Heliyon 10 (2024) e37799

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Modelling the heterogeneity of tourist spending in a mature destination: An approach through infinite mixture ☆

E. Gómez-Déniz, N. Dávila-Cárdenes *

Department of Quantitative Methods, University of Las Palmas de G.C. and TIDES Institute, Campus Universitario de Tafira, Las Palmas de Gran Canaria, 35017, Las Palmas, Spain

ARTICLE INFO

Keywords:
Empirical and non empirical Bayes
Heterogeneous expenditure
Multivariate Pareto distribution
Tourism

ABSTRACT

Identifying tourists' preferences is essential for stakeholders to provide better products and services. Among the tools to classify such choices, expenditure segmentation is valuable for separating tourist groups with shared interests. The underlying idea of the (infinite) mixture model is that tourists spend on a specific activity depending on their preferences. However, the propensity to consume may be based on or influenced by the group to which the tourist belongs. Thus, such a tendency could increase depending on homogeneity or heterogeneity. This paper uses a compound distribution mixture to model the expenditure heterogeneity. The resulting mixture model derives from a multivariate Pareto (Lomax) distribution that is easy to implement and includes zero value in its support since it is empirically proven that a tourist's expenditure on some activity can be zero. Results show that once the spending on transport has been carried out, tourists prefer to spend more on food than other activities. Conditioned to the expense carried out on food, the mean expenditure on leisure activities is more significant than on transport. Finally, tourists would prefer to spend more on food than on transportation once they decide to spend on other activities.

1. Introduction

The United Nations World Tourism Organization (UNWTO) defines tourism expenditure as the amount paid by visitors or on their behalf to acquire goods, services, and valuables for their use or to give away during tourism trips. In addition, payments made in advance or after the trip for those services received are also considered. For tourism operators, it is essential to identify visitors' target markets. Segmenting the market helps to identify such targets because it is possible to study how the different activities tourists undertake during their stay affect their total expenditure. Several authors analyze expenditure, such as [1], who examined the effects of socio-demographic, travel-related, and psychographic variables on travel expenditures. Their results, looking for travel expenditure patterns based on multiple independent variables, found that income and trip-related characteristics were the most influential variables affecting tourism expenditures. As explained by [2], after reviewing twenty studies on segmentation, the authors affirm that tourism expenditure is a segmentation variable with tremendous potential as it helps to identify issues and challenges

☆ EGD and NDC have been partially funded by grant PID2021-127989OB-I00 (Ministerio de Economía y Competitividad, Spain). EDG is also partially funded by grant TUR-RETOS2022-075 (Ministerio de Industria, Comercio y Turismo).

* Corresponding author.
E-mail addresses: emilio.gomez-deniz@ulpgc.es (E. Gómez-Déniz), nancy.davila@ulpgc.es (N. Dávila-Cárdenes).
URLs: https://dmc.ulpgc.es/emilio-gomez-deniz.html (E. Gómez-Déniz), https://dmc.ulpgc.es/nancy-davila.html (N. Dávila-Cárdenes).

https://doi.org/10.1016/j.heliyon.2024.e37799
Received 5 March 2024; Received in revised form 21 August 2024; Accepted 10 September 2024

Available online 1 October 2024

that impede the growth of this area and its policy translation. As a result, stakeholders may provide better products, services, or tailor-made products; consequently, destination managers can develop efficient policies and marketing campaigns [3].

There have been different approaches to facing the market segmentation in tourism [4] make use of three statistical models (logit, OLS, and Tobit) to identify and examine the differences or similarities in analyzing the impact on expenditure levels and patterns of visits to festivals, considering visitors' socio-demographic and festival experience-related variables. Also [5], evaluate the usefulness of an expenditure-based segmentation as a relevant instrument of tourism policy to increase the economic benefits from travelers; they do it using decision trees and linear regression statistical techniques to explore which variables might be considered as the best predictors of travel expenditure. [6] focus on total spending and expenses on accommodation, food, shopping, and car rental and use OLS, quantile regression, and instrumental variable techniques. The results obtained suggest that the impact of socio-demographic and trip-related variables on tourist expenditures goes far beyond the mean effect.

In a tourism market context, segmentation is a technique to divide a heterogeneous group into subgroups with homogeneous features. The basic idea is that different groups of people with specific needs configure the market. As an example, [7] use constraints scales to the wine tourism context, a factor-cluster segmentation approach generated homogenous subgroups, and the analysis of variance tests indicated that preference and intentions to visit wine regions were significantly different among the clusters.

This article studies factors that fundamentally explain spending segmentation based on food, transportation, and activities in a tourist destination. We start from the idea that tourists who visit a particular region (the Canary Islands) do not constitute a homogeneous group but rather that each of them, or certain groups, have specific characteristics that differentiate them from the others and, therefore, will have different preferences. The tool for mixing probability distributions allows for modelling the heterogeneity of a group of individuals. Hence, it constitutes the basis of the proposed modelling. The final objective will be to understand tourist groups and their behaviour regarding expenses such as food, transportation, and activities in a specific tourist destination (expenditure segmentation). This measure will offer a series of results on spending that can provide tourism management authorities with information that allows them to organize themselves or consider new projects that may be more promising. Consequently, this organization can impact employment policies and local income derived from the tourism sector. In addition, given the spending in another activity, conditional expectations of the expenditure in one activity are obtained easily, which can be helpful for predicting purposes. Our results suggest that, based on the cost of transportation, the mean expenditure on food is greater than the mean expenditure on other activities. Furthermore, the spending on transport has been carried out; tourists prefer to spend more on food than other activities. In addition, once they decide on food, tourists spend more on other activities than transportation. Finally, tourists would prefer to spend more on food than on transportation once they choose to spend on different activities. Finally, an expression for the posterior mean of the expenditure given the sample information is provided in closed form and written, giving information about the heterogeneity of the group of tourists.

Empirically, tourists' spending in a given destination can be zero on specific activities, such as transportation, food, or other various activities, and shows a positive correlation between the different types of expenditures. This work considers a multivariate survival distribution derived from an inverse gamma mixture of exponential distributions to model the heterogeneity of the behaviour of additional payments carried out by the tourist in a destination. The result mixture model derives from a multivariate Pareto (Lomax) distribution that is easy to implement and includes zero value in its support, which brings a different approach to the more used in the literature and is distinct from the used by [8] who examine and discuss expenditure-based segmentation of tourists through a finite mixture modelling. Additionally, covariates are incorporated naturally and easily. In the statistical literature, the multivariate Pareto distribution considered here has been studied in the field of reliability by [9]. See also [10], [11], [12] and [13].

This paper is organized as follows. First, section 2 motivates this work and deals with the model proposed. Then, a numerical experiment is developed in Section 3, and finally, in Section 4, we present some concluding remarks.

