Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 52 (2) (2023), 529 – 545
DOI : 10.15672/hujms.820849

Research Article

Comparing discrete Pareto populations under a fixed effects model

M. Baratnia, A.H. Rezaei Roknabadi, M. Doostparast*

Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

Abstract

The discrete Pareto distribution can be considered as a lifetime distribution and then is widely used in practice. It follows the power law tails property which makes it as a candidate model for natural phenomena. This paper deals with comparison of discrete Pareto populations by proposing a non-linear fixed effects model. Estimators for the factor effects are derived in explicit expressions. Stochastic properties of the estimators are studied in details. A test for assessing the homogeneity of populations is proposed. Illustrative examples are also given. The proposed model is an alternative model for analyzing data sets in which the linear models have poor performance.

Mathematics Subject Classification (2020). 62-07, 62F10

Keywords. Non-linear models, one-way classification, fixed effects, estimation, discrete Pareto population

1. Introduction

Analyzing the effect of a factor on a response variable is a traditional purpose in statistics, and also known as *one-way classification*. If the response variable follows the normal distribution, then one-way analysis of variance (ANOVA) is usually implemented; See, e.g., [7, 14, 16] and references therein. Alternative models for non-normal distributed responses have been proposed in the literature. For example, Kruskal and Wallis [11] suggested a rank based method for one-way ANOVA. This method, known as Kruskal-Wallis test, does not estimate the factor effects. So, it is not applicable for cases in which researchers need to estimate factor effects. Nelder and Wedderburn [15] proposed theory of *generalized linear models* (GLMs). The GLMs assume that the response distribution belongs to the *exponential family* of distributions. Also, they investigate the factor effects only on a function of response mean, known as *link function*. But in practice, researchers may face with some cases in which the response distribution does not belong to the exponential family of distributions or one may wish to model another characteristic of the response distribution, such as minimum bound parameter or reversed hazard function. Therefore, the GLMs are also restrictive; See, e.g., [1] and [2].

*Corresponding Author.
Email addresses: m-baratnia@mail.um.ac.ir (M. Baratnia), rezaei@um.ac.ir (A.H. Rezaei Roknabadi), doustparast@um.ac.ir (M. Doostparast)
Received: 04.11.2020; Accepted: 06.10.2022

Let X be a discrete random variable with support $S_X = \{\alpha, \alpha+1, \cdots\}$ and the probability mass function (PMF)

$$f_X(t) = \alpha^\theta \left(\frac{1}{t^\theta} - \frac{1}{(t+1)^\theta} \right), \qquad t \in S_X, \qquad \alpha \in \mathbb{N}, \qquad \theta > 0, \tag{1.1}$$

where (α, θ) are parameters and $\mathbb{N}$ stands for the set of the natural numbers. The random variable X is called *discrete Pareto* (DP) variable. One may construct this random variable from the continuous Pareto distribution. More precisely, if Y is a continuous random variable with the probability density function

$$f_Y(t) = \frac{\theta \alpha^\theta}{t^{\theta+1}}, \qquad t \geq \alpha, \qquad \alpha \in \mathbb{N}, \qquad \theta > 0,$$

then $X = \lfloor Y \rfloor$ has the DP distribution with the PMF (1.1), where $\lfloor b \rfloor$ indicates the largest integer less than or equal to b; See, e.g., [3]. The DP distribution (1.1) has a *power-law* tail as

$$P(X \geq t) \simeq t^{-\theta}, \qquad t \to \infty. \tag{1.2}$$

On the other hand, many random phenomena in practice, such as biology, chemistry, computer science, economics, finance and etc, have empirical distributions with power-law tails as in (1.2). Therefore, various stochastic phenomena may follow the DP distribution; See, e.g., [3, 9, 10].

When α in (1.1) is unknown, the DP distribution does not belong to the exponential family of distributions. Then, GLMs cannot be used for data sets with DP responses. Moreover, there is no parametric method in the literature for classification of DP random variables. Therefore, in this paper, we focus on comparison of multiple DP populations, that is when the response variable follows the DP distribution (1.1). For motivation, suppose that a manager of a publishing company wants to compare m different types of cartridges. To do this, she may compare the number of sheets printed by each of the cartridges. If the number of printed sheets follows the DP distribution, then she can compare these cartridges using the proposed approach in this paper.

In order to provide the new model for data analyses of classified discrete observations, consider the following notation throughout the paper:

1. m (≥ 2) is the number of varying levels (or classes) of the factor variable;
2. For $1 \leq i \leq m$, n_i (≥ 1) stands for the number of observations in the i-th class and $N := \sum_{i=1}^m n_i$ is the total number of observations;
3. For $1 \leq i \leq m$ and $1 \leq j \leq n_i$, Y_{ij} is the response variable for the j-th individual in the i-th class. Also, $Y_{i,(1)} := \min_{1 \leq j \leq n_i} \{Y_{ij}\}$ and $\mathbf{Y} := [\mathbf{Y}_1, \cdots, \mathbf{Y}_m]^T$, where $\mathbf{Y}_i := [Y_{i1}, \cdots, Y_{in_i}]^T$ for $1 \leq i \leq m$;
4. "IID" stands for independent and identically distributed;
5. $DP(\alpha, \theta)$ indicates the DP distribution with the PMF (1.1);
6. $J_B(t) = 1$ if $t \in B$ and 0 otherwise;
7. The lowercase letters of random variables denote the corresponding realizations;
8. For any integers $u \geq 0$ and $v \geq 1$, $B_u := \{u, u+1, \cdots\}$ and $\bar{B}_v := \{1, 2, \cdots, v\}$.
9. $\perp$ stands for the statistical independence.

Now, let us assume that

$$Y_{i1}, \cdots, Y_{in_i} \overset{IID}{\sim} DP(\alpha_i, \theta), \qquad 1 \leq i \leq m, \tag{1.3}$$

$$(Y_{i1}, \cdots, Y_{in_i}) \perp (Y_{j1}, \cdots, Y_{jn_j}), \qquad 1 \leq i \neq j \leq m, \tag{1.4}$$

where $\alpha_i \in \mathbb{N}$ is the fixed effect of the i-th class, for $1 \le i \le m$, and θ is a common parameter. From Equation (1.3), the likelihood function (LF) of the available data for the i-th class is

$$L_i(\alpha_i, \theta; \mathbf{y}_i) = \prod_{j=1}^{n_i} f_i(y_{ij})$$
$$= \alpha_i^{\theta n_i} J_{\bar{B}_{y_{i,(1)}}}(\alpha_i) \prod_{j=1}^{n_i} \left(\frac{1}{y_{ij}^{\theta}} - \frac{1}{(y_{ij}+1)^{\theta}} \right), \qquad 1 \le i \le m, \tag{1.5}$$

where f_i is the PMF of the observations in the i-th class, for $1 \le i \le m$. The independence assumption (1.4) and Equation (1.5) yield the LF of the available data as

$$L(\alpha_1, \cdots, \alpha_m, \theta; \mathbf{y}) = \prod_{i=1}^{m} L_i(\alpha_i, \theta; \mathbf{y}_i)$$
$$= \prod_{i=1}^{m} \left(\alpha_i^{\theta n_i} J_{\bar{B}_{y_{i,(1)}}}(\alpha_i) \prod_{j=1}^{n_i} \left(\frac{1}{y_{ij}^{\theta}} - \frac{1}{(y_{ij}+1)^{\theta}} \right) \right). \tag{1.6}$$

Then, the log-likelihood function (log-LF) is

$$l(\alpha_1, \cdots, \alpha_m, \theta; \mathbf{y}) = \log L(\alpha_1, \cdots, \alpha_m, \theta; \mathbf{y})$$
$$= \sum_{i=1}^{m} \left(\theta n_i \log \alpha_i + \log J_{\bar{B}_{y_{i,(1)}}}(\alpha_i) + \sum_{j=1}^{n_i} \log \left(\frac{1}{y_{ij}^{\theta}} - \frac{1}{(y_{ij}+1)^{\theta}} \right) \right). \tag{1.7}$$

In this paper, we analyze the available data $\mathbf{y}$ on the basis of the respective log-LF (1.7). Thus, the rest of the paper is organized as follows: In Section 2, it is assumed that the parameter θ in (1.7) is known. Explicit expressions for the maximum likelihood estimators (MLEs) of the model parameters $\alpha_1, \cdots, \alpha_m$ are derived. Stochastic properties of the derived MLEs are studied in details. A procedure for testing the homogeneity of populations is also proposed. In Section 3 and by assuming that θ is unknown, inferences about the model parameters are studied. Finally, in Section 4, findings of the previous sections are assessed numerically. Section 5 is dedicated to conclusion.

