frac{\alpha\sigma}{\tau}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right) -\right.$$

$$\left.\frac{\tau}{\sigma}\phi\left(\frac{\sigma\delta}{\tau}\right)\phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\alpha\sigma\delta}{\tau}\right)-\delta^2\sigma\phi\left(\frac{\sigma\delta}{\tau}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\alpha\sigma\delta}{\tau}\right)\right] \quad (A11)$$

$$\Lambda_{\sigma\sigma} = -\frac{n}{\tau^2} + \sum_{i=1}^{n}\frac{\frac{\delta}{\tau}\phi\left(\frac{\alpha\sigma}{\tau}\right)\left(\frac{1}{\sigma^2}+\frac{\alpha\sigma}{\tau^2}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right)+\frac{\alpha^3}{\sigma\tau^2}\left[\delta\phi\left(\frac{\sigma\delta}{\tau}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\alpha\sigma\delta}{\tau}\right)-\phi(z_{ci})\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)\right]}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)}$$

$$+\sum_{i=1}^{n}\frac{\frac{2\alpha}{\tau\sigma^2}\left(1-\frac{1}{\sigma}z_{ci}-\frac{1}{\tau}\right)\left(z_{ci}+\frac{2\sigma}{\tau}\right)\phi(z_{ci})\phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)-\frac{1}{\sigma^2}\phi(z_{ci})\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)\left[2\left(z_{ci}+\frac{\sigma}{\tau}\right)-\left(z_{ci}+\frac{2\sigma}{\tau}\right)^2z_{ci}\right]}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)}$$

$$-\frac{2}{\sigma^2}\sum_{i=1}^{n}w_i + \frac{1}{\sigma^2}\sum_{i=1}^{n}w_i^2, \quad (A12)$$

where:

$$w_i = \frac{\frac{\delta}{\tau}\sigma\left(\frac{\alpha\sigma}{\tau}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right) - \left(z_{ci}+\frac{2\sigma}{\tau}\right)\phi(z_{ci})\Phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)}$$

$$\begin{aligned}
\Lambda_{\sigma\alpha} = \quad & -\frac{\delta}{\sigma\tau}\phi\left(\frac{\alpha\sigma}{\tau}\right)\sum_{i=1}^{n}\frac{\left[\left(\frac{\delta^2}{\alpha^3}-\frac{\alpha\sigma^2}{\tau^2}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right)+\left(2\delta z_{ci}+\frac{\sigma\delta^3}{\alpha^3}\right)\phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right)\right]}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} \\
& +\frac{1}{\sigma}\sum_{i=1}^{n}\frac{\left(z_{ci}+\frac{2\sigma}{\tau}\right)\left(z_{ci}+\frac{\sigma}{\tau}\right)\phi(z_{ci})\phi\left(\alpha z_{ci}+\frac{\alpha\sigma}{\tau}\right)}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} + \sum_{i=1}^{n}P_\alpha(x_i;\theta_1)P_\sigma(x_i;\theta_1)
\end{aligned} \tag{A13}$$

$$\begin{aligned}
\Lambda_{\alpha\alpha} = \quad & -\frac{\delta}{\tau}\phi\left(\frac{\alpha\sigma}{\tau}\right)\sum_{i=1}^{n}\frac{\left[\left(\frac{2\delta^2}{\alpha^2}-\frac{\sigma^2}{\tau^2}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right)+\frac{\delta}{\alpha}\left(2z_{ci}+\frac{\sigma\delta^2}{\alpha^3\tau}\right)\phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\sigma\delta}{\tau}\right)\right]}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} + \sum_{i=1}^{n}P_\alpha^2(x_i;\theta_1) \\
& +\frac{\delta^2}{\sigma}\phi\left(\frac{\sigma\delta}{\tau}\right)\sum_{i=1}^{n}\frac{\phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\alpha\sigma\delta}{\tau}\right)\left[\frac{\delta}{\alpha^2}\left(2+\frac{\sigma^2\delta^2}{\alpha^2\tau^2}\right)-2\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\alpha\sigma\delta}{\tau}\right)\left(z_{ci}+\frac{\sigma\delta^2}{\tau\alpha^2}\right)\right]}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)} \\
& +\frac{\delta^3\sigma}{\alpha\tau}\phi\left(\frac{\sigma\delta}{\tau}\right)\sum_{i=1}^{n}\frac{\left[\frac{1}{\alpha}\left(2-\frac{\sigma^2\delta^4}{\alpha^2\tau^2}\right)\Phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\alpha\sigma\delta}{\tau}\right)+2\delta\left(z_{ci}+\frac{\sigma\delta^2}{\alpha^2\tau}\right)\phi\left(\sqrt{1+\alpha^2}z_{ci}+\frac{\alpha\sigma\delta}{\tau}\right)\right]}{\Phi_B\left(z_{ci},\frac{\sigma\delta}{\tau};-\delta\right)}.
\end{aligned} \tag{A14}$$

