

Article

An Over and Underdispersed Biparametric Extension of the Waring Distribution

Valentina Cueva-López [†], María José Olmo-Jiménez [†] and José Rodríguez-Avi ^{*,†}

Department of Statistics and Operations Research, University of Jaén, 23071 Jaén, Spain; vcueva@ujaen.es (V.C.-L.); mjolmo@ujaen.es (M.J.O.-J.)

* Correspondence: jravi@ujaen.es; Tel.: +34-953212207

† These authors contributed equally to this work.

Abstract: A new discrete distribution for count data called extended biparametric Waring (*EBW*) distribution is developed. Its name is related to the fact that, in a specific configuration of its parameters, it can be seen as a biparametric version of the univariate generalized Waring (*UGW*) distribution, a well-known model for the variance decomposition into three components: randomness, liability and proneness. Unlike the *UGW* distribution, the *EBW* can model both overdispersed and underdispersed data sets. In fact, the *EBW* distribution is a particular case of a *UGW* distribution when its first parameter is positive; otherwise, it is a particular case of a Complex Triparametric Pearson (*CTP*) distribution. Hence, this new model inherits most of their properties and, moreover, it helps to solve the identification problem in the variance components of the *UGW* model. We compare the *EBW* with the *UGW* by a simulation study, but also with other over and underdispersed distributions through the Kullback-Leibler divergence. Additionally, we have carried out a simulation study in order to analyse the properties of the maximum likelihood parameter estimates. Finally, some application examples are included which show that the proposed model provides similar or even better results than other models, but with fewer parameters.

Keywords: count data distribution; goodness of fit; overdispersion; underdispersion



Citation: Cueva-López, V.; Olmo-Jiménez, M.J.; Rodríguez-Avi, J. An Over and Underdispersed Biparametric Extension of the Waring Distribution. *Mathematics* **2021**, *9*, 170. <https://doi.org/10.3390/math9020170>

Received: 3 December 2020
Accepted: 8 January 2021
Published: 15 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The univariate generalized Waring (*UGW*) is a triparametric distribution for overdispersed count data that has been studied by [1–4], among others. The interest of the *UGW* distribution lies in the decomposition of its variance into three components, randomness, liability and proneness, which allows us to get a deeper knowledge of the nature of data variability, that is, how and why data vary. Whereas the Poisson distribution provides the simplest answer to this issue (pure chance), any one-step Poisson mixture distributions assume that there are only two sources of variability (for example, the negative binomial or *NB* distribution which is a Poisson-Gamma mixture).

For this reason, the *UGW* distribution and the related regression model [5–7] have been widely applied for modelling overdispersed count data sets in different fields, such as lexicology [8], the number of authors in scientific articles [9], the evolution of the number of links in the World Wide Web [10], accident theory [11], clustered data [12], sources of variance in motor vehicle crash analysis [13], completeness errors in geographic data sets [14] or agriculture [15].

However, the *UGW* distribution has a serious drawback related to the variance decomposition. Since its first two parameters are interchangeable in the expression of the probability mass function (pmf), it is difficult to determine which component refers to liability or proneness. There are in the literature some suggestions available to avoid this problem. Ref. [1] recommends choosing the values of liability and proneness according to the researcher experience; Ref. [11] proposes the calculus of a bivariate version of the Waring distribution and [5] solves the problem using additional information provided by

covariates through a regression model. In all cases the solution of the indetermination needs external information that is not always available.

Several extensions have been developed, such as the extended Waring distribution or *GHDI* [4], the Stuttering generalized Waring distribution [16] and the bivariate generalized Waring distribution [11], but they do not manage to solve the identification problem aforementioned.

In this paper, we study a specific biparametric distribution within the Gaussian hypergeometric distributions (*GHD*) family [17] and we propose it as an extension of the *UGW* distribution but with only two parameters. The proposed model does not only perform similar to the *UGW* distribution for overdispersed data sets but also solves the identification problem of the variance components. Moreover, the way in which the extension is carried out also allows for modelling underdispersed count datasets, since it can be seen as a particular case of a complex triparametric Pearson (*CTP*) distribution [18,19] although, in this case, the result of decomposition of the variance is not verified because the model cannot be expressed as a Poisson mixture. Thus, this extension—that we will call extended biparametric Waring (henceforward *EBW*) distribution—inherits the good properties of the *UGW* and *CTP* distributions.

The rest of the paper is laid out as follows. Section 2 is devoted to defining the *EBW* distribution and to exploring its main probabilistic properties. In Section 3 some estimation methods are described and the properties of the maximum likelihood estimators are analysed by a simulation study. In Section 4 we compare the *EBW* distribution with some other biparametric over- and underdispersed distributions. Some examples of application to real over- and underdispersed data that illustrate the versatility of the proposed model are included in Section 5. Finally, in Section 6, some conclusions of the current research are presented.

2. The Extended Bivariate Waring Distribution

2.1. Definition

The *GHD* family, generated by the ${}_2F_1(\alpha, \beta; \gamma, \lambda)$ function

$${}_2F_1(\alpha, \beta; \gamma; \lambda) = \sum_{x=0}^{\infty} \frac{(\alpha)_x(\beta)_x}{(\gamma)_x} \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

with $(\gamma)_x = \frac{\Gamma(\gamma+x)}{\Gamma(\gamma)}$, $\alpha, \beta \in \mathbb{C}$ and $\gamma, \lambda \in \mathbb{R}$, arises as a solution of the difference equation

$$G(x)f(x + 1) - L(x)f(x) = 0, \quad x = 0, 1, 2, \dots \tag{1}$$

where G and L are quadratic polynomials with real coefficients, $G(x) = (\gamma + x)(x + 1)$ and $L(x) = \lambda(x - \alpha)(x - \beta)$ [20]. When convergence, positivity and normalization conditions are verified, the solution $f(x)$ is the pmf of as a discrete distribution, that is

$$f(x) = P(X = x) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + x)\Gamma(\beta + x)}{{}_2F_1(\alpha, \beta; \gamma; \lambda)} \frac{\lambda^x}{\Gamma(\gamma + x) x!} \tag{2}$$

and the probability generating function (pgf)

$$G(t) = {}_2F_1(\alpha, \beta; \gamma; \lambda t) / {}_2F_1(\alpha, \beta; \gamma; \lambda), \quad t \in \mathbb{R}.$$

It is important to point out that the first three parameters of the *GHD* are the roots of the polynomials $L(x)$ and $G(x)$ (except the sign of γ).

A thorough classification of the *GHD* family in terms of the parameters can be seen in [4]. In the aforementioned paper, a detailed study of the *GHD* when α, β and γ are positive real numbers and $0 < \lambda \leq 1$ (denoted by type I) is made. The case when α and β are conjugate complex numbers, $\gamma > 0$ and $0 < \lambda \leq 1$ (denoted by type II distributions) has been studied in [18,19,21]. Type VII distributions, a finite case which may be seen as a

generalization of the beta-binomial model, have been addressed by [22]. Likewise, the case when $\lambda = 1$ has been analysed by [23,24], among others.

In this paper, we focus on the case in which both *GHD* of type I and type II with $\lambda = 1$ converge. Thus, $L(x)$ in Equation (1) has a real double root, that is, $\alpha = \beta$. Then, the solution of Equation (1) is given in terms of a ${}_2F_1(\alpha, \alpha; \gamma; 1)$ function leading to a highly versatile biparametric discrete distribution with infinite range which is formalized in the following definition. From now on we will call it *EBW*, the acronym of *Extended Bivariate Waring*, distribution. Later on we will explain the nomenclature chosen for this new distribution.

