مجلة التجارة والتمويل

https://caf.journals.ekb.eg/

كلية التجارة – جامعة طنطا

العدد : الاول

مارس ٢٠٢٢

MULTILEVEL MODELS: AN APPLICATION TO LONGITUDINAL DATA IN A MEDICAL INVESTIGATION

Lobna Eid AL-Tayeb

Associate Prof. of statistic

Faculty of Commerce, Al-Azhar University (Girls' Branch), Egypt
Corresponding author's email: lobnaalatyeb@azhar.edu.eg

Abstract

During the last years, multilevel models have been used frequently to analyse longitudinal data, especially in the field of education. This search presents an application of this methodology in medical field, and the advantages of these models over classical methods are discussed. The motivation stems from an anesthesiology problem in which it is desired to study the influence of two anesthetic induction devices on blood pressure. The application highlights the potential of multilevel models for longitudinal data analysis.

Keywords: Multilevel models, longitudinal data, Medical investigation, regression analysis.

1. INTRODUCTION

Longitudinal models have become increasingly popular in recent years because of their power to test theoretically derived hypotheses by modelling within-person processes with repeated measures.

Anaesthesiologists continue to be concerned about airway management during surgery. The endotracheal tube and the laryngeal mask are two of the most commonly utilised anesthetic induction devices. The recommended procedure for airway control is tracheal intubation, which anesthesiologists regularly perform; nonetheless, its successful use is not always practicable. The laryngeal mask thus emerges as a viable alternative to endotracheal intubation in circumstances when it is not possible. The laryngeal mask has gained popularity in recent years. Medical literature states that it is one of the least risky; thus it is expected to be used routinely in patients who do not

have contraindications. As a result, studies on the effectiveness and usefulness of this anesthetic induction technique are of particular interest.

The risk of problems during surgery is reduced when blood pressure is stable. The purpose of this study was to compare the effects of the endotracheal tube and the laryngeal mask on blood pressure at three different points during the surgical procedure. This study's problem is classified as a longitudinal or repeated measurements study [6]. The investigation begins with a typical approach to data, exposing the issues that come with its application. The problem is then approached using a multilevel modelling approach [3], which demonstrates how it may explore complicated forms of variation between people. At present, multilevel modeling in the analysis of longitudinal data in scientific investigations of different branches of knowledge is notable. In Egypt, however, the application of this technique is not widely used in the medical literature [1]. This work proposes using multilevel models in an anesthesiology problem and tries to present to researchers the potential of these models in bio statistical research.

2. MATERIALS AND METHODS

The data presented in this work come from a case study. N = 99 patients who had been electively scheduled to undergo surgery were included. These were randomly divided into two groups to receive two types of anesthetic induction, either through the laryngeal mask (Device I) or the endotracheal tube (Device II). There are $n_1 = 50$ patients in the first group, and in the second group, $n_2 = 49$ patients. The primary interest in this research is to study the influence of the type of device on the behaviour of the systolic blood pressure (SBP) and diastolic blood pressure (DBP) of patients.

SBP and DBP were measured in each patient at three different times:

T_0: five minutes before anaesthetic induction.

T_1: immediately after placing the device in the airway.

T_2: five minutes after the device is placed.

Table 1 shows the mean responses of the SBP and DBP on each of the three occasions (T_0, T_1 and T_2) for each group defined by the type of device used. Observe the differences between the means for each group after the device was applied.

