On the distribution of nutritional status transition probabilities through a hierarchical beta mixture model

Luis Rojo-González
luis.rojo.g@usach.cl
University of Santiago of Chile (USACH), Faculty of
Engineering, Program for the Development of
Sustainable Production Systems (PDSPS), Santiago, Chile
University of Santiago of Chile (USACH), Faculty of
Engineering, Department of Industrial Engineering,
Santiago, Chile

Óscar C. Vásquez
oscar.vasquez@usach.cl
University of Santiago of Chile (USACH), Faculty of
Engineering, Program for the Development of
Sustainable Production Systems (PDSPS), Santiago, Chile
University of Santiago of Chile (USACH), Faculty of
Engineering, Department of Industrial Engineering,
Santiago, Chile

ABSTRACT

This work addresses the problem of estimating the probability density function for each nutritional status considering the sex and the kind of transition a person suffers between two consecutive periods. To this end, an optimization scheme is proposed based on a hierarchical model considering previously known information as input to estimate mixture densities through the expectation maximization algorithm. Besides, a goodness-of-fit metric is shown based on the shared area of the actual density obtained from a Dirichlet distribution and its close relation with the Beta distribution. Thus, several experiments are performed showing a goodness-of-fit of 93.1%, at least.

KEYWORDS

Nutritional status, Hierarchical mixture model, Beta distribution, Expectation Maximization

1 INTRODUCTION

Within the non-communicable diseases (NCDs), one of the most important ones is obesity, which is a nutritional status with other three -main- possible classifications: i) Underweight, ii) Normal, iii) Overweight, being obesity the last one. In particular, this disease has been widely studied in literature due to its close relation with serious health risks (comorbidities) [17] as well as high associated costs to face it [7, 9]. Even more, analyzing the data including the income level of countries, it shows an increasing behavior depending on the income level of the country [24, 27]. Regardless this fact, the exactly reverse effect can be observed when the income is analyzed at household level instead of countries. It is well known that the higher the country income the higher the inequality income is; thus, the household acquisition power is different for each person/family in the same country, thereby a lower income is generally associated with worse dietary profile [2, 23] and therefore with higher obesity presence and diet-related illnesses [16]. Although the prevalence[1] of nutritional statuses is very important from an economic point of view, it is more a public health challenge. For instance, people that have not a normal nutritional status are associated with an increased risk of anovulatory infertility [3] whereas people with overweight or obesity are more likely to develop type 2 diabetes

[1] and/or coronary artery disease [8]. Thus, the estimation of the nutritional statuses must be an outstanding concern for any government since it is expensive as well as dangerous for public health.

In literature, there have been several studies addressing the estimation of this phenomenon using the so-called transition probabilities, being transition rates or transference rates are often used as synonyms, which are defined as the probability or rate to move from one state of a categorical risk factor to another [25]. Although the results depend on these variables, estimation guidelines are not properly stated and hence different results might be obtained under the same experimental settings. Shumei Sun Guo et al. [13] fit logistic models in order to estimate the probability to develop overweight and/or obesity (using the body mass index as metric) until 35 years old considering a longitudinal study with Men and Women aged between 3 and 20 years. González-Parra et al. [12] develop an homogenous network epidemic model using the fact that NCDs are influenced by the environment, e.g. an obesogenic environment, assuming the waiting times (transition probabilities) are Gamma distributed. J. van de Kassteele et al. [25] formulate a linear programming model based on the well-known transportation problem to estimate the transition probabilities between any nutritional status. Zachary J. Ward et al. [26] perform a longitudinal study using the National Health and Nutrition Examination Survey to estimate the (transition) probability children aged 2 years have to be normal weight or obese at 35-years-old. Recently, Ávalos et al. [29] proposed a mathematical model based on cross-sectional data in Chile to estimate the transition probabilities from a body mass index to another on the assumption they depend on the age group of the person and just one transition is allowed between two consecutive years. Even although a work we found uses a distribution over the transition probabilities, this does not accomplish two important conditions, the sum up to 1 nor the domain of the probabilities [0, 1]. Thus, we propose a hierarchical approach to estimate the distribution of transition probabilities based on the literature review performed by Ávalos [28] where transition probabilities are reported. To achieve the proposed goal, we state the problem in Section 2 and how we address it in Section 3 where mathematical expressions are developed showing a general scheme to handle the estimation problem. Then, the obtained results are shown in Section 4 and thereafter analyzed to get insights for future works in Section 5.