Definition 1. A random variable X following a $EBW(\alpha, \gamma)$ distribution is defined by the following pmf

$$P(X = x) = \frac{\Gamma(\gamma - \alpha)^2}{\Gamma(\alpha)^2\Gamma(\gamma - 2\alpha)} \frac{\Gamma(\alpha + x)^2}{\Gamma(\gamma + x)} \frac{1}{x!}, \quad x = 0, 1, \dots \tag{3}$$

where $\alpha \in \mathbb{R}$ and $\gamma > \max(0, 2\alpha)$.

The mean, μ , and variance, σ^2 , of X are

$$\mu = \frac{\alpha^2}{\gamma - 2\alpha - 1}, \quad \sigma^2 = \frac{\alpha^2(\gamma - \alpha - 1)^2}{(\gamma - 2\alpha - 1)^2(\gamma - 2\alpha - 2)} = \mu \frac{\mu + \gamma - 1}{\gamma - 2\alpha - 2} \tag{4}$$

so it is necessary that $\gamma > 2\alpha + 1$ and $\gamma > 2\alpha + 2$ to guarantee the existence of μ and σ^2 , respectively. In general, it can be proved that $\gamma > 2\alpha + m$ to guarantee the existence of the m -th raw moment.

2.2. Properties

To study the properties of the *EBW* distribution we will distinguish among $\alpha > 0$ and $\alpha < 0$ but $\alpha \notin \mathbb{Z}^-$.

2.2.1. $\alpha > 0$

It is necessary that $\gamma > 2\alpha$, so we consider another parametrization of the distribution in terms of α and $\rho = \gamma - 2\alpha > 0$. Then, the expression of the pmf given in Equation (3) is now

$$P(X = x) = \frac{\Gamma(\alpha + \rho)^2}{\Gamma(\alpha)^2\Gamma(\rho)} \frac{\Gamma(\alpha + x)^2}{\Gamma(2\alpha + \rho + x)} \frac{1}{x!}, \quad x = 0, 1, \dots \tag{5}$$

and the expressions in Equation (4) reduce to

$$\mu = \frac{\alpha^2}{\rho - 1}, \quad \sigma^2 = \frac{\alpha^2(\alpha + \rho - 1)^2}{(\rho - 1)^2(\rho - 2)} = \mu \frac{\mu + 2\alpha + \rho - 1}{\rho - 2}. \tag{6}$$

To guarantee the existence of the m -th raw moment it is necessary that $\rho > m$.

Theorem 1. The $EBW(\alpha, \rho)$ distribution with $\alpha, \rho > 0$ is a $UGW(\alpha, \alpha, \rho)$ distribution.

Proof. Considering $\alpha = \beta > 0$ and $\lambda = 1$ in Equation (2) and applying that

$${}_2F_1(\alpha, \beta; \gamma; 1) = \frac{\Gamma(\gamma)\Gamma(\gamma - \alpha - \beta)}{\Gamma(\gamma - \alpha)\Gamma(\gamma - \beta)},$$

it is easy to see that the pmf given in Equation (5) coincides with that of a $UGW(\alpha, \alpha, \rho)$ distribution. \square

Hence, our model may be seen as a biparametric case of a *UGW* distribution when $\alpha > 0$. As a consequence, it inherits the properties of the *UGW* distribution which are listed below:

1. It can be obtained from a two-step Poisson mixture:

- $X|\lambda \sim \mathcal{P}(\lambda)$
- $\lambda|\alpha, v \sim \text{Gamma}(\alpha, v)$ with density

$$f(\lambda|\alpha, v) = \frac{1}{\Gamma(\alpha)v^\alpha} \lambda^{\alpha-1} e^{-\lambda/v}, \quad \lambda > 0, \quad \alpha, v > 0$$

Therefore, $X|\alpha, v \sim \text{NB}(\alpha, v)$ with pmf

$$f(x|\alpha, v) = \frac{1}{x!} \frac{\Gamma(x + \alpha)}{\Gamma(\alpha)} \left(\frac{1}{1 + v}\right)^\alpha \left(\frac{v}{1 + v}\right)^x, \quad x = 0, 1, \dots$$

- $v|\alpha, \rho \sim \text{Beta}(\alpha, \rho)$ with density

$$f(v|\alpha, \rho) = \frac{\Gamma(\alpha + \rho)}{\Gamma(\alpha)\Gamma(\rho)} v^{\alpha-1} (1 + v)^{\alpha+\rho}, \quad v > 0, \quad \alpha, \rho > 0$$

2. Since the *EBW* distribution with $\alpha > 0$ is a Poisson mixture, it is always overdispersed.
3. It converges to $\mathcal{P}(\mu)$ when ρ and $\alpha^2 \rightarrow \infty$ with the same order of convergence.
4. As a consequence of the mixture, the variance of X can be split into three components known as randomness, liability and proneness, respectively:

$$\sigma^2 = \frac{\alpha^2}{\rho - 1} + \frac{\alpha^2(\alpha + 1)}{(\rho - 1)(\rho - 2)} + \frac{\alpha^3(\alpha + \rho - 1)}{(\rho - 1)^2(\rho - 2)}. \tag{7}$$

Since we have got rid of one of the first two parameters of the *UGW* distribution, the indetermination problem with regard to the components of the variance [4] disappears in the biparametric model and, therefore, it is not necessary to provide additional information when determining the partition of the variance.

In order to know the effect of each parameter on the variance components of the *EBW* model we consider the proportion of variance explained by each one, that is:

$$1 = \frac{(\rho - 1)(\rho - 2)}{(\alpha + \rho - 1)^2} + \frac{(\alpha + 1)(\rho - 1)}{(\alpha + \rho - 1)^2} + \frac{\alpha}{(\alpha + \rho - 1)}$$

Figure 1 shows the evolution of the variance partition percentages for each component considering α fixed and ρ variable (low and high values) and then, ρ fixed and α variable (low and high values). In the first column (α fixed), we can observe that the greater ρ is, the more important is the proneness. In the second column (ρ fixed), the greater α is, the more important is the randomness. Otherwise, if α and ρ increase with the same convergence order, the proneness has a lower limit in 50% of the variance, whereas the other two parts tend to 25% each one.

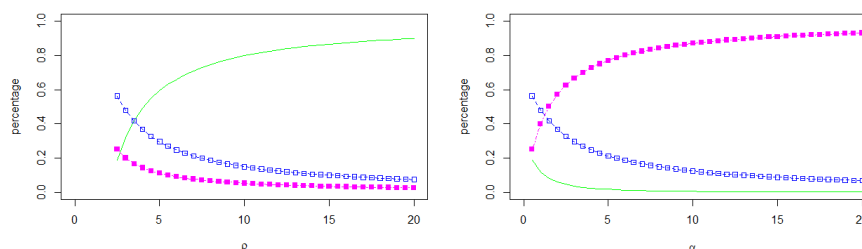


Figure 1. Cont.

