Article

Monte Carlo Simulation of a Modified Chi Distribution Considering Asymmetry in the Generating Functions: Application to the Study of Health-Related Variables

Nuria Ortigosa [1], Marcos Orellana-Panchame [1,2], Juan Carlos Castro-Palacio [3,*], Pedro Fernández de Córdoba [1] and J. M. Isidro [1]

1   Instituto Universitario de Matemática Pura y Aplicada, Universitat Politècnica de València, 46022 Valencia, Spain; nuorar@upv.es (N.O.); marcos.orellana@unah.edu.hn (M.O.-P.); pfernandez@mat.upv.es (P.F.d.C.); joissan@mat.upv.es (J.M.I.)
2   Departamento de Matemática, Valle de Sula, Universidad Nacional Autónoma de Honduras (UNAH), Sector Pedregal, 21102 San Pedro Sula, Honduras
3   Departamento de Ingeniería Eléctrica, Electrónica, Automática y Física Aplicada, Universidad Politécnica de Madrid, 28012 Madrid, Spain
*   Correspondence: juancas@upvnet.upv.es

Abstract: Random variables in biology, social and health sciences commonly follow skewed distributions. Many of these variables can be represented by exGaussian functions; however, in practice, they are sometimes considered as Gaussian functions when statistical analysis is carried out. The asymmetry can play a fundamental role which can not be captured by central tendency estimators such as the mean. By means of Monte Carlo simulations, the effect of a small asymmetry in the generating functions of the chi distribution is studied. To this end, the k generating functions are taken as exGaussian functions. The limits of this approximation are tested numerically for the practical case of three health-related variables: one physical (body mass index) and two cognitive (verbal fluency and short-term memory). This work is in line with our previous works on a physics-inspired mathematical model to represent the reaction times of a group of individuals.

Keywords: skewed distribution; ex-gaussian; numerical study

1. Introduction

The chi and chi-squared distributions are well-known continuous probability distributions widely used in Applied Statistics [1–5]. The fact that they can be generated by a set of Gaussian-distributed random variables makes them amenable to simulations. We devoted a previous work to study the percentile ratios in a chi distribution [6].

A chi distribution of $k = 3$ degrees of freedom is found in physics to model the velocities of the independent particles of an ideal gas in thermodynamic equilibrium. Similarly, a chi-squared distribution models the energies of the particles in the same physical system. Another typical case from physics is the Rayleigh distribution (chi of $k = 2$ degrees of freedom).

In one of our previous works [7], we found a new interesting application of the chi distribution of $k = 3$ degrees of freedom, that is, of the Maxwell–Boltzmann (MB) distribution. In reference [8] we proved that the reaction times (RTs) of children responding independently to visual stimuli in a short time (hundreds of milliseconds) and without exchanging information are correlated. We interpreted this fact as an experimental evidence for the existence of a system of individuals (or collective). In order to gain insights into this correlation, we developed a physics-inspired mathematical model in reference [7] to represent these correlations. In fact, we could elucidate a correspondence between a system of particles and a group of correlated individuals.

We are rather interested in the conceptual modelling of different situations using the chi distribution. In this respect, in our recent work we have been studying the limits within which the chi distribution can still represent well probability distributions originated from generating functions which are not necessarily Gaussians of equal variances. For example, in reference [9], we studied the limits of the chi modelling for the case of unequal variances in the generating Gaussians. We also proposed a discrete model and an ansatz to calculate the only parameter of this distribution as a function of the unequal variances of the generating Gaussian.

In line with our previous works [7,9], we have extended here the simulation study of the chi distribution for the case of asymmetric generating functions. In this respect, the exGaussian function is a simple, flexible and intuitive function that can be used to represent a skewed distribution as it results from the convolution between the Gaussian and exponential decay functions. The convolution between two functions can be easily simulated as the sum of the respective randomly generated variables. Two practical examples which can be represented by exGaussians are the reaction time distributions in Experimental Psychology [10–17] or the peaks in Chromatography [18].

In this paper we will carry out Monte Carlo simulations to study the distribution Z that originates from combining k generating functions with certain asymmetry (Z_j), as $\left(\sum_{j=1}^{k} Z_j^2\right)^{1/2}$, and will evaluate the result by means of a fit to a chi distribution. The level of asymmetry is considered by using exGaussians as generating functions. Our aim is to explore the level of asymmetry for which the fit to chi distribution can be still considered reasonably good for practical applications. Our approach is useful to model multiple situations in health and social sciences where random variables commonly follow asymmetrical distributions. In this respect, an example involving health-related variables is also included in this work.

2. Methodology

2.1. Generalities on the Chi Distribution

The chi distribution is a continuous probability distribution of a random variable defined as

$$chi = \left(\sum_{j=1}^{k} X_j^2\right)^{1/2} \tag{1}$$

where each of the X_j, $j = 1, \ldots, k$, is a Gaussian-distributed independent random variable. Each one of the k variables X_j follows a Gaussian distribution with mean zero and variance to achieve unity. In Ref. [9] we analysed the case of different values of the variance for each Gaussian component j. In the present paper we propose to study the case in which the X_j components deviate from exact Gaussianity, instead exhibiting a certain degree of asymmetry. We propose to model the deviation of each component from pure Gaussianity, considering them as exGaussian distributions. The exGaussian distribution [10] is given by

$$f(x; \mu, \sigma, \tau) = \frac{1}{2\tau} \exp\left[\frac{1}{2\tau}\left(2\mu + \frac{\sigma^2}{\tau} - 2x\right)\right] \text{erfc}\left(\frac{\mu + \frac{\sigma^2}{\tau} - x}{\sqrt{2}\,\sigma}\right), \tag{2}$$

where erfc is the complementary error function. The above $f(x; \mu, \sigma, \tau)$ is the result of convoluting the pure Gaussian

$$g(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2\right] \tag{3}$$

with the exponential distribution

$$h(x; \tau) = \frac{1}{\tau} \exp\left(-\frac{x}{\tau}\right), \tag{4}$$

where μ and σ are the mean and standard deviation of the Gaussian component and τ the decay constant of the exponential. Let us recall that, if a random variable X is distributed according to (3), and a random variable Y is distributed according to (4), then the sum $X + Y$ is distributed according to the exGaussian (2). In this way, the parameters μ, σ are *not* the mean and the standard deviation of the exGaussian distribution (2), but those of its Gaussian component (3). Thus the parameter τ is a measure of the skewness of the exGaussian distribution, i.e., of its deviation from pure Gaussianity. The value $\tau = 0$ corresponds to a pure symmetric Gaussian [9].