2 STATEMENT OF THE PROBLEM

Let $\mathcal{I}$, $\mathcal{J}$ and $\mathcal{K}$ be the set of nutritional status, the set of transitions and the set of sexes, respectively. Thus, let's denote $X_{i,j,k}$ a positive variable that denotes the transition probabilities for

---

[1]Prevalence of a disease is defined as the net number of individuals that have the disease in a particular moment.

each nutritional status $i \in \mathcal{I}$, transition $j \in \mathcal{J}$ and sex $k \in \mathcal{K}$. Thus, we denote $Y_{j,k}$ that represents the probabilities associated to all set of nutritional status $i \in \mathcal{I}$ for a given transition $j \in \mathcal{J}$ and sex $k \in \mathcal{K}$ such that $Y_{j,k} = \bigcup_{i \in \mathcal{I}} X_{i,j,k}$. Besides, we denote $T_{i,k}$ and Z_k as collections of transition probabilities for a given nutritional status and sex, and the collection of transition probabilities for each sex $k \in \mathcal{K}$ such that $T_{i,k} = \bigcup_{j \in \mathcal{J}} X_{i,j,k}$ and $Z_k = \bigcup_{j \in \mathcal{J}} Y_{j,k} = \bigcup_{i \in \mathcal{I}} T_{i,k}$. Thereby, given the nature of probabilities, we must impose two constraints to be satisfied:

   i) Domain $[0, 1]$: Such as probabilities are defined, they must belong to a bounded interval satisfying $0 \leq X_{i,j,k} \leq 1$, $\forall i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}$.

   ii) Sum up to one: Let a person with nutritional status $i \in \mathcal{I}$ and sex $k \in \mathcal{K}$. Then, this person has a certain probability associated to each transition $j \in \mathcal{J}$, nevertheless we know that the sum over the transitions must satisfies $\sum_{j \in \mathcal{J}} X_{i,j,k} = 1$, $\forall i \in \mathcal{I}, k \in \mathcal{K}$.

Besides, we denote f as the probability density function (*PDF*) associated to transition probabilities such as $X_{i,j,k} \sim f(X_{i,j,k}|\theta_{i,j,k})$, $\forall i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}$, where $\theta_{i,j,k}$ represents the set of parameters that governs the *PDF*. Thus, we denote g and h as the *PDF*s associated to a certain transition $j \in \mathcal{J}$ and to a certain nutritional status $i \in \mathcal{I}$, such as $Y_{j,k} \sim g(Y_{j,k}|\eta_{j,k})$, $T_{i,k} \sim q(T_{i,k}|\tau_{i,k})$ and $Z_k \sim h(Z_k|\psi_k)$. The latter relationships can be drawn such as Figure 1 shows.

Figure 1: Illustration of the relationships between the stated variables over the sets for each sex $k \in \mathcal{K}$.

Thus, the problem is to estimate the *PDF* for each nutritional status $i \in \mathcal{I}$, transition $j \in \mathcal{J}$ and sex $k \in \mathcal{K}$ subject to the aforementioned conditions.

3 HIERARCHICAL BETA MIXTURE MODEL

We address the estimation of the aforementioned *PDF*s using a linear combination of them using the expectation maximization algorithm (EM) [6] to maximize the log-likelihood of a mixture model under a hierarchical scheme.

3.1 Mixture model and hierarchical scheme

Given these constraints, it is possible to propose that the *PDF*s associated to each transition $j \in \mathcal{J}$ (g) corresponds to a Beta distribution such as $Y_{j,k} \sim \text{Beta}(Y_{j,k}|\alpha_{j,k}, \beta_{j,k})$, $\forall k \in \mathcal{K}$ since they belong to interval $[0, 1]$; whereas the collection of the associated distribution to that variable that satisfy the sum up to one condition ($T_{i,k}$) corresponds to a Dirichlet distribution such as $Z_k \sim \text{Dirichlet}(\xi_{j,k})$, where $\xi_{j,k} \in \mathbb{R}^{|\mathcal{J}|}$, $\forall k \in \mathcal{K}$.

Fortunately, these two distributions are closely related such that a Dirichlet distribution is the multivariate generalization of the Beta distribution, where the relationship is given by the following equivalences

$$\alpha_{j,k} = \xi_{j,k}, \quad \beta_{j,k} = \sum_{\ell \in \mathcal{J} \setminus \{j\}} \xi_{\ell,k} \qquad \forall j \in \mathcal{J}, k \in \mathcal{K}$$

3.1.1 Beta mixture model. Thus, as the probabilities of each nutritional status $i \in \mathcal{I}$ are also Beta distributed we have that $\theta_{i,j,k} = (\alpha'_{i,j,k}, \beta'_{i,j,k})$; then, each *PDF* corresponding to each combination of transition $j \in \mathcal{J}$ and sex $k \in \mathcal{K}$ is given by the mixture model defined as

$$g(Y_{j,k}|\alpha_{j,k}, \beta_{j,k}) = \sum_{i \in \mathcal{I}} w_{i,j,k} f(Y_{j,k}|\alpha'_{i,j,k}, \beta'_{i,j,k}) \qquad (1)$$

subject to
$$\sum_{i \in \mathcal{I}} w_{i,j,k} = 1 \qquad (2)$$

$$w_{i,j,k} \geq 0 \qquad \forall i \in \mathcal{I} \quad (3)$$

Nonetheless, this problem is highly non-linear, thereby the approach we propose to address this issue is by using a hierarchical scheme using the EM algorithm simulating data from the Dirichlet distribution satisfying the sum up to one constraint for each nutritional status $i \in \mathcal{I}$.