2. Likelihood inference without nuisance parameter

In this section, suppose that $\theta > 0$ in (1.7) is known. So, we use the symbol $l(\alpha_1, \cdots, \alpha_m; \mathbf{y})$ instead of $l(\alpha_1, \cdots, \alpha_m, \theta; \mathbf{y})$ for the log-LF (1.7).

2.1. Estimation

To derive the maximum likelihood estimates, we first assume that the parameters $\alpha_1, \cdots, \alpha_m$ take real values. The partial derivative of the log-LF (1.7) with respect to the parameter α_i is

$$\frac{\partial l(\alpha_1, \cdots, \alpha_m; \mathbf{y})}{\partial \alpha_i} = \frac{\theta n_i}{\alpha_i} > 0, \qquad 1 \le i \le m.$$

This means that $l(\alpha_1, \cdots, \alpha_m; \mathbf{y})$ is increasing in α_i, for $1 \le i \le m$. Therefore, the unique MLE of α_i, denoted by $\hat{\alpha}_i$, exists and is given by

$$\hat{\alpha}_i = Y_{i,(1)}, \qquad 1 \le i \le m. \tag{2.1}$$

Theorem 2.1. *For $1 \le i \le m$, the MLE of α_i in Equation (2.1) has the following properties:*

[i] $\hat{\alpha}_i \sim DP(\alpha_i, \theta n_i)$. *Moreover, $\hat{\alpha}_1, \cdots, \hat{\alpha}_m$ are independent;*

[ii] *For* $k \le \theta n_i - 1$, $E\left(\hat{\alpha}_i^k\right)$ *exists and*

$$E\left(\hat{\alpha}_i^k\right) = \alpha_i^k + \sum_{t=\alpha_i}^{\infty} \frac{\alpha_i^{\theta n_i}}{(t+1)^{\theta n_i - k}} \left(1 - \left(\frac{t}{t+1}\right)^k\right).$$

Proof. From [3, Proposition 3.1] and the independence assumption (1.4), part [i] is concluded. For $1 \le i \le m$ and $k \le \theta n_i - 1$, one can see that

$$\begin{aligned} E\left(\hat{\alpha}_i^k\right) &= \sum_{t=\alpha_i}^{\infty} t^k f_{\hat{\alpha}_i}(t) \\ &= \sum_{t=\alpha_i}^{\infty} \alpha_i^{\theta n_i} \left(\frac{1}{t^{\theta n_i - k}} - \frac{t^k}{(t+1)^{\theta n_i}} \pm \frac{1}{(t+1)^{\theta n_i - k}}\right) \\ &= \alpha_i^k \sum_{t=\alpha_i}^{\infty} \alpha_i^{\theta n_i - k} \left(\frac{1}{t^{\theta n_i - k}} - \frac{1}{(t+1)^{\theta n_i - k}}\right) + \sum_{t=\alpha_i}^{\infty} \frac{\alpha_i^{\theta n_i}}{(t+1)^{\theta n_i - k}} \left(1 - \left(\frac{t}{t+1}\right)^k\right) \qquad (2.2) \\ &= \alpha_i^k + \sum_{t=\alpha_i}^{\infty} \frac{\alpha_i^{\theta n_i}}{(t+1)^{\theta n_i - k}} \left(1 - \left(\frac{t}{t+1}\right)^k\right), \end{aligned}$$

which proves [ii]. Notice that, for $k < \theta n_i - 1$

$$\sum_{t=\alpha_i}^{\infty} \frac{\alpha_i^{\theta n}}{(t+1)^{\theta n - k}} \left(1 - \left(\frac{t}{t+1}\right)^k\right) \le \sum_{t=\alpha_i}^{\infty} \frac{\alpha_i^{\theta n}}{(t+1)^{\theta n - k}} < \infty, \qquad 1 \le i \le m,$$

Also, for $k = \theta n_i - 1$

$$\sum_{t=\alpha_i}^{\infty} \frac{\alpha_i^{\theta n_i}}{(t+1)^{\theta n_i - k}} \left(1 - \left(\frac{t}{t+1}\right)^k\right) = \sum_{t=\alpha_i}^{\infty} \frac{\alpha_i^{\theta n_i}\left((t+1)^{\theta n_i - 1} - t^{\theta n_i - 1}\right)}{(t+1)^{\theta n_i}}. \qquad (2.3)$$

Since $\sum_{t=\alpha_i}^{\infty} t^{-2} < \infty$ and

$$\lim_{t\to\infty} \frac{\alpha_i^{\theta n_i}\left((t+1)^{\theta n_i - 1} - t^{\theta n_i - 1}\right)}{(t+1)^{\theta n_i}} \times t^2 = \alpha_i^{\theta n_i} > 0,$$

by the well known *limit comparison test* [8, Theorem 5.2.5. page 147], the series (2.3) is convergent. Therefore, Equation (2.2) holds and $E\left(\hat{\alpha}_i^k\right)$ exists for $k \le \theta n_i - 1$. □

For $1 \le i \le m$, let

$$b_i(n_i) := \sum_{t=\alpha_i}^{\infty} \left(\frac{\alpha_i}{t+1}\right)^{\theta n_i}, \qquad d_i(n_i) := \sum_{t=\alpha_i}^{\infty} \frac{\alpha_i^{\theta n_i}(2t+1)}{(t+1)^{\theta n_i}}.$$

Therefore,

$$\lim_{n_i\to\infty} b_i(n_i) = 0, \qquad \lim_{n_i\to\infty} d_i(n_i) = 0, \qquad 1 \le i \le m.$$

From Theorem 2.1 ([ii]), we have for $n_i \ge 2\theta^{-1}$

$$E(\hat{\alpha}_i) = \alpha_i + b_i(n_i),$$

Also, for $n_i \ge 3\theta^{-1}$, we obtained variance and mean squared error (MSE) of $\hat{\alpha}_i$ as

$$\begin{aligned} Var(\hat{\alpha}_i) &= d_i(n_i) - b_i(n_i)\left(b_i(n_i) + 2\alpha_i\right), \\ MSE(\hat{\alpha}_i) &= d_i(n_i) - 2b_i(n_i)\alpha_i. \end{aligned}$$

Then, $\hat{\alpha}_i$ is asymptotically unbiased. Moreover, $MSE(\hat{\alpha}_i) \to 0$ as $n_i \to \infty$, i.e., $\hat{\alpha}_i$ converges to α_i in $\mathcal{L}^2$, and hence $\hat{\alpha}_i$ is consistent for α_i; See, e.g., [6, page 208]. In the next lemma, and for $1 \le i \le m$, it is proved that a general form of $\hat{\alpha}_i - \alpha_i$ degenerates at zero.

Lemma 2.2. *Let $(\Theta, \mathcal{F}, P)$ be a common probability space for the responses, and $\mathcal{Z}_0$ denotes the set of all measurable functions, say $h(t)$, such that $h(t) = 0$ almost surely has the unique solution $t = 0$. Then, for any $h \in \mathcal{Z}_0$ and $1 \leq i \leq m$, the random variable $h(\hat{\alpha}_i - \alpha_i)$ tends to 0 as n_i goes to infinity.*

Proof. For any $h \in \mathcal{Z}_0$ and $1 \leq i \leq m$, we have

$$\lim_{n_i \to \infty} P\left(h(\hat{\alpha}_i - \alpha_i) = 0\right) = \lim_{n_i \to \infty} P\left(\hat{\alpha}_i - \alpha_i = 0\right) = \lim_{n_i \to \infty} 1 - \left(\frac{\alpha_i}{\alpha_i + 1}\right)^{\theta n_i} = 1.$$

Then $h(\hat{\alpha}_i - \alpha_i)$ is asymptotically degenerated. □

Theorem 2.3. *For $1 \leq i \leq m$, $\hat{\alpha}_i$ is a sufficient statistic for α_i. Also, if $n_i > \theta^{-1}$, then $\hat{\alpha}_i$ is boundedly complete.*

