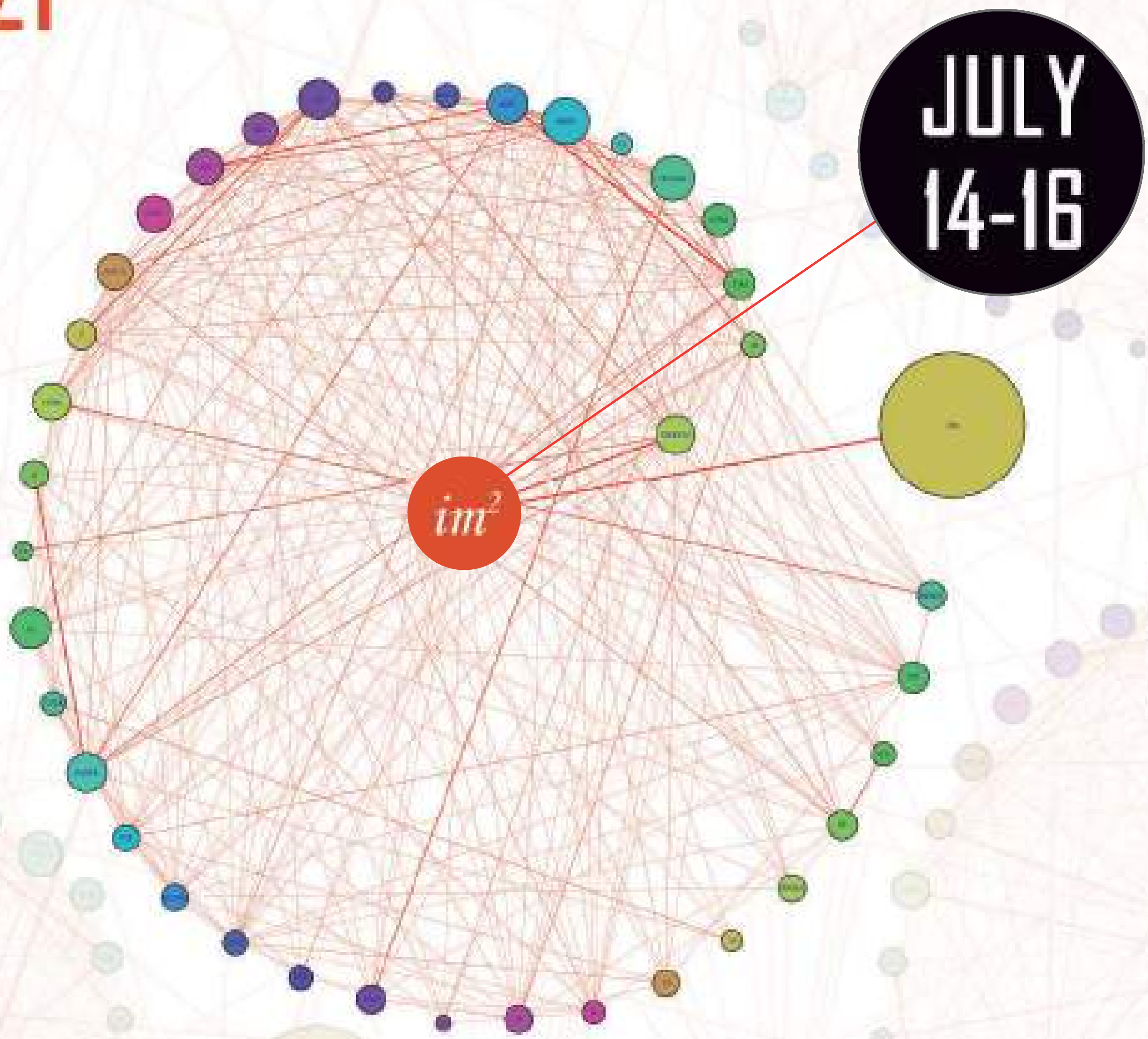
MODELLING FOR ENGINEERING & HUMAN BEHAVIOUR 2021