3.1.2 Estimation through EM algorithm. We can rewrite expression (1) in order to get the likelihood l such that for each transition $j \in \mathcal{J}$ and sex $k \in \mathcal{K}$, we have that

$$l(Y_{j,k}|\alpha_{j,k}, \beta_{j,k}) = \prod_{t=1}^{n_{j,k}} \sum_{i \in \mathcal{I}} w_{i,j,k} f(Y_{j,k}|\alpha'_{i,j,k}, \beta'_{i,j,k}) \qquad (4)$$

where $n_{j,k}$ represents the length of the observed values (data) for a given transition $j \in \mathcal{J}$ and sex $k \in \mathcal{K}$. Then, the log-likelihood L is given by

$$L(Y_{j,k}|\alpha_{j,k}, \beta_{j,k}) = \sum_{t=1}^{n_{j,k}} \ln \left(\sum_{i \in \mathcal{I}} w_{i,j,k} f(Y_{j,k}|\alpha'_{i,j,k}, \beta'_{i,j,k}) \right) \qquad (5)$$

thereby, for each nutritional status $j \in \mathcal{J}$ and sex $k \in \mathcal{K}$ we have to tackle the non-linear optimization problem given by $\max_{j,k} L(\cdot)$ subject to expressions (2) and (3).

So far, we have stated the optimization problem about the estimation of the parameters that govern the associated distribution. However, we have to set the scheme in order to implement the EM algorithm and some metrics to control the estimation process.

3.1.3 Scheme. Regardless of the problem and the approach to solve it has been stated already, we have to take into account certain considerations regarding the most likely scenarios motivated by the high non-linearity of the model. Thus, we focus on the second stage of the EM algorithm where we can use as many suitable optimization algorithms as possible. Nonetheless, the solution might lie in a bad local maximum no matter how many iterations were in the settings. Therefore, and in concordance with the later analysis, it is useful to consider a rule that helps the algorithm to avoid local maximum looking for a better solution.

This issue has been widely studied especially when the estimation of neural networks arose. Indeed, we propose a transformation stage following the momentum term proposed by Rumelhart & McClelland [21] to improve the gradient descent method. In particular, when a local optimal point is found and it is not possible to improve after several trials, we consider a perturbation of the estimated parameter generating a random number between two values around the unit, i.e., randomly expand or contract the solution. Otherwise, if the algorithm gets a better solution the process continues while the stopping criteria, time limit, number of iterations and tolerance, have not been achieved. The general
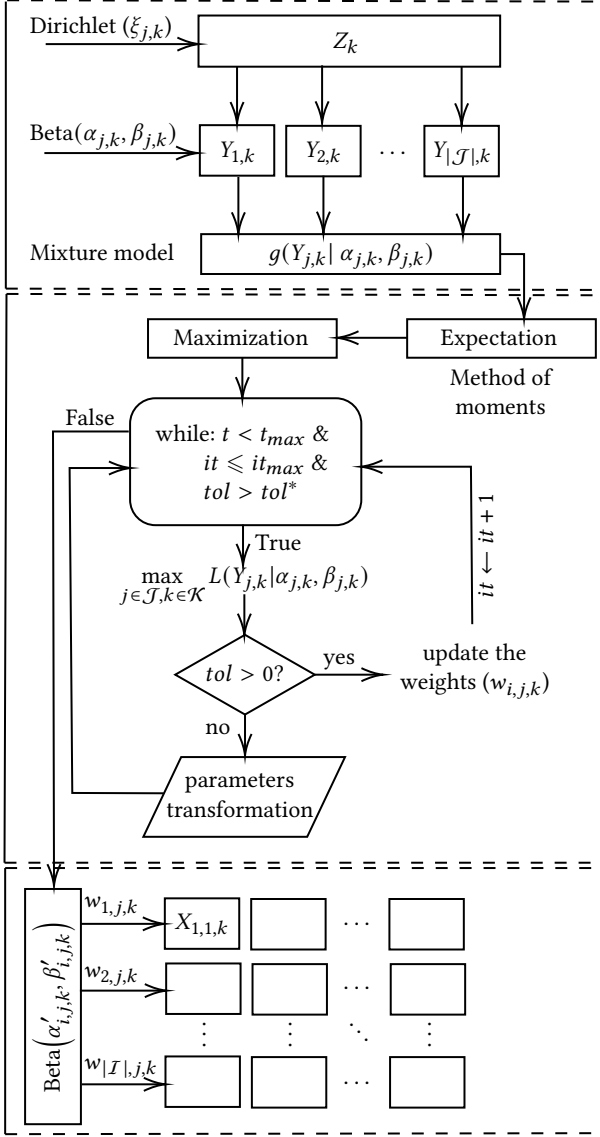
Figure 2: Hierarchical model estimation process workflow. Notice that the workflow depicts the hierarchical model, the expectation maximization algorithm and the mixtures densities in the first, second and third rectangle, respectively.

diagram of the estimation process is set out by in Figure 2, where the hierarchical scheme is depicted.