Proof. One can rewrite the LF (1.5) as follows:

$$L_i(\alpha_i, \theta; \mathbf{y}_i) = A_i\left(y_{i,(1)}, \alpha_i\right) H_i(\mathbf{y}_i), \qquad 1 \leq i \leq m,$$

where $A_i\left(y_{i,(1)}, \alpha_i\right) := \alpha_i^{\theta n_i} J_{\bar{B}_{y_{i,(1)}}}(\alpha_i)$ and $H_i(\mathbf{y}_i) := \prod_{j=1}^{n_i}\left(y_{ij}^{-\theta} - (y_{ij}+1)^{-\theta}\right)$ are non-negative functions. Therefore, the LF (1.5) can be written as product of two factors. The factor H_i does not depend on α_i and the other factor A_i depends on α_i as well $\mathbf{y}_i$ only through $y_{i,(1)}$. Hence, from the well known *factorization theorem* [12, page 35], $\hat{\alpha}_i = Y_{i,(1)}$ is sufficient for α_i. In addition, for an arbitrary bounded function, say $g(.)$, let us assume that

$$0 = E\left(g(\hat{\alpha}_i)\right) = \sum_{t=\alpha_i}^{\infty} g(t)\alpha_i^{\theta n_i}\left(\frac{1}{t^{\theta n_i}} - \frac{1}{(t+1)^{\theta n_i}}\right), \qquad \alpha_i \in \mathbb{N}. \tag{2.4}$$

Multiplying $\alpha_i^{-\theta n_i} \neq 0$ on the both sides of (2.4) concludes

$$0 = \sum_{t=\alpha_i}^{\infty}\left(\frac{g(t)}{t^{\theta n_i}} - \frac{g(t)}{(t+1)^{\theta n_i}}\right). \tag{2.5}$$

Since $g(.)$ is bounded, there is $M > 0$ such that $|g(t)| < M$ for all t. Also, since $\sum_{t=1}^{\infty} t^{-p} < \infty$ for $p > 1$, we have for $n_i > \theta^{-1}$ (or equivalently $\theta n_i > 1$)

$$\sum_{t=\alpha_i}^{\infty} \frac{g(t)}{t^{\theta n_i}} < \sum_{t=\alpha_i}^{\infty} \frac{M}{t^{\theta n_i}} < M\sum_{t=1}^{\infty} t^{-\theta n_i} < \infty,$$
$$\sum_{t=\alpha_i}^{\infty} \frac{g(t)}{(t+1)^{\theta n_i}} < \sum_{t=\alpha_i}^{\infty} \frac{M}{(t+1)^{\theta n_i}} < M\sum_{t=1}^{\infty} t^{-\theta n_i} < \infty.$$

Then,

$$\sum_{t=\alpha_i}^{\infty}\left(\frac{g(t)}{t^{\theta n_i}} - \frac{g(t)}{(t+1)^{\theta n_i}}\right) = \sum_{t=\alpha_i}^{\infty} \frac{g(t)}{t^{\theta n_i}} - \sum_{t=\alpha_i}^{\infty} \frac{g(t)}{(t+1)^{\theta n_i}},$$

and therefore Equation (2.5) yields

$$\begin{aligned} 0 &= \sum_{t=\alpha_i}^{\infty} \frac{g(t)}{t^{\theta n_i}} - \sum_{t=\alpha_i}^{\infty} \frac{g(t)}{(t+1)^{\theta n_i}} \\ &= \frac{g(\alpha_i)}{\alpha_i^{\theta n_i}} + \sum_{t=\alpha_i+1}^{\infty} \frac{g(t)}{t^{\theta n_i}} - \sum_{t=\alpha_i+1}^{\infty} \frac{g(t-1)}{t^{\theta n_i}}, \end{aligned}$$

or equivalently

$$\frac{g(\alpha_i)}{\alpha_i^{\theta n_i}} + \sum_{t=\alpha_i+1}^{\infty} \frac{g(t) - g(t-1)}{t^{\theta n_i}} = 0, \qquad \alpha_i \in \mathbb{N}. \tag{2.6}$$

Equation (2.6) implies

$$\sum_{t=\alpha_i+k+1}^{\infty}\frac{g(t)-g(t-1)}{t^{\theta n_i}} = g(\alpha_i)\left(\frac{1}{(\alpha_i+1)^{\theta n_i}}-\frac{1}{(\alpha_i)^{\theta n_i}}\right) + g(\alpha_i+1)\left(\frac{1}{(\alpha_i+2)^{\theta n_i}}-\frac{1}{(\alpha_i+1)^{\theta n_i}}\right)+\cdots + g(\alpha_i+k-1)\left(\frac{1}{(\alpha_i+k)^{\theta n_i}}-\frac{1}{(\alpha_i+k-1)^{\theta n_i}}\right) - \frac{g(\alpha_i+k)}{(\alpha_i+k)^{\theta n_i}}, \qquad \alpha_i, k\in\mathbb{N}. \tag{2.7}$$

Since Equation (2.6) is established for any $\alpha_i \in \mathbb{N}$, we have

$$\frac{g(\alpha_i+1)}{(\alpha_i+1)^{\theta n_i}}+\sum_{t=\alpha_i+2}^{\infty}\frac{g(t)-g(t-1)}{t^{\theta n_i}}=0. \tag{2.8}$$

Then, Equation (2.7) for $k=1$ and Equation (2.8) conclude

$$\frac{g(\alpha_i+1)}{(\alpha_i+1)^{\theta n_i}}+g(\alpha_i)\left(\frac{1}{(\alpha_i+1)^{\theta n_i}}-\frac{1}{(\alpha_i)^{\theta n_i}}\right)-\frac{g(\alpha_i+1)}{(\alpha_i+1)^{\theta n_i}}=0,$$

which results $g(\alpha_i)=0$. Similarly, from Equation (2.6)

$$\frac{g(\alpha_i+2)}{(\alpha_i+2)^{\theta n_i}}+\sum_{t=\alpha_i+3}^{\infty}\frac{g(t)-g(t-1)}{t^{\theta n_i}}=0. \tag{2.9}$$

Therefore, Equation (2.7) for $k=2$ and Equation (2.9) imply

$$\frac{g(\alpha_i+2)}{(\alpha_i+2)^{\theta n_i}}+g(\alpha_i)\left(\frac{1}{(\alpha_i+1)^{\theta n_i}}-\frac{1}{(\alpha_i)^{\theta n_i}}\right)+g(\alpha_i+1)\left(\frac{1}{(\alpha_i+2)^{\theta n_i}}-\frac{1}{(\alpha_i+1)^{\theta n_i}}\right)-\frac{g(\alpha_i+2)}{(\alpha_i+2)^{\theta n_i}}=0,$$

which concludes $g(\alpha_i+1)=0$, since $g(\alpha_i)=0$. By proceeding in this way and by induction, one can see that $g(\alpha_i+k)=0$ for any $k\in\mathbb{N}$. This means that $P\left(g\left(\hat{\alpha}_i\right)=0\right)=1$, i.e., $\hat{\alpha}_i$ is boundedly complete for α_i. □

In sequel, we provide a $100(1-\gamma)\%$ confidence interval (CI) for α_i, $1\le i\le m$. To do this, the following lemma is essential.

Lemma 2.4. *[4, page 434] Let X be a discrete statistic with support S_X and the cumulative distribution function $F_X(x;\mu)=P(X\le x\mid\mu)$. Also, let $0<\gamma_1,\gamma_2,\gamma<1$ are fixed values where $\gamma=\gamma_1+\gamma_2$. Suppose that for any $x\in S_X$, $\mu_L(x)$ and $\mu_U(x)$ can be defined as follows:*

i. *If $F_X(x;\mu)$ is a decreasing function of μ for each x, define $\mu_L(x)$ and $\mu_U(x)$ by*

$$P\left(X\le x\mid\mu_U(x)\right)=\gamma_1, \qquad P\left(X\ge x\mid\mu_L(x)\right)=\gamma_2.$$

ii. *If $F_X(x;\mu)$ is an increasing function of μ for each x, define $\mu_L(x)$ and $\mu_U(x)$ by*

$$P\left(X\ge x\mid\mu_U(x)\right)=\gamma_1, \qquad P\left(X\le x\mid\mu_L(x)\right)=\gamma_2.$$

Then, the random interval $(\mu_L(X),\mu_U(X))$ is a $100(1-\gamma)\%$ CI for μ.