On the other hand, the probability density function corresponding to (1) is given by

$$f(x;k) = 2^{1-\frac{k}{2}} \left[\Gamma\left(\frac{k}{2}\right) \right]^{-1} x^{k-1} \exp\left(-\frac{x^2}{2}\right). \tag{5}$$

In the case when all the variances of the k generating Gaussians take a value different from one, then the above $f(x;k)$ generalises to [9]

$$f(x;B,k) = 2^{1-\frac{k}{2}} B^{-\frac{k}{2}} \left[\Gamma\left(\frac{k}{2}\right) \right]^{-1} x^{k-1} \exp\left(-\frac{x^2}{2B}\right), \tag{6}$$

where B is related to the variance of distribution (6). The chi distribution as stated in (1) describes a k-dimensional ideal gas of free, independent particles. In this latter case, the k variances are all equal. The cumulative distribution function corresponding to (6) is given by

$$F(x;B,k) = 1 - \left[\Gamma\left(\frac{k}{2}\right) \right]^{-1} \Gamma\left(\frac{k}{2}, \frac{x^2}{2B}\right) \tag{7}$$

where $\Gamma(a,b)$ is the upper incomplete gamma function [19].

A particularly interesting case of the above appears in the statistical mechanics of ideal gases [20]. This is the case of a chi distribution with $k = 3$ degrees of freedom. Then the random variables X_j are the three components v_x, v_y, v_z of the velocities of the particles. These components are Gaussian-distributed, and their modulus $(v_x^2 + v_y^2 + v_z^2)^{1/2}$ is distributed according to (6). This special case, called the Maxwell–Boltzmann distribution [21,22], is such that all three component distributions are centered around $v_j = 0$, and the three variances are all equal (and proportional to the temperature of the gas). A k-dimensional ideal gas would be represented by (6).

2.2. Monte Carlo Simulations for Non-Gaussian Generating Distributions Z_j

In this paper we will perform Monte Carlo simulations to generate one random variable Z obtained as $Z = \sqrt{\sum_{j=1}^{k} Z_j^2}$. Each one of the Z_j is a random variable whose probability density function resembles a Gaussian but however has some degree of asymmetry $\gamma = 2\tau^3/(\sigma^2 + \tau^2)^{3/2}$ (for the exGaussian [10]). In this work we considered $\gamma < 1.7$ (see Figure 1, and Tables 1 and 2). This asymmetry will be implemented considering distributions such as the exGaussians presented in (2). As we have stated above, a random variable following an exGaussian can be simulated by summing a random variable following a Gaussian distribution and another random variable following an exponential decay distribution. We will first simulate k generating exGaussians, each one with a vanishing value of the mean and standard deviations σ_j all equal to one. A total of 10^6 random numbers were generated to obtain the probability distribution of the variable Z. The generating exGaussian random variables (Z_j) will be chosen to have different levels of asymmetry. All fittings are performed by using the non-linear fitting algorithm of Levenberg–Marquardt [23,24]. We used the FORTRAN 90 programming language to make all calculations. The machine epsilon is 2.220446×10^{-16} for the "double precision" real type. The same methodology as in our previous article [9] to study the modified chi distribution for the case of unequal variances in the generating Gaussians was followed. The cases of $k = 3$ and $k = 5$ degrees of freedom are developed hereafter in a general way.
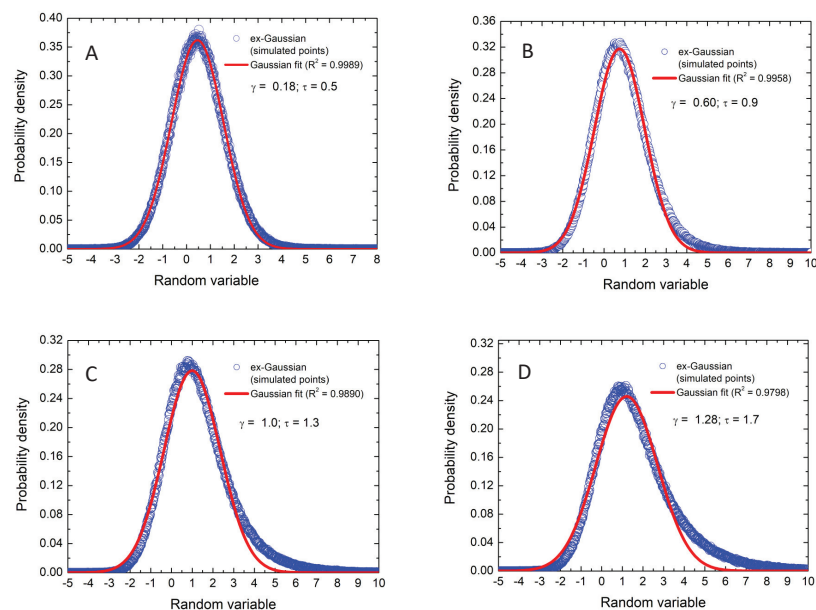
Figure 1. Gaussian fittings to exGaussian functions generated from Monte Carlo simulations. Several increasing values for the asymmetry were chosen, that is, $\gamma = 0.18$ (panel (**A**)), $\gamma = 0.60$ (panel (**B**)), $\gamma = 1.0$ (panel (**C**)), and $\gamma = 1.28$ (panel (**D**)).