3.2 Goodness-of-fit

Regardless a local maximum is computed by the proposed optimization scheme, another issue arises when a bad local maximum is obtained and, even more, the non-linearity of the model does not necessarily imply that the best solution found is also the best fit. Thereby, we have to show how an expected result could be in the first place and how to pick a suitable measuring metric in order to get the goodness-of-fit of the solution.

We propose to assess the performance of the maximization process considering the shared area between the actual curve and the mixture model as a goodness-of-fit measure. This area, however, can be difficult to compute due to there might be several
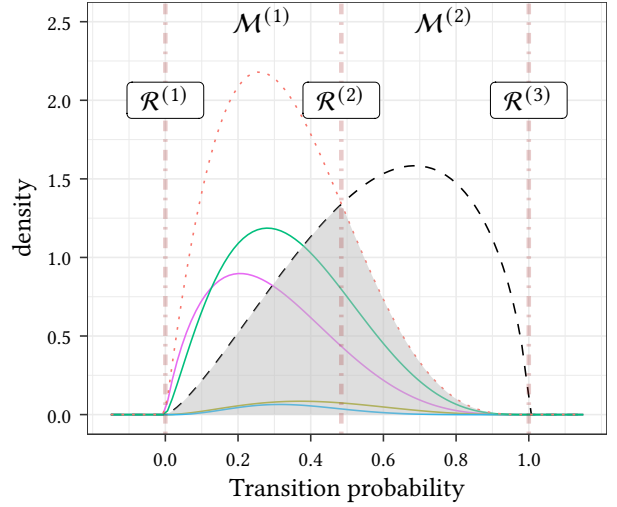

Figure 3: Illustration of a mixture model (dotted line) for a particular nutritional status $j \in \mathcal{J}$ and sex $k \in \mathcal{K}$ considering four mixture densities and the marginal beta distribution (black dashed line) with parameters 2.38 and 1.63, where the common area $A_{j,k}$ is depicted by the grey zone under both curves. Besides, it shows the roots and the intervals.

points where both curves intersect to each other. To this end, it is necessary to introduce $\mathcal{R}_{j,k}$ as the set of roots obtained from equation (1) ordered in an increasing order where the first and the last element are $\mathcal{R}_{j,k}^{(1)} = 0$ and $\mathcal{R}_{j,k}^{(|\mathcal{R}|)} = 1$, respectively, for each transition $i \in \mathcal{I}$ and sex $k \in \mathcal{K}$. Consequently, we consider the set of intervals between two consecutive roots where each interval is denoted by $\mathcal{M}_{j,k}^{(r)} = \{Y_{j,k} : \mathcal{R}_{j,k}^{(r)} \leq Y_{j,k} \leq \mathcal{R}_{j,k}^{(r)}\}$ such as for each transition $j \in \mathcal{J}$ and sex $k \in \mathcal{K}$ we have that $\bigcup_{r \in \mathcal{R}_{j,k} \setminus \{|\mathcal{R}_{j,k}|\}} \mathcal{M}_{j,k}^{(r)} = [0, 1]$ and $\mathcal{M}_{j,k}^{(r)} \bigcap \mathcal{M}_{j,k}^{(r+1)} = \mathcal{R}_{j,k}^{(r+1)}$, $\forall r \in \mathcal{R}_{j,k} \setminus \{|\mathcal{R}_{j,k}|\}$. Thus, for each mixture model obtained in each combination of transition $j \in \mathcal{J}$ and sex $k \in \mathcal{K}$ the goodness-of-fit is given by the common area $A_{j,k}$ as follows

$$A_{j,k} = \sum_{r \in \mathcal{R}_{j,k} \setminus \{|\mathcal{R}_{j,k}|\}} \left(\int_{Y_{j,k} \in \mathcal{M}_{j,k}^{(r)}} \Omega_{j,k}^{(r)}(Y_{j,k}) dY_{j,k} \right) \quad (6)$$

where

$$\Omega_{j,k}^{(r)}(Y_{j,k}) = \min_{Y_{j,k} \in \mathcal{M}_{j,k}^{(r)}} \left(g(Y_{j,k}), \sum_{i \in \mathcal{I}} w_{i,j,k} f(Y_{j,k}) \right)$$

Notice that in the ideal scenario the perfect fit implies $A_{j,k} = 1$ and $|\mathcal{R}_{j,k}| \to \infty \, \forall j \in \mathcal{J}, k \in \mathcal{K}$, i.e., both the actual density and the mixture model coincide over the domain $[0, 1]$. An illustration of a solution to this problem using four components into the mixture model is as Figure 3 shows.