In Theorem 2.1, it is proved that $\hat{\alpha}_i\sim DP\left(\alpha_i,\theta n_i\right)$, for $1\le i\le m$. Therefore

$$F_{\hat{\alpha}_i}(t;\alpha_i)=1-\frac{\alpha_i^{\theta n_i}}{(1+t)^{\theta n_i}}, \qquad t\in B_{\alpha_i}, \qquad 1\le i\le m. \tag{2.10}$$

Obviously, $F_{\hat{\alpha}_i}(t;\alpha_i)$ in (2.10) is a decreasing function of α_i for any $t\in B_{\alpha_i}$. Hence, according to Lemma 2.4, an equi-tail $100(1-\gamma)\%$ CI for α_i is

$$I_\theta\left(\alpha_i;\gamma\right):=\left(\hat{\alpha}_i\left(\frac{\gamma}{2}\right)^{1/\theta n_i},(1+\hat{\alpha}_i)\left(1-\frac{\gamma}{2}\right)^{1/\theta n_i}\right), \qquad 1\le i\le m. \tag{2.11}$$

Remark 2.5. For $1 \le i \le m$, we have $\alpha_i \in \mathbb{N}$ in Equation (1.3). Therefore, one should modify the proposed CI in Equation (2.11) as follows:

$$I_\theta(\alpha_i;\gamma) := \left[\left\lceil \hat{\alpha}_i \left(\frac{\gamma}{2}\right)^{1/\theta n_i} \right\rceil, \left\lfloor (1+\hat{\alpha}_i)\left(1-\frac{\gamma}{2}\right)^{1/\theta n_i} \right\rfloor\right], \qquad 1 \le i \le m, \tag{2.12}$$

where $\lceil . \rceil$ denotes the ceiling function; See, e.g., [8].

Remark 2.6. By the well known *Bonferroni inequality*, a joint $100(1-\gamma)\%$ confidence set for $\alpha_1, \cdots, \alpha_m$ is given by $I_\theta(\alpha_1;\gamma/m) \otimes \cdots \otimes I_\theta(\alpha_m;\gamma/m)$ where "$\mathcal{A} \otimes \mathcal{B}$ stands for the Cartesian product of two sets $\mathcal{A}$ and $\mathcal{B}$.

2.2. Homogeneity testing

In this subsection, we propose a procedure for testing the homogeneity hypothesis of populations, that is $H_0 : \alpha_1 = \cdots = \alpha_m$ against the alternative hypothesis

$$H_1 : \exists i \ne j, \qquad i, j \in \{1, \cdots, m\}, \qquad \text{such that} \qquad \alpha_i \ne \alpha_j.$$

Therefore, if the null hypothesis is rejected, then effects of the factor variable are meaningful. Suppose that α is the common value of α_is under the null hypothesis H_0. Under H_0, the LF (1.6) simplifies to

$$L(\alpha;\mathbf{y}) = \alpha^{\theta N} J_{\bar{B}_{y_{(1),(1)}}}(\alpha) \prod_{i=1}^{m}\prod_{j=1}^{n_i}\left(\frac{1}{y_{ij}^{\theta}} - \frac{1}{(y_{ij}+1)^{\theta}}\right), \tag{2.13}$$

where $Y_{(1),(1)} := \min_{1\le i\le m}\{Y_{i,(1)}\}$. Similar to Subsection 2.1, the LF (2.13) is increasing in α, and hence $\hat{\alpha} := Y_{(1),(1)}$ is the MLE of α. Also, under the homogeneity hypothesis, we have

$$P(\hat{\alpha} > t) = P\left(Y_{(1),(1)} > t\right) = \prod_{i=1}^{m}\prod_{j=1}^{n_i} P(Y_{ij} > t) = \prod_{i=1}^{m}\prod_{j=1}^{n_i}\frac{\alpha^{\theta}}{(t+1)^{\theta}} = \frac{\alpha^{\theta N}}{(t+1)^{\theta N}},$$

which means that $\hat{\alpha} \sim DP(\alpha, \theta N)$. Moreover

$$\hat{\alpha} = Y_{(1),(1)} = \min_{1\le i\le m}\{Y_{i,(1)}\} = \min_{1\le i\le m}\{\hat{\alpha}_i\}.$$

For testing the null hypothesis H_0 against the alternative H_1, we use generalized likelihood ratio test (GLRT). The GLRT statistic is

$$\Lambda = \frac{\sup_{\Omega_0} L(\alpha_1,\cdots,\alpha_m;\mathbf{y})}{\sup_{\Omega} L(\alpha_1,\cdots,\alpha_m;\mathbf{y})} = \prod_{i=1}^{m}\left(\frac{\hat{\alpha}}{\hat{\alpha}_i}\right)^{\theta n_i},$$

where $\Omega = \{(\alpha_1,\cdots,\alpha_m) \mid \alpha_i \in \mathbb{N},\ 1 \le i \le m\}$ is the whole parameter space and $\Omega_0 = \{(\alpha_1,\cdots,\alpha_m) \mid \alpha_1 = \cdots = \alpha_m \in \mathbb{N}\}$ is the parameter space under H_0. The null hypothesis H_0 is rejected if Λ is too small. Therefore, the GLRT function is as follows:

$$\varphi(\mathbf{y}) = \begin{cases} 1, & \Lambda < c, \\ \eta, & \Lambda = c, \\ 0, & \Lambda > c, \end{cases}$$

where $0 < c \le 1$ and $0 \le \eta \le 1$ are constants and determined at the significance level γ by

$$\sup_{\alpha\in\mathbb{N}} E_{H_0}(\varphi(\mathbf{Y})) = \gamma.$$

Notice that, Λ tends to 1 as $\hat{\alpha}$ goes to infinity. Therefore, $\sup_{\alpha\in\mathbb{N}} E_{H_0}(\varphi(\mathbf{Y})) = 1$ and hence the GLRT does not exist.

In order to overcome this limitation, we suggest to consider a subclass of the homogeneity

Table 1. Values of c^* in (2.14) at the significant level $\gamma = 0.05$.

θ		0.5				1				2			
α^*	n \\ m	2	3	4	5	2	3	4	5	2	3	4	5
1	1	0.0093	0.0020	0.0005	0.0001	0.0100	0.0027	0.0008	0.0002	0.0156	0.0069	0.0017	0.0008
	2	0.0100	0.0028	0.0007	0.0002	0.0156	0.0059	0.0021	0.0008	0.0625	0.0625	0.0123	0.0039
	3	0.0131	0.0037	0.0011	0.0004	0.0370	0.0156	0.0046	0.0029	1.0000	1.0000	0.0156	0.0156
	4	0.0156	0.0069	0.0017	0.0008	0.0625	0.0625	0.0123	0.0123	1.0000	1.0000	1.0000	1.0000
	5	0.0312	0.0077	0.0032	0.0014	0.0312	0.0312	0.0312	0.0312	1.0000	1.0000	1.0000	1.0000
	6	0.0370	0.0156	0.0046	0.0029	1.0000	1.0000	0.0156	0.0156	1.0000	1.0000	1.0000	1.0000
2	1	0.0090	0.0018	0.0004	0.0001	0.0093	0.0023	0.0006	0.0002	0.0131	0.0035	0.0010	0.0003
	2	0.0096	0.0020	0.0006	0.0001	0.0123	0.0035	0.0009	0.0003	0.0256	0.0123	0.0039	0.0016
	3	0.0102	0.0029	0.0007	0.0002	0.0156	0.0060	0.0020	0.0007	0.0878	0.0156	0.0077	0.0077
	4	0.0123	0.0037	0.0011	0.0003	0.0256	0.0123	0.0039	0.0024	0.0390	0.0390	0.0390	0.0390
	5	0.0141	0.0041	0.0015	0.0005	0.0312	0.0173	0.0041	0.0041	1.0000	0.0173	0.0173	0.0173
	6	0.0190	0.0046	0.0020	0.0007	0.0878	0.0156	0.0077	0.0077	1.0000	1.0000	1.0000	1.0000
3	1	0.0087	0.0018	0.0004	0.0001	0.0099	0.0019	0.0006	0.0001	0.0105	0.0026	0.0008	0.0003
	2	0.0094	0.0020	0.0005	0.0001	0.0123	0.0028	0.0007	0.0002	0.0198	0.0055	0.0021	0.0008
	3	0.0093	0.0024	0.0006	0.0002	0.0156	0.0042	0.0014	0.0004	0.0317	0.0083	0.0056	0.0022
	4	0.0100	0.0030	0.0009	0.0002	0.0198	0.0063	0.0021	0.0008	0.1001	0.0168	0.0100	0.0100
	5	0.0137	0.0033	0.0010	0.0003	0.0185	0.0074	0.0024	0.0014	0.0563	0.0563	0.0563	0.006
	6	0.0156	0.0034	0.0011	0.0004	0.0317	0.0083	0.0056	0.0015	0.0317	0.0317	0.0317	0.0317
4	1	0.0082	0.0019	0.0004	0.0001	0.0093	0.0020	0.0005	0.0001	0.0108	0.0026	0.0007	0.0002
	2	0.0086	0.0019	0.0005	0.0001	0.0100	0.0026	0.0006	0.0002	0.0123	0.0043	0.0016	0.0004
	3	0.0104	0.0024	0.0005	0.0002	0.0119	0.0035	0.0009	0.0003	0.0230	0.0077	0.0031	0.0014
	4	0.0107	0.0024	0.0006	0.0002	0.0123	0.0051	0.0013	0.0004	0.0390	0.0114	0.0065	0.0019
	5	0.0113	0.0028	0.0008	0.0002	0.0173	0.0057	0.0021	0.0007	0.1074	0.0173	0.0115	0.0115
	6	0.0130	0.0035	0.0011	0.0003	0.0230	0.0077	0.0024	0.0012	0.0687	0.0687	0.0077	0.0077

hypothesis, that is $H_0^* : \alpha_1 = \cdots = \alpha_m = \alpha^*$ where $\alpha^* \in \mathbb{N}$ is a given value. In this case, the GLRT function simplifies to

$$\varphi^*(\mathbf{y}) = \begin{cases} 1, & \Lambda^* < c^*, \\ \eta^*, & \Lambda^* = c^*, \\ 0, & \Lambda > c^*, \end{cases} \tag{2.14}$$

where $0 < c^* \leq 1$ and $0 \leq \eta^* \leq 1$ are constants and

$$\Lambda^* = \prod_{i=1}^{m} \left(\frac{\alpha^*}{\hat{\alpha}_i} \right)^{\theta n_i}. \tag{2.15}$$