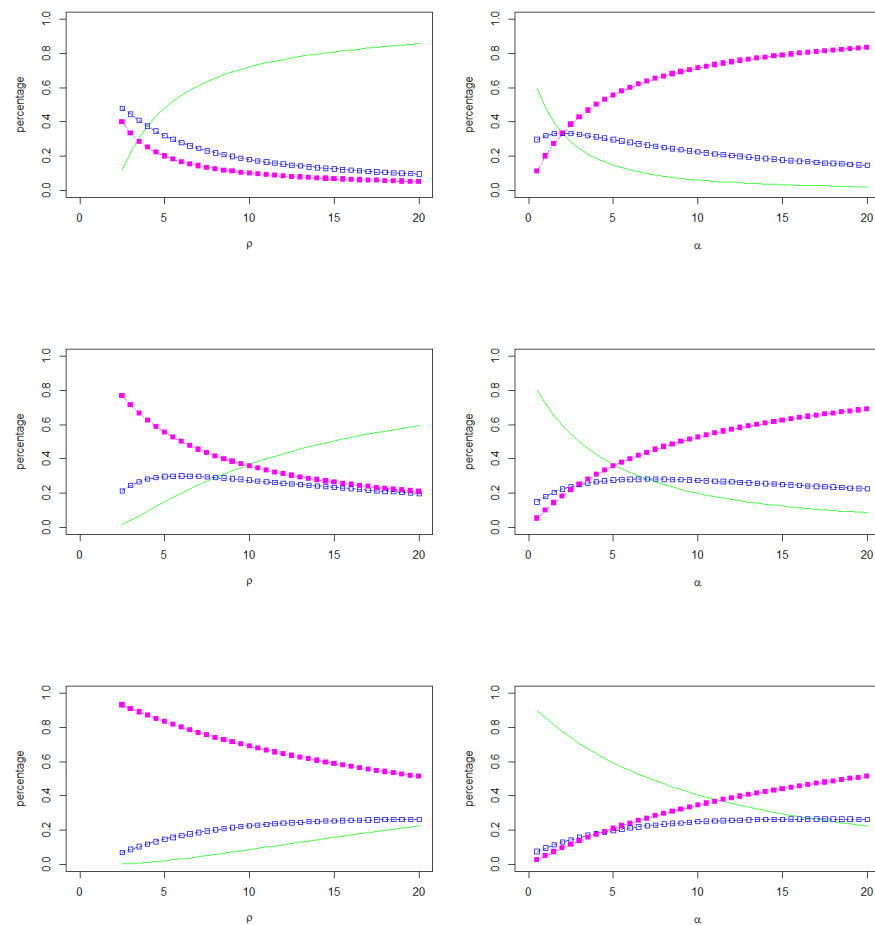


Figure 1. Percentages of the extended biparametric Waring (*EBW*) variance components: randomness (green solid line), liability (blue dashed line) and proneness (purple dotted line). Column 1 has values of $\alpha = 0.5, 1, 5$ and 20 , respectively, for ρ from 2.1 to 20 . Column 2 has values of $\rho = 2.5, 5, 10$ and 20 , respectively, for α from 0.1 to 20 .

Due to the structure of the *UGW* distribution in which the first two parameters are interchangeable and appear in a multiplicative form in the pmf, moments and decomposition of the variance, the maximum likelihood estimates of its first two parameters are usually almost equal. In fact, given a *UGW*(a, k, ρ) distribution, there exists an *EBW*(α, ρ) distribution with $\alpha = \sqrt{ak}$ and the same parameter ρ , that is very close to the former. This can be seen in Figure 2 where the maximum Kullback-Leibler (KL) divergence for more details see [25] between the two models has been calculated for several values of a, k and ρ . We have considered the same scale in all the graphs in order to compare them. We observe that: (1) the divergence increases as k is separated from a ; (2) the difference between a and k is less relevant as ρ increases; (3) in any case, the divergence is very small. Hence, the *EBW* type I distribution has the property of providing in many cases a similar fit but with one more degree of freedom. In general, the *EBW* distribution is able to provide acceptable fits for data simulated from a *UGW* distribution. This implies that, in most cases, there exists a *EBW* model reasonably similar to the *UGW*, but with the advantage of having fewer parameters. To show this fact we have simulated $M = 1000$ samples of size $N = 100, 300$ and 500 from a *UGW* distribution with several values of its parameters and, for each sample, we have obtained the corresponding *EBW* and *UGW* fits. All the estimates have been computed by the maximum likelihood method. We have implemented our own functions in R [26]: the pmf of the *EBW* and *UGW* distributions and the fitting function for

both models. To do the latter, we have used the optim function of the stats package with the L-BFGS-B method [27], since it allows box constraints, and considering as initial values the estimates provided by the method of moments (see Section 3 for more details).

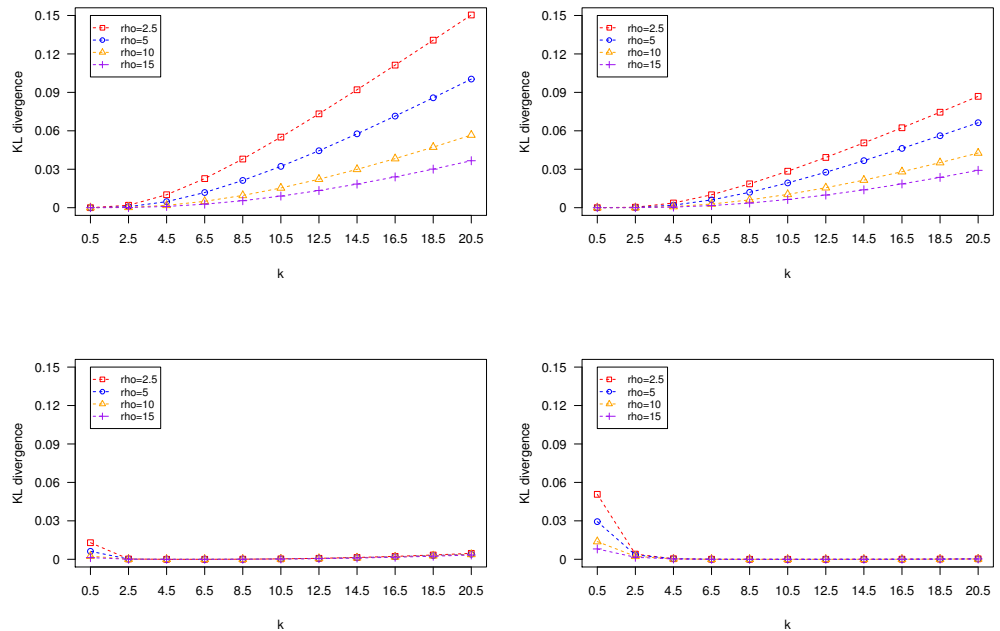


Figure 2. Maximum Kullback Leibler divergence between the univariate generalized Waring ($UGW(a, k, \rho)$) distribution and the $EBW(\sqrt{ak}, \rho)$ distribution with $a = 0.5, 1, 5$ and 10 (from left to right).

For each group of 1000 samples we have computed the percentage of EBW fits achieved as well as the percentage of these fits which are better than the corresponding UGW fit in two senses: the AIC value and the χ^2 -goodness of fit test. Specifically, for the EBW fits achieved we have computed the percentage of them whose AIC value is less than that of the UGW fit and the percentage of p -values in the χ^2 -goodness of fit test greater than 0.01, 0.05 and 0.1 (that is, the null hypothesis *data comes from a EBW model* cannot be rejected). These results appear in Tables 1 and 2. We have also carried out the Kolmogorov-Smirnov test for discontinuous distributions [28] using the ks.test function of the dgof package in R, but in all the cases the p -values are greater than 0.1.

Table 1. Percentage of: (1) EBW fits achieved for UGW generated data; (2) EBW fits with less AIC value than the corresponding UGW fit.