2.3. Measurement of Goodness of Fit between Data and Models

Statistical analysis was performed to measure the quality of the proposed model with respect to the data, including both simulations and real health-related variables. The goodness of fit of the model with respect to the data was first studied by the coefficient of determination R^2. It quantifies the percentage of data variance that can be explained by the model [25] with values in the range [0, 1] representing from a null fit to a perfect fit.

In addition, a non-parametric test was assessed to quantify the equality between the two continuous probability distributions under comparison in each case: the ex-Gaussians, and the distributions for the simulated or real data. Kolmogorov–Smirnov (KS) distance is the maximum vertical distance between the cumulative density functions (CDFs) of the simulated/real data, and the model [26,27]. This statistic is sensitive to differences in both location and shape of the CDFs [28,29]. We also checked the associated *p*-value to check whether both distributions can be considered to follow the same distribution (i.e., the null hypothesis is true).

Finally, quantile–quantile (Q-Q) plots are also depicted to study the goodness of fit between data and models [30]. Each point of the plot (x, y) corresponds to one of the quantiles of the first distribution (simulated/real data) which is compared against the same quantile of the second distribution (the model). Thus, points in the Q-Q plot lie approximately on the line $y = x$ when data follow the same distribution. We used these probability plots to confirm that both probability distributions (simulated/real data and the model) had good fitting agreement.

3. Results and Discussion

Figure 1 shows the probability densities of the exGaussian random variables (Z_j) obtained for four different values of the parameter τ and therefore of the asymmetry γ in the generating exGaussians. The coefficient of determination (R^2) of a Gaussian fit (red solid line) has also been included. The larger the value of γ the lower the values of R^2, as expected. For the sake of clarity, quantile–quantile plots of these probability densities are also shown in Figure 2. The good quality of the fittings can be observed since the points lie approximately in a line. This is also remarked with the low values of Kolmogorov–Smirnov

distances, and the *p*-values higher than 0.5 which show strong acceptance of the null
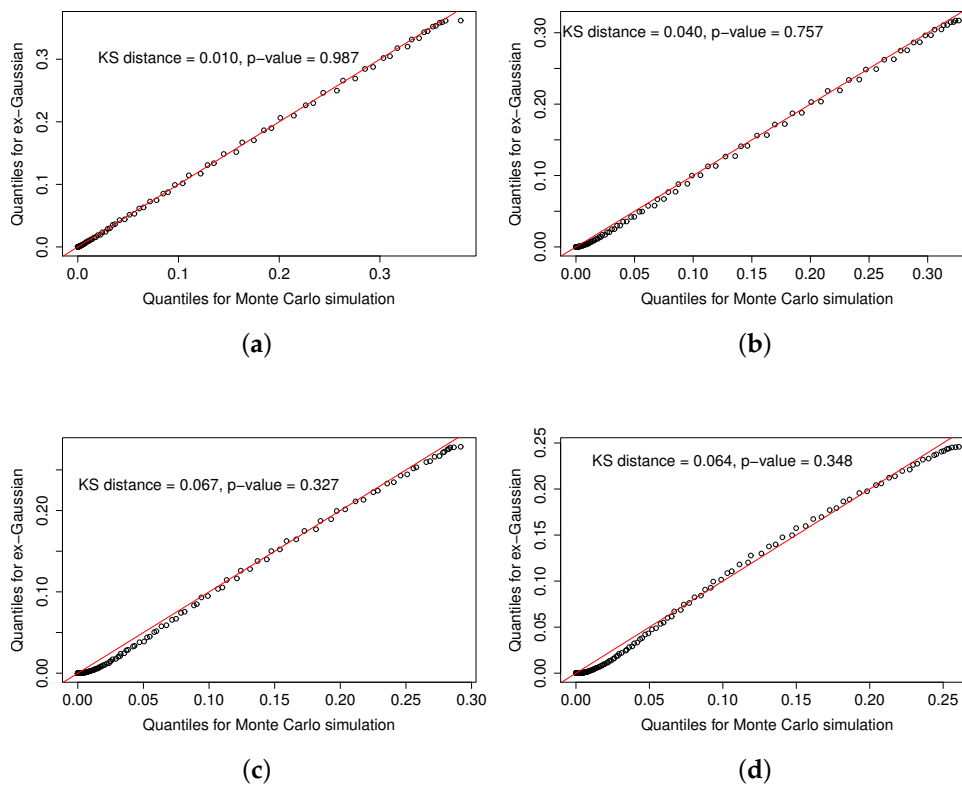hypothesis: data follow the same distribution.

Figure 2. Quantile–quantile plots of the Gaussian fittings to exGaussian functions generated from
Monte Carlo simulations for the different asymmetry values shown in Figure 1, that is, $\gamma = 0.18$
(panel (**a**)), $\gamma = 0.60$ (panel (**b**)), $\gamma = 1.0$ (panel (**c**)), and $\gamma = 1.28$ (panel (**d**)). Each panel also includes
the Kolmogorov–Smirnov distance and the associated *p*-value.