2. Motivation and model specification

Empirically, it has been proven that tourists' expenditure on activities can be zero. Thus, a distribution with support in $[0,\infty)$ will better model this expenditure. In this case, the simplest case for modelling the expenditure is judged to have an exponential distribution with mean $\lambda_i > 0$. Here, i is a subscript representing the mean expenditure in an activity i, $i = 1,\dots,k$. A random variable $X \geq 0$ is said to have an exponential distribution with mean $\lambda_i > 0$, we will write $X \sim \mathcal{E}(\lambda_i)$, if the probability density function (pdf) of X is given by $f(x;\lambda_i) = (1/\lambda_i)\exp(-x/\lambda_i)$, $x \geq 0$. Suppose now that the effect of being attracted by other activities different from the ith activity is to either increase or decrease the parameter λ_i by a common positive factor Θ, so the mean expenditure in the different activities changes to $\lambda_i\Theta$, $i = 1,\dots,k$. For business people who benefit from spending, an optimistic view would suggest $\Theta > 1$. The parameter Θ is assumed to be unknown, a typical situation where the population to be studied is heterogeneous. Contrarily, this parameter is fixed for a homogeneous population and can be estimated from the observed data. In our present representation, we will assume that this parameter is an unknown quantity varying from one activity to another. Then, a description of the collective's structure regarding the expenditure on different activities is needed. We need a probabilistic representation with which a random activity falls in a portion of Θ from the activities described by a distribution $g(\Theta)$.

The mixture of distributions is significant in modelling non-homogeneous populations that are very common in the tourism market. For instance, tourists in a specific destination, such as the Canary Islands, can proceed from different regions or countries, such as Spanish from the mainland, Germans, British, etc. Therefore, their behaviour concerning expenditures such as food, transport, and activities in a given tourist spot could differ. The underlying idea of the (infinite) mixture model is that tourists spend on a specific activity depending on their preferences. However, the propensity to consume may be based on or influenced by the group to which the tourist belongs. Thus, such a tendency could increase depending on homogeneity or heterogeneity. For example, a tourist may

Table 1
Illustration of the heterogeneity of the expending of the populations of tourists.

Tourists \ Activity	1	2	…	k
1	x_{11}	x_{21}	…	x_{k1}
2	x_{12}	x_{22}	…	x_{k2}
⋮	⋮	⋮	⋱	⋮
n	x_{1n}	x_{2n}	…	x_{kn}
Θ	θ_1	θ_2	…	θ_k

prefer a specific activity, such as visiting a museum. Consequently, his relationships with other tourists he has contact with may also influence the latter to wish to see the museum.

Initially, we considered a group of similar but somewhat heterogeneous tourists. It is assumed that detailed prior statistics are available from this pool; in particular, the mean of the expenditure in every one of the activities should be computed as $\bar{x}_i = (1/n)\sum_{j=1}^n x_{ij}$, i.e., the average value of the expenditure in the activity i.

Table 1 illustrates the comments above with k activities and n tourists. The tourist j makes an expenditure of x_{ij} in the activity i. Furthermore, the tourist tends to expend in an activity i measured by the parameter θ_i. From a Bayesian point of view, a prior density can be used to describe the heterogeneity of the collection of activities, say $g(\Theta)$, which tells us how this propensity is distributed through the collection of activities. In this case, θ_i are considered particular realizations of the random variable Θ.

2.1. Model

We initially assume independent and identically distributed exponential distribution for the joint expenditure in the k activities. That is, $f(x_1,\dots,x_k) = \Theta^{-k}\prod_{i=1}^n \lambda_i^{-1}\exp(-\Theta^{-1}\sum_{i=1}^k x_i/\lambda_i)$, $i=1,\dots,k$. Now, the model is introduced by the mixture of independent multivariate exponential distributions. Since $\Theta > 0$, a reasonably flexible and analytically tractable is the inverse gamma distribution. We assume that Θ follows an inverse gamma distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$. Here, λ_i is a parameter associated with the expenditure related to the activity i and θ a shared parameter with the expenditure of the collection of activities. Then, the unconditional distribution of $(x_1,\dots,x_k)$ is given by

$$\int_0^\infty f(x_1,\dots,x_k|\Theta)g(\Theta)\,d\Theta = \int_0^\infty \prod_{i=1}^k \lambda_i^{-1}\Theta^{-k}\exp\left(-\frac{1}{\Theta}\sum_{i=1}^k \frac{x_i}{\lambda_i}\right)g(\Theta)\,d\Theta.$$

Some straightforward computations provide the unconditional joint multivariate distribution, which results

$$f(x_1,\dots,x_k) = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\beta^k\prod_{i=1}^k \lambda_i}\left(1-\sum_{i=1}^k \frac{x_i}{\beta\lambda_i}\right)^{-\alpha-k}, \quad x_1,\dots,x_k \ge 0. \tag{1}$$

Denoting now $\sigma_i = \beta\lambda_i$, $i=1,\dots,k$, we have that (1) can be rewritten as

$$f(x_1,\dots,x_k) = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\prod_{i=1}^k \sigma_i}\left(1+\sum_{i=1}^k \frac{x_i}{\sigma_i}\right)^{-\alpha-k}, \quad x_1 \ge 0,\dots,x_k \ge 0, \tag{2}$$

with $\alpha > 0$ and $\sigma_i > 0$, $i=1,\dots,k$.

This multivariate distribution given in (2) can be also obtained by assuming $Y_1,\dots,Y_k$ and Y_α be mutually independent inverse gamma random variables with distributions $Y_i \sim \mathcal{IG}(1,1)$, $i=1,\dots,k$ and $Y_\alpha \sim \mathcal{IG}(\alpha,1)$, $\alpha>0$, where $\mathcal{IG}(\alpha,\beta)$ represents the inverse gamma distribution with parameters $\alpha>0$ and $\beta>0$. In this case, the multivariate dependent Pareto distribution (2) can be written by the following stochastic representation,

$$(X_1,\dots,X_k)' = \left(\sigma_1\frac{Y_1}{Y_\alpha},\dots,\sigma_k\frac{Y_k}{Y_\alpha}\right)',$$

where $\sigma_i > 0$, $i=1,\dots,k$. For this family, the associated copula is the Pareto copula or Clayton copula. See [14] for details.

The survival function corresponding to the pdf given in (2) results

$$\bar{F}(x_1,\dots,x_k) = \left(1+\sum_{i=1}^k \frac{x_i}{\sigma_i}\right)^{-\alpha}, \tag{3}$$

having Pareto type II (Lomax) marginals. A distribution analog to the one provided in (3), but sharing the parameter σ by all the marginals, has been used recently in actuarial statistics.