	Achieved EBW Fits			<AIC		
	100	300	500	100	300	500
$UGW(a, k, \rho)$						
(0.5, 1, 2.5)	94.4	93.4	95.9	92.6	89.3	86.4
(0.5, 10, 2.5)	99.8	100	100	53.4	25.1	9.9
(0.5, 10, 20)	95.2	93.8	94.4	99.1	95.5	89
(1.5, 3, 2.5)	100	100	100	88.3	88.8	87.2
(1.5, 3, 25)	97.9	99	98.5	100	100	99.9
(1.5, 20, 25)	100	100	100	96.9	82.3	74.3
(4, 6, 2.5)	99.9	100	100	90	89.8	91.1
(4, 6, 10)	100	100	100	96.3	91.1	92.3
(4, 6, 50)	95.8	96.5	97.6	100	99.9	99.8

Table 2. Percentage of samples that come from a *EBW* model at the 1%, 5% and 10% significance levels according to the χ^2 -goodness of fit test.

	<i>p</i> -Value > 0.01			<i>p</i> -Value > 0.05			<i>p</i> -Value > 0.1		
	<i>N</i>			<i>N</i>			<i>N</i>		
<i>UGW</i> (<i>a, k, ρ</i>)	100	300	500	100	300	500	100	300	500
(0.5, 1, 2.5)	98.9	97.9	97.4	95.4	94.1	93.3	91.6	91	88.8
(0.5, 10, 2.5)	92.9	93	95.9	86.1	86.8	90.5	80.5	81.6	86.1
(0.5, 10, 20)	99.2	98.6	97.2	95.4	95.3	93.3	89.2	89.6	87.1
(1.5, 3, 2.5)	95.6	92.3	91.2	91.2	87.5	84.3	86.4	83.3	77.6
(1.5, 3, 25)	100	97.8	98.1	95.9	94.3	93.7	88.9	87.6	87.6
(1.5, 20, 25)	99	98.5	98.5	95.5	94.3	92.9	90.9	89.7	88.9
(4, 6, 2.5)	88.7	83.9	81.2	82.8	77	73.9	78	72.7	70
(4, 6, 10)	97.8	96.8	95.4	93.7	91.4	91.4	90.3	86.3	87.8
(4, 6, 50)	98.1	96.1	96.6	94	91	91.8	89.8	87.3	87

2.2.2. Case $\alpha < 0$ but $\alpha \notin \mathbb{Z}^-$

It can be seen as a particular case of a *CTP* distribution [18,21], which arises when the polynomial $L(x)$ in Equation (1) has conjugate complex roots $\alpha = a + ib$ and $\beta = a - ib$; specifically, we have the following result.

Theorem 2. *If $\alpha < 0$ the $EBW(\alpha, \gamma)$ distribution with $\gamma > 0$ is a $CTP(\alpha, 0, \gamma)$ distribution.*

Proof. The proof is straightforward since the pmf of the $CTP(\alpha, 0, \gamma)$ with $\alpha \in \mathbb{R}$ and $\gamma > 0$, see for instance [21], coincides with the pmf of the $EBW(\alpha, \gamma)$ given in Equation (3). □

This result is also true when $\alpha > 0$. So, a $UGW(\alpha, \alpha, \rho) \equiv CTP(\alpha, 0, 2\alpha + \rho)$.

At this point we can justify the name chosen for the model proposed. On the one hand, when $\alpha > 0$ the model may be seen as a biparametric case of the *UGW* distribution, which is always overdispersed and that may replace it with fewer parameters; on the other hand, when $\alpha < 0$ the model can be underdispersed, so it may be considered as an underdispersed extension of a biparametric *UGW* distribution.

Once again, the proposed distribution inherits the properties of another distribution, in this case of the *CTP* distribution that we next summarize:

1. If $\frac{(\alpha-1)^2}{\gamma-2\alpha+1} \in \mathbb{Z}$, the distribution has two consecutive modes in the values

$$\frac{(\alpha - 1)^2}{\gamma - 2\alpha + 1} - 1 = \frac{\alpha^2 - \gamma}{\gamma - 2\alpha + 1} \quad \text{and} \quad \frac{(\alpha - 1)^2}{\gamma - 2\alpha + 1}.$$

Otherwise the distribution is unimodal with mode in 0 if $\alpha^2 < \gamma$ or in $\lceil \frac{(\alpha-1)^2}{\gamma-2\alpha+1} \rceil$, where $\lceil \cdot \rceil$ symbolises the integer part. Hence, the pmf is *J*-shaped or bell-shaped.

2. It may be underdispersed, equidispersed or overdispersed. Specifically:
 - It is underdispersed when $\alpha \leq -1$ or when $-1 < \alpha < -0.5$ and $\gamma > \frac{3\alpha^2 + 4\alpha + 1}{2\alpha + 1}$.
 - It is equidispersed when $-1 < \alpha < -0.5$ and $\gamma = \frac{3\alpha^2 + 4\alpha + 1}{2\alpha + 1}$.
 - It is overdispersed when $\alpha \geq -0.5$ or when $-1 < \alpha < -0.5$ and $\gamma < \frac{3\alpha^2 + 4\alpha + 1}{2\alpha + 1}$.
3. A sufficient condition to be infinitely divisible (i.d.) is that $\alpha > -0.5$ and $\gamma > \alpha^2 / (1 + 2\alpha)$. So, if $\alpha \leq -0.5$ the *EBW* distribution is not i.d. As a consequence, an underdispersed *EBW* cannot be i.d. since a necessary condition to be underdispersed is $\alpha < -0.5$.

4. It converges to the:

- $\mathcal{P}(\mu)$ when γ and $\alpha^2 \rightarrow \infty$ with the same order of convergence.
- Normal distribution, $\mathcal{N}(\mu, \sigma)$, when γ and α have the same order of convergence.

The CTP distribution cannot be expressed as a mixture, so in the EBW with $\alpha < 0$ there is no a result of variance decomposition.

3. Estimation

3.1. Methods for Obtaining Estimators

We can estimate the two parameters of the EBW distribution using the method of moments and the maximum likelihood estimation method.

To apply the method of moments, we first solve the equations given in Equation (4). To this end we substitute $\gamma - 2\alpha - 1 = \alpha^2/\mu$ in the equation of σ^2 . Then,

$$\sigma^2 = \mu \frac{\mu^2 + \alpha^2 + 2\alpha\mu}{\alpha^2 - \mu},$$

which is equivalent to $\alpha^2(\sigma^2 - \mu) - 2\mu^2\alpha - \mu(\mu^2 + \sigma^2) = 0$. Then, replacing μ and σ^2 by their sample counterparts, \bar{x} and s^2 , and solving the equation there are two possible estimates for α by the method of moments:

$$\begin{aligned} \hat{\alpha}_1 &= \frac{\bar{x}^2 + \sqrt{\bar{x}^4 + \bar{x}(s^2 - \bar{x})(\bar{x}^2 + s^2)}}{s^2 - \bar{x}} \\ \hat{\alpha}_2 &= \frac{\bar{x}^2 - \sqrt{\bar{x}^4 + \bar{x}(s^2 - \bar{x})(\bar{x}^2 + s^2)}}{s^2 - \bar{x}} \end{aligned}$$

It is clear that if data exhibit overdispersion, then $\hat{\alpha}_1 > 0$ and $\hat{\alpha}_2 < 0$. On the other hand, if data are underdispersed both $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are negative. Estimated α , the estimate of γ is calculated as $\hat{\gamma} = \hat{\alpha}^2/\bar{x} + 2\hat{\alpha} + 1$. Hence, there are also two possible estimates for γ with the only restriction of being positive, which it is true when:

- $0 < \bar{x} < 1$ or
- $\bar{x} > 1$ and $\hat{\alpha} < -\bar{x} - \sqrt{\bar{x}(\bar{x} - 1)}$ or $\hat{\alpha} > -\bar{x} + \sqrt{\bar{x}(\bar{x} - 1)}$.