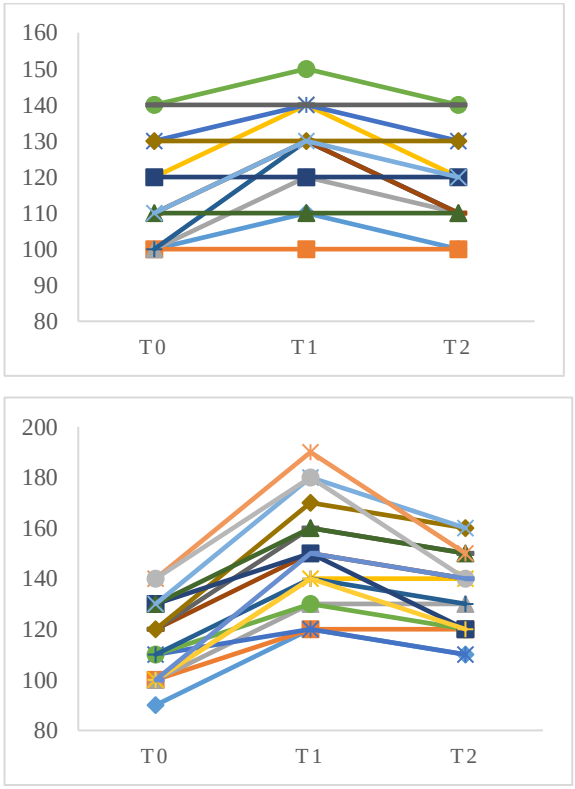
Table 1: Mean of SBP and DBP

	T_0	SBP T_1	T_2	T_0	DBP T_1	T_2
Dispositivo I (n_1=50)	122.86	127.96	120.41	79.39	80	76.12
Dispositivo II (n_2=49)	120.4	154.2	145.2	75.6	103.2	95.6
Total (N=99)	121.62	141.21	132.93	77.47	91.72	85.96

An exploratory analysis of the data indicated a wide variation in SBP and DBP responses among patients in each group. In this study, a similar pattern of responses was observed between the two variables considered. By way of illustration, Figure 1 (a and b) shows, for each group, the SBP profiles on the three occasions. Note the variability of the measurement values between patients. Note that once device I is applied, the SBP seems to increase slightly, but five minutes later, it tends to reestablish (Fig. 1.a). However, it can be observed that the SBP increases markedly when the device II is applied (Figure 1.b), and five minutes later it still does not recover the initial values.

The analysis can be done from two different approaches. In the first, the problem is analyzed from the classical point of view, using the standard

regression model, without distinguishing between the patients. The use of multilevel models for longitudinal data [3] is proposed in the second section, which considers the correlation pattern of repeated measures. A two-level hierarchical structure is formed when observations recorded over time are nested within selected individuals from a population of interest. The lowest level, or level 1, is the variation of reactions among people over time, whereas level 2 is the mean response among individuals. After that, the model can be treated as a hierarchical system of regression equations [2].



(a) Device I $n_1(50)$ (b) Device II $n_2(49)$

Figure 1: SBP profiles on the 3 occasions for each patient.

These models treat individuals as a random sample drawn from a larger population. This approach allows modelling not only the fixed parameters of the model, but also the random effects. Inferences about the variation among all individuals in the population are made using the random sample [7].

3. RESULTS AND DISCUSSION

In the first part of the study, the SBP and DBP variables were analyzed separately since the analysis observed a similar behaviour in the two variables under study, in this research, only the results associated with the SBP are presented to illustrate the univariate case. Finally, a multivariate model is used to study the two variables simultaneously.

3.1 Regression model for repeated measures.

There are two SBP readings for each patient after the device is placed. Let Y_{ij} and the SBP for patient i $(i = 1, \ldots, 99)$ on occasion j $(j = 1,2)$ and let $T_{ij} = \begin{cases} 0 & if\ j = 1 \\ 1 & if\ j = 2 \end{cases}$

Thus, the mean behaviour of SBP among all patients can be studied by the equation:

$$y_{ij} = \beta_o + \beta_1 T_{ij} + e_{ij} \quad (1)$$

where the coefficient β_o is the mean SBP of a patient immediately after applying the device and β_1 measures the mean increase in SBP of a patient, five minutes after placing the device. In this model, only one error term appears (e_{ij}) that represents the difference in the SBP of each patient on

each occasion, concerning the mean SBP for all patients. Note that the model does not separate variability due to occasion from that due to patients.