Marginal means, variance, covariance, and correlation are given by,

$$\mathbb{E}(X_i) = \frac{\sigma_i}{\alpha-1}, \quad i=1,2,\dots,k, \ \ \alpha>1,$$

$$var(X_i) = \frac{\alpha\sigma_i^2}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2,$$
$$cov(X_i, X_j) = \frac{\sigma_i\sigma_j}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2, \ i \neq j,$$
$$\rho(X_i, X_j) = \alpha^{-1}, \quad \alpha > 2, \ i \neq j,$$

respectively.

2.2. The role of the covariates

Most models explaining the tourism demand use expenditure as a dependent variable. These models are supported by covariates related to socioeconomic level, nationality, age, job, income, length of travel, type of travel, vacation accommodation, group travel, and loyalty to the destination. See for instance, [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28] and [29].

When covariates should be introduced into the model, it is convenient to rewrite (3) as

$$\bar{F}(x_1, \dots, x_k) = \left(1 + \frac{1}{\alpha-1}\sum_{i=1}^{k}\frac{x_i}{\mu_i}\right)^{-\alpha}, \quad \alpha > 1 \tag{4}$$

ensuring that $E(X_i) = \mu_i > 0$, $i = 1, \dots, k$. In this case, the corresponding pdf obtained from (4) is given by

$$f(x_1, \dots, x_k) = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)(\alpha-1)^k \prod_{i=1}^{k}\mu_i}\left(1 + \frac{1}{\alpha-1}\sum_{i=1}^{k}\frac{x_i}{\mu_i}\right)^{-\alpha-k}, \quad \alpha > 1. \tag{5}$$

Now, let $\boldsymbol{z}_i = (z_{1i}, \dots, z_{pi})'$ be a vector of p covariates associated with the ith observation, $i = 1, 2 \dots, k$, a vector of linearly independent regressors that determine the expenditure at every action (food, sport, transport, etc.). For the ith observation, the model takes the form,

$$(X_1, \dots, X_k) \sim \mathrm{MP}(\theta_i, \alpha),$$
$$\log(\mu_i) = \boldsymbol{z}_i'\boldsymbol{\delta},$$

where $\mathrm{MP}(\mu_i, \alpha)$ represents the multivariate Pareto distribution with pdf given in (5) and $\boldsymbol{\delta} = (\delta_1, \dots, \delta_p)'$ is the corresponding vector of regression coefficients. Observe that the log link assumed ensures that μ_i falls within the interval $(0, \infty)$.

The marginal distributions obtained from (5) are given by

$$f(x_j) = \frac{\alpha}{(\alpha-1)\mu_j}\left(1 + \frac{x_j}{(\alpha-1)\mu_j}\right)^{-\alpha-1}, \quad j = 1, \dots, k, \ \alpha > 1, \tag{6}$$

which correspond to normalized Pareto type II (Lomax) distributions.

3. Data, testing and forecasting results

The database contained 38759 observations from the 2019 Canary Islands Tourist Expenditure Survey, carried out by the Canary Islands Institute of Statistics (ISTAC). The study sample corresponds to personal interviews with tourists on departure and provides quarterly information about their total expenditure in the Canary Islands. Appendix A of this work displays the parameters' numerical estimation using the maximum likelihood method.

3.1. Data

The model proposed in this study divides tourists' destination expenditure is divided into three segments, as provided in the data base. The first includes daily spending per person and day on transport, either on car hire or public transport. The second refers to expenditure on food and beverages. Finally, the third segment relates to activities that tourists may engage in during their visit, i.e., tourists may be interested in cultural activities, sports, health, shopping, etc. Besides being essential, these activities positively influence spending at the destination. Still, they also have a multiplier effect on the rest of the local economy and allow the design of policies that favour employment and the local income derived from tourism. For example, [6] refers to the case of tourists spending money in restaurants; these themselves will invest a part of the money earned on the acquisition of more food and beverages, another part on transport, heating, and lighting, accounting, and other services. The three segments considered allow us to define a three-dependent variables model.

Concerning the covariates, we consider those included in the database and detailed below. Depending on the nature of the variables, some were categorized, such as income, occupational status, and the option of booking their holidays in advance. The main reason for including all the covariates we use is related to the nature of our study and the heterogeneity that characterizes tourism analyzed along the previous epigraphs of the paper. As we analyze expenditures in food, transport, and activities segments

Table 2
Descriptive statistics of the dependent variables.

Variable	Mean	Stand. Dev.	Min	Max
transport	119.137	154.755	0	6678
food	300.369	457.373	0	16240
activity	187.634	370.683	0	18000.4

at the destination, characteristics such as age and gender may give rise to differences in the spending choice election. Depending on age and gender, expenses may focus more on one segment than the others. A priori, older people may have more propensity to spend on food than on activities. Regarding nationality, for instance, Germans, English, and Spanish have different profiles of tourists; being in a Spanish destination and depending on the reason for travelling (holiday or not), their election about spending may vary; maybe Spanish prefer food more than activities, and German people could order leisure activities among their spending preferences. The length of stay may also define the type of expenditures or preferences. Papers like [28], [30] analyze the aggregate spending and model it using the length of stay as a variable to be explained. Concerning travelling season is essential because the Canary Islands, a European territory, defines its high season in winter due to the mild temperatures that characterize its location in winter. Regarding repeating a destination, a tourist who decides to repeat knows the destination and reorients their expenditure preferences compared to when travelling for the first time. Finally, the accommodations category offers tourists several possibilities; living in a hotel may reduce their expenditures on food at the destination in cases like a full-included stay; otherwise, staying in an apartment or when travelling in a group and sharing expenses may reduce expenditures on food or transport, as renting a car. Next, we detail the covariates and their definitions.

- Age. (taken in logarithm). Age ranges from 16 to 97 years.
- Gender. The dichotomous variable that takes 1 = male; and 0 = Female.
- Nationality. Four dummy variables are used for the respective quarters of 2019 to study seasonality: TRIM1, TRIM2, and TRIM3.
- Length of stay (taken in logarithm): Trip duration measured on the number of nights in the Canary Islands.
- Quarterly seasonality. Four dummy variables are used for the respective quarters of 2019 to study seasonality: TRIM1, TRIM2, and TRIM3.
- Repetition. A dummy variable takes the value 1 when tourists have previously visited the Canary Islands and 0 otherwise.
- Occupational status, job: This variable contains the following categories. Salaried employee, a family business, or paid internship = 1. Self-employed / business person without employees = 2. Business person with at most ten employees = 3. Business person with ten or more employees = 4. Unemployed = 5. Students or people doing unpaid internships = 6. Retired, early retired, or has left working = 7. Permanently and permanently disabled = 8. People in compulsory military service or community service = 9. Housework, childcare, or dependents people = 10. Another type of inactivity (person of independent means, voluntary work, etc.) = 11.
- Income: This item in the survey is an ordered categorical variable, not a continuous one It takes the following values: = 1, from 12000 to 24000 €; = 2, from 24001 to 36000 €; = 3, from 36001 to 48000 €; = 4, from 48001 to 60000 €; = 5, from 60001 to 72000 €; = 6, from 72001 to 84000 €; and = 7, higher than 84000 €.
- Accommodation category (aloj. cat): This is a dummy variable that takes the value one if the tourist accommodation is in a hotel and 0 otherwise
- Holiday: This is also a dummy variable that takes the value one if the tourist is on vacation and 0 otherwise.
- Booking in advance (book adv). In this variable the following modalities are distinguished: 1 = booking the arrival day; 2 = booking in 1-15 days before the trip; 3 = 16 − 30 days before; 4 = 1 − 2 months; 5 = 3 − 6 months; 6 = More than 6 months; 7 = Do not know.
- Group. It refers to the number of people travelling together.
- Sun and Beach: This variable indicates the reason for visiting the Islands. It would take one if tourists came to the Canary Islands exclusively attracted by the sun and beaches, and 0 otherwise.