Figure 3 includes the asymmetry of the generating exGaussian random variables (Z_j)
as a function of τ for $\mu = 0$ (mean value) and $\sigma = 1$ (standard deviation). It can be seen
that almost a constant value of the asymmetry γ is reached when τ increases. Values of
asymmetry in the linear region of this curve were chosen for this work ($0.18 < \gamma < 1.67$).
A linear fit for $0.18 < \gamma < 1.67$ yields a coefficient of determination of 0.97.
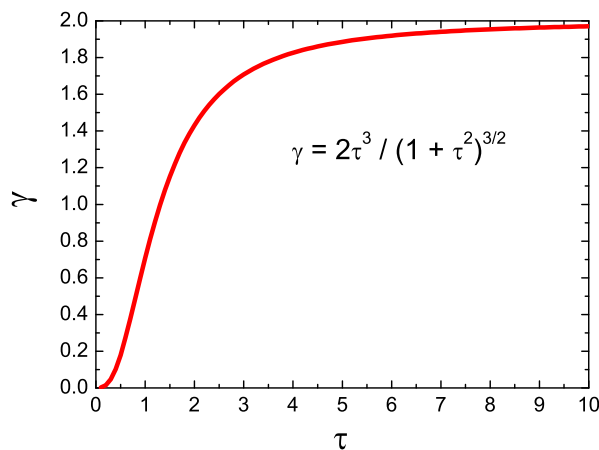

$$\gamma = 2\tau^3 / (1 + \tau^2)^{3/2}$$

Figure 3. Asymmetry of the exGaussian distribution as a function of τ for $\sigma = 1$ and $\mu = 0$.

The resulting distributions of the variable Z for three and five generating exGaussians, and for different values of the asymmetry were fitted by using chi distributions of $k = 3$ and $k = 5$ degrees of freedom, respectively. All the generating exGaussians were centered at and divided by the corresponding mode in order to standardise the resulting distribution. The mean asymmetry is calculated over the asymmetry values within each of the sets shown in Tables 1 and 2. The values of the coefficient of determination R^2 are represented in Figure 4 versus the mean asymmetry for several increasing values of asymmetry. A total of 10^6 random numbers were used to obtain the probability distributions to which chi distributions are fitted.

Table 1. Results for different levels of asymmetry when considering $k = 3$ generating ex-Gaussians. The columns in order show: the set number, the values of τ for each of the generating exGaussians (τ_1, τ_2, and τ_3), the corresponding percentage difference between the smallest and the largest value of τ (e_τ), the mean asymmetry among the three generating exGaussians ($\langle \gamma \rangle$), the calculated parameter (B_{calc}), the fitted parameter (B_{fit}), the mean coefficient of determination ($\langle R^2 \rangle$), and the mean percentage difference between B_{calc} and B_{fit} over the values within the set ($\langle e_B \rangle$).

Set	τ_1	τ_2	τ_3	e_τ (%)	$\langle \gamma \rangle$	B_{calc}	B_{fit}	$\langle R^2 \rangle$	$\langle e_B \ (\%) \rangle$
1	0.5	0.8	1.45	97.44	0.67	5.27	5.00	0.9863	4.84
	0.5	1	1.45	97.44		5.12	4.87		
	0.5	1.2	1.45	97.44		5.26	5.04		
2	0.7	0.9	1.4	66.67	0.73	4.55	4.31	0.9885	4.89
	0.7	1	1.4	66.67		4.14	3.94		
	0.7	1.2	1.4	66.67		4.25	4.07		
3	0.9	0.95	1.2	28.57	0.74	4.14	3.93	0.9891	4.74
	0.9	1	1.2	28.57		3.97	3.79		
	0.9	1.1	1.2	28.57		4.02	3.84		
4	1	1.5	2.8	94.74	1.26	3.44	3.60	0.9462	4.06
	1	2	2.8	94.74		3.26	3.51		
	1	2.5	2.8	94.74		3.47	3.48		
5	1.3	1.5	2.5	63.16	1.32	3.52	3.58	0.9439	3.43
	1.3	2	2.5	63.16		3.32	3.49		
	1.3	2.3	2.5	63.16		3.59	3.72		
6	1.7	1.8	2.1	21.05	1.38	3.02	3.28	0.9412	8.68
	1.7	1.9	2.1	21.05		3.01	3.28		
	1.7	2	2.1	21.05		2.94	3.24		

Table 2. Results for different levels of asymmetry when considering $k = 5$ generating exGaussians. The columns in order show: the set number, the values of τ for each of the generating exGaussians (τ_1, τ_2, τ_3, τ_4, and τ_3), the corresponding percentage difference between the smallest and the largest value of τ (e_τ), the mean asymmetry among the five generating exGaussians ($\langle \gamma \rangle$), the calculated parameter B_{calc}, the fitted parameter (B_{fit}), the mean coefficient of determination ($\langle R^2 \rangle$), and the mean percentage difference between B_{calc} and B_{fit} over the values within the set ($\langle e_B \rangle$).

Set	τ_1	τ_2	τ_3	τ_4	τ_5	e_τ (%)	$\langle\gamma\rangle$	B_{calc}	B_{fit}	$\langle R^2\rangle$	$\langle e_B(\%)\rangle$
1	0.5	0.8	0.9	1.0	1.45	97.44	0.74	4.55	4.50	0.9744	1.61
	0.5	1.0	1.1	1.2	1.45			4.48	4.53		
	0.5	1.2	1.3	1.4	1.45			4.69	4.83		
2	0.7	0.9	1.0	1.2	1.4	66.67	0.80	4.15	4.14	0.9753	1.15
	0.7	1.0	1.1	1.2	1.4			3.91	3.94		
	0.7	1.2	1.3	1.35	1.4			4.05	4.15		
3	0.9	0.95	1.0	1.1	1.2	28.57	0.77	3.85	3.84	0.9788	0.72
	0.9	1.0	1.1	1.15	1.2			3.77	3.80		
	0.9	1.1	1.15	1.16	1.2			3.79	3.84		
4	1.0	1.5	1.6	1.8	2.8	94.74	1.35	3.21	3.68	0.8951	18.01
	1.0	2.0	2.2	2.4	2.8			3.14	3.88		
	1.0	2.5	2.6	2.7	2.8			4.02	4.88		
5	1.3	1.5	1.8	2.0	2.5	63.16	1.39	3.19	3.84	0.8954	22.55
	1.3	2.0	2.1	2.2	2.5			3.05	3.90		
	1.3	2.3	2.35	2.4	2.5			3.30	4.24		
6	1.7	1.8	1.9	2.0	2.1	21.05	1.40	2.89	3.70	0.8979	25.61
	1.7	1.9	1.95	2.0	2.1			2.90	3.72		
	1.7	2.0	2.05	2.08	2.1			2.84	3.73		