Also, at the significance level γ, the quantities c^* and η^* in (2.14) are determined by

$$E_{H_0}(\varphi^*(\mathbf{Y})) = \gamma. \tag{2.16}$$

For balanced data ($n_i = n$ for $1 \leq i \leq m$) and for some selected values of α^*, m, n, θ, the quantities c^* and η^* are obtained from Equation (2.16) at the significance level $\gamma = 0.05$ by conducting a Monte Carlo simulation. The derived results are shown in Tables 1 and 2.

Remark 2.7. Under the homogeneity hypothesis, $\hat{\alpha}$ is the MLE of the factor effects. So, if $H_0 : \alpha_1 = \cdots = \alpha_m$ is correct, then we greatly expect that $\alpha_1 = \cdots = \alpha_m \simeq \hat{\alpha}$. Therefore, we propose to consider α^* in (2.15) as the observed value of $\hat{\alpha}$ and then test $H_0^* : \alpha_1 = \cdots = \alpha_m = \alpha^*$ instead of $H_0 : \alpha_1 = \cdots = \alpha_m$.

3. Likelihood inference with nuisance parameter

In this section, the parameter $\theta > 0$ in (1.7) is considered as an unknown nuisance parameter.

Table 2. Values of η^* in (2.14) at the significant level $\gamma = 0.05$.

	θ	0.5				1				2			
α^*	n \ m	2	3	4	5	2	3	4	5	2	3	4	5
1	1	0.0020	0.0020	0.0020	0.0020	0.0060	0.0020	0.0040	0.0020	0.1270	0.0080	0.1445	0.0356
	2	0.0060	0.0020	0.0020	0.0020	0.1629	0.0059	0.0039	0.0120	0.1866	0.0410	0.2617	0.5650
	3	0.0688	0.0155	0.0020	0.03170	0.1029	0.0997	0.2936	0.0114	0.0210	0.0046	0.6790	0.5450
	4	0.1421	0.0236	0.1319	0.0335	0.1998	0.0456	0.2172	0.0048	0.0425	0.0358	0.0338	0.0308
	5	0.0448	0.0319	0.1327	0.0019	0.6352	0.4118	0.2474	0.1454	0.0487	0.0486	0.0460	0.0452
	6	0.0747	0.1408	0.2754	0.0039	0.0205	0.0058	0.6651	0.5887	0.0496	0.0489	0.0490	0.0489
2	1	0.0020	0.0020	0.0020	0.0020	0.0040	0.0020	0.0040	0.0020	0.01760	0.0040	0.0118	0.0020
	2	0.0060	0.0020	0.0020	0.0020	0.0645	0.0020	0.0217	0.0060	0.2234	0.0425	0.1150	0.0525
	3	0.0020	0.0040	0.0159	0.0040	0.1613	0.0256	0.0212	0.0812	0.0731	0.2979	0.4735	0.1565
	4	0.1081	0.0259	0.0374	0.0040	0.0815	0.0644	0.1230	0.0081	0.5098	0.2741	0.1909	0.1074
	5	0.0411	0.1069	0.0230	0.0496	0.2092	0.0996	0.4325	0.0627	0.0109	0.8908	0.7366	0.4848
	6	0.0523	0.1743	0.0114	0.0553	0.0844	0.2435	0.3974	0.1255	0.0347	0.0275	0.0159	0.0148
3	1	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0279	0.0100	0.0060	0.0040
	2	0.0040	0.0020	0.0020	0.0020	0.0240	0.0020	0.0059	0.002	0.0466	0.0078	0.0431	0.0816
	3	0.0040	0.0020	0.0020	0.0040	0.0050	0.0699	0.0311	0.0330	0.2040	0.2785	0.1209	0.0149
	4	0.0350	0.0020	0.0020	0.0040	0.0301	0.0524	0.0408	0.0203	0.0398	0.1772	0.3245	0.0909
	5	0.0078	0.0039	0.0417	0.0020	0.2434	0.1362	0.1434	0.0930	0.3154	0.1632	0.0586	0.3198
	6	0.0016	0.1317	0.0200	0.0519	0.1294	0.3000	0.1199	0.2426	0.7205	0.4499	0.3004	0.2073
4	1	0.0020	0.0020	0.0020	0.0020	0.0040	0.0040	0.0020	0.0020	0.0118	0.0179	0.0020	0.0080
	2	0.0020	0.0020	0.0020	0.0020	0.0298	0.0040	0.0040	0.0020	0.0758	0.0259	0.0233	0.0060
	3	0.0059	0.0020	0.0020	0.0020	0.0169	0.0115	0.0059	0.0020	0.0807	0.1257	0.1195	0.0521
	4	0.0040	0.0020	0.0060	0.0020	0.2052	0.0052	0.0337	0.0060	0.0097	0.1404	0.0598	0.1807
	5	0.0134	0.0339	0.008	0.0040	0.2414	0.1490	0.0060	0.0039	0.0526	0.1218	0.2244	0.0197
	6	0.0158	0.0436	0.0040	0.0080	0.1061	0.1506	0.0915	0.0239	0.2124	0.0938	0.3716	0.0768

3.1. Estimation

To obtain the MLEs of the parameters $\alpha_1, \cdots, \alpha_m$ and θ, we use the well-known *profile likelihood* method; See, e.g., [5]. As in the previous section, the log-LF (1.7) is increasing in α_i, for $1 \le i \le m$. Similar to Subsection 2.1, it is concluded that the unique MLEs of $\alpha_1, \cdots, \alpha_m$ exist and given by

$$\hat{\alpha}_i = Y_{i,(1)}, \qquad 1 \le i \le m. \tag{3.1}$$

Upon substituting $\alpha_i = Y_{i,(1)}$, $1 \le i \le m$, into Equation (1.7), the profile log-LF reads

$$\begin{aligned} l_p(\theta; \mathbf{y}) &:= \sum_{i=1}^{m} \left(\theta n_i \log y_{i,(1)} + \sum_{j=1}^{n_i} \log \left(\frac{1}{y_{ij}^{\theta}} - \frac{1}{(y_{ij}+1)^{\theta}} \right) \right) \\ &= \sum_{i=1}^{m} \left(\theta \left(n_i \log y_{i,(1)} - \sum_{j=1}^{n_i} \log y_{ij} \right) + \sum_{j=1}^{n_i} \log \left(1 - \left(\frac{y_{ij}}{y_{ij}+1} \right)^{\theta} \right) \right). \end{aligned} \tag{3.2}$$

Now, we proceed with two cases.

Case I: All equal observations within classes. Suppose that

$$Y_{ij} = Y_{i,(1)}, \qquad 1 \le j \le n_i, \qquad 1 \le i \le m. \tag{3.3}$$

Therefore, the profile log-LF (3.2) simplifies to

$$l_p(\theta; \mathbf{y}) = \sum_{i=1}^{m} n_i \log \left(1 - \left(\frac{y_{i,(1)}}{y_{i,(1)}+1} \right)^{\theta} \right).$$

Obviously, $l_p(\theta;\mathbf{y})$ is continuous for $\theta\in(0,\infty)$ and

$$l'_p(\theta;\mathbf{y}) := \frac{\partial}{\partial\theta} l_p(\theta;\mathbf{y}) = -\sum_{i=1}^{m} n_i \times \frac{\left(\frac{y_{i,(1)}}{y_{i,(1)}+1}\right)^{\theta}\log\left(\frac{y_{i,(1)}}{y_{i,(1)}+1}\right)}{1-\left(\frac{y_{i,(1)}}{y_{i,(1)}+1}\right)^{\theta}} > 0.$$

Therefore, $l_p(\theta;\mathbf{y})$ is strictly increasing in θ. Hence, given Equation (3.3), the MLE of θ does not exist.