Some descriptive of the dependent variables considered here are displayed in Table 2.

3.2. Estimated results

Since the prior distribution, $g(\Theta)$, is unknown, we estimate the model's unknown parameters from the tourists' data, defined in the literature as empirical Bayes. We have fitted the observed values of the three dependent variables by using the marginal distributions provided in equation (6) and the multivariate Pareto model in equation (5), via maximum likelihood method and the results obtained are shown in Table 3. As we can see, all the estimated parameters are statistically significant, attending to the standard errors (in parentheses) and the estimated values of the parameters. Here, the values of the negative of the log-likelihood function (NLL) are also provided.

The population correlation between variables obtained from the estimated value of α under the model with covariates results in 0.1962, which is near to the empirical correlation between the variables, which are as follows: $\rho(\text{transport}, \text{food}) = 0.123$, $\rho(\text{transport}, \text{activity}) = 0.095$ and $\rho(\text{food}, \text{activity}) = 0.173$. The population correlation between variables obtained using the estimated value of α under the model without including covariates is 0.285. Using the estimated values of the parameters provided in Table 3,

Table 3
Estimation results of the model without including covariates.

	Multivariate	Transport	Food	Activity
Parameter	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
$\hat{\alpha}$	3.513 (0.053)	4.839 (0.190)	1.898 (0.044)	1.565 ($< 10^{-3}$)
$\hat{\mu}_1$	130.226 (0.931)	119.943 (0.804)		
$\hat{\mu}_2$	309.152 (2.186)		349.881 (6.010)	
$\hat{\mu}_3$	174.511 (1.206)			243.488 ($< 10^{-3}$)
Observations	38759	38759	38759	38759
NLL	718497.067	223336.460	257213.738	234858.719

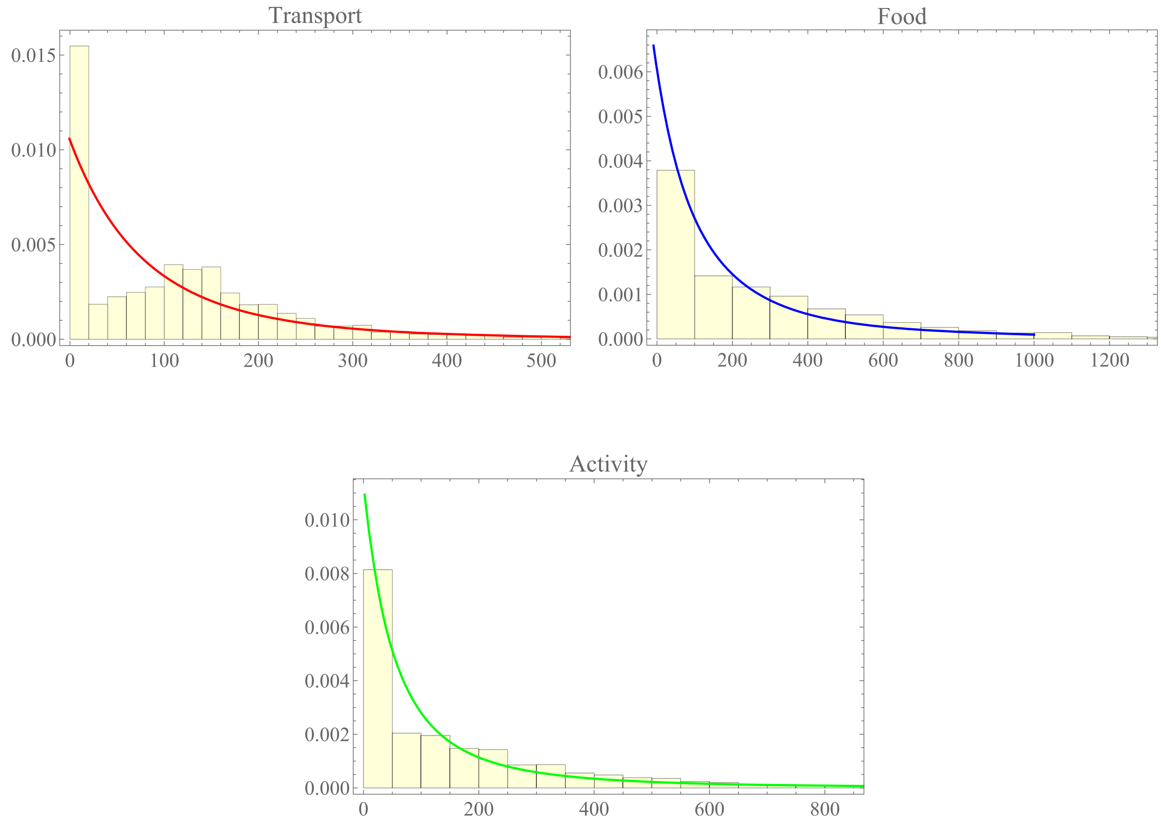

Fig. 1. Empirical histogram and fitted pdf (thick line) for the three dependent variables: transport (red), food (blue) and activity (green).

Fig. 1 shows the graph of the histogram and theoretical densities for the three studied types of expenditure. The empirical data's patterns reflect the suitability of the proposed distribution for modelling this type of data. Table 4 presents the estimated effects of the proposed model. In general, the marginal obtained impacts are statistically significant.

3.3. *Discussion including covariates*

Having studied the standard model without covariates, we now move on to study the effects that the different covariates used have on the different expenditure items. That is, we move on to study the model with covariates.

Table 4 presents the estimated effects of the proposed model. In general, the marginal obtained impacts are statistically significant.

The effect of income is always positive and significant on all three segments of tourist expenditure at the destination, i.e., the higher the income levels, the higher the expenditure. This result is consistent with the results of previous studies as in [22], [25], [27], [29] and [31], these authors also explain that higher income levels are positively related to higher levels of tourism expenditure. For [21], the socio-economic trait of tourists with the most significant impact on their daily expenditure is income. They affirm that tourists with high income spend 50% more than those who declare low income.