A change in the τ values of the generating exGaussians leads to a change in their variances (S^2) as they depend on this parameter as $S^2 = \sigma^2 + \tau^2$ [10]. In Figure 5, the values of B_{fit} (the fitted B parameter of the chi distribution (6)) are compared with the B value as calculated from $B_{calc} = [(S_1 + S_2 + S_3)/3]^2 - \langle\gamma\rangle^2$. This expression is an extended version of the ansatz defined in our previous reference [9], $B = [(S_1 + S_2 + S_3)/3]^2$, but for the case when a small asymmetry is present. It should be noticed that the exGaussian parameters involved in the calculation of B_{calc} (i.e., σ and τ) should be divided by the corresponding mode of each generating Gaussian.

The results shown in Figures 4 and 5 are summarised in Tables 1 and 2, respectively. For very low values of asymmetry (<0.8), the difference between B_{calc} and B_{fit} (i.e., the error e_B) remains very small (<6%) and R^2 is reasonably good (>0.97). However, for asymmetry values larger than 1.2 the results for e_B are higher and R^2 lower, especially for $k = 5$, where R^2 is below 0.9 and e_B higher than 18%.
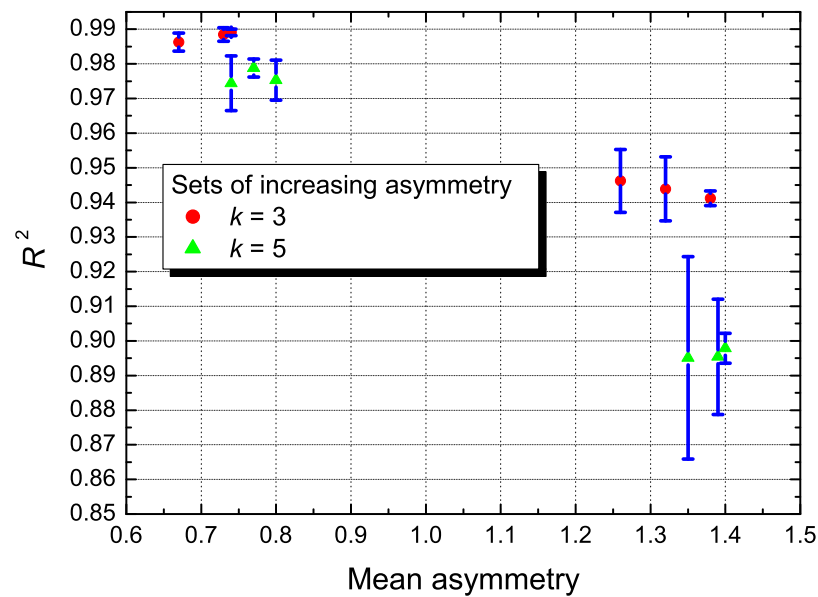
Figure 4. Coefficient of determination (R^2) from fitting a chi distribution (of $k = 3$ and $k = 5$) to the simulated distribution (Z) as a function of the mean asymmetry (within each set defined in Tables 1 and 2).
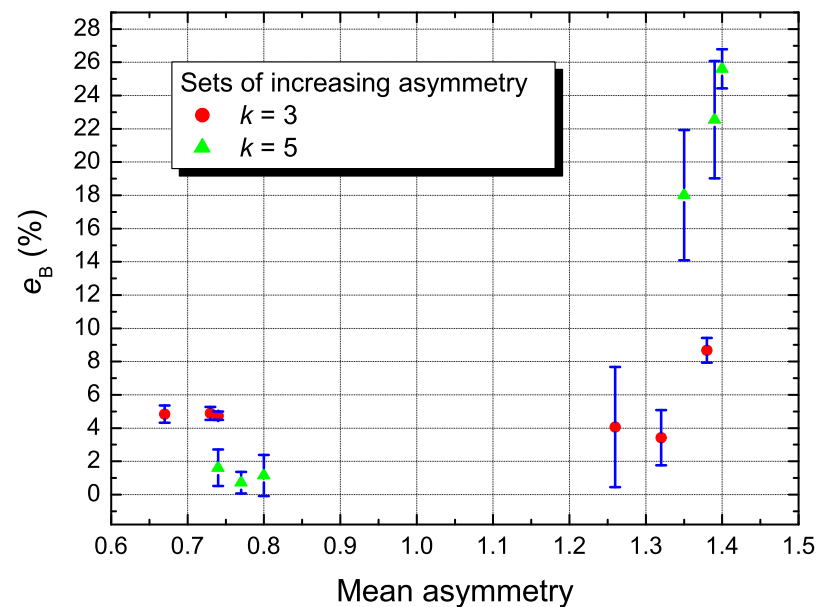

Figure 5. Percentage difference (e_B) between the fitted (B_{fit}) and calculated parameter (B_{calc}) of the chi distribution (of $k = 3$ and $k = 5$) when used to fit the simulated distribution (Z). The error bars represent the standard deviations over the e_B values within each set.