Edited by

Juan Ramón Torregrosa Juan Carlos Cortés Antonio Hervás Antoni Vidal Elena López-Navarro

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

im²
Instituto Universitario de Matemática Multidisciplinar

Modelling for Engineering & Human Behaviour 2021

València, July 14th-16th, 2021

This book includes the extended abstracts of papers presented at XXIII Edition of the Mathematical Modelling Conference Series at the Institute for Multidisciplinary Mathematics *Mathematical Modelling in Engineering & Human Behaviour*.

I.S.B.N.: 978-84-09-36287-5

November 30th, 2021
Report any problems with this document to imm@imm.upv.es.

Edited by: I.U. de Matemàtica Multidisciplinar, Universitat Politècnica de València.
J.R. Torregrosa, J-C. Cortés, J. A. Hervás, A. Vidal-Ferràndiz and E. López-Navarro

im2
Instituto Universitario
de Matemática Multidisciplinar

Contents

Modeling excess weight in Spain by using deterministic and random differential equations

Julia Calatayud ♭,[1]

(♭) Departament de Matemàtiques, Universitat Jaume I,
12071 Castellón, Spain.

Abstract

In this work, the aim is to model the data from the Spanish National Health Survey (ENSE) 2017, which gathers the percentage of overweight and obese adults in Spain along the last three decades. A compartmental system of differential equations is employed, based on the classification “normal weight”, “overweight” and “obese”. It is assumed homogeneous mixing, non–constant population, and social transmission of excess weight due to peer pressure. The model is randomized by incorporating a discrete and uncorrelated Gaussian error (frequentist regression), random parameters and errors (Bayesian inference), and a Gaussian white noise perturbation into the derivatives (Itô stochastic differential equation). In all those cases, inverse parameter estimation is conducted. Some remarkable results are obtained. For example, the long–term behavior of the system shows that 37% and 24% of Spanish adults will be overweight and obese in the long run, respectively. The sensitivity analyses from the different strategies agree and suggest that prevention strategies are more important than treatment strategies to control adulthood obesity. This methodology and the results are based on the recent papers [1] and [2].

1 Introduction

Yet most neglected, today excess weight causes more than 2.8 million deaths each year and a significant economic burden. Paradoxically coexisting with malnutrition, the prevalence of obesity has nearly tripled since 1975. In Spain, 7% health cost is due to excess weight. Due to these facts, and the likely higher impact in the future, it is necessary to address the problem of obesity. Mathematical models are an effective tool to understand the past history, forecast, and propose targeted measures.

The classification between “normal weight”, “overweight” and “obesity” refers to BMI (weight divided by height squared): $\mathrm{BMI} < 25$, $25 \leq \mathrm{BMI} < 30$, $30 \leq \mathrm{BMI}$. (Malnutrition is not considered.) When there are groups in the population, a compartmental model can be considered. The flux between compartments is studied by means of ordinary differential equations, with no spatial effects, and homogeneous mixing:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X),$$

[1]calatayj@uji.es

$$X^T(t) = (x_1(t), \dots, x_m(t)), \quad F^T(x) = (f_1(x), \dots, f_m(x)),$$

$$x_i(t) = \text{number of individuals in compartment } i \text{ at instant } t.$$

See [3, 4].
Some medical and sociological studies suggest that excess weight may be transmitted through persons, similarly to an infectious disease. Random networks (graphs) were analyzed in [5]. The authors concluded that the probability of becoming obese augments by 57% with an obese friend. Similarly, if your partner is obese, the chance of obesity increases by 37%. On the other hand, according to [6], excess weight is transmitted by imitation and social pressure. There are some models that consider imitation and social pressure for different phenomena: alcohol, tobacco, telecommunications, drugs, etc. [7–9].
The Spanish National Health Survey (ENSE) is a massive scale survey that collects transversal data on health regarding the resident population in Spain. It started in 1987. The last survey appeared in 2017, with 23089 interviews. See `https://www.mscbs.gob.es/estadEstudios/estadisticas/encuestaNacional/encuesta2017.htm`. The percentage of obesity has increased from 7.4% in 1987 to 17.4% in 2017. Among children, it is maintained around 10%. The aim is to model the prevalence of overweight and obesity from 1987 till 2017 in Spain, among the adulthood Spanish population (older than 18 years old). We base on the papers [1, 2], and the reader is referred to them for detailed information.