Case II: Unequal observations within classes. In this case, we find the following theorem.

Theorem 3.1. *Assume that Equation* (3.3) *does not hold for at least one* i, $1\le i\le m$. *Then, the unique MLE of* θ, *denoted by* $\hat{\theta}$, *exists and satisfies the following equation:*

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i}\log\left(\frac{y_{i,(1)}}{y_{ij}}\right) = \sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{y_{ij}^{\hat{\theta}}\log\left(\frac{y_{ij}}{y_{ij}+1}\right)}{(y_{ij}+1)^{\hat{\theta}}-y_{ij}^{\hat{\theta}}}. \tag{3.4}$$

Proof. The profile log-LF (3.2) is continuous for $\theta\in(0,\infty)$. Also, for $1\le i\le m$, one can see that $n_i\log y_{i,(1)}\le\sum_{j=1}^{n_i}\log y_{ij}$. Hence

$$\lim_{\theta\to0^+} l_p(\theta;\mathbf{y}) = -\infty, \qquad \lim_{\theta\to\infty} l_p(\theta;\mathbf{y}) = -\infty.$$

Therefore, $l_p(\theta;\mathbf{y})$ has at least one global maximum point, i.e. $\hat{\theta}$ exists. Moreover, the profile log-LF (3.2) concludes the following profile likelihood equation:

$$l'_p(\theta;\mathbf{y}) = \sum_{i=1}^{m}\left(n_i\log y_{i,(1)} - \sum_{j=1}^{n_i}\log y_{ij} - \sum_{j=1}^{n_i}\frac{\left(\frac{y_{ij}}{y_{ij}+1}\right)^{\theta}\log\left(\frac{y_{ij}}{y_{ij}+1}\right)}{1-\left(\frac{y_{ij}}{y_{ij}+1}\right)^{\theta}}\right) = 0. \tag{3.5}$$

Clearly, $\hat{\theta}$ is a solution for Equation (3.5). Equation (3.4) is derived from Equation (3.5). Now, we prove that Equation (3.5) has a unique solution. To do this, note that

$$\frac{\partial^2}{\partial\theta^2} l_p(\theta;\mathbf{y}) = -\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{\left(\frac{y_{ij}}{y_{ij}+1}\right)^{\theta}\left(\log\left(\frac{y_{ij}}{y_{ij}+1}\right)\right)^2}{\left(1-\left(\frac{y_{ij}}{y_{ij}+1}\right)^{\theta}\right)^2} < 0.$$

Therefore, the function $l'_p(\theta;\mathbf{y})$ is strictly decreasing in θ. Since

$$\lim_{\theta\to0^+} l'_p(\theta;\mathbf{y}) = \infty, \qquad \lim_{\theta\to\infty} l'_p(\theta;\mathbf{y}) = \sum_{i=1}^{m}\sum_{j=1}^{n_i}\log\left(\frac{y_{i,(1)}}{y_{ij}}\right) < 0,$$

Equation (3.5) has exactly one solution which means that $\hat{\theta}$ is unique. □

Remark 3.2. If $\hat{\theta}$ exists, then an equi-tail $100(1-\gamma)\%$ CI for α_i is given by $I_{\hat{\theta}}(\alpha_i;\gamma)$, $1\le i\le m$.

3.2. Homogeneity testing

If α be the common value of α_is under the null hypothesis $H_0 : \alpha_1 = \cdots = \alpha_m$, we saw that the LF (1.6) simplifies to Equation (2.13). But, here $\theta > 0$ is unknown. Similarly, under the homogeneity hypothesis, one can see that $\hat{\alpha} = Y_{(1),(1)}$. Also, as in Subsection 3.1 (Case I), if

$$Y_{ij} = Y_{(1),(1)}, \qquad 1 \le j \le n_i, \qquad 1 \le i \le m, \tag{3.6}$$

then the MLE of θ does not exist. Notice that, if Equation (3.6) holds, then there is no reason to reject H_0. Otherwise, the unique MLE of θ will be derived from the following equation:

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i} \log\left(\frac{y_{(1),(1)}}{y_{ij}}\right) = \sum_{i=1}^{m}\sum_{j=1}^{n_i} \frac{y_{ij}^{\hat{\theta}} \log\left(\frac{y_{ij}}{y_{ij}+1}\right)}{(y_{ij}+1)^{\hat{\theta}} - (y_{ij})^{\hat{\theta}}}. \tag{3.7}$$

So, we propose to test $H_0 : \alpha_1 = \cdots = \alpha_m$ by following steps:

Step 1. Derive $\hat{\theta}_{H_0}$ from Equation (3.7);

Step 2. Given $\theta = \hat{\theta}_{H_0}$, use the procedure described in Subsection 2.2.

4. Numerical evaluation

In this section, three examples are given to illustrate the obtained results. In practice, researchers can consider the following steps in dealing with a real problem.

Step 1. By procedures described in Subsection 3.1, estimate the parameters θ and α_i, $1 \le i \le m$.

Step 2. If $\hat{\theta}$ exists, then check the basic assumptions (1.3) and (1.4). For checking the assumption (1.3), one my use a goodness of fit measure such as P-P plots, Pearson's chi-squared test and etc.
When $\hat{\theta}$ does not exist, if Equation (3.6) holds, then do not reject the homogeneity hypothesis. Otherwise, the proposed method has no response.

Step 3. If the basic assumptions are approved, then by procedure described in Subsection 3.2 test the homogeneity hypothesis $H_0 : \alpha_1 = \cdots \alpha_m$.
When the basic assumptions are not approved, the proposed method has no response.

First, we give two simulated example in which θ is known.

Example 4.1. By the statistical software $\mathcal{R}$ version 3.4.2, we generated 10^4 IID experiments according to the following cases for three populations.

	Population 1		Population 2		Population 3	
Case-Design	n_1	Distribution	n_2	Distribution	n_3	Distribution
1-Balanced	5	$DP(3,1)$	5	$DP(1,1)$	5	$DP(3,1)$
2-Balanced	5	$DP(1,1)$	5	$DP(1,1)$	5	$DP(1,1)$
3-Unbalanced	3	$DP(3,1)$	5	$DP(1,1)$	8	$DP(3,1)$
4-Unbalanced	3	$DP(1,1)$	5	$DP(1,1)$	8	$DP(1,1)$

Descriptive statistics for $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\alpha}_3$ have been reported in Table 3. Given $\theta = 1$ and according to the procedure described in Subsection 2.2, we tested $H_0^* : \alpha_1 = \alpha_2 = \alpha_3 = \alpha^*$ instead of $H_0 : \alpha_1 = \alpha_2 = \alpha_3$, where α^* is the observed value of $\hat{\alpha}$; See, Remark 2.7. Therefore, at the significance level $\gamma = 0.05$, we derived the following test functions:

Table 3. Descriptive statistics in Example 4.1.

Case	Estimator	Min.	Median	Mean	Max.	Bias	Var.
1	$\hat{\alpha}_1$	3	3	3.3897	24	0.3897	0.8981
	$\hat{\alpha}_2$	1	1	1.0359	6	0.0359	0.0478
	$\hat{\alpha}_3$	3	3	3.3832	20	0.3832	0.9202
2	$\hat{\alpha}_1$	1	1	1.0395	8	0.0395	0.0557
	$\hat{\alpha}_2$	1	1	1.0376	6	0.0376	0.0490
	$\hat{\alpha}_3$	1	1	1.0368	8	0.0368	0.0519
3	$\hat{\alpha}_1$	3	3	4.0713	51	1.0713	5.1097
	$\hat{\alpha}_2$	1	1	1.0388	6	0.0388	0.0513
	$\hat{\alpha}_3$	3	3	3.1324	12	0.1324	0.1925
4	$\hat{\alpha}_1$	1	1	1.1892	16	0.1892	0.4114
	$\hat{\alpha}_2$	1	1	1.0390	5	0.0390	0.0521
	$\hat{\alpha}_3$	1	1	1.0055	4	0.0055	0.0063

Table 4. Estimated overall power and type I error for the proposed approach for homogeneity testing in Example 4.1.