Using the MLE method we have to maximize the log-likelihood function. Thus, if $\mathbf{x}' = (x_1, \dots, x_n)$ is a sample of size n , the expression of the log-likelihood function is:

$$\begin{aligned} \ln L_{x_1, \dots, x_n}(\alpha, \rho) &= \sum_{i=1}^n [2 \ln \Gamma(\alpha + x_i) - \ln \Gamma(\rho + 2\alpha + x_i)] \\ &\quad - n[2 \ln \Gamma(\alpha) - 2 \ln \Gamma(\rho + \alpha) + \ln \Gamma(\rho)], \end{aligned} \tag{8}$$

when $\alpha > 0$, using the parametrization given in Section 2.2.1, or

$$\begin{aligned} \ln L_{x_1, \dots, x_n}(\alpha, \gamma) &= \sum_{i=1}^n [2 \ln \Gamma(\alpha + x_i) - \ln \Gamma(\gamma + x_i)] \\ &\quad - n[2 \ln \Gamma(\alpha) - 2 \ln \Gamma(\gamma - \alpha) + \ln \Gamma(\gamma - 2\alpha)], \end{aligned} \tag{9}$$

in another case. Both expressions can be maximized using numerical methods. In particular, we have used the $L - BFGS - B$ method implemented in the optim function of the MASS package in R. This method allows box constraints on the parametric space, so we can impose $\rho > 0$ or $\gamma > 0$ in Equations (8) and (9), respectively. We consider the estimates obtained by the method of moments as initial values, in such a way that we maximize Equation (8) if $\hat{\alpha} > 0$ or Equation (9) in another case.

3.2. Properties of the Estimators

We have carried out a simulation study in order to analyse the performance of the estimates of the model parameters. Specifically, we have simulated $M = 1000$ samples of size $N = 500$ of the *EBW* distribution and we have fitted the *EBW* model for each sample using the MLE method described in the previous section.

We have considered two scenarios: $\alpha > 0$, in which case the *EBW* distribution is always overdispersed, and $\alpha < 0$, in which case the *EBW* distribution can be under- and overdispersed. In all cases the values of the parameters satisfy the conditions for the existence of μ and σ^2 .

Results of the simulation procedure are shown in Table 3. Thus, Column 1 contains the mean bias and the s.d., in brackets (* indicates a significant bias at 5% level based on a normal 95% confidence interval, given that there are 1000 observations). Column 2 shows the average of the mean square error (MSE) of the parameter estimates and Column 3 the percentage of simulations in which the parameter estimate does not differ significantly at 5% from the true value, known as coverage.

We have only included low values of ρ and γ because the higher these values are compared with α , the lower the mean and the variance are. In fact, if these parameters tend to infinity, holding α fixed, the *EBW* distribution degenerates into 0. In addition, if both α and ρ (or γ) are high, the *EBW* is similar to the Poisson or the Normal distribution.

In general, we can deduce that:

- If $\alpha > 0$ the estimates are biased to the right, but the bias decreases as α increases, holding ρ fixed. The opposite happens with the bias of ρ , which increases as ρ increases.
- If $\alpha < 0$ the estimates are also biased, those for α to the left and for γ to the right, but the bias disappears as α decreases ($\alpha < -1$). Holding γ fixed, the bias decreases as α decreases and the same happens for γ .
- The average MSE is low for both parameter estimates, although this measure increases as ρ (or γ) increases since the estimates accuracy and precision decrease.
- Regarding the coverage, it approaches 95%, the confidence level considered, so it shows the validity of the inference made.

Table 3. Mean bias and s.d. in brackets (* indicates a statistically significant bias at 5% level), average mean square error (MSE) and coverage for *EBW* fits.

$\alpha > 0$				$\alpha < 0$			
Parameters	Bias (s.d.)	MSE	Coverage	Parameters	Bias (s.d.)	MSE	Coverage
$\alpha = 0.5$	0.02 (0.10) *	0.02	96.3	$\alpha = -0.75$	-0.01 (0.05) *	0.00	95.1
$\rho = 2.1$	0.23 (0.80) *	1.30	94.8	$\gamma = 0.75$	0.02 (0.11) *	0.02	95.2
$\alpha = 1$	0.01 (0.10) *	0.02	95.5	$\alpha = -0.75$	-0.01 (0.08) *	0.01	97.1
$\rho = 2.1$	0.07 (0.36) *	0.27	96.3	$\gamma = 1.5$	0.07 (0.35) *	0.26	95.3
$\alpha = 1.5$	0.01 (0.13) *	0.03	94.2	$\alpha = -0.75$	-0.02 (0.14) *	0.05	96.8
$\rho = 2.1$	0.04 (0.29) *	0.16	94.5	$\gamma = 3$	0.29 (1.34) *	5.01	94
$\alpha = 1.5$	0.01 (0.14) *	0.04	94.4	$\alpha = -1.5$	0.01 (0.09)	0.01	94.2
$\rho = 2.5$	0.06 (0.39) *	0.30	94.4	$\gamma = 0.75$	0.00 (0.13)	0.03	94.1
$\alpha = 1.5$	0.03 (0.29) *	0.06	96.1	$\alpha = -2.5$	-0.00 (0.10)	0.02	94.5
$\rho = 3.5$	0.14 (0.70) *	1.00	96.6	$\gamma = 0.75$	0.01 (0.20)	0.07	94.2

4. Comparison with Other Count Data Distributions

Next we study the differences and similarities between the *EBW* and other well-known biparametric discrete distributions for count data using again the KL divergence. Specifically, we consider the distributions *NB*, Complex Biparametric Pearson or *CBP* [19], which is a particular case of the *CTP* distribution, Generalized Poisson or *GP* [29,30], COM-Poisson or *CMP* [31,32] and Hyper-Poisson or *HP* [33]. The first two are suitable

only for overdispersed data, whereas the other three can cope with both underdispersed and overdispersed data, although the *GP* has finite range in the underdispersed case.

We focus on the overdispersed scenario since the underdispersed one, for being the *EBW* distribution a particular case of the *CTP* distribution, was already carried out by [21].

To compute the KL divergence between the *EBW* distribution and the above-mentioned distributions (and vice versa), we have considered several values of μ and σ^2 , with $\sigma^2 > \mu$, and then we have obtained the corresponding values of the parameters of each distribution (see Appendix A). For the *CMP* and *HP* distributions it should be taken into account that not all the combinations of μ and σ^2 are possible; empirically there seems to be an upper limit for σ^2 in $\mu(\mu + 1)$. Thus, the values of the KL divergence are shown in Figure 3.

In general, we can observe that in an overdispersed scenario the most distant models from the *EBW* distribution are the *CBP* and *HP* distributions and the closest ones to the *EBW* distribution are the *GP* and *NB* distributions. On the other hand, in an underdispersed scenario the *HP* distribution, which is very similar to the *CMP* distribution, is the closest one [21]. Nevertheless, these distances in relation to the *EBW* distribution are really small, which implies that the performance of these distributions is very similar.