4 COMPUTATIONAL EXPERIMENTS

The computational experiments are implemented in R [19] 4.0.0 programming language and was executed in NLHPC infrastructure (National Laboratory for High Performance Computing) considering a time limit of 600 seconds, the maximum number of
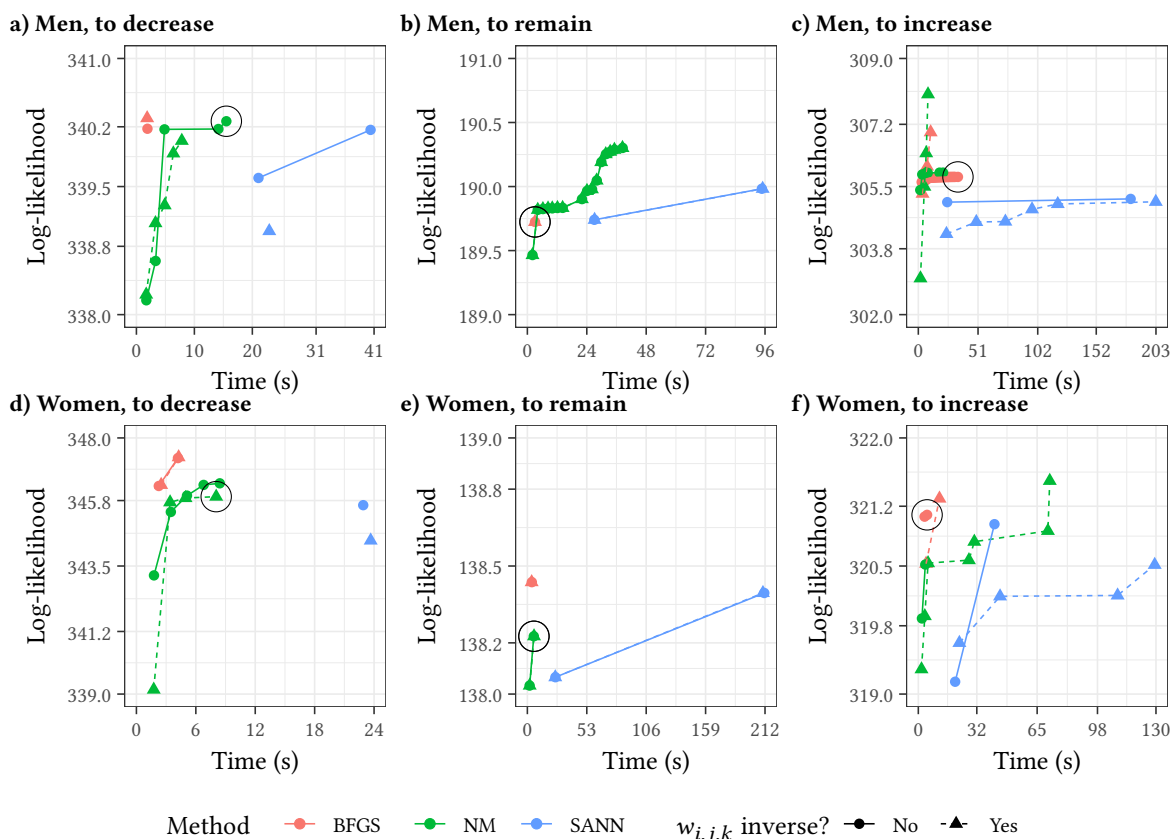
Figure 4: Log-likelihood value throughout time (in seconds) according to the proposed estimation scheme using the optimization algorithms Broyden, Fletcher, Goldfarb and Shanno (BFGS), Nelder-Mead (NM) and Simulated Annealing (SANN). Notice that the black circle indicates where the best goodness-of-fit (common area) was obtained.