Case	1	2	3	4
Percentage	98.17	1.39	99.60	2.58

For balanced design:

$$\text{If } \hat{\alpha}=1 \rightarrow \varphi^*(\mathbf{y}) = \begin{cases} 1, & \Lambda^* < 0.0312 \\ 0.4118, & \Lambda^* = 0.0312 \\ 0, & \Lambda^* > 0.0312 \end{cases}$$

$$\text{If } \hat{\alpha}=2 \rightarrow \varphi^*(\mathbf{y}) = \begin{cases} 1, & \Lambda^* < 0.0173 \\ 0.0996, & \Lambda^* = 0.0173 \\ 0, & \Lambda^* > 0.0173 \end{cases}$$

$$\text{If } \hat{\alpha}=3 \rightarrow \varphi^*(\mathbf{y}) = \begin{cases} 1, & \Lambda^* < 0.0074 \\ 0.1362, & \Lambda^* = 0.0074 \\ 0, & \Lambda^* > 0.0074 \end{cases}$$

$$\text{If } \hat{\alpha}=4 \rightarrow \varphi^*(\mathbf{y}) = \begin{cases} 1, & \Lambda^* < 0.0057 \\ 0.1490, & \Lambda^* = 0.0057 \\ 0, & \Lambda^* > 0.0057 \end{cases}$$

For unbalanced design:

$$\text{If } \hat{\alpha}=1 \rightarrow \varphi^*(\mathbf{y}) = \begin{cases} 1, & \Lambda^* < 0.0312 \\ 0.4207, & \Lambda^* = 0.0312 \\ 0, & \Lambda^* > 0.0312 \end{cases}$$

$$\text{If } \hat{\alpha}=2 \rightarrow \varphi^*(\mathbf{y}) = \begin{cases} 1, & \Lambda^* < 0.0116 \\ 0.0992, & \Lambda^* = 0.0116 \\ 0, & \Lambda^* > 0.0116 \end{cases}$$

$$\text{If } \hat{\alpha}=3 \rightarrow \varphi^*(\mathbf{y}) = \begin{cases} 1, & \Lambda^* < 0.0080 \\ 0.0139, & \Lambda^* = 0.0080 \\ 0, & \Lambda^* > 0.0080 \end{cases}$$

$$\text{If } \hat{\alpha}=4 \rightarrow \varphi^*(\mathbf{y}) = \begin{cases} 1, & \Lambda^* < 0.0058 \\ 0.0099, & \Lambda^* = 0.0058 \\ 0, & \Lambda^* > 0.0058 \end{cases}$$

The overall power and type I error of the proposed approach for homogeneity testing are reported in Table 4. According to the obtained results, the MLEs have positive bias and the overall type I errors are less than the nominal level 0.05. Also, in view of power, the proposed approach for homogeneity testing is powerful in both balanced and unbalanced considered designs. These findings confirm the accuracy and precision of the proposed method in Subsection 2.2. ■

Example 4.2. In this example, we used statistical software $\mathcal{R}$ version 3.4.2 to study the empirical coverage of the CI for the parameter α_1 in Equation (2.12), as an illustration. To do this, we let $\theta = 1$ and investigated two different cases. In the first, n is fixed and m varies while in the later n varies and m is fixed. Precisely, we consider the following cases:

Case A. In this case, we assumed that n is fixed, $\alpha \in \{1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 100\}$ and $m \in \{2, 3, 4, 5, 10, 20\}$. Then, we generated IID samples with size $n = 10$ for m DP populations $DP(\alpha_1, 1), \cdots, DP(\alpha_m, 1)$ where $\alpha_1 = \cdots = \alpha_m = \alpha$. Finally,

Table 5. The empirical coverage of the proposed CI in Example 4.2 - Case A.

m \ α	1	2	3	4	5	10	20	30	40	50	100
2	99.89	98.05	99.45	98.34	99.19	98.25	97.69	97.82	97.43	97.52	97.58
3	99.94	98.35	99.37	98.13	99.14	98.14	97.55	97.86	97.62	97.77	97.51
4	99.87	98.33	99.30	98.13	99.27	98.21	97.64	97.70	97.61	97.67	97.33
5	99.91	98.07	99.37	98.10	98.99	98.39	97.51	98.00	97.66	97.57	97.58
10	99.82	98.03	99.41	98.22	99.19	98.11	97.48	97.94	97.54	97.63	97.47
20	99.85	98.37	99.41	98.35	99.00	98.11	97.24	97.66	97.75	97.57	97.47
50	99.87	98.26	99.32	98.32	99.12	98.13	97.57	97.94	97.70	97.77	97.54

we calculated 95% CI for the parameter α_1 from Equation (2.12). We iterated this procedure 10^4 times. The empirical coverages of the CIs are given in Table 5.

Case B. In this case, we assumed that m is fixed and $\alpha \in \{1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 100\}$. Then, we generated IID samples with size $n \in \{1, 2, 3, 4, 5, 10, 20, 50\}$ for $m = 5$ DP populations $DP(\alpha_1, 1), \cdots, DP(\alpha_m, 1)$ where $\alpha_1 = \cdots = \alpha_m = \alpha$. Eventually, we calculated 95% CI for the parameter α_1 from Equation (2.12). We iterated this procedure 10^4 times. The empirical coverages of the CIs are given in Table 6.

From Tables 5 and 6, the empirical coverage of the proposed CI is greater than the nominal coverage 95% in both cases A and B. Therefore, the proposed CI in Equation (2.12) is conservative. Also, for given m, the empirical coverages are usually increasing in n (sample size) and decreasing in the underlying parameter α. ■

Table 6. The empirical coverage of the proposed CI in Example 4.2 - Case B.

n \ α	1	2	3	4	5	10	20	30	40	50	100
1	97.46	97.42	97.49	97.64	97.69	97.54	97.36	97.41	94.47	95.52	95.77
2	97.78	97.89	97.51	97.93	97.57	97.29	97.56	97.53	97.41	97.37	95.75
3	98.67	97.68	97.97	97.50	97.86	97.56	97.67	97.53	97.24	96.98	97.68
4	98.87	98.84	97.75	98.32	97.86	97.67	97.38	97.57	97.45	97.76	97.45
5	99.55	98.94	98.65	98.43	98.24	97.13	97.55	97.62	97.32	97.45	97.54
10	99.87	98.09	99.38	98.55	98.98	98.12	97.58	97.74	97.49	97.59	97.86
20	100.00	99.98	99.68	98.80	99.91	99.38	98.74	98.24	98.53	98.14	97.78
50	100.00	100.00	100.00	99.99	99.99	99.07	99.07	98.95	99.12	97.70	97.92

Example 4.3. Suppose that, the number of defective tires produced by three production lines of a factory in the past seven days are as in Table 7. We want to answer the following qustion: "Is there a significant difference among the minimum number of defective tires produced by the production lines?" To do this, we applied the proposed model in Sections 2 and 3. Accordingly, Equation (3.1) concluded that

$$\hat{\alpha}_1 = 5, \qquad \hat{\alpha}_2 = 7, \qquad \hat{\alpha}_3 = 5.$$

Therefore, the unique MLE of θ exists and derived from Equation (3.4) as $\hat{\theta} = 1.6068$. For checking the basic assumption (1.3), we used the P-P plots. Then, for the i-the class, $1 \leq i \leq m$, we drew the sample cumulative probabilities against the cumulative probabilities of $DP\left(\hat{\alpha}_i, \hat{\theta}\right)$. The obtained results are given in Figure 1. From Figure 1, there is no reason to reject the basic assumption (1.3). So, we considered the procedure described in Subsection 3.2. From Equation (3.7), we derived $\hat{\theta}_{H_0} = 1.3609$. Therefore,

Table 7. The number of defective tires produced by three production lines of a factory.

Line \ Day	1	2	3	4	5	6	7
Line 1	5	6	12	5	26	7	10
Line 2	14	9	9	32	11	7	21
Line 3	7	13	6	10	5	5	30

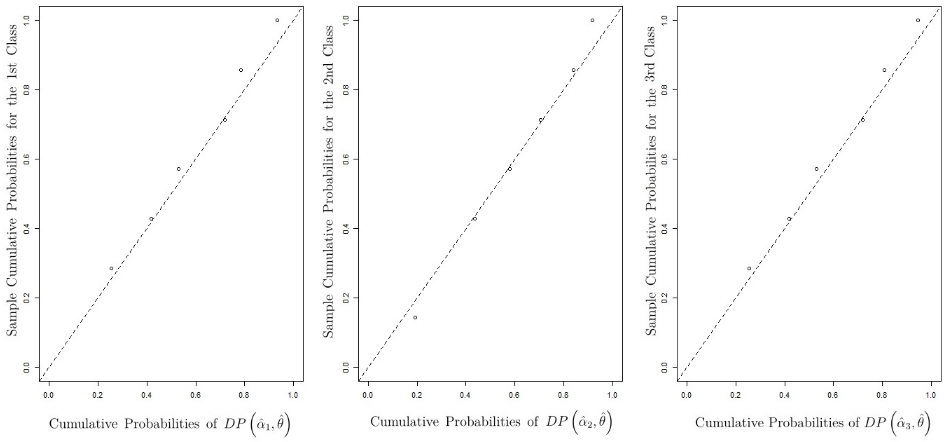

Figure 1. The P-P plots in Example 4.3.