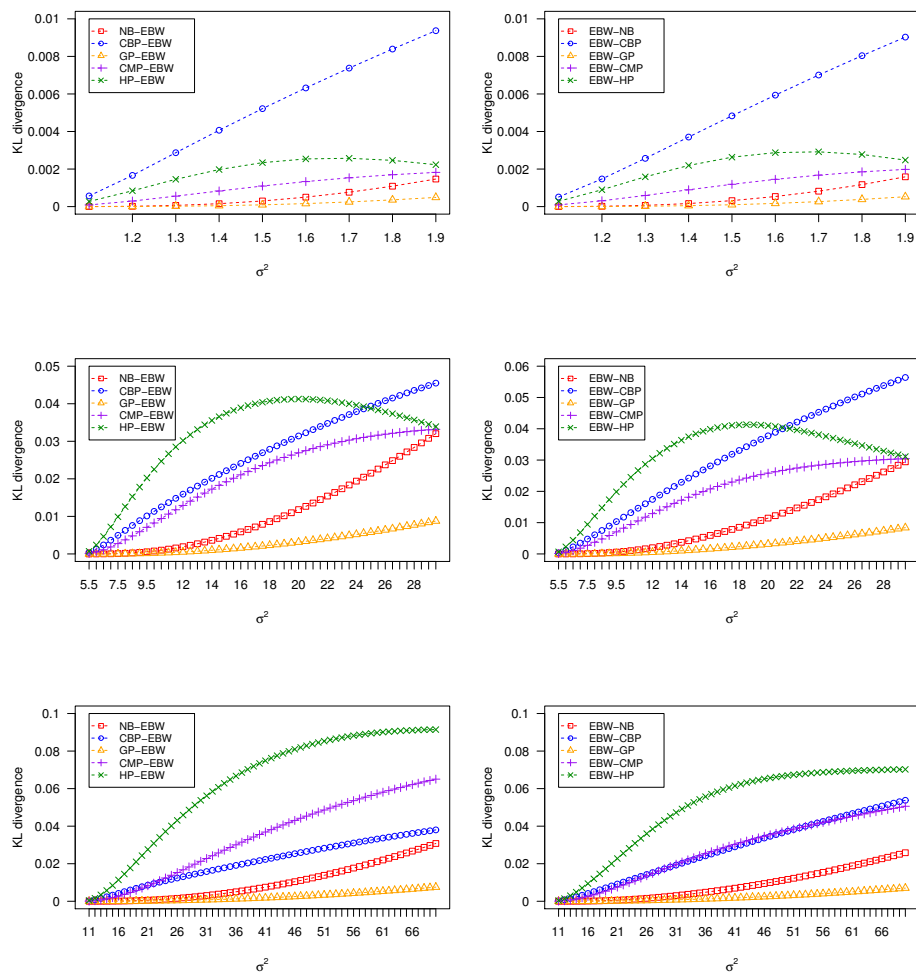


Figure 3. Kullback Leibler divergence between the negative binomial (*NB*), Complex Biparametric Pearson (*CBP*), Generalized Poisson (*GP*), COM-Poisson (*CMP*), Hyper-Poisson (*HP*) and *EBW* distributions (and vice versa) in an overdispersed scenario. Rows 1–3 have $\mu = 1, \mu = 5$ and $\mu = 10$, respectively.

5. Examples

In this section we use the *EBW* distribution to fit both over- and underdispersed real data and we compare this fit with those obtained from other discrete distributions.

5.1. Overdispersed Data: Number of Some Sports Facilities by Municipality in Andalusia, 2015

We consider the variable *X*: number of some sports facilities by municipality in Andalusia, in 2015. Data have been directly obtained from the System of Multiterritorial Information of Andalusia (SIMA) of the Junta de Andalucía [34]. This category includes all sports facilities in the municipality except sports complexes, sports courts, pelota courts (frontons) and pools. A description of these data appears in Table 4, which contains the mean, variance, Aggregation Index (AI), quartiles and maximum.

Table 4. Descriptive statistics for data in examples.

	\bar{x}	s^2	AI	Q ₁	Q ₂	Q ₃	Max
Facilities	3.57	25.89	7.26	1	2	4	55
Syllables −1	1.58	1.17	0.74	1	2	2	5

We will model these data by the following distributions: *EBW*, *NB*, *GP*, *CBP*, *UGW* and *CMP*. *AIC* values, statistics and *p*-values corresponding to the χ^2 -goodness of fit test are shown in Table 5. We can see that the best fit is that provided by the *EBW* distribution. The Wald test supports this statement since the null hypothesis $a = k$ (the first two parameters of the *UGW* distribution are equal) cannot be rejected: the statistic value is -1.31×10^{-6} and the corresponding *p*-value is 1. With the likelihood ratio test (*LRT*) we come to the same conclusion ($LRT = 2.36 \times 10^{-10}$ and *p*-value $\simeq 1$).

Table 5. *AIC* values and χ^2 -goodness of fit test for data about some sports facilities.

	<i>EBW</i>	<i>NB</i>	<i>GP</i>	<i>CBP</i>	<i>UGW</i>	<i>CMP</i>
<i>AIC</i>	3532.7	3579.2	3548.7	3555.1	3534.7	3579.86
χ^2	10.981	38.407	18.897	36.047	10.98	37.620
d.f.	15	15	15	16	14	15
<i>p</i> -value	0.7540	0.0008	0.2184	0.0029	0.6876	0.0010

Table 6 shows the observed and expected frequencies for the *EBW*, *NB*, *GP* and *CBP* fits. Figure 4 shows graphically the frequencies for the values between 0 and 10 sport facilities. We can see that, in general, the *EBW* distribution fit is really accurate (the greater Pearson residual is 1.55 for the interval 18–21). In the other side, the remaining distributions considered provide worse fits, in special in reference to the lowest values of the variable (high Pearson residuals in 0, 1 and 2).

Additionally, we can calculate the three components of the variance. The percentage of data variability due to randomness, liability and proneness is 14.25%, 32.83% and 52.92%, respectively. We can observe that randomness does not play a very important role with respect to the total variability of data and that the most important component is proneness, which refers to specific and internal conditions instead of general conditions of the municipality (external), although liability is also remarkable. The idiosyncrasy of a municipality explains more than 50% of the variability in the number of some sports facilities, whereas shared conditions have less influence, but also noteworthy, on this variability.

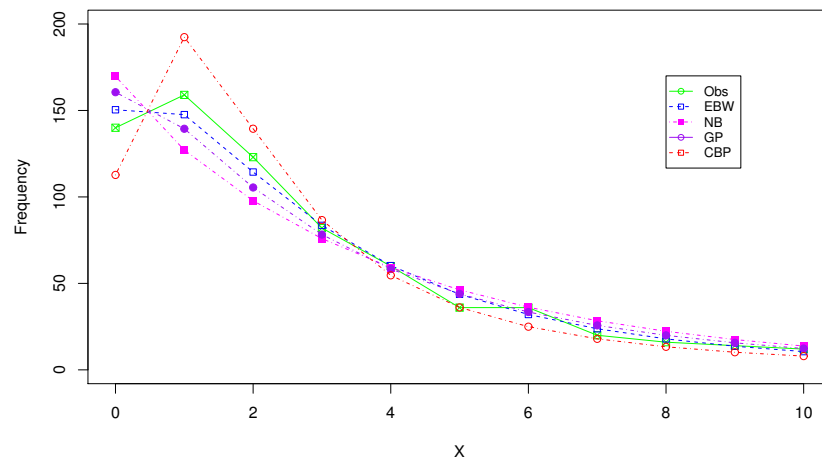


Figure 4. Observed and expected frequencies for data about other sports facilities (from 0 to 10).

Table 6. Observed and expected frequencies for data about other sports facilities.