The specialists affirm that the variation of the SBP between the individuals, after the placement of the device, depends fundamentally on their initial values. In the following model, the SBP is included before applying the device (initial value), as a new variable:

$$y_{ij} = \beta_o + \beta_1 T_{ij} + \beta_2 I_i + e_{ij} \quad (2)$$

where I_i represents the initial value of the SBP for patient i, centered around the average initial value for all patients. If you want to introduce the effect of the device, you should consider a model of the form:

$$y_{ij} = \beta_o + \beta_1 T_{ij} + \beta_2 I_i + \beta_3 D_i + e_{ij} \quad (3)$$

where D_i represents group i, defined by the type of device (coded as 0 for the laryngeal mask and 1 for the endotracheal tube).

None of the models presented so far has taken into account the characteristics of the patients to explain the behavior of the SBP once the device is placed, however, a problem of the exact nature as the one indicated above persists, regarding the variability: the estimated coefficients are the same for all individuals. In general, this model considers the effects of the variables associated with individuals to be fixed. Another drawback is that these models do not take into account the correlation structure between the occasions; however, the SBP on the second occasion could be influenced by

the value on the previous occasion, and therefore, they cannot be considered as independent observations, a fundamental assumption to estimate the parameters of a classical regression model.

3.2. Multilevel models

To try to describe the associations between variables more clearly and subsequently make predictions, it is necessary to study the structure of the measurements taken on each occasion [7]. Longitudinal data can be considered a hierarchical structure, where repeated measures are nested within individuals. In this way, repeated measures are considered level-1 units, and individuals, level-2 units. In some studies an additional level can be formed, considering groups of individuals [5].

The results obtained by applying this approach to the study data are presented below. In the two-level hierarchical models, a level-1 model is proposed for each i unit at level-2 ($i = 1,\ldots,90$). Thus, for the example data, the regression equation at level 1 is:

$$y_{ij} = \beta_{o1} + \beta_{1i}T_{ij} + e_{ij} \quad (4)$$

where e_{ij} represents the variability of the SBP between the occasions of individual i. It is known that the patients studied are a random sample of a specific population and that the SBP on each occasion varies between patients. This variation is expressed by the inclusion of a random term u_{0i}, which represents the deviation of the SBP of patient i on the second occasion, concerning the mean values of the population; I mean:

$$\beta_{01} = \gamma_{00} + u_{0\mathrm{i}}$$

(5)

$$\beta_{1I} = \gamma_{I0}$$

(6)

Substituting these last two expressions in the previous regression equation and rearranging terms, the multilevel model is obtained:

$$\gamma_{ij} = \gamma_{00} + \gamma_{10} T_{ij} + u_{0\mathrm{i}} + e_{ij}$$

(7)

The coefficient γ_{00} It is interpreted as the average SBP of the patient population, at the time of applying the device and γ_{10} measures the mean increase in SBP, five minutes later. Note that now the model has a fixed component $(\gamma_{00} + \gamma_{10} T_{ij})$ and a random component $(u_{0\mathrm{i}} + e_{ij})$. In total, in this model four parameters must be estimated: the two fixed (γ_{00} & γ_{10}) and the two variances $(var(u_{0\mathrm{i}}) = \sigma^2_{u_0}\ and\ var(e_{ij}) = \sigma^2_e)$, called random parameters. Table 2 presents the estimated parameters of model (1) and the following three fitted models. The estimates of the parameter σ^2_e are not presented because they do not have a helpful interpretation for the example. The negative sign of the estimated parameter $\hat{\gamma}_{10}$ confirms what is expected, that is, the mean SBP decreases five minutes after the device is placed. The estimated variance $\sigma^2_{u_0} = 347{,}440$ is an indicator of the variation between patients. This value is statistically significant, suggesting that the SBP varies between individuals immediately after the device is placed.

