Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v10n1p146

Dual balance correction in repeated measures ANOVA with missing data

By Alexeyeva

Published: 26 April 2017

Electronic Journal of Applied Statistical Analysis
Vol. 10, Issue 01, April 2017, 146-159
DOI: 10.1285/i20705948v10n1p146

Dual balance correction in repeated measures ANOVA with missing data

Nina Alexeyeva*

Department of Statistical Modeling
Faculty of Mathematics and Mechanics
St. Petersburg State University
Universitetskiy pr. 28, 198504, St. Petersburg, Russia

Published: 26 April 2017

The repeated measures ANOVA is being used in the study of longitudinal data. In its standard form, the method requires full data. In practice, this is not always possible. In this paper, we consider an approach, in which the missing data are neither ignored nor are imputed. The incomplete data result in biased estimations of the time mean. This bias consists of the individual and the group components which are constructed using sequences of the regular stochastic matrices. The group adjustment is necessary in cases of the different proportion of missing values in compared groups.

keywords: ANOVA Repeated Measures, missing data, centralization of model, the regular stochastic matrix, covariance matrix of errors, the cross-averaging procedure.

1 Introduction

A major problem in the analysis of clinical trials using a repeated measures ANOVA (Fitzmaurice G., Davidian M., Verbeke G., Molenberghs G., 2008) is missing data caused by the patients' dropping out of the study before completion. For example, a randomized placebo-controlled trial was performed to test the efficacy of oral naltrexone (NL) with or without fluoxetine (FR) for preventing relapse to heroin addiction (Krupitsky E., Zvartau E., Verbitskaya E., Alexeyeva N. at al., 2012). Some patients received placebo (PL) instead of the drugs. Therefore, four groups were compared: the full drug (NL+FR),

*Corresponding author: nina.alekseeva@spbu.ru

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

the only naltrexone (NL+PL), the only fluoxetine (PL+FR) and the placebo-placebo (PL+PL). The full data were available for only 10% of the year-round monitored patients. Usually, cases with incomplete data are excluded from analysis or the missing values are imputed in some way (Rubin D. B., 1976), (Hedden S.L., Woolson R.F., Carter R.E. et al., 2009). Thus in (Myers W.R., 2000) five strategies of missing data handling are discussed, whereby it is stated that the nonignorable missing-data model is not trivial. This problem is however partly solved in (Alexeyeva N., 2013), (Ufliand A., Alexeyeva N., 2014) based on the ergodicity of cross-mean sequences in cases of the same frequency of missing data appearance across compared groups. The ergodic method for individual corrections of the time mean is applied to data analysis in clinical cardiology (Alexeyeva N.P., Tatarinova A.A., Bondarenko B.B. et al., 2011). The present paper is intended to generalize the above-mentioned result to the case of non-uniformity of missing values across the groups and to simplify the covariance matrix of errors.

2 The repeated measures analysis of variance and missing data

Consider the classical model of repeated measures analysis (Afifi A., Azen S., 1972) of variance (ANOVA)

$$x_{ijt} = \mu + \alpha_i + e^1_{ij} + \beta_t + \gamma_{it} + e_{ijt}, \tag{1}$$

where x_{ijt} is the data of the jth individual from the ith group at the time moment t, μ is the general mean, α_i is the group effect, β_t is the time effect, γ_{it} is the group and the time interaction effect, $e^1_{ij} \sim N(0, \sigma_1^2)$ is the error, caused by the individuals variety and $e_{ijt} \sim N(0, \sigma^2)$ is the general model error. All errors are assumed to be independent. Let the number of groups be equal to I, the number of individuals in the group is equal to ν_i and the number of time points is equal T. Let M_{it} be the set of individuals from group i, who have complete data at the time t; and let m_{it} denote its cardinality, $\sum\limits_{t=1}^{T} m_{it} = m_{i\cdot}$, $\sum\limits_{i=1}^{I} m_{it} = m_{\cdot t}$, $\sum\limits_{t=1}^{T} m_{t\cdot} = m_{\cdot\cdot}$. Let N_{ij} be the set of time points of the individual number j from group i and we call n_{ij} its cardinality. In order to obtain the unique solutions of the systems of linear equations by means of LSM (Least Square Method) for parameters estimation (Scheffe H., 1999), the partial plan was considered:

$$\sum_{i=1}^{I} \frac{\alpha_i m_{i\cdot}}{m_{\cdot\cdot}} = 0, \quad \sum_{t=1}^{T} \frac{\beta_t m_{\cdot t}}{m_{\cdot\cdot}} = 0, \quad \sum_{i=1}^{I} \frac{\gamma_{it} m_{it}}{m_{\cdot\cdot}} = 0, \quad \sum_{t=1}^{T} \frac{\gamma_{it} m_{it}}{m_{\cdot\cdot}} = 0. \tag{2}$$

In order to estimate parameters, the model (2) was divided into two parts: $x_{ijt} = z_{ij} + y_{ijt}$, where $\mathbb{E} z_{ij} = \mu + \alpha_i$, $\mathbb{E} y_{ijt} = \beta_t + \gamma_{it}$. When the data are complete, z_{ij} is just the time mean $x_{ij\cdot}$. In case of missing data, the time mean becomes

$$x_{ij\cdot} = \frac{1}{n_{ij}} \sum_{t \in N_{ij}} x_{ijt},$$

and the mathematical expectation $\mathbb{E}x_{ij\cdot}$ is not equal to $\mu+\alpha_i$. To solve this problem, we propose to introduce two corrections: group and individual. Denote

$$e^1_{\cdot\cdot}(t)=\frac{1}{m_{\cdot t}}\sum_{i=1}^{I}\sum_{j\in M_{it}}e^1_{ij}\,,\quad \alpha(t)=\frac{1}{m_{\cdot t}}\sum_{i=1}^{I}m_{it}\alpha_i,\quad \beta_{\cdot}(i)=\frac{1}{m_{i\cdot}}\sum_{t=1}^{T}m_{it}\beta_t \tag{3}$$

sums depending on points $t=1,\dots,T$ and groups $i=1,\dots,I$. Then consider expressions through the parameters of different types of averages:

- the time mean

$$x_{ij\cdot}=\frac{1}{n_{ij}}\sum_{t\in N_{ij}}(\mu+\alpha_i+e^1_{ij}+\beta_t+\gamma_{it}+e_{ijt})=\mu+\alpha_i+\frac{1}{n_{ij}}\sum_{t\in N_{ij}}(\beta_t+\gamma_{it})+e^1_{ij}+e_{ij\cdot}\,,$$

  where $e_{ij\cdot}=\frac{1}{n_{ij}}\sum\limits_{t\in N_{ij}}e_{ijt}\,;$

- the space mean $x_{\cdot\cdot t}$

$$x_{\cdot\cdot t}=\frac{1}{m_{\cdot t}}\sum_{i=1}^{I}\sum_{j\in M_{it}}(\mu+\alpha_i+e^1_{ij}+\beta_t+\gamma_{it}+e_{ijt})=$$

$$=\mu+\underbrace{\frac{1}{m_{\cdot t}}\sum_{i=1}^{I}m_{it}\alpha_i}_{\alpha(t)}+\beta_t+\underbrace{\frac{1}{m_{\cdot t}}\sum_{i=1}^{I}m_{it}\gamma_{it}}_{=0}+e_{\cdot\cdot t}+\underbrace{\frac{1}{m_{\cdot t}}\sum_{i=1}^{I}\sum_{j\in M_{it}}e^1_{ij}}_{e^1_{\cdot\cdot}(t)}\,;$$

- the space mean in ith group $x_{i\cdot\cdot}$

$$x_{i\cdot\cdot}=\frac{1}{m_{i\cdot}}\sum_{t=1}^{T}\sum_{j\in M_{it}}(\mu+\alpha_i+e^1_{ij}+\beta_t+\gamma_{it}+e_{ijt})=$$

$$=\mu+\alpha_i+\underbrace{\frac{1}{m_{i\cdot}}\sum_{t=1}^{T}m_{it}\beta_t}_{\beta_{\cdot}(i)}+\underbrace{\frac{1}{m_{i\cdot}}\sum_{t=1}^{T}m_{it}\gamma_{it}}_{=0}+e^1_{i\cdot}+e_{i\cdot\cdot},$$

$$\text{where } e^1_{i\cdot}=\frac{1}{m_{i\cdot}}\sum_{t=1}^{T}\sum_{j\in M_{it}}e^1_{ij}=\frac{1}{m_{i\cdot}}\sum_{j=1}^{\nu_i}n_{ij}e^1_{ij},\quad e_{i\cdot\cdot}=\frac{1}{m_{i\cdot}}\sum_{t=1}^{T}\sum_{j\in M_{it}}e_{ijt}\,;$$

- the general mean

$$x_{\cdot\cdot\cdot}=\frac{1}{m_{\cdot\cdot}}\sum_{i=1}^{I}m_{i\cdot}x_{i\cdot\cdot}=\frac{1}{m_{\cdot\cdot}}\sum_{i=1}^{I}m_{i\cdot}(\mu+\alpha_i+\beta_{\cdot}(i)+e^1_{i\cdot}+e_{i\cdot\cdot})=$$

$$=\mu+\underbrace{\frac{1}{m_{\cdot\cdot}}\sum_{i=1}^{I}m_{i\cdot}\alpha_i}_{=0}+\underbrace{\frac{1}{m_{\cdot\cdot}}\sum_{i=1}^{I}m_{i\cdot}\beta_{\cdot}(i)}_{=0}+e^1_{\cdot\cdot}+e_{\cdot\cdot\cdot}\,.$$

2.1 Group correction

Consider the model parameters as vectors $a = (\alpha_1, \dots, \alpha_I)^T$, $b = (\beta_1, \dots, \beta_I)^T$, $g_i = (\gamma_{i1}, \dots, \gamma_{iT})^T$, $i = 1, 2, \dots, I$. Denote differences between the means

$$L = (x_{\cdot\cdot 1} - x_{\dots}, \dots, x_{\cdot\cdot T} - x_{\dots})^T, \quad K = (x_{1\cdot\cdot} - x_{\dots}, \dots, x_{I\cdot\cdot} - x_{\dots})^T , \tag{4}$$

where the differences between the means are given by

$$x_{i\cdot\cdot} - x_{\dots} = \alpha_i + \beta_\cdot(i) + e^1_{i\cdot} + e_{i\cdot\cdot} - e^1_{\cdot\cdot} - e_{\dots}, \tag{5}$$

$$x_{\cdot\cdot t} - x_{\dots} = \alpha(t) + \beta_t + e^1_{\cdot\cdot}(t) + e_{\cdot\cdot t} - e^1_{\cdot\cdot} - e_{\dots} ,$$

and the stochastic matrices M and N

$$M = M_{(I,T)} = \begin{bmatrix} \frac{m_{11}}{m_{1\cdot}} & \dots & \frac{m_{1T}}{m_{1\cdot}} \\ \vdots