2 Deterministic model

According to the INE, the Spanish population has not been constant from 1987 until 2017. A logistic model is fitted:

$$T(t) = \text{Spanish adults (rescaled by ten million)}, \quad t = 0 \text{ is } 1987;$$

$$T'(t) = \mu T(t)\left(1 - \frac{T(t)}{K}\right) \Rightarrow$$

$$\hat{\mu} = 0.0491843, \quad \hat{K} = 4.47700; \quad T(t) = \frac{12.3654}{2.76197 + 1.71503\,\mathrm{e}^{-0.0491843t}}.$$

Let $S(t)$ and $O(t)$ be the number of overweight and obese adults at year t since 1987 ($t = 0$). These are rescaled by ten million. As described in the Introduction section, there is a social pressure to become overweight. The model is:

$$\begin{cases} S'(t) = \mu S(t)\left(1 - \frac{T(t)}{K}\right) + \beta \frac{T(t)-S(t)-O(t)}{T(t)}\left[S(t) + O(t)\right] - (\rho + \gamma)S(t) + \epsilon O(t), & t \geq 0, \\ O'(t) = \mu O(t)\left(1 - \frac{T(t)}{K}\right) + \gamma S(t) - \epsilon O(t), & t \geq 0, \\ S(0) = S_0 = 0.902400, \\ O(0) = O_0 = 0.208680, \end{cases}$$

where $\mu = 0.0491843$ and $K = 4.47700$. The model may also be written with three compartments ($N = T - S - O$):

$$\begin{cases} N'(t) = \mu N(t)\left(1 - \frac{T(t)}{K}\right) - \beta \frac{N(t)}{T(t)}\left[S(t) + O(t)\right] + \rho S(t), \\ S'(t) = \mu S(t)\left(1 - \frac{T(t)}{K}\right) + \beta \frac{N(t)}{T(t)}\left[S(t) + O(t)\right] - (\rho + \gamma)S(t) + \epsilon O(t), \\ O'(t) = \mu O(t)\left(1 - \frac{T(t)}{K}\right) + \gamma S(t) - \epsilon O(t), \\ N(0) = N_0 = T(0) - S_0 - O_0, \\ S(0) = S_0, \\ O(0) = O_0. \end{cases}$$

These systems are well-posed (existence, uniqueness and positivity). The parameters are interpreted as rates. These are assumed to be time independent, because there have not been strong measures to face excess weight.
The mean-square error is minimized:

$$\sum_{i=1}^{11} \left(S(t_i|\beta,\gamma,\epsilon,\rho) - s_i\right)^2 + \sum_{i=1}^{11} \left(O(t_i|\beta,\gamma,\epsilon,\rho) - o_i\right)^2 .$$

This yields the parameter estimates

$$\hat{\beta} = 0.368989,\ \hat{\gamma} = 0.0222886,\ \hat{\epsilon} = 0.0344076,\ \hat{\rho} = 0.240838.$$

The MAPE (mean absolute percentage error) is:

$$\frac{100}{22}\left(\sum_{i=1}^{11} \frac{|S(t_i|\hat{\beta},\hat{\gamma},\hat{\epsilon},\hat{\rho}) - s_i|}{s_i} + \sum_{i=1}^{11} \frac{|O(t_i|\hat{\beta},\hat{\gamma},\hat{\epsilon},\hat{\rho}) - o_i|}{o_i}\right) = 3.17554.$$

According to Lewis' scale, the model is very accurate.
A sensitivity analysis consists in finding the parameter with the highest impact in the model, to control and mitigate the problem. The analysis may be based on MAPEs (remove parameter and recompute the MAPE) and on differentiation (local variation). It is derived that prevention strategies are preferred over treatment strategies.
Asymptotically, it holds $S(\infty) = 1.64077$ and $O(\infty) = 1.06286$. That is, the prevalences are 36.65% and 23.74%, respectively.