In model (1), a variability between patients is considered, for the effect of the occasion at the time of placing the device, however, the effect of the

occasion, five minutes after applying the device, is modeled as constant for all the patients. To allow the variation of the effect of the third occasion between the different patients, the error term u_{Ii} is introduced, considering the coefficient β_1 also as random, that is, now:

$$\beta_{1I} = \gamma_{I0} + u_{\mathrm{Ii}}$$

(8)

The new multilevel model is then:

$$\gamma_{ij} = \gamma_{00} + \gamma_{10}T_{ij} + u_{0\mathrm{i}} + u_{1i}T_{ij} + e_{ij}$$

(9)

As can be deduced from Table 2, there is no strong evidence of the supposed variation of the effect associated with the third occasion. The change in -2 * log-likelihood from 1718.337-1717.844 = 0.493 (compared to a χ^2 distribution with 2df) is small. So the rate of change in SBP on the third occasion does not vary from patient to patient. To explain the variation of the SBP on the second occasion, it is proposed to include the variable I in the regression equation for β_{0i}:

$$\beta_{0i} = \gamma_{00} + \gamma_{01}I_i + u_{0\mathrm{i}}$$

(10)

The multilevel model now takes the form:

$$\gamma_{ij} = \gamma_{00} + \gamma_{10}T_{ij} + \gamma_{10}I_i + u_{0\mathrm{i}} + e_{ij}$$

(11)

In this model, the hypothesis is corroborated that the variation in SBP between individuals, after placing the device, depends essentially on its

initial values, since a significant change is obtained in the -2 * log-likelihood. (1717.844-1661.430 = 56.414). Another indicator for this dependency is the change in the estimated variance (Table 2).

Finally, the device effect is introduced, so that now:

$$\beta_{0i} = \gamma_{00} + \gamma_{01} I_i + u_{0\mathrm{i}} \quad (12)$$

Substituting, the multilevel model is obtained:

$$\gamma_{ij} = \gamma_{00} + \gamma_{10} T_{ij} + \gamma_{10} I_i + \gamma_{02} D_i + u_{0\mathrm{i}} + e_{ij} \quad (13)$$

The value of 27.56 units for γ_{02} in Table 2 reflects the difference between the two groups of devices. Note that this value indicates how much the estimated mean SBP of device group II is higher than that of device group I. Also note the significant decrease in the value of the estimated variance (from 261.99 to 71.42). The random parameter at level 2 corresponds to the estimate of the population variance of the intercept. Assuming model (7) as valid and considering the variance estimate at level 2, a 95% confidence interval was obtained for the intercept of 127.30+ $1.96\sqrt{71.42}$[110.74,143.56] units. Based on experiences from previous experiments, it is expected that the SBP will tend to reset 5 minutes after the device has been placed. This is reflected in the negative value of the estimated covariance $\sigma^2_{u_0 u_1}$(Table 2), which implies that patients with a higher SBP on the second occasion tend to show a more significant reduction on the third occasion.

Table 2: Estimated parameters and their corresponding standard errors for models 1, 2, 3 and 4.

Models for SBP / Fixed parameters	1	2	3	4
γ_{00} (Const)	141.22 (1.87)	141.22 (1.93)	141.21 (1.63)	127.30 (1.04)
γ_{10} (T)	-8.28 (2.65)	-8.28 (2.64)	-8.28 (2.30)	-8.28 (1.20)
γ_{01} (I)	-	-	0.726 (0.09)	0.83 (0.04)
γ_{02} (D)	-	-	-	27.56 (1.20)
Random Parameters Level 2				
$\sigma^2_{u_0}$	347.44 (34.57	371.98 (52.87)	261.99 (26.32)	71.42 (7.17)
$\sigma^2_{u_1}$	-	1.73 (0.02)	-	-
$\sigma^2_{u_0u_1}$	-	-25.31 (1.96)	-	-
(=-2 * log-likelihood)	1718.337	1717.844	1661.43	1403.087

A graphical analysis of the normality of the residuals (Figure 2) shows that this assumption is not violated.