X	Observed	Expected			
		EBW	NB	GP	CBP
0	140	150.41	169.71	160.56	112.72
1	159	147.58	127.08	139.42	192.40
2	123	114.40	97.78	105.46	139.44
3	82	83.53	75.90	78.25	86.61
4	60	60.27	59.18	58.32	54.66
5	36	43.68	46.27	43.91	36.11
6	36	32.02	36.23	33.43	24.96
7	20	23.78	28.41	25.71	17.94
8	16	17.90	22.30	19.96	13.33
9	14	13.66	17.51	15.63	10.18
10	12	10.56	13.76	12.32	7.95
11	4	8.26	10.82	9.78	6.34
12	9	6.53	8.51	7.81	5.14
13	6	5.21	6.70	6.27	7.75
14	5	7.63	5.27	5.05	
15	5		7.42	7.42	5.49
16	4	5.12			
17	1				
18	0		6.20	6.76	5.67
19	2	6.07			
20	1		5.95	8.94	
21	0				5.72
22–24	2				
25–33	4	8.40			5.11
34–55	4				7.48
		$\hat{\alpha} = 3.147$ (0.181)	$\hat{\theta} = 0.948$ (0.064)	$\hat{\lambda} = 1.535$ (0.059)	$\hat{b} = 1.727$ (0.097)
		$\hat{\rho} = 3.800$ (0.383)	$\hat{\mu} = 3.566$ (0.151)	$\hat{\theta} = 0.570$ (0.018)	$\hat{\gamma} = 1.746$ (0.141)

5.2. Underdispersed Data: Turkish Poem

We consider data about the word length (in terms of number of syllables) in the turkish poem *Gidisat* by Ercüment Behzat Lâv available in [35]. Following these authors, the count for 1 is treated as a count for 0, and in general the count for the response variable X is treated as $X - 1$, as though the data are generated by adding 1 to the distribution. These data exhibit underdispersion with a variance-mean ratio of 0.74 (see Table 4). Table 7 contains the parameter estimates, their standard errors (in parenthesis), the AIC , the observed and expected frequencies and the corresponding Pearson χ^2 test for each one of the models

that copes with underdispersion, that is, *EBW*, *CTP*, *CMP* and *HP* (the *GP* distribution has been excluded because it is of finite range).

CTP and *EBW* fits provide practically the same results. In fact, $b = 0$ using the Wald test ($z_{exp} = 2.3 \times 10^{-5}$ and p -value ≈ 1) and the LRT ($\chi^2_{exp} = 0$ and p -value = 1). Observed and expected frequencies for each fit are represented in Figure 5 (the *CTP* distribution has been suppressed). Although the three fits are very similar and really good, the *EBW* distribution fit is the most accurate taking into account the expected frequencies.

Table 7. Parameter estimates, standard errors (in parenthesis), observed and expected frequencies, *AIC* and χ^2 test for fits to data about the word length of a Turkish poem.

<i>X</i>	<i>Observed</i>	<i>Expected</i>			
		<i>EBW</i>	<i>CTP</i>	<i>CMP</i>	<i>HP</i>
1	64	61.24	61.24	59.69	61.20
2	131	136.23	136.23	141.87	145.24
3	122	121.68	121.68	118.70	112.60
4	61	56.93	56.93	53.92	52.17
5	13	15.27	15.27	15.88	17.27
≥ 6	3	2.66	2.66	3.94	5.55
		$\hat{\alpha} = -10.530$ (2.144)	$\hat{\alpha} = -10.530$ (2.158)	$\hat{\lambda} = 2.377$ (0.276)	$\hat{\gamma} = 0.485$ (0.099)
		$\hat{\gamma} = 49.843$ (24.257)	$\hat{b} = 0.001$ (14.254)	$\hat{v} = 1.506$ (0.137)	$\hat{\lambda} = 1.151$ (0.104)
			$\hat{\gamma} = 49.843$ (24.416)		
	<i>AIC</i>	1158.3	1160.3	1160.7	1164.4
	χ^2 - statistic	1.000	1.000	2.914	6.014
	<i>p</i> -value	0.801	0.606	0.405	0.111

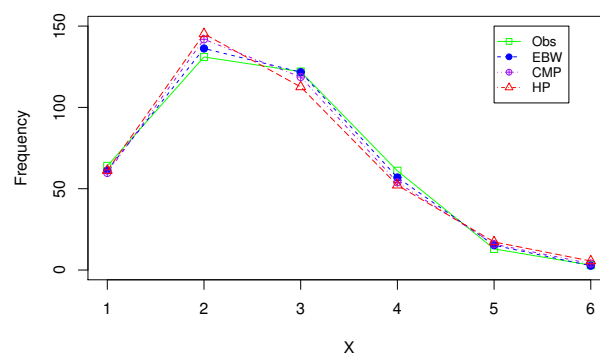


Figure 5. Observed and expected frequencies for data about the number of syllables of a Turkish poem.

6. Conclusions

The *EBW* distribution is a very flexible biparametric discrete distribution that allows for modelling a wide variety of over and underdispersed count datasets. There are other biparametric distributions that can also cope with over and underdispersion such as the *GP*, *CMP* or *HP* distributions, but the *EBW* distribution is more general because its pmf and moments can be explicitly obtained in terms of the parameters. In this paper we have proposed this new model to fits data from different fields of knowledge, that shows the versatility of this model in respect with its possible application. In addition, when the first parameter of the *EBW* distribution is positive, it allows to split the variance into three uniquely determined components. This property avoids the problem of indeterminacy

present in the *UGW* distribution. In consequence, and taking into account this property, the *EBW* distribution is more adequate than other biparametric discrete distributions for modelling overdispersed data in which the non-random part of the variance has two components, none of them negligible.

Furthermore, when the first parameter is a negative integer the *EBW* distribution has finite range and it is underdispersed. Something similar happens with the *GP* distribution that also has finite range but only in the underdispersed case, whereas the *EBW* distribution can also be underdispersed with infinite range.

Author Contributions: Data curation, J.R.-A.; Formal analysis, V.C.-L., M.J.O.-J. and J.R.-A.; Investigation, V.C.-L., M.J.O.-J. and J.R.-A.; Methodology, V.C.-L. and M.J.O.-J.; Software, M.J.O.-J. and J.R.-A.; Supervision, J.R.-A.; Writing—original draft, J.R.-A., V.C.-L. and M.J.O.-J.; writing—review and editing, J.R.-A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data of example 1 is available in <http://www.juntadeandalucia.es/institutodeestadisticaycartografia/sima/index2.htm> and the example 2 is included in the article.

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

Appendix A. Obtaining the Parameters in Terms of μ and σ^2

For the *EBW* distribution there is a pair of solutions for α and γ from Equation (4):

$$\alpha_1 = \frac{\mu^2 + \sqrt{\mu^4 + \mu(\sigma^2 - \mu)(\mu^2 + \sigma^2)}}{\sigma^2 - \mu}, \quad \gamma_1 = \frac{\alpha_1}{\mu} + 2\alpha_1 + 1 \quad (\text{A1a})$$

$$\alpha_2 = \frac{\mu^2 - \sqrt{\mu^4 + \mu(\sigma^2 - \mu)(\mu^2 + \sigma^2)}}{\sigma^2 - \mu}, \quad \gamma_2 = \frac{\alpha_2}{\mu} + 2\alpha_2 + 1 \quad (\text{A1b})$$

It can be shown that if the *EBW* distribution is overdispersed, $\alpha_1, \gamma_1 > 0$ and $\alpha_2 < 0$, but $\gamma_2 > 0$ if $\mu < 1$. If the *EBW* distribution is underdispersed, both α_1 and α_2 are negative, but:

- $\gamma_1 > 0$ when $\mu < 1$ and $\sigma^2 > \mu(1 - \mu)$ or when $\mu \geq 1$ and $\sigma^2 > \frac{\mu - \mu^2 + \sqrt{\mu^3(\mu - 1)}}{2}$
- $\gamma_2 > 0$ when $\mu < 1$ and $\sigma^2 > \mu(1 - \mu)$.