Regarding personal characteristics, estimates of age, gender, and nationality on the different segments of destination expenditure are carried out. For example, the estimated coefficients of the age variable (in logs) are significant and positive for tourists' expenses on transport and meals, a result that is in line with those obtained by [17], [28]. However, the effect is negative for spending on activities, suggesting that older tourists spend less on activities in the destination. Concerning gender, significant differences between

Table 4
Heterogeneous expenditure: empirical Bayes estimation.

	Transport	Food	Activity
Variable	Estimate (SE)	Estimate (SE)	Estimate (SE)
log(age)	0.186 (0.016)***	0.103 (0.016)***	-0.032 (0.016)*
gender	0.004 (0.010)	0.076 (0.008)***	0.023 (0.010)**
Germans	0.213 (0.015)***	-0.357 (0.016)***	-0.271 (0.016)***
Spanish	0.009 (0.015)	0.112 (0.015)***	-0.038 (0.015)**
British	-0.322 (0.013)***	0.069 (0.013)***	-0.059 (0.014)***
log(length of stay)	0.324 (0.008)***	0.753 (0.011)***	0.585 (0.013)***
trim 1	-0.074 (0.014)***	-0.006 (0.013)	-0.071 (0.015)***
trim 2	-0.001 (0.014)	$6.3 * 10^{-4}$ (0.014)	0.045 (0.015)**
trim 3	0.185 (0.014)***	$-7.6 * 10^{-4}$ (0.013)	0.116 (0.015)***
repetition	0.087 (0.011)***	-0.197 (0.012)***	0.034 (0.011)**
job	-0.015 (0.002)***	-0.001 (0.002)	-0.025 (0.002)***
income	0.033 (0.005)***	0.081 (0.005)***	0.015 (0.005)**
aloj cat	0.004 (0.011)	1.178 (0.011)***	-0.123 (0.010)***
holiday	0.609 (0.033)***	0.056 (0.026)**	0.689 (0.033)***
book adv	-0.013 (0.004)**	0.001 (0.004)	0.029 (0.004)***
group	0.101 (0.004)***	0.107 (0.004)***	0.157 (0.004)***
sun & beach	0.078 (0.011)***	0.070 (0.010)***	0.141 (0.012)***
constant	2.516 (0.072)***	2.602 (0.071)***	2.813 (0.074)***
α	5.095 (0.089)***		
Observations	38759		
NLL	707236.789		

*** indicates 1% significance level.
** indicates 5% significance level.
* indicates 10% significance level.

women and men are found in food and activity expenditure, with men estimated to be more likely to spend than women. These results are consistent with those obtained by [5], [32], and [33] in the sense that male tourists spend more at the destination than female tourists. Tourists of different nationalities behave differently at the destination regarding spending segments. For example, there is a significant and positive effect on German tourists' transport expenditure but a negative effect on food and activity. On the other hand, the Spanish are more likely to spend more on eating and less on activities, as the British do regarding the expenditure on eating and activities. However, their spending shows a negative effect on transport. The occupational status variable significantly and negatively affects all three expenditure segments. It is a categorical variable that moves from active professional to non-active professional categories. Thus, the tourist's spending capacity decreases, moving up on this scale. In the literature, there are numerous studies where the occupational level and tourist expenditure are related, [1,28–30,34] and [35], even as [31] explain, when the income of tourists is not known, the occupational level has sometimes been used as a proxy variable. Concerning the variables that refer to holiday characteristics, as in [36] repetition, or loyalty, to the destination determines tourism expenditure [36]. The coefficients estimated by the model for the repetition variable have positive and significant effects on spending on transport and activities at the destination; however, having made previous visits is negative and statistically significant for the food expenditure segment. In their model, [17] proposed that tourists repeating have positive effects on spending in bars, cafes, and car parks; however, it was not significant for restaurant spending. The coefficients for the length of stay (in logs) variable are positive and statistically significant for all three segments of destination expenditure. Consequently, the longer the stay in the tourist destination, the higher the spending in the three studied segments. These results coincide with those obtained by [29]. For the accommodation category, aloj. cat, the estimated effects are statistically significant, positive for food expenditure and negative for activities. For instance, in their estimates, [6] obtained that staying in better hotels negatively impacts shopping; these authors interpret it as a trade-off between accommodation and shopping expenditure. On the other hand, the effect of accommodation type on transport expenditure is not significant, which could be explained by hotel tourists being less likely to rent vehicles and engage in activities, [30]. When the motivation for the trip is holiday and sun and beach holidays, the effects on all three segments of destination spending are significant and positive. These results are in line with [28] and [37]. The positive impact on the three segments of destination expenditure suggests that this is a mature tourist destination with a wide range of alternatives to sun and beaches. Concerning the variable booking in advance, it positively impacts activities and negatively affects transport expenditure; as explained by [38], when tourists rent a car in advance, they usually get better prices, reducing their spending on transport at the destination. The estimated effects on destination expenditure on transport, food, and activities for the group travel variable, group, are significant and positive. Tourists organize and plan more activities when travelling in a group. Holidays in a group are cheaper because they reduce costs per person per day, making it attractive to do more activities, [21]. For [20], small to medium-sized groups such as families, groups of friends, or relatives are most suitable for sharing costs. Finally, quarterly seasonality is incorporated through dummy variables. The fourth quarter, TRIM4 (October-December), is the reference period. TRIM1 (January-March) negatively impacts transport and activities; however, TRIM2 (April-June) is only significant and negative for extra activities; finally, TRIM3 (June-September) presents results that are significant and positive for transport and activities. For no quarter is the expenditure on meals significant.

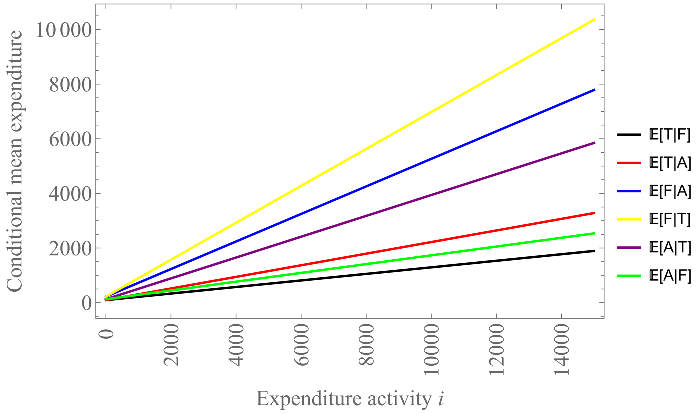

Fig. 2. Conditional expectations given the different activities.