EM iterations of 200, a tolerance of 10E-9 and the maximum number of iterations without improvement of 50. Also, we consider the perturbation as a uniform random number between 0.85 and 1.15.
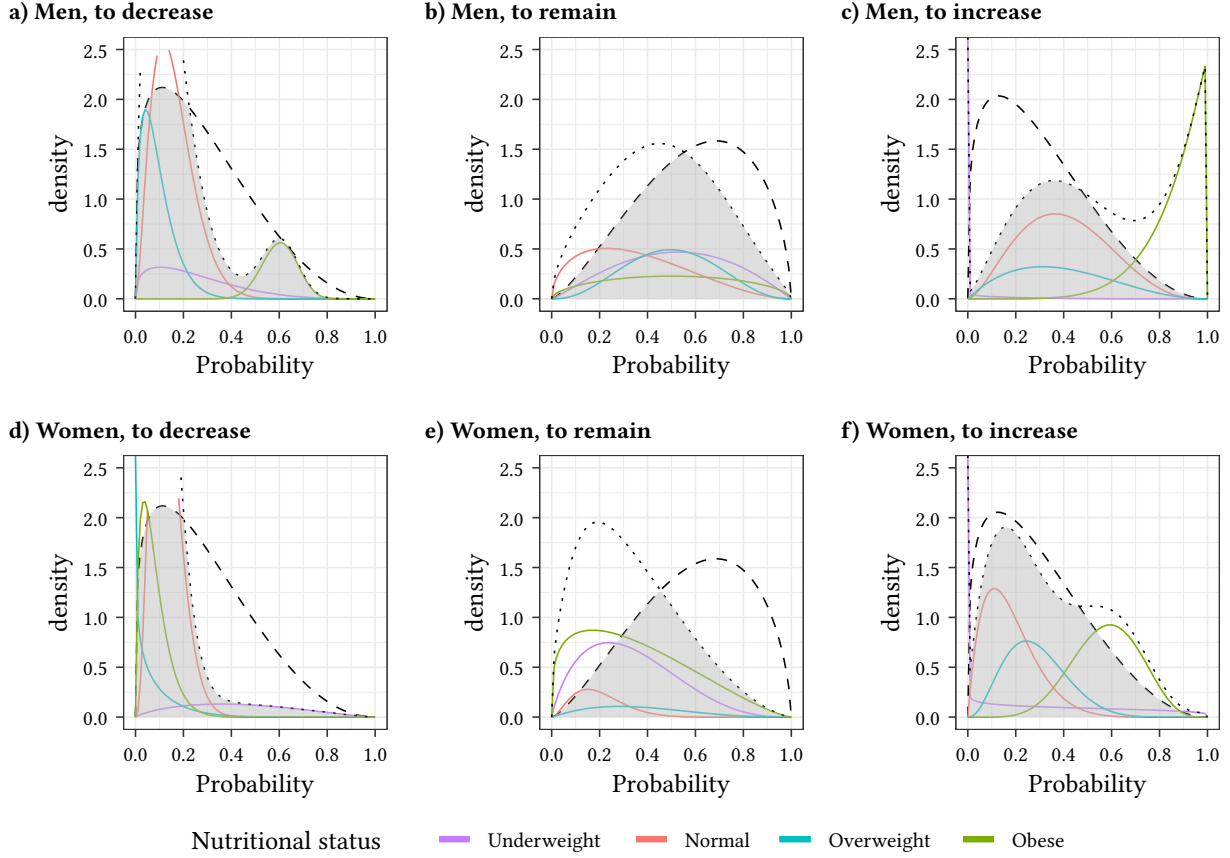
To compute the log-likelihood we consider the values provided by Ávalos [28] to estimate the maximum likelihood estimators of the Dirichlet distribution using the *sirt* package [20] and thereof using them to simulate random numbers such that $n_{j,k} = 1000$, $\forall j \in \mathcal{J}, k \in \mathcal{K}$. Then, we use the *maxLik* package [14] to solve the maximization problem considering three optimization algorithms:

(1) The BFGS algorithm proposed by Broyden [4], Fletcher [10], Goldfarb [11] and Shanno [22]; considering 200 iterations.
(2) The NM algorithm proposed by Nelder & Mead [18]; considering the reflection factor equal to 1, the contraction factor equal to 0.5 and the expansion factor equal to 2.
(3) The SANN (Simulated Annealing) algorithm proposed by Kirkpatrick, Gelatt and Vecchi [15] and Černỳ [5]; considering the starting temperature for the cooling schedule equal to 10 and the number of function evaluations at each temperature equal to 10.

Besides, we also address missing values related issues given those infeasible transitions, e.g. decreasing or increasing when a person has underweight or obesity, respectively; in this regard, we use the number data to compute the starting mixture weights such as they are considered to be directly or inversely proportional to the presence of missing values in the input.

Figure 4 shows the evolution of the log-likelihood considering all the settings involved in the experiment throughout time highlighting that point where the best goodness-of-fit value was obtained. It is clear to see that both log-likelihood and goodness-of-fit are not closely related, which is why there is no guarantee of choosing one over the other would give the best result because they are equally important considering the relevance of the log-likelihood and the interpretability and the suitableness of the estimated parameters. Thereby, we can realize that in almost all the estimated models the best results were obtained by using directly proportional starting weights but for those Women that decrease their current nutritional status.