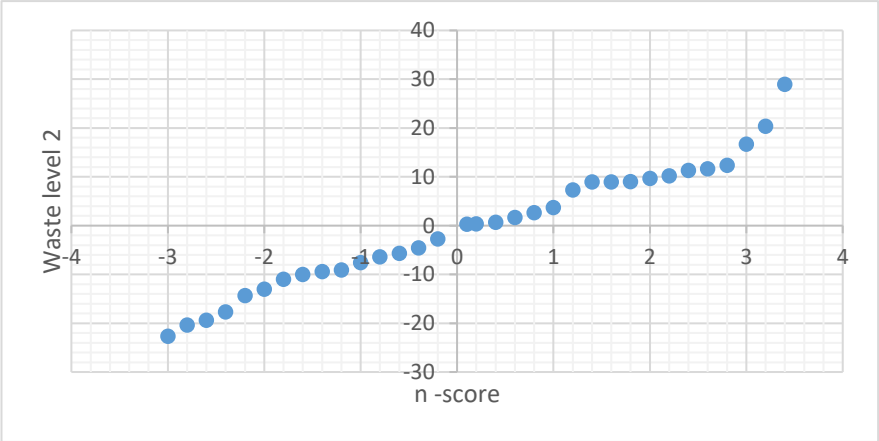


Figure 2: Graph of the normality of the residuals at level 2.

3.3. Multivariate multilevel models

The multilevel model can be extended to incorporate multidimensional responses, taking the structure of a 3-level model. The different dimensions are recognized as units at level 1, nested within occasions, and units at level 2, nested within individuals at level 3. Differences between and within individuals are treated in the same way as the univariate case and the

differences between the dimensions are assumed to be fixed. The example in this article assumes that the model parameterization for each dimension is the same, but this is not a necessary restriction. Trying to find out the existence of correlation between the dimensions, a graph (Figure 3) of the SBP against the DBP was made, which corresponds to a positive correlation.

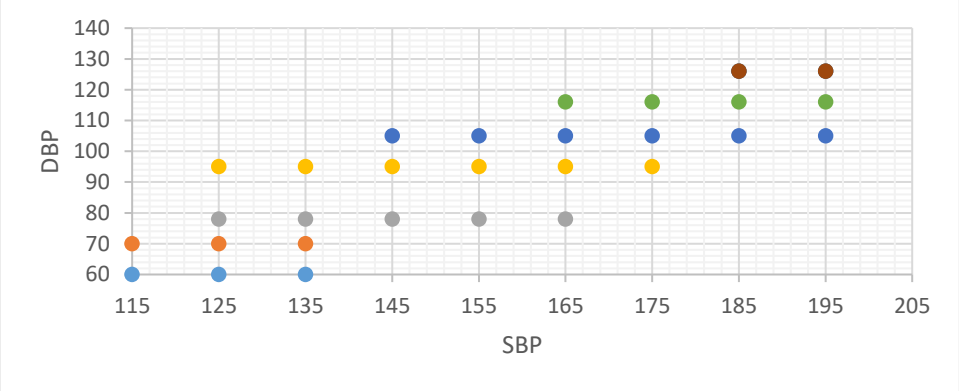


Figure 3: Relationship between SBP and DBP

A model (7) has allowed us a substantial interpretation of the phenomenon under study, we have chosen it to illustrate the extension to the multivariate case.

In the multivariate case, γ_{ijl} denotes the response of the Ith variable (l = 1,2) for the jth measurement of the ith patient and the model is:

$$\gamma_{ij} = \sum_{k=1}^{2}\{\gamma_{00k} + \gamma_{10k}T_{ijk} + \gamma_{10k}I_{ik} + \gamma_{02k}D_{ijk} + u_{0ik} + e_{ijk}\}z_{ijk} \quad (14)$$

where z_{ijk} is worth 1 when $k = l$, and is worth 0 otherwise.

The fixed parameters γ_{00}, γ_{10}, γ_{01}, and γ_{02} represent the population mean effects for the responses in the k-th variable. The estimates of the model parameters are presented in Table 3.