3 Frequentist regression model

The model presents errors that are not controllable. If $(S(t|\beta,\gamma,\epsilon,\rho), O(t|\beta,\gamma,\epsilon,\rho))$ is the deterministic solution, then consider the random variables

$$S_i = S(t_i|\beta,\gamma,\epsilon,\rho) + \mathcal{E}_i^S,$$

$$O_i = O(t_i|\beta,\gamma,\epsilon,\rho) + \mathcal{E}_i^O,$$

$i = 1,\ldots,11$. The datums are realizations. The terms $\mathcal{E}_i^S$, $\mathcal{E}_i^O$ are random errors with zero expectation and constant variance σ^2, all of them identically distributed and independent. The normality of residuals is accepted by Kolmogorov-Smirnov and Shapiro-Wilk tests. Levene's test renders equality of variances.
Non-linear regression $Y = g(q) + \mathcal{E}$, $q = (\beta,\gamma,\epsilon,\rho)$, is based on [10, chapter 7], as a generalization of linear regression. The Jacobian matrix $X(q) = Jg(q)$ gives a linear model. From the theory of linear models, the estimates are:

$$\hat{\sigma}^2 = 0.00103513, \quad \hat{Q} = (\text{statistic of } q) \approx \text{Normal}(q, \Sigma_{\hat{Q}}),$$

$$\hat{q} = \text{deterministic estimator}, \quad \hat{\Sigma}_{\hat{Q}} = \hat{\sigma}^2 (X(\hat{q})^\top X(\hat{q}))^{-1}.$$

The sensitivity analysis is based on MAPEs (as before) and t-tests. The t-tests are conducted from

$$(t\text{-Student})_i = \frac{\hat{q}_i}{\hat{\sigma}\sqrt{(X(\hat{q})^\top X(\hat{q}))^{-1}_{ii}}},$$

with the interpretation

$$(t\text{-Student})^2 = \frac{R^2_{\text{new}} - R^2_{\text{old}}}{(1 - R^2_{\text{old}})/(n - n_p)}.$$

The conclusions coincide with the previous section.
The asymptotic estimates (95% prediction intervals and expectations) for overweight and obesity are the following: $[34.22\%, 39.08\%]$ and $[13.68\%, 33.80\%]$, 36.65% and 23.74%. Observe the wide prediction interval when extrapolating.

4 Bayesian model

The parameters q are considered as random variables. A Bayesian model [10, chapter 8] is the following:

$$Y|q,\sigma^2 \sim \prod_{i=1}^{22} \text{Normal}\left(g_i(q),\sigma^2\right),$$

$$q \sim \pi(q) = \pi(\beta)\pi(\gamma)\pi(\epsilon)\pi(\rho),\ \sigma^2 \sim \pi(\sigma^2).$$

The Bayes' formula for the posterior distribution of the parameters is the following:

$$\pi(q,\sigma^2|y) = \frac{\pi(y|q,\sigma^2)\pi(q)\pi(\sigma^2)}{\int_{-\infty}^{\infty} \pi(y|q,\sigma^2)\pi(q)\pi(\sigma^2)\,\mathrm{d}^4 q\,\mathrm{d}\sigma^2}.$$

Instead of quadratures, Markov Chain Monte Carlo algorithms are used, such as Metropolis and Adaptive Metropolis. Each iteration requires the numerical resolution of the model. The Adaptive Metropolis algorithm is the most efficient, as the covariance matrix of the proposal distribution is changed on the fly.
The sensitivity analysis is based on the BIC and derivations (Savage-Dickey density ratio). The conclusions coincide with the preceding section.
The asymptotic estimates (0.95 credible intervals and expectations) for overweight and obesity are the following: $[35.1\%, 38.0\%]$, $[21.6\%, 26.7\%]$, 36.6% and 24.0%.