References

1. Hohle, R.H. Inferred components of reaction times as functions of foreperiod duration. *J. Exp. Psychol.* **1965**, *69*, 382–386. [CrossRef]
2. Tejo, M.; Niklitschek-Soto, S.; Marmolejo-Ramos, F. Fatigue-life distributions for reaction time data. *Cogn. Neurodyn.* **2018**, *12*, 351–356. [CrossRef] [PubMed]
3. Tejo, M.; Araya, H.; Niklitschek-Soto, S.; Marmolejo-Ramos, F. Theoretical models of reaction times arising from simple-choice tasks. *Cogn. Neurodyn.* **2019**, *13*, 409–416. [CrossRef] [PubMed]
4. Leiva, V.; Tejo, M.; Guiraud, P.; Schmachtenberg, O.; Orio, P. Modeling neural activity with cumulative damage distributions. *Biol. Cybern.* **2015**, *109*, 421–433. [CrossRef] [PubMed]
5. Martínez-Flórez, G.; Barranco-Chamorro, I.; Bolfarine, H.; Gómez, H. Flexible Birnbaum–Saunders distribution. *Symmetry* **2019**, *11*, 1305. [CrossRef]
6. Iriarte, Y.; Varela, H.; Gómez, H.; Gómez, H. A Gamma-Type distribution with applications. *Symmetry* **2020**, *12*, 870. [CrossRef]
7. McGill, W.J. Stochastic latency mechanisms. In *Handbook of Mathematical Psychology*; Luce, R.D., Bush, R.R., Galanter, E., Eds.; Wiley: New York, NY, USA, 1963; Volume 1, pp. 309–359.
8. Christie, L.S.; Luce, R.D. Decision structure and time relations in simple choice behavior. *Bull. Math. Biophys.* **1956**, *18*, 89–112. [CrossRef]
9. Azzalini, A. A class of distributions which includes the normal ones. *Scand. J. Stat.* **1985**, *12*, 171–178.
10. Azzalini, A.; Capitanio, A. *The Skew-Normal and Related Families*, 1st ed.; Cambridge University Press: Cambridge, UK, 2013.
11. Ma, Y.; Genton, M.G. Flexible class of skew-symmetric distributions. *Scand. J. Stat.* **2004**, *31*, 459–468. [CrossRef]
12. Grushka, E. Characterization of exponentially modified Gaussian peaks in chromatography. *Anal. Chem.* **1972**, *44*, 1733–1738. [CrossRef] [PubMed]
13. Durrans, S.R. Distributions of fractional order statistics in hydrology. *Water Resour. Res.* **1992**, *28*, 1649–1655. [CrossRef]
14. Pewsey, A.; Gómez, H.W.; Bolfarine, H. Likelihood-based inference for power distributions. *Test* **2012**, *21*, 775–789. [CrossRef]
15. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H. Skew-normal alpha-power model. *Statistics* **2014**, *48*, 1414–1428. [CrossRef]
16. Rotnitzky, A.; Cox, D.R.; Bottai, M.; Robins, J. Likelihood-based inference with singular information matrix. *Bernoulli* **2000**, *6*, 243–284. [CrossRef]

17. Azzalini, A.; Capitanio, A. Statistical Applications of the Multivariate Skew Normal Distribution. *J. R. Stat. Soc.* **1999**, *61*, 579–602. [CrossRef]

18. Ansorge, U.; Kiefer, M.; Khalid, S.; Grassl, S.; König, P. Testing the theory of embodied cognition with subliminal words. *Cognition* **2010**, *116*, 303–320. [CrossRef] [PubMed]

19. Schotanus-Dijkstra, M.; ten Klooster, P.; Drossaert, C.; Pieterse, M.; Bolier, L.; Walburg, J.; Bohlmeijer, E. Validation of the Flourishing Scale in a sample of people with suboptimal levels of mental well-being. *BMC Psychol.* **2016**, *4*, 12. [CrossRef] [PubMed]

20. Akaike, H. A new look at statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]

21. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]