3.4. Prediction via empirical Bayes method

Since the conditional distribution of $X_i|X_j = x_j$ is obtained in closed-form expression it is possible to get, from model (5), the conditional expectations (regression line of X_i on X_j) which are linear. They are given by equation (7),

$$\mathbb{E}(X_i|X_j = x_j) = \frac{\sigma_i(x_j + \sigma_j)}{\alpha\sigma_j}, \quad i, j = 1, \dots, k, \; i \neq j, \tag{7}$$

and therefore, the variation of expenditure's mean in the activity i concerning the spending in the activity j is measured through $\sigma_i/(\alpha\sigma_j)$.

Given the values provided in Table 3 (multivariate case), we have plotted the conditional means given in equation (7) for the three activities considered, which are displayed in Fig. 2. The following inequalities can be appreciated in the graph, $\mathbb{E}(A|T) < \mathbb{E}(F|T)$, $\mathbb{E}(A|F) > \mathbb{E}(T|F)$ and $\mathbb{E}(F|A) > \mathbb{E}(T|A)$. Roughly speaking, results indicate that, conditioned to the cost of transportation, the mean expenditure on food is greater than the mean expenditure on other activities. On the other hand, conditioned to the expense carried out on food, the mean expenditure on other activities is more significant than that carried out on transport; and finally, conditioned to the mean expenditure on other activities, the mean expenditure on food is greater than the mean expenditure on transport. In other words, once the spending on transport has been carried out, tourists seem to prefer to spend more on food than other activities. In addition, once they decide to spend money on food, tourists seem to spend more on activities other than transportation. Finally, tourists would prefer to spend more on food than on transportation once they decide to spend on other activities.

3.5. Prediction via non empirical Bayes method

For activity i, the a priori mean expenditure is given by $\mathbb{E}(X_i|\Theta) = \lambda_i\Theta$. The expenditure expected by the collective of activities results $\mathbb{E}(\lambda_i\Theta) = \sigma_i/(\alpha - 1)$, $\alpha > 1$. Finally, after observing a sample of size n given by $\tilde{x}_i = (x_{i1}, \dots, x_{in})$, we get that the posterior mean of the expenditure given the sample information results equation (8)

$$\mathbb{E}(\lambda_i\theta|\tilde{x}_i) = \frac{\sigma_i + n\bar{x}_i}{\alpha + n - 1} = [1 - \varphi(\alpha, n)]\,\mathbb{E}(\lambda_i\Theta) + \varphi(\alpha, n)\bar{x}_i, \tag{8}$$

where $\bar{x}_i = (1/n)\sum_{j=1}^{n} x_{ij}$ and $\varphi(\alpha, n) = n/(\alpha + n - 1)$. The weight factor $\varphi(\alpha, n)$, a value between 0 and 1, can be interpreted as the amount of belief attached to the tourist's experience of expenditure in activity i and the collateral experiences. Straightforward computations (see the Appendix B) provide that $\varphi(\alpha, n) = n/(n + k)$, where $k = \mathbb{E}[var(X_i|\lambda_i\Theta)]/var[\mathbb{E}(X_i|\lambda_i\Theta)]$. Here, $var[\mathbb{E}(X_i|\lambda_i\Theta)]$ is a measure of the heterogeneity of the expenditure in the collection of activities. In contrast, the quantity $\mathbb{E}[var(\mathbb{E}(X_j|\lambda_i\Theta)]$ provides a global measure of the dispersion of these means among all activities. If the tourist group were utterly homogeneous, setting $\varphi(\alpha, n) = 0$ would be reasonable. At the same time, if the group were heterogeneous, it would be helpful to set $\varphi(\alpha, n) = 1$. Using intermediate values is appropriate because individual and group history helps infer future individual behaviour. For additional properties about this posterior mean, see [39].

Observe that the inverse gamma distribution is conjugate concerning the exponential one. A prior distribution is conjugate to an experiment when the prior distribution is so related to the conditional distribution that the posterior distribution is the same type as the prior. In this case, if $X_i \sim \mathcal{E}(\lambda_i\Theta)$ and $\Theta \sim \mathcal{IG}(\alpha, \beta)$, then the posterior distribution of $\lambda_i\Theta$ given the sample information $\tilde{x}_i$ results again an inverse gamma distribution with the updated parameters α^*, β^*, where $\alpha^* = \alpha + n - 1$ and $\beta^* = \beta + \sum_{j=1}^{n} x_{ij}/\lambda_i$. Thus, the resulting posterior mean given in equation (8) is direct. It is necessary to recall that for a large sample n, the posterior expectation given in equation (8) coincides with the sample mean. Thus, the Bayesian methodology is more appropriate for small values of n.

Example 1. Suppose, as above, that the expenditure in activity i follows an exponential distribution with mean $\lambda_i\Theta$. Assume $\lambda_i = 1$ for all $i = 1, 2$, i.e., two activities, A and B. Now, the practitioner elicits an inverse gamma prior to Θ. In practice, this model can

Table 5
Posterior mean expenditure given the observed sample mean $\bar{x}_A$ and $\bar{x}_B$.

Activity			
$\bar{x}_A$	Posterior mean (A)	$\bar{x}_B$	Posterior mean (B)
10	38.88	100	188.88
20	44.44	200	244.44
30	50.00	300	300.00
40	55.55	400	355.55
50	61.11	500	411.11
60	66.66	600	466.66
70	72.22	700	522.22
80	77.77	800	577.77
90	83.33	900	633.33
100	88.88	1000	688.88

be interpreted as follows: Θ represents a tourist's propensity to expend in some activity, and the inverse gamma prior reflects how that propensity is distributed within the population of tourists. Suppose now that the practitioner assumes that the most frequent expenditure (mode) in activity A is about 50 € while it is 200 € in the second activity, B. These quantities were assumed by recording the expenditure corresponding to $n = 5$ tourists in each activity. With this prior information, an inverse gamma prior with shape and scale parameters of $(\alpha, \beta) = (5, 300)$ for the first activity and $(\alpha, \beta) = (5, 1200)$ seems appropriate. Recall that the model of the inverse gamma distribution is given by $\beta/(\alpha + 1)$. Using equation (4), the correlation between these two activities results in 0.25, and the prediction of the posterior mean of the expenditure given some sample means are displayed in Table 5, which were obtained by using equation (8).

Here, $\varphi(\alpha, n) = n/(\alpha + n - 1) = 0.55$ gives us a heterogeneous population of tourists in their expenditure behaviour. Furthermore, the values of $\mathbb{E}[var(X_i|\lambda_i\Theta)]$ and $var[\mathbb{E}(X_i|\lambda_i\Theta)]$ result also very large, confirming the fact of a heterogeneous population.