In order to illustrate the presented work with real data, three health-related variables were chosen from the seventh wave of SHARE (Survey of Health, Ageing and Retirement in Europe) (released 17 December 2020) [31–33]. The three variables in this example were: one related to physical health (body mass index–BMI), and two related to cognitive frailty (verbal fluency, measured as the number of animals named within a minute; and short-term memory, measured as the number of words the participant was able to repeat from a 10-word list). We considered a sample formed by 1503 participants from several European

countries. We chose those three variables to illustrate results since they were numeric and not categorical, in which case the proposed representation may be compromised.

The empirical probability distributions of these variables were fitted by exGaussian functions (Figure 6A–C) with a good coefficient of determination. For the sake of clarity, quantile–quantile plots of these probability densities are also shown in Figure 7, where it is depicted that both probability distributions match to a straight line. In addition, the good quality of the fittings can be visually checked with the small Kolmogorov–Smirnov distance values, while *p*-values are higher than 0.05, showing no-significant differences.

Figure 6D shows the probability distribution of the random variable $Z = \left[\sum_{j=1}^{3}(Z_j)^2\right]^{1/2}$ and the corresponding fit to a MB distribution (see (6) for $k = 3$). The variables Z_j stand for body mass index, verbal fluency, and short-term memory, which were standardised [7,9]. The parameters, uncertainties, and coefficients of determination (R^2) from the exGaussian and MB fittings are included in Table 3. This new variable, Z, combines the values of the three health-related variables in a unique value for each individual, which can be considered as a new index able to characterise each individual in the sample. The MB-like distribution in Figure 6D models the probability distribution of Z in the sample. Thus, the entire sample can be modelled by only one parameter, namely, the parameter B of the MB distribution. We illustrated this methodology for three variables but it can be extended for any number of k variables. The methodology developed in this work can have potential applications in diverse areas, for instance, to model health-related [34] and psychological variables [16,35].
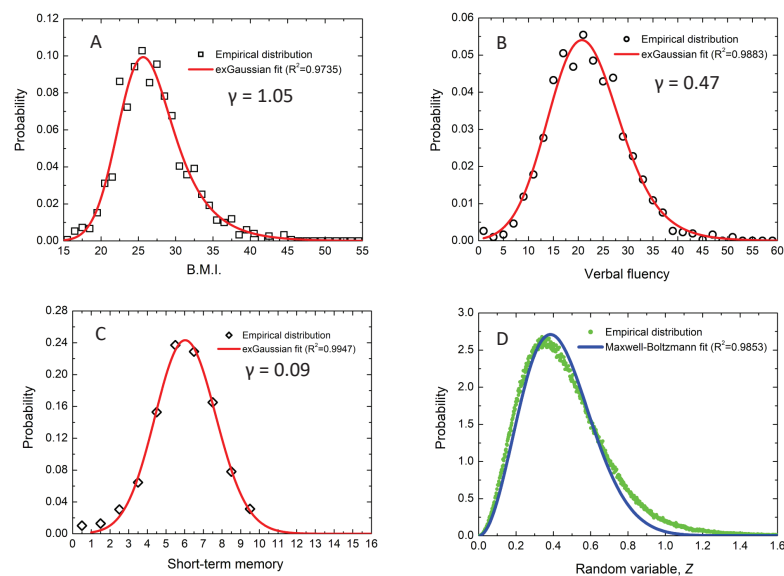

Figure 6. Experimental probability distributions (open symbols) and exGaussian fittings (solid lines) to body mass index (panel (**A**)), verbal fluency (panel (**B**)), and short-term memory (panel (**C**)) variables. The random variable in panel (**D**) represents $Z = \left[\sum_{j=1}^{3} Z_j^2\right]^{1/2}$. The asymmetry values γ (panels **A**–**C**) and the coefficients of determination (R^2) (in all panels) are also included. The corresponding MB curve along the fitting is included in panel (**D**).
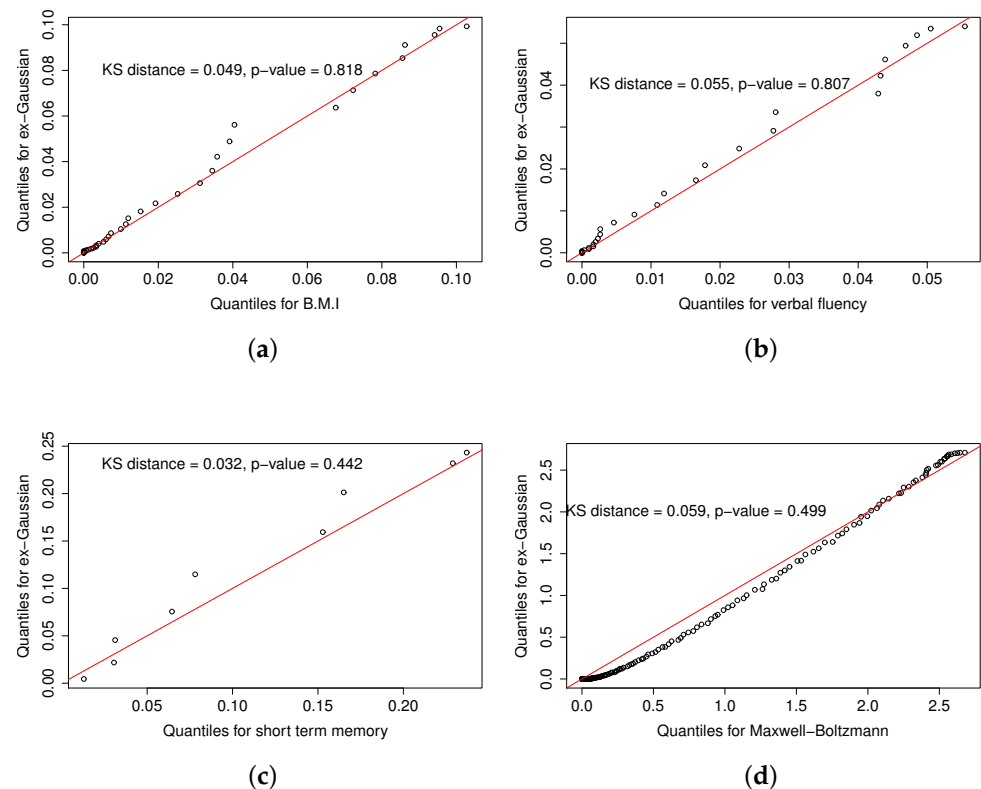
Figure 7. Quantile–quantile plots of the exGaussian fittings and the empirical probability distributions of the three health-related variables shown in Figure 6: (**a**) BMI, (**b**) verbal fluency, (**c**) short-term memory, and the corresponding fit to a MB distribution (**d**). Each panel also shows the Kolmogorov–Smirnov distance and the associated *p*-value.