As a consequence, if $\mu \geq 1$, the only possible solution is that given in Equation (A1a) for both cases (over- and underdispersed).

Regarding the rest of the models, the expressions of their parameters in terms of μ and σ^2 can be seen in [21].

References

1. Irwin, J.O. The generalized Waring distribution applied to accident theory. *J. R. Stat. Soc. Ser. A* **1968**, *131*, 205–225. [CrossRef]
2. Xelakaki, E. Infinite divisibility, completeness and regression properties of the univariate generalized Waring distribution. *Ann. Inst. Stat. Math.* **1983**, *35*, 279–289. [CrossRef]
3. Xelakaki, E. The univariate generalized Waring distribution in relation to accident theory: Proneness, spells or contagion? *Biometrics* **1983**, *39*, 887–895. [CrossRef]
4. Rodríguez-Avi, J.; Conde-Sánchez, A.; Sáez-Castillo, A.J.; Olmo-Jiménez, M.J. A new generalization of the Waring distribution. *Comput. Stat. Data Anal.* **2007**, *51*, 6138–6150. [CrossRef]
5. Rodríguez-Avi, J.; Conde-Sánchez, A.; Sáez-Castillo, A.J.; Olmo-Jiménez, M.J.; Martínez-Rodríguez, A.M. A generalized Waring regression model for count data. *Comput. Stat. Data Anal.* **2009**, *53*, 3717–3725. [CrossRef]
6. Hilbe, J.M. *Negative Binomial Regression*; Cambridge University Press: Cambridge, UK, 2011.
7. Vélchez-López, S.; Sáez-Castillo, A.J.; Olmo-Jiménez, M.J. GWRM: An R Package for Identifying Sources of Variation in Overdispersed Count Data. *PLoS ONE* **2016**, *11*, e0167570. [CrossRef]
8. Tesitelova, H. On the role of nouns in the lexical statistics. *Prague Stud. Math. Linguist.* **1967**, *2*, 121–131.

9. Ajiferuke, I. A probabilistic model for the distribution of authorships. *J. Am. Soc. Inf. Sci.* **1991**, *42*, 279–289. [[CrossRef](#)]
10. Levene, M.; Fenner, T.; Loizou, G.; Wheeldon, R. A stochastic model for the evolution of the web. *Comput. Netw.* **2002**, *39*, 277–287. [[CrossRef](#)]
11. Xekalaki, E. The bivariate generalized Waring distribution and its application to Accident Theory. *J. R. Stat. Soc. Ser. A* **1984**, *147*, 488–498. [[CrossRef](#)]
12. Grunwaldm, G.K.; Bruce, S.L.; Jiang, L.; Strand, M.; Rabinovitch, N. A statistical model for under- or overdispersed clustered and longitudinal count data. *Biom. J.* **2011**, *53*, 578–594. [[CrossRef](#)] [[PubMed](#)]
13. Peng, Y.; Lord, D.; Zou, Y. Applying the generalized Waring model for investigating sources of variance in motor vehicle crash analysis. *Accid. Anal. Prev.* **2014**, *73*, 20–26. [[CrossRef](#)] [[PubMed](#)]
14. Ariza-López, F.J.; Rodríguez-Avi, J. Estimating the count of completeness errors in geographic data sets by means of a generalized Waring regression model. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 1394–1418. [[CrossRef](#)]
15. Huete-Morales, M.D.; Marmolejo-Martín, J.A. The Waring Distribution as a Low-Frequency Prediction Model: A Study of Organic Livestock Farms in Andalusia. *Mathematics* **2020**, *8*, 2025. [[CrossRef](#)]
16. Panaretos, J.; Xekalaki, E. Extension of the Application of Conway-Maxwell-Poisson Models: Analyzing Traffic Crash Data Exhibiting Underdispersion. *Risk Anal.* **1986**, *4*, 313–318.
17. Johnson, N.L.; Kemp, A.W.; Kotz, S. *Univariate Discrete Distributions*, 3rd ed.; Wiley: New York, NY, USA, 2005.
18. Rodríguez-Avi, J.; Conde-Sánchez, A.; Sáez-Castillo, A.J.; Olmo-Jiménez, M.J. A triparametric discrete distribution with complex parameters. *Stat. Pap.* **2004**, *45*, 81–95. [[CrossRef](#)]
19. Rodríguez-Avi, J.; Olmo-Jiménez, M.J. A regression model for overdispersed data without too many zeros. *Stat. Pap.* **2017**, *58*, 749–773. [[CrossRef](#)]
20. Jordan, C. *Calculus on Finite Differences*; Chelsea Publishing Company: London, UK, 1965.
21. Olmo-Jiménez, M.J.; Rodríguez-Avi, J.; Cueva-López, V. A review of the CTP distribution: A comparison with other over- and underdispersed count data models. *J. Stat. Comput. Sim.* **2018**, *88*, 2684–2706. [[CrossRef](#)]
22. Rodríguez-Avi, J.; Conde-Sánchez, A.; Sáez-Castillo, A.J.; Olmo-Jiménez, M.J. A generalization of the Beta-Binomial distribution. *J. R. Stat. Soc. Ser. C* **2007**, *56*, 51–61. [[CrossRef](#)]
23. Sibuya, M. Generalized hypergeometric, digamma and trigamma distributions. *Ann. Inst. Statist. Math.* **1979**, *31*, 373–390. [[CrossRef](#)]
24. Sibuya, M.; Shimizu, R. Classification of the generalized hypergeometric family of distributions. *Keio Sci. Technol. Rep.* **1981**, *34*, 1–38.
25. Burnham, K.P.; Anderson, D.R. *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*, 2nd ed.; Springer: New York, NY, USA, 2002.
26. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.
27. Byrd, R.H.; Lu, P.; Nocedal, J.; Zhu, C. A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.* **1995**, *16*, 1190–1208. [[CrossRef](#)]
28. Arnold, T.B.; Emerson, J.W. Nonparametric Goodness-of-Fit Tests for Discrete Null Distributions. *R J.* **2011**, *3*, 34–39. [[CrossRef](#)]
29. Consul, P.C. *Generalized Poisson Distributions: Properties and Applications*; Marcel Dekker: New York, NY, USA, 1989.
30. Joe, H.; Zhu, R. Generalized Poisson Distribution: The Property of Mixture of Poisson and Comparison with Negative Binomial Distribution. *Biom. J.* **2005**, *45*, 219–229. [[CrossRef](#)] [[PubMed](#)]
31. Conway, R.W.; Maxwell, W.L. A queuing model with state dependent service rates. *J. Ind. Eng.* **1962**, *12*, 132–136.
32. Sellers, K.F.; Borle, S.; Shmueli, G. The COM-Poisson model for count data: A survey of methods and applications. *Appl. Stoch. Model. Bus. Ind.* **2012**, *28*, 104–116. [[CrossRef](#)]
33. Bardwell, G.E.; Crow, E.L. A two parameter family of hyper-Poisson distributions. *J. Am. Stat. Assoc.* **1964**, *54*, 133–141. [[CrossRef](#)]
34. Instituto de Estadística y Cartografía de Andalucía. SIMA: Sistema de Información Multiterritorial de Andalucía. Available online: <http://www.juntadeandalucia.es/institutodeestadisticaycartografia/sima/index2.htm> (accessed on 2 January 2021).
35. Wimmer, G.; Köhler, R.; Grotjahn, R.; Altmann, G. Towards a Theory of Word Length Distribution. *J. Quant. Linguist.* **1994**, *1*, 98–106. [[CrossRef](#)]