4. Conclusions and limitations

Expenditure segmentation represents an essential instrument for understanding tourism expenditure behaviour. The descriptive data from the analyzed survey indicates that the average expenditure per tourist during their stay in the Canary Islands was 607,140 €. The distribution of such spending was as follows: 20% (119,137 €) on transportation, 49% (300,369 €) on food, and the remaining 31% (187,634 €) was spent on various activities related to tourism and vacations. According to the spending segments considered at the destination, the most important was food, which represented almost half of the tourist spending. In addition, spending on food is positively correlated with transportation and activities. This paper empirically shows that tourist's expenditure on any activities can be zero, so a distribution with support in $[0, \infty)$ is used to model this expenditure. On the other hand, the mixture of distributions is significant in modelling non-homogeneous populations that are very common in the tourism market. The underlying idea of the (infinite) mixture model is that tourists spend on a specific activity depending on their preferences. However, the propensity to consume may be based on or influenced by the group to which the tourist belongs. Thus, such a tendency could increase depending on homogeneity or heterogeneity.

In the model proposed in this study, tourists' destination expenditure is segmented into three, and covariates are included.

The first includes daily spending per person and day on transport, the second refers to expenditure on food and beverages, and the third relates to activities that tourists may engage in during their visit. Results show that:

- Once the spending on transport has been carried out, tourists prefer to spend more on food than other activities.
- However, they spend more on different activities than on transportation once spending on food was carried out.
- Finally, tourists would prefer to spend more on food than on transport once they decide to spend on other activities.

Besides being essential, these activities positively influence spending at the destination. They also have a multiplier effect on the rest of the local economy and allow the design of policies that favour employment and the local income derived from tourism. Moreover, this result engages entrepreneurs in different segments because they may study the feedback favouring mutual revenue growth.

It is worth highlighting two fundamental limitations of this work. The first limitation is technical and is based on the correlation between each pair of marginal variables being the same. However, the estimated value for the parameter that controls the correlation has given rise to a correlation similar to the empirical correlation between the different segments of the expenditure considered. Introducing more complex modelling, which makes analysis and calculations more difficult, is the only way to include a more flexible correlation.

On the other hand, the variables used to study the spending behaviour in the three segments considered are not exhaustive. All those available 2019 Canary Islands Tourist Expenditure Survey have been used. Other types of spending segments should be taken into account. Fundamentally, they have the expense of activities separated by the different activities that this expense contemplates. In this sense, and continuing with the available data, it would be interesting to incorporate, for instance, the spending structure attached to LGTBI+ tourism, which has shown enormous strength in the last decade. As empirically proven, such tourism spends much more than the classic tourism that has visited us since the 60s of the previous century.

CRediT authorship contribution statement

E. Gómez-Déniz: Writing – original draft, Supervision, Software, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **N. Dávila-Cárdenes:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Emilio Gómez-Déniz and Nancy Dávila-Cárdenes have been partially funded by grant PID2021-127989OB-I00 (Ministerio de Economía y Competitividad, Spain). Emilio Gómez-Déniz is also partially funded by grant TUR-RETOS2022-075 (Ministerio de Industria, Comercio y Turismo).

Data availability

Data will be made available on request.

Acknowledgements

We thank the reviewer and the editor for their valuable comments, which undoubtedly helped to improve the presentation of our results.

Appendix A

A.1. Estimation of the parameters

For a sample $\tilde{x} = \{x_{ij}; i = 1, \dots, k; j = 1, \dots, n\}$ of size n, the log-likelihood obtained from model (5) is given by,

$$\ell(\tilde{x}; \Delta) = n \left[\log\Gamma(\alpha + k) - \log\Gamma(\alpha) - k\log(\alpha - 1) - \sum_{i=1}^{k} \log\mu_i \right] - (\alpha + k) \sum_{j=1}^{n} \log\left(1 + \frac{1}{\alpha - 1} \sum_{i=1}^{k} \frac{x_{ij}}{\mu_i}\right), \tag{9}$$

where $\Delta = (\alpha, \mu_1, \dots, \mu_k)$ is the vector of parameters that have to be estimated.

The partial derivatives which provide the normal equations are obtained from (9) and result,

$$\frac{\partial\ell(\tilde{x}; \Delta)}{\partial\alpha} = n\left(\psi(\alpha + k) - \psi(\alpha) - \frac{k}{\alpha - 1}\right) - k\sum_{j=1}^{n} \log\left(1 + \frac{1}{\alpha - 1}\sum_{i=1}^{k} \frac{x_{ij}}{\mu_i}\right) + (\alpha + k)\sum_{j=1}^{n} \frac{\sum_{i=1}^{k} x_{ij}/\mu_i}{(\alpha - 1)^2 + \sum_{i=1}^{k} x_{ij}/\mu_i}, \tag{10}$$

$$\frac{\partial\ell(\tilde{x}_i; \Delta)}{\partial\mu_i} = -\frac{n}{\mu_i} + (\alpha + k)\frac{\sum_{i=1}^{k} x_{ij}/\mu_i^2}{\alpha - 1 + \sum_{i=1}^{k} x_{ij}/\mu_i}, \quad i = 1, \dots, k, \tag{11}$$

where $\psi(\cdot)$ is the logarithmic derivative of the gamma function.

For the model with covariates the vector to be estimated is given by $\Delta = (\alpha, \delta_1, \dots, \delta_p)$. The partial derivative with respect to the parameter α is similar to expression (10) but with μ_i replaced by $\mu_i = \exp(\boldsymbol{z}_i'\boldsymbol{\delta})$. Equation (11) should be replaced by

$$\frac{\partial\ell(\tilde{x}; \Delta)}{\partial\delta_j} = \frac{\partial\ell(\tilde{x}_i; \Delta)}{\partial\mu_i} z_j\mu_i, \quad j = 1, \dots, p.$$

Appendix B

We have that

$$var\left[\mathbf{E}(X_i|\Theta)\right] = \mathbf{E}[(\lambda_i\Theta)^2] - \left[\mathbf{E}(\lambda_i\Theta)\right]^2 = \lambda_i^2\mathbf{E}(\Theta^2) - \left(\frac{\lambda_i\beta}{\alpha - 1}\right)^2$$

$$= \frac{\lambda_i^2 \beta^2}{(\alpha-1)^2(\alpha-2)}.$$

On the other hand, we have that

$$\mathbf{E}[var(X_i|\Theta)] = \mathbf{E}(\lambda_i^2\Theta^2) = \frac{\lambda_i^2 \beta^2}{(\alpha-1)(\alpha-2)}.$$

Thus,

$$\frac{n\, var[\mathbf{E}(X_i|\Theta)]}{n\, var[\mathbf{E}(X_i|\Theta)] + \mathbf{E}[var(X_i|\Theta)]} = \frac{n}{n+\alpha-1}.$$