Table 8. Findings of the Kruskal-Wallis test in Example 4.3.

Test statistic	Degrees of freedom	p-value
2.7315	2	0.2552

at the significance level $\gamma = 0.05$, we obtained the following test function for testing $H_0^* : \alpha_1 = \alpha_2 = \alpha_3 = 5$.

$$\varphi^*(\mathbf{y}) = \begin{cases} 1, & \Lambda^* < 0.0114, \\ 0.0877, & \Lambda^* = 0.0114, \\ 0, & \Lambda^* > 0.0114. \end{cases}$$

Using this test function, the homogeneity hypothesis H_0^* is accepted, since $\Lambda^* = 0.0405$. Hence, there is no significant difference among the minimum number of defective tires produced by the production lines. For comparison, findings of the Kruskal-Wallis test are also reported in Table 8. So, the final result of the proposed method is in accordance with the result of the Kruskal-Wallis test. ■

Example 4.4. McClave and Dietrich [13]: Studies conducted at the University of Melbourne indicate that there may be a difference between the pain thresholds of blonds and brunettes. Men and women of various ages were divided into four categories according to hair color: light blond, dark blond, light brunette, and dark brunette. Each person in the experiment was given a pain threshold score based on his or her performance in a pain sensitivity test. Results of the experiment are as in Table 9. Here, we want to determine whether the minimum of the pain threshold is equal among the four groups? To do this,

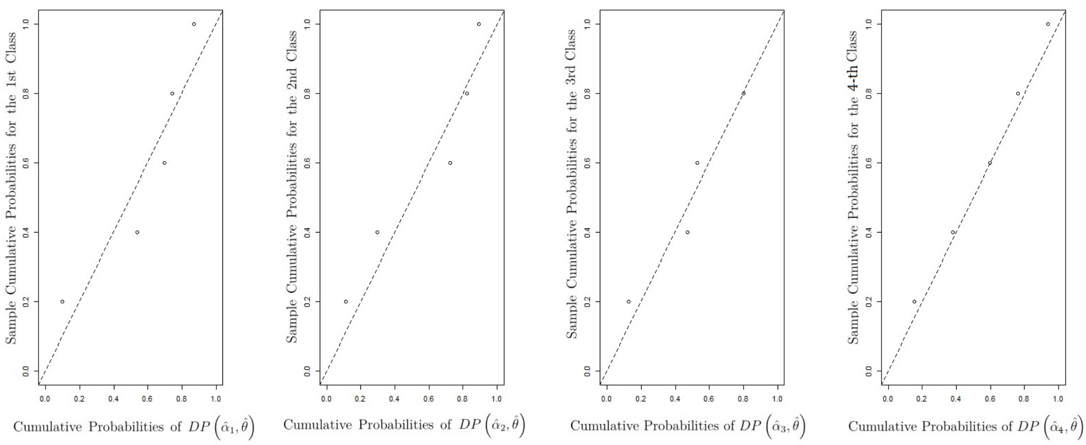

Figure 2. The P-P plots in Example 4.4.

we employed the proposed method in the paper. So, we derived from Equation (3.1)

$$\hat{\alpha}_1 = 48, \qquad \hat{\alpha}_2 = 41, \qquad \hat{\alpha}_3 = 37, \qquad \hat{\alpha}_4 = 30.$$

Then, we concluded from Equation (3.4) that $\hat{\theta} = 5.0043$. Similar to Example 4.3, we used the P-P plots to examine the basic assumption (1.3). The obtained results are given in Figure 2. From Figure 2 and without rigor, the basic assumption (1.3) is not violated. Therefore, we considered the procedure described in Subsection 3.2. From Equation (3.7), we derived $\hat{\theta}_{H_0} = 2.2223$. Hence, for testing $H_0^* : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 30$ at the significance level $\gamma = 0.05$, the proposed test function is derived as

$$\varphi^*(\mathbf{y}) = \begin{cases} 1, & \Lambda^* < 0.0008, \\ 0.0020, & \Lambda^* = 0.0008, \\ 0, & \Lambda^* > 0.0008. \end{cases}$$

Finally, the homogeneity hypothesis H_0^* is rejected, since $\Lambda^* = 2.5994 \times 10^{-5}$. This means that, the minimum of the pain threshold is different among the four groups. Moreover, findings of the Kruskal-Wallis test are given in Table 10. Then, the obtained inference from the proposed method is similar to the result of the Kruskal-Wallis test. ■

Table 9. Pain thresholds of blonds and brunettes in Example 4.4.

Hair color	The amount of pain				
Light blond	62	60	71	55	48
Dark blond	63	57	52	41	43
Light brunette	42	50	41	37	
Dark brunette	32	39	51	30	35

Table 10. Findings of the Kruskal-Wallis test in Example 4.4.

Test statistic	Degrees of freedom	p-value
10.5890	3	0.0142

5. Conclusion

This paper dealt with the one-way classification analysis when the response variable follows the DP distribution. It is an alternative method for some cases in which the classical models can not be used. The classical models such as one-way ANOVA and GLMs investigate the factor effects on a function of the response mean, while the proposed model in this paper investigates the factor effects on the minimum bound parameter of the response distribution. In the proposed model, one can also estimate the factor effects. However, the well known Kruskal-Wallis test does not estimate the factor effects.
The derived results may be extended in some directions. For example, other discrete distributions such as discrete uniform and telescopic family of distributions are worth for consideration. The general case

$$Y_{i1}, \cdots, Y_{in_i} \overset{IID}{\sim} DP\left(\alpha_i, \theta_i\right), \qquad 1 \leq i \leq m.$$

and multivariate discrete responses are interesting. Extensions with random sample sizes are also feasible. Moreover, a challenging problem in practice is to determine how to check the basic assumptions of the model when the sample sizes are small.

Acknowledgment. The authors express their sincere thanks to the Editor-in-Chief, the Associate Editor and the anonymous reviewers for their many useful comments and suggestions on an earlier version of this manuscript which resulted in this improved version of the manuscript.

References

[1] M. Baratnia and M. Doostparast, *One-way classification with random effects: A reversed-hazard-based approach*, J. Comput. Appl. Math. **349**, 60–69, 2019.
[2] M. Baratnia and M. Doostparast, *A random effects model for comparing pareto populations*, Comput Ind Eng **147**, 106612, 2020.
[3] A. Buddana and T.J. Kozubowski, *Discrete pareto distributions*, Stoch. Qual **29**, 143–156, 2014.
[4] G. Casella and R.L. Berger, *Statistical Inference, 2nd Edition*, Thomson Learning, 2002.
[5] S.R. Cole, H. Chu and S. Greenland, *Maximum likelihood, profile likelihood, and penalized likelihood: A primer*, Am. J. Epidemiol. **179**, 252–260, 2014.
[6] A. Gut, *Probability: A Graduate Course, 2nd Edition*, Springer, 2013.
[7] J. Jiang, *Linear and Generalized Linear Mixed Models and Their Applications*, Springer, 2007.
[8] A.I. Khuri, *Advanced Calculus with Applications in Statistics: 2nd Edition Revised and Expanded*, Wiley, 2003.
[9] T.J. Kozubowski, A.K. Panorska and M.L. Forister, *A discrete truncated pareto distribution*, Stat. Methodol. **26**, 135–150, 2015.
[10] H. Krishna and P.S. Pundir, *Discrete burr and discrete pareto distributions*, Stat. Methodol. **6**, 177–188, 2009.
[11] W.H. Kruskal and W.A. Wallis, *Use of ranks in one-criterion variance analysis*, J Am Stat Assoc **47**, 583–621, 1952.
[12] E.L. Lehmann and G. Casella, *Theory of Point Estimation, 2nd Edition*, Springer, 1998.
[13] J.T. McClave and F.H. Dietrich, *Statistics*, Dellen Publishing, 1991.
[14] C.E. McCulloch, S.R. Searle and J.M. Neuhaus, *Generalized, Linear, and Mixed Models, 2nd Edition*, Wiley, 2008.
[15] J.A. Nelder and R.W.M. Wedderburn, *Generalized linear models*, J R Stat Soc Ser A Stat Soc **135**, 370–384, 1972.

[16] S.R. Searle, G. Casella and C.E. McCulloch, *Variance Components*, Wiley, 1992.