Then, even although the obtained results show excellent performance in terms of goodness-of-fit, Figure 5 shows this metric is not conclusive by itself. However, for both sexes that remain their current nutritional status, the obtained results show low kurtosis in comparison with other transitions, which is the major problem of other models where there is one density, at least, that has high kurtosis, i.e., it concentrates the most of the density around a narrow interval. Therefore, despite the goodness-of-fit shows excellent values, the suitableness of the obtained results are not correct at all but for those individuals that remain their current nutritional status, especially for Men. Nonetheless, this contradictory analysis is given by the unbalanced weight of each mixture density and, moreover, the important outstanding differences between the parameters $\alpha'_{i,j,k}$ and $\beta'_{i,j,k}$, where it is highlighted the case $\alpha'_{i,j,k} < 1$ (see Figure 5 c), d) and f)), where this parameter takes values equal to 0.89, 0.73 and 0.84, respectively).

Figure 5: Mixture model of the best obtained result based on the goodness-of-fit for each combination of sex and transition. The long-dashed line represents the actual density whereas the dashed line is the mixture density. Besides, the parameters for each _PDF_ associated to each nutritional status and their respective weights in the respective mixture model are shown in the table.

			Underweight			Normal			Overweight			Obese			Marginal	
Sex	Transition	$A_{j,k}$	$\alpha'_{1,j,k}$	$\beta'_{1,j,k}$	$w_{1,j,k}$	$\alpha'_{2,j,k}$	$\beta'_{2,j,k}$	$w_{1,j,k}$	$\alpha'_{3,j,k}$	$\beta'_{3,j,k}$	$w_{1,j,k}$	$\alpha'_{4,j,k}$	$\beta'_{4,j,k}$	$w_{1,j,k}$	$\alpha_{j,k}$	$\beta_{j,k}$
Men	To decrease	96.9	1.38	4.34	0.12	2.78	14.57	0.52	1.68	16.27	0.25	26.62	17.76	0.10	1.26	3.08
Men	To remain	97.4	2.18	2.06	0.30	1.55	2.83	0.28	3.52	3.55	0.24	1.56	1.56	0.17	2.38	1.63
Men	To increase	97.6	0.89	4.63	0.01	2.69	3.95	0.42	2.01	3.21	0.17	6.33	1.02	0.40	1.29	3.01
Women	To decrease	94.9	1.79	2.43	0.08	3.82	25.18	0.58	0.73	7.71	0.09	1.68	19.37	0.25	1.27	3.10
Women	To remain	93.1	1.86	3.77	0.36	2.86	11.41	0.07	2.07	3.71	0.05	1.28	2.37	0.52	2.40	1.64
Women	To increase	96.7	0.84	1.22	0.09	2.01	9.22	0.33	3.44	8.56	0.24	6.90	5.05	0.34	1.29	3.04

5 CONCLUSION

In this work, we propose a hierarchical beta mixture model to estimate the probability density functions associated to each nutritional status according to the sex and the transition a person could suffer between two consecutive periods based on the imposed conditions and the relationship between the Dirichlet and Beta distribution. Thereby, several computational experiments are carried out following the proposed optimization scheme using the expectation maximization algorithm representing the goodness-of-fit through the shared area between the actual density (marginal Beta) and the mixture model. The obtained results show excellent goodness-of-fit values but they also set out the best log-likelihood (objective function) does not imply the best goodness-of-fit. Thus, it is possible to get bad local maximum points due to the high non-linearity of the model, which are the cases of the transitions to decrease and to increase, for both sexes. Likewise, it is possible to see these -bad- local maximums have high kurtosis values and high differences in the weight of each

nutritional status on the mixture model. Finally, it is important to remark that the optimization process does not use the time limit for computing the results due to it cannot improve the obtained solution according to the proposed transformation step.

For future works, we propose to increase the grid of experiments and to construct a hybrid scheme about the optimization algorithms and studying bad local optimal maximum regarding the characterization of an actual good maximum following a multi-criteria framework.

REFERENCES

[1] David B. Allison and Suzanne E. Saunders. 2000. Obesity in North America: an overview. *Medical Clinics of North America* 84, 2 (2000), 305–332.

[2] May A. Beydoun and Youfa Wang. 2008. Do nutrition knowledge and beliefs modify the association of socio-economic factors and diet quality among US adults? *Preventive medicine* 46, 2 (2008), 145–153.

[3] Chrysoula Boutari, Panagiotis D. Pappas, Gesthimani Mintziori, et al. 2020. The effect of underweight on female and male reproduction. *Metabolism* (2020), 154229.

[4] Charles George Broyden. 1970. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA Journal of Applied Mathematics* 6, 1 (1970), 76–90.

[5] Vladimír Černỳ. 1985. Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of optimization theory and applications* 45, 1 (1985), 41–51.

[6] Arthur P Dempster, Nan M Laird, and Donald B Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 39, 1 (1977), 1–22.