5 Model based on Itô stochastic differential equations

A model is the following:

$$\begin{cases} \mathrm{d}S(t) = \left[\mu S(t)\left(1-\frac{T(t)}{K}\right) + \beta\frac{T(t)-S(t)-O(t)}{T(t)}\left[S(t)+O(t)\right] - (\rho+\gamma)S(t) + \epsilon O(t)\right]\mathrm{d}t + \sigma_1 \mathrm{d}B_1(t), \\ \mathrm{d}O(t) = \left[\mu O(t)\left(1-\frac{T(t)}{K}\right) + \gamma S(t) - \epsilon O(t)\right]\mathrm{d}t + \sigma_2 \mathrm{d}B_2(t), \\ S(0) = S_0, \\ O(0) = O_0. \end{cases}$$

Here $B_1(t)$ and $B_2(t)$ are independent standard Brownian motions. Gaussian white noise random perturbations (formal derivatives of Brownian motions) have been added into the response derivatives. It may be proved that the model possesses a unique solution on $[0,\infty)$ [11, theorem 3.6]. Strictly speaking, the response processes may take negative values, which does not make sense. Nonetheless, the probability of such occurrence is, for the typical values of the parameters, negligible.
From the Euler-Maruyama discretization, several strategies are utilized for estimating the parameters, based on the moments method and maximum likelihood estimation.
The sensitivity analysis is based on MAPEs. The same conclusions are obtained.

References

[1] Calatayud, J., Jornet, M., Mathematical modeling of adulthood obesity epidemic in Spain using deterministic, frequentist and Bayesian approaches. *Chaos, Solitons & Fractals*, 140(110179), 2020.

[2] Calatayud, J., Jornet, M., Modeling of adulthood obesity in Spain using Itô-type stochastic differential equations. *Chaos, Solitons & Fractals*, 145(110786), 2021.

[3] Murray, J.D., Mathematical Biology. Berlin, Springer, 1993.

[4] Brauer, F., Compartmental models in epidemiology. In: Mathematical Epidemiology. Berlin, Springer, 2008. pp. 19–79.

[5] Christakis, N.A., Fowler, J.H., The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.*, 357:370–379, 2007.

[6] Blanchower, D.G., Oswald, A.J., Landeghem, B.V., Imitative obesity and relative utility. *J. Eur. Econ. Assoc.*, 7:528–38, 2009.

[7] Santonja, F.J., Sánchez, E., Rubio, M., Morera, J.L., Alcohol consumption in Spain and its economic cost: a mathematical modeling approach. *Math. Comput. Model.*, 52(7–8):999–1003, 2010.

[8] Casabán, M.C., Cortés, J.C., Navarro-Quiles, A., Romero, J.V., Roselló, M.D., Villanueva, R.J., A comprehensive probabilistic solution of random SIS-type epidemiological models using the Random Variable Transformation technique. *Commun. Nonlinear Sci. Numer. Simulat.*, 32:199–210, 2016.

[9] Cervelló, R, Cortés, J.C., Santonja, F.J., Villanueva, R.J., The dynamics over the next few years of the Spanish mobile 480 telecommunications market share: A mathematical modelling approach. *Math. Comp. Model. Dyn.*, 20(6):557–565, 2014.

[10] Smith, R.C., Uncertainty Quantification: Theory, Implementation, and Applications. Philadelphia, SIAM, 2013.

[11] Mao, X, Stochastic Differential Equations and Applications. Elsevier, 2007.