Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Economics and Human Biology 3 (2005) 1–16
ECONOMICS AND HUMAN BIOLOGY
http://www.elsevier.com/locate/ehb

Analysis of periodic fluctuations of the height of Swedish soldiers in 18th and 19th centuries

Marek Brabec

Department of Biostatistics and Informatics, National Institute of Public Health, Srobarova 48, Praha 10, 10 042, Czech Republic

Received 18 August 2004; received in revised form 19 August 2004; accepted 19 August 2004

Abstract

This paper investigates the periodicity of the adult height of Swedish soldiers of the 18th and 19th centuries using spectral analysis. The height data are left truncated due to the enforcement of minimum height requirement. Hence, we use a truncated regression model using maximum likelihood estimation. We isolate the various frequency components, assess their importance, and perform sensitivity analysis by means of fitting several alternative models.

JEL classification: I31; N15

Keywords: Height cycles; Spectral analysis; Physical stature; Anthropometric history

1. Introduction and methods

The paper analyzes heights of adult Swedish soldiers measured in 18th and 19th centuries. The 18 000 records were collected by Richard Steckel and Lars Sandberg in the late 1970s (Heintel et al., 1997; Sandberg and Steckel, 1988). Properties of these data are complicated by the enforcement of a minimum height requirement (MHR) to be eligible for the military (Heintel, 1996; Komlos, in press). After the MHR was estimated and data below it discarded, more than 17 000 records remain. Birth years span 1711–1864. Because there are sufficient annual observations, one can explore the periodicity of the mean height

E-mail address: mbrabec@szu.cz.

1570-677X/$ – see front matter © 2004 Published by Elsevier B.V.
doi:10.1016/j.ehb.2004.08.001

time series. Because birth location is available for many soldiers, it is possible to analyze the differences between heights of urban and rural-born men.

Investigation of periodic properties of height series is typically based on spectral analysis. Standard approach uses periodogram and its smoothers (Brockwell and Davis, 1991), decomposing a stationary time series into a sum of sine and cosine components of various frequencies. Then, one investigates the spectrum estimates in order to assess the importance of various frequency contributions. That is the plot of amplitudes against frequencies (or its smoothed version).

Application of this approach in the present context is complicated by several specific data features, however. A relatively minor complication arises from the fact that annual sample sizes vary substantially, so that amount of information coming from different years can differ quite considerably (and for a few birth years there are no data). Therefore, classical techniques of simple and computationally efficient periodogram estimation are not appropriate here. Nevertheless, the problem can be solved easily by introducing a regression model with trigonometric terms. It can handle both varying weights of different years (different number of observations per year), and possible nonstationarity of the series (e.g. by inclusion of polynomial, for instance linear) and trigonometric terms. If the data represented a random sample from the general population of interest, a regression model can be as follows:

$$\begin{aligned} &Y_{tij} = \mu + \alpha_i + \beta t + \sum_{k=1}^{F} (\delta_{1k} \cos(2\pi t f_k) + \delta_{2k} \sin(2\pi t f_k)) + \varepsilon_{tij} \\ &\varepsilon_{tij} \sim N(0, \sigma^2) \end{aligned} \tag{1}$$

where Y_{tij} is the height of j-th man from i-th birth location, born in year t ($t = 1$ in 1711). Resulting estimates of α_1–α_i can be used to test hypotheses about differences of mean heights between birth locations. Estimate of β can be used for test of a linear trend. The presence of the linear term can also be viewed as a precaution against possible non-stationarity coming from linear time trend. Most importantly, the estimates of δ_{11}–δ_{1F}, δ_{21}–δ_{22} can be used to assess periodic properties of the height series. Namely, the squared amplitude $\gamma_k^2 = \delta_{1k}^2 + \delta_{2k}^2$ can be plotted against f_k to assess relative importance of different frequencies (or periods $1/f_k$, in years). Contributions of various frequencies can be tested. Moreover, the resulting plot of γ_k^2 against f_k can be smoothed to get a clearer idea of the power spectral density shape.

A more series complication arises due to the MHR. Instead of Y_{tij}, we have Y'_{tij}, the left-truncated version of Y_{tij} available as data. Due to the truncation, it is not possible to use model (1) directly. Instead, we have to take the truncation into account:

$$\begin{aligned} &Y'_{tij} = Y_{tij}, \quad \text{if } Y_{tij} \geq \tau_t \\ &T_{tij} \text{ remains unobserved when } Y_{tij} < \tau_t \end{aligned}$$

where τ_t is the truncation point.

$$\begin{aligned} &Y_{tij} = \mu + \alpha_i + \beta t + \sum_{k=1}^{F} (\delta_{1k} \cos(2\pi t f_k) + \delta_{2k} \sin(2\pi t f_k)) + \varepsilon_{tij} \\ &\varepsilon_{tij} \sim N(0, \sigma^2) \end{aligned} \tag{2}$$