[7] B. Detournay, F. Fagnani, M. Phillippo, et al. 2000. Obesity morbidity and health care costs in France: an analysis of the 1991-1992 Medical Care Household Survey. *International Journal of Obesity* 24, 2 (2000), 151–155.

[8] Robert H. Eckel and Ronald M. Krauss. 1998. American Heart Association call to action: obesity as a major risk factor for coronary heart disease. *Circulation* 97, 21 (1998), 2099–2100.

[9] Eric A. Finkelstein, Ian C. Fiebelkorn, and Guijing Wang. 2003. National Medical Spending Attributable To Overweight And Obesity: How Much, And Who's Paying? Further evidence that overweight and obesity are contributing to the nation's health care bill at a growing rate. *Health Affairs* 22, 1 (2003), 219–226.

[10] Roger Fletcher. 1970. A new approach to variable metric algorithms. *The computer journal* 13, 3 (1970), 317–322.

[11] Donald Goldfarb. 1970. A family of variable-metric methods derived by variational means. *Mathematics of computation* 24, 109 (1970), 23–26.

[12] Gilberto González-Parra, L. Acedo, Rafael-J. Villanueva Micó, and Abraham J. Arenas. 2010. Modeling the social obesity epidemic with stochastic networks. *Physica A: Statistical Mechanics and its Applications* 389, 17 (2010), 3692–3701.

[13] Shumei Sun Guo, Wei Wu, William Cameron Chumlea, et al. 2002. Predicting overweight and obesity in adulthood from body mass index values in childhood and adolescence. *The American journal of clinical nutrition* 76, 3 (2002), 653–658.

[14] Arne Henningsen and Ott Toomet. 2011. maxLik: A package for maximum likelihood estimation in R. *Computational Statistics* 26, 3 (2011), 443–458.

[15] Scott Kirkpatrick, C. Daniel Gelatt, and Mario P. Vecchi. 1983. Optimization by simulated annealing. *science* 220, 4598 (1983), 671–680.

[16] Shiriki Kumanyika. 2006. Nutrition and chronic disease prevention: priorities for US minority groups. *Nutrition reviews* 64, suppl_1 (2006), S9–S14.

[17] National Institutes of Health. 1998. Clinical guidelines for the identification, evaluation, and treatment of overweight and obesity in adults - The Evidence Report. *Obesity Reviews* 6, 2 (1998), 51–209.

[18] John A Nelder and Roger Mead. 1965. A simplex method for function minimization. *The computer journal* 7, 4 (1965), 308–313.

[19] R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

[20] Alexander Robitzsch. 2020. *sirt: Supplementary Item Response Theory Models*. https://CRAN.R-project.org/package=sirt R package version 3.9-4.

[21] David E Rumelhart, James L McClelland, PDP Research Group, et al. 1988. *Parallel distributed processing*. Vol. 1. IEEE Massachusetts.

[22] David F Shanno. 1970. Conditioning of quasi-Newton methods for function minimization. *Mathematics of computation* 24, 111 (1970), 647–656.

[23] Jeffery Sobal and Albert J. Stunkard. 1989. Socioeconomic status and obesity: a review of the literature. *Psychological bulletin* 105, 2 (1989), 260.

[24] D. Talukdar, S. Seenivasan, A. J. Cameron, et al. 2020. The association between national income and adult obesity prevalence: Empirical insights into temporal patterns and moderators of the association using 40 years of data across 147 countries. *PloS One* 15, 5 (2020), e0232236.

[25] J. van de Kassteele, R. T. Hoogenveen, P. M. Engelfriet, et al. 2012. Estimating net transition probabilities from cross-sectional data with application to risk factors in chronic disease modeling. *Statistics in Medicine* 31, 6 (2012), 533–543.

[26] Zachary J. Ward, Michael W. Long, Stephen C. Resch, et al. 2017. Simulation of growth trajectories of childhood obesity into adulthood. *The New England Journal of Medicine* 377 (2017), 2145–2153.

[27] World Health Organization. 2017. Prevalence of obesity among adults, BMI ≥ 30, crude Estimates by World Bank income group. https://www.nimh.nih.gov/health/statistics/what-is-prevalence.shtml. Online; accessed 11 sept 2020.

[28] Denisse Ávalos. 2015. *Modelamiento dinámico del índice de masa corporal de la población en Chile*. M Sc. Eng. mention in Industrial Engineering. Universidad de Santiago de Chile.

[29] Denisse Ávalos, Cristóbal Cuadrado, Jocelyn Dunstan, Javier Moraga-Correa, Luis Rojo-González, Nelson Troncoso, and Óscar Vásquez. 2021. Mathematical Model for Estimating Nutritional Status of the Population with Poor Data Quality in Developing Countries: The Case of Chile. In *Proceedings of the 10th International Conference on Operations Research and Enterprise Systems - Volume 1: ICORES,*. INSTICC, SciTePress, 408–415.