References

[1] Y. Wang, P. Rompf, D. Severt, N. Peerapatdit, Examining and identifying the determinants of travel expenditure patterns, Int. J. Tour. Res. 8 (5) (2006) 333–346.
[2] T. Vinnciombe, P. Sou, Market segmentation by expenditure: an underutilized methodology in tourism research, Tour. Rev. 69 (2) (2014) 122–136.
[3] J. Tang, T.J. Ying, S. Ye, Chinese pet owners traveling with pets: motivation-based segmentation, J. Hosp. Tour. Manag. 50 (2022) 31–43.
[4] S. Kim, B. Prideaux, K. Chon, A comparison of results of three statistical methods to understand the determinants of festival participants' expenditures, Int. J. Hosp. Manag. 29 (2010) 297–307.
[5] B. Svensson, P. Moreno, D. Martín, Understanding travel expenditure by means of market segmentation, Serv. Ind. J. 31 (10) (2011) 1683–1698.
[6] C. Santos, J. Cabral, An analysis of visitors' expenditures in a tourist destination: OLS, quantile regression and instrumental variable estimators, Tour. Econ. 18 (3) (2012) 555–576.
[7] M. Cho, M.A. Bonn, R.A. Brymer, A constraint-based approach to wine tourism market segmentation, J. Hosp. Tour. Res. 41 (4) (2017) 415–444.
[8] R. Mortazavi, M. Lundberg, Expenditure-based segmentation of tourists taking into account unobserved heterogeneity: the case of Venice, Tour. Econ. 26 (3) (2020) 474–499.
[9] T.K. Nayak, Multivariate Lomax distribution: properties and usefulness in reliability theory, J. Appl. Probab. 24 (1987) 170–177.
[10] K.V. Mardia, Multivariate Pareto distributions, Ann. Math. Stat. 33 (1962) 1008–1015.
[11] B.C. Arnold, Pareto Distributions, International Cooperative Publishing House, Silver Spring, MD, 1983.
[12] A. Chiragiev, Z. Landsman, Multivariate flexible Pareto model: dependency structure, properties and characterizations, Stat. Probab. Lett. 79 (16) (2009) 1733–1743.
[13] B.C. Arnold, Pareto Distributions, second ed., Chapman & Hall/CRC Monographs on Statistics & Applied Probability, 2015, Boca Ratón, FL.
[14] J.M. Sarabia, E. Gómez-Déniz, F. Prieto, V. Jordá, Risk aggregation in multivariate dependent Pareto distributions, Insur. Math. Econ. 71 (2016) 154–163.
[15] E. Aguiló, C. Juaneda, Tourist expenditure for mass tourism markets, Ann. Tour. Res. 27 (3) (2000) 624–637.
[16] P. Fredman, Determinants of visitor expenditures in mountain tourism, Tour. Econ. 14 (2) (2008) 297–311.
[17] R. Craggs, P. Schofield, Expenditure-based segmentation and visitor profiling at the Quays in Salford, UK, Tour. Econ. 15 (1) (2009) 243–260.
[18] D. Wang, M. Davidson, A review of micro-analyses of tourist expenditure, Curr. Issues Tour. 8 (5) (2010) 333–346.
[19] C. Marcussen, Determinants of tourist spending in cross-sectional studies and at Danish destinations, Tour. Econ. 17 (2011) 833–855.
[20] C. Thrane, E. Farstad, Domestic tourism expenditures: the non-linear effects of length of stay and travel party size, Tour. Manag. 32 (2011) 46–52.
[21] A. García-Sánchez, E. Fernández-Rubio, M. Collado, Daily expenses of foreign tourists, length of stay and activities: evidence from Spain, Tour. Econ. 19 (3) (2013) 613–630.
[22] J. Brida, M. Disegna, L. Osti, Visitors' expenditure behaviour at cultural events: the case of Christmas markets, Tour. Econ. 19 (5) (2013) 1173–1196.
[23] B. Zheng, Y. Zhang, Household expenditures for leisure tourism in the USA, 1996 and 2006, Int. J. Tour. Res. 15 (2) (2013) 197–208.
[24] C. Thrane, Modelling micro-level tourism expenditure: recommendations on the choice of independent variables, functional form and estimation technique, Tour. Econ. 20 (1) (2014) 51–60.
[25] E. Marrocu, R. Paci, A. Zara, Micro-economic determinants of tourist expenditure: a quantile regression approach, Tour. Manag. 50 (2015) 13–30.
[26] M. Disegna, L. Osti, Tourists' expenditure behaviour: the influence of satisfaction and the dependence of spending categories, Tour. Econ. 22 (1) (2016) 5–30.
[27] E. Aguiló, J. Rosselló, M. Vila, Length of stay and daily tourist expenditure: a joint analysis, Tour. Manag. Perspect. 21 (2017) 10–17.
[28] E. Gómez-Déniz, J. Pérez-Rodríguez, Modelling distribution of aggregate expenditure on tourism, Econ. Model. 78 (2019) 293–308.
[29] E. Gómez-Déniz, J. Pérez-Rodríguez, J. Boza-Chirino, Modelling tourist expenditure at origin and destination, Tour. Econ. 26 (3) (2020) 437–460.
[30] E. Gómez-Déniz, N. Dávila-Cárdenes, J. Boza-Chirino, Modelling expenditure in tourism using the log-skew normal distribution, Curr. Issues Tour. (2021), https://doi.org/10.1080/13683500.2021.1960282.
[31] S. Park, M. Woo, J.L. Nicolau, Determinant factors of tourist expenses, J. Travel Res. 59 (2) (2020) 267–280.
[32] C. Thrane, Jazz festival visitors and their expenditures: linking spending patterns to musical interest, J. Travel Res. 40 (3) (2002) 281–286.
[33] M. Saayman, A. Saayman, Determinants of spending: an evaluation of three major sporting events, Int. J. Tour. Res. 14 (2) (2012) 124–138.
[34] S.S. Jang, B. Bai, G.S. Hong, J.T. O'Leary, Understanding travel expenditure patterns: a study of Japanese pleasure travelers to the United States by income level, Tour. Manag. 25 (3) (2004) 331–341.
[35] J.L. Nicolau, F.J. Mas, Heckit modelling of tourist expenditure: evidence from Spain, Int. J. Serv. Ind. Manag. 16 (3) (2005) 271–293.
[36] F. Dayour, C.A. Adongo, F. Taale, Determinants of backpackers' expenditure, Tour. Manag. Perspect. 17 (2016) 36–43.
[37] J. Mitra, C. Wilson, S. Managi, P. Kler, P. Prayaga, U. Khanal, What determines whale watching tourists' expenditure? A study from Hervey Bay, Australia, Tour. Econ. 25 (7) (2019) 1134–1141.
[38] E. Gómez-Déniz, J. Boza-Chirino, N. Dávila-Cárdenes, Tourist tax to promote rentals of low-emission vehicles, Tour. Econ. 27 (7) (2021) 1461–1481.
[39] W. Ericson, A note on the posterior mean of a population mean, J. R. Stat. Soc., Ser. B, Methodol. 31 (2) (1969) 332–334.