Table 3. Parameters (μ, σ, and τ), uncertainties ($\Delta\mu$, $\Delta\sigma$, and $\Delta\tau$), and coefficient of determination (R^2) from the exGaussian fitting of the analysed variables. In the last two rows, the results for the MB fitting are included. The fitted parameter (B_{fit}) is compared with the calculated parameter (B_{calc}, with the ansatz introduced in this work) yielding a percentage difference of $e_B = 7.17$ %.

Variable	μ	$\Delta\mu$	σ	$\Delta\sigma$	τ	$\Delta\tau$	R^2
B.M.I	23.274	0.235	2.8314	0.1971	3.8610	0.4428	0.9735
Verbal fluency	17.063	0.485	6.1891	0.2964	4.8408	0.7148	0.9883
Short term mem.	5.4995	0.4108	1.5475	0.1351	0.58533	0.48266	0.9947
	B_{fit}	ΔB_{fit}	R^2	B_{calc}	$e_B(\%)$		
MB parameter	0.07370	0.00025	0.9853	0.07918	7.17		

4. Conclusions

The influence of the asymmetry in the chi distribution was investigated by means of fitting this function to the distribution resulting by taking three and five generating exGaussian functions. The results indicate that, for very small asymmetries ($\gamma < 0.8$) in the generating functions, good values for the coefficient of determination ($R^2 > 0.97$) are still obtained when the simulated distribution is fitted with a chi function for both $k = 3$ and $k = 5$. The results for $k = 3$ are also good for asymmetries larger than 1.2 while they worsen for $k = 5$. We also extend the ansatz proposed in [9] to include small asymmetries. As a practical example to illustrate the results of the Monte Carlo simulations, three health-related non-dichotomic variables (body mass index, verbal fluency and short-term memory) were studied. These variables were combined by taking the square root of the sum

of their squares. The resulting new variable can be fitted by a Maxwell–Boltzmann (MB) distribution. Thus, the entire sample can be characterised by a one-parameter distribution, namely, B. The values of the MB variable can be considered as a new index Z able to characterise each individual in the sample. In this article, we chose three variables but this methodology can be extended to any number of variables that can be combined into a single scalar which is the variable of the resulting chi distribution.

Author Contributions: Conceptualisation and methodology design: N.O., M.O.-P., J.C.C.-P., P.F.d.C. and J.M.I.; data analysis and simulations: N.O., M.O.-P., J.C.C.-P., P.F.d.C. and J.M.I. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Montgomery, D.C.; Runger, G.C. *Applied Statistics and Probability for Engineers*, 6th ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2014.
2. Pearson, K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1900**, *50*, 157–175. [CrossRef]
3. Fisher, R. On the Interpretation of χ^2 from Contingency Tables, and the Calculation of P. *J. R. Stat. Soc.* **1922**, *85*, 87–94 [CrossRef]
4. Fisher, R.A. The Conditions Under Which χ^2 Measures the Discrepancy Between Observation and Hypothesis. *J. R. Stat. Soc.* **1924**, *87*, 442–450
5. Bolboacă, S.D.; Jäntschi, L.; Sestraş, A.F.; Sestraş, R.E.; Pamfil, D.C. Pearson-Fisher Chi-Square Statistic Revisited. *Information* **2011**, *2*, 528–545. [CrossRef]
6. Castro-Palacio, J.C.; Fernández-de-Córdoba, P.; Isidro, J.M.; Navarro-Pardo, E.; Selva-Aquilar, R. Aguilar. Percentile study of Chi distribution. Application to response time data. *Mathematics* **2020**, *8*, 514. [CrossRef]
7. Castro-Palacio, J.C.; Navarro-Pardo, E.; Isidro, J.M.; Sahu, S.; Fernández-de-Córdoba, P. Brain reaction times. Linking individual and collective behavior through Physics modeling. *Symmetry* **2021**, *13*, 451. [CrossRef]
8. Iglesias-Martínez, M.E.; Hernaiz-Guijarro, M.; Castro-Palacio, J.C.; Fernández-de-Córdoba, P.; Isidro, J.M.; Navarro-Pardo, E. Machinery Failure Approach and Spectral Analysis to Study the Reaction Time Dynamics over Consecutive Visual Stimuli: An Entropy-Based Model. *Mathematics* **2020**, *8*, 1979. [CrossRef]
9. Castro-Palacio, J.C.; Isidro, J.M.; Navarro-Pardo, E.; Velázquez-Abad, L.; Fernández-de-Córdoba, P. Monte Carlo Simulation of a Modified Chi Distribution with Unequal Variances in the Generating Gaussians. A Discrete Methodology to Study Collective Response Times. *Mathematics* **2021**, *9*, 77. [CrossRef]
10. Moret-Tatay, C.; Gamermann, D.; Navarro-Pardo, E.; Fernández-de-Córdoba, P. ExGUtils: A Python Package for Statistical Analysis with the ex-Gaussian Probability Density. *Front. Psycol.* **2018**, *9*, 612. [CrossRef]
11. Ratclif, R.; Murdock, B.B. Retrieval processes in recognition memory. *Psychol. Rev.* **1976**, *83*, 190. [CrossRef]
12. Gmehlin, D.; Fuermaier, A.B.M.; Walther, S.; Debelak, R.; Rentrop, M.; Westermann, C.; Sharma, A.; Tucha, L.; Koerts, J.; Tucha, O.; et al. Intraindividual variability in inhibitory function in adults with ADHD. An ex-Gaussian approach. *PLoS ONE* **2014**, *9*, e112298. [CrossRef] [PubMed]
13. Adamo, N.; Hodsoll, J.; Asherson, P.; Buitelaar, J.K.; Kuntsi, J. Ex-Gaussian, frequency and reward analyses reveal specificity of reaction time fluctuations to ADHD and not autism traits. *J. Abnorm. Child Psychol.* **2019**, *47*, 557. [CrossRef] [PubMed]
14. Xavier Castellanos, F.; Sonuga-Barke, E.J.S.; Scheres, A.; Di Martino, A.; Hyde, C.; Walters, J.R. Varieties of Attention-Defficit/Hyperactivity Disorder-Related Intra-Individual Variability. *Biol. Psychiatry* **2005**, *57*, 1416. [CrossRef]
15. Navarro-Pardo, E.; Navarro-Prados, A.B.; Gamermann, D.; Moret-Tatay, C. Differences between young and old university students on a lexical decision task: Evidence through an ex-gaussian approach. *J. Gen. Psychol.* **2013**, *140*, 251–268. [CrossRef] [PubMed]