Table 3: Estimated parameters and their i.e., for the multivariate model of SBP and DB

Models for SBP / Fixed parameters	SBP (k = 1)	DBP (k = 2)
γ_{00}	127.29 (1.04)	79.49 (0.77)
γ_{10}	-8.28 (1.20)	-5.76 (0.87)
γ_{01}	0.83 (0.04)	0.75 (0.05)
γ_{02}	27.55 (1.21)	24.19 (0.89)
Random Parameters Level 3		
$\sigma^2_{u_{0k}}$	71.42 (7.17)	24.19 (3.82)
$\sigma^2_{u_0u_1}$		20.96 (1.08)

The analysis provided evidence of significant differences between the levels of SBP (and DBP) between the different subjects, after the device was placed. As can be deduced from the estimates presented in Table 3, the initial values of the SBP are, to some extent, responsible for this difference. Still, the type of device contributes much more to the variability. If the negative value of the estimate of γ_{10} decreasing over time is observed for both variables, it also follows that there is a trend

Note that the values for the SBP reflect approximately the same behaviour as those for the DBP. The estimate of the covariance between the intercepts of both dimensions, according to Table 3, is $\sigma^2_{u_{01}u_{02}}$= 20.96. Using a multivariate model has made it possible to make this estimate.

The estimate of the correlation between patients ($\frac{20.96}{\sqrt{71.42\times24.19}}$= 0.81) suggests that patients who on average have a high DBP, also have a high SBP.

4. CONCLUSIONS

Longitudinal data in this research can be considered within a hierarchical structure. They can be explained in a more realistic way through multilevel models, where each individual is represented by their own regression equation. The use of this type of model takes into account individual heterogeneity and the correlation structure of the data. In multilevel models, the fixed components of the model are studied and the variance components. This allows inferences to be made about population effects. All of the above places multilevel models in an advantageous position concerning classical regression models.

Regarding the example, it was corroborated that there are differences in the levels of the SBP and the DBP between the individuals. The differences are mainly due to the type of device used. For the patients who received the laryngeal mask, the SBP and DBP remained almost unchanged during the period in which the records were made, which could be interpreted that the use of the laryngeal mask appears to be safe has little effect in blood pressure.

Using the proposed multivariate model allowed studying the influence of two variables with the same response number in each dimension. Still, the multivariate formulation can be more flexible, allowing different numbers of responses in each dimension and the fact that individuals, for example, can have an answer in one dimension where the corresponding answer in another dimension is missing.

The advantages of multilevel modelling over traditional methods are undeniable, although also in many situations, in which the complex

structure outlined here is not present, the use of classical regression techniques will suffice.

References

[1] Adzrago, D., Kavi, L., Ezeugoh, R. I., & Osafo-Darko, B. (2022). Temporal Patterns of Cigarette Smoking and Its Associated Covariates: a Multilevel Longitudinal Data Analysis. *Global Social Welfare*, 1-9.

[2] BEACON, H. J., & THOMPSON, S. G. (1996). Multi-level models for repeated measurement data: application to quality of life data in clinical trials. *Statistics in Medicine*, *15*(24), 2717-2732.

[3].Hair Jr, Joseph F., and Luiz Paulo Fávero. (2019) "Multilevel modeling for longitudinal data: concepts and applications." *RAUSP management journal* 54 , 459-489.

[4] Holden, J. E., Kelley, K., & Agarwal, R. (2008). Analyzing change: a primer on multilevel models with applications to nephrology. *American journal of nephrology*, *28*(5), 792-801.

[5] Goldstein, Harvey, William Browne, and Jon Rasbash. (2002) "Multilevel modelling of medical data." *Statistics in medicine* 21.21, 3291-3315.

[6] Kwok, O. M., Underhill, A. T., Berry, J. W., Luo, W., Elliott, T. R., & Yoon, M. (2008). Analyzing longitudinal data with multilevel models: an example with individuals living with lower extremity intra-articular fractures. *Rehabilitation psychology*, *53*(3), 370.

[7] Mwangi, M., Verbeke, G., Njagi, E. N., Mwalili, S., Ivanova, A., Bukania, Z. N., & Molenberghs, G. (2021). Improved longitudinal data analysis for cross-over design settings, with a piecewise linear mixed-effects model. *Communications in Statistics: Case Studies, Data Analysis and Applications*, *7*(3), 413-431.