16. Moret-Tatay, C.; Moreno-Cid, A.; de Lima Argimon, I.I.; Quarti Irigaray, T.; Szczerbinski, M.; Murphy, M.; Vázquez-Martínez, A.; Vázquez-Molina, J.; Sáiz-Mauleón, B.; Navarro-Pardo, E.; et al. The effects of age and emotional valence on recognition memory: An ex-Gaussian components analysis. *Scand. J. Psychol.* **2014**, *55*, 420–426. [CrossRef]

17. Hernáiz-Guijarro, M.; Castro-Palacio, J.C.; Navarro-Pardo, E.; Isidro, J.M.; Fernández-de-Córdoba, P. A probabilistic classification procedure based on response time analysis towards a quick pre-diagnosis of student's attention deficit. *Mathematics* **2019**, *7*, 473. [CrossRef]

18. Li, J.A. Simplified Exponentially Modified Gaussian Function for Modeling Chromatographic Peaks. *J. Chromatogr. Sci.* **1995**, *33*, 568–572. [CrossRef]

19. Gradshteyn, I.; Ryzhik, I. *Table of Integrals, Series and Products*; Academic Press: Boston, MA, USA, 2007.

20. Tolman, R. *The Principles of Statistical Mechanics*; Dover Publications Inc.: New York, NY, USA, 2003.

21. Boltzmann, L. Studies on the balance of kinetic energy between moving material points. *Wiener Berichte* **1868**, *58*, 517–560.

22. Fodor, M.P.; Bolboacă, S.D.; Jäntschi, L. Distribution of Molecules by Kinetic Energy Revisited. *Bull. UASVM Hortic.* **2013**, *70*, 10–18.

23. Levenberg, K. A Method for the Solution of Certain Non-Linear Problems in Least Squares. *Q. Appl. Math.* **1944**, *2*, 164–168. [CrossRef]

24. Marquardt, D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441. [CrossRef]

25. Devore, J.L. *Probability and Statistics for Engineering and the Sciences*; Cengage Learning: Boston, MA, USA, 2011.

26. Conover, W.J. *Practical Nonparametric Statistics*; Wiley: New York, NY, USA, 1999.

27. Kolmogoroff, A. Confidence Limits for an Unknown Distribution Function. *Ann. Math. Stat.* **1941**, *12*, 461–463. [CrossRef]

28. Jäntschi, L. A Test Detecting the Outliers for Continuous Distributions Based on the Cumulative Distribution Function of the Data Being Tested. *Symmetry* **2019**, *11*, 835. [CrossRef]

29. Jäntschi, L. Detecting Extreme Values with Order Statistics in Samples from Continuous Distributions. *Mathematics* **2020**, *8*, 216. [CrossRef]

30. Wilk, M.B.; Gnanadesikan, R. Probability plotting methods for the analysis of data. *Biom. Biom. Trust* **1968**, *55*, 1–17. [CrossRef]

31. Börsch-Supan, A. *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 7*; Release Version: 7.1.1; SHARE-ERIC: München, Germany, 2020.

32. Bergmann, M.; Scherpenzeel, A.; Börsch-Supan, A. *SHARE Wave 7 Methodology: Panel Innovations and Life Histories*; Munich Center for the Economics of Aging (MEA): München, Germany, 2019.

33. Börsch-Supan, A.; Brandt, M.; Hunkler, C.; Kneip, T.; Korbmacher, J.; Malter, F.; Schaan, B.; Stuck, S.; Zuber, S. Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *Int. J. Epidemiol.* **2013**, *42*, 992–1001. [CrossRef]

34. Fernandez-de-las-Penas, C.; Fernandez-Munoz, J.J.; Palacios-Cena, M.; Navarro-Pardo, E. Sleep disturbances in tension-type headache and migraine. *Ther. Adv. Neurol. Disord.* **2017**, *11*, 1–6. [CrossRef]

35. Viguer, P.; Carlos Melendez, J.; Valencia, S.; Cantero, M.-J.; Navarro, E. Grandparent-Grandchild Relationships from the Children's Perspective: Shared Activities and Socialization Styles. *Span. J. Psychol.* **2010**, *13*, 708–717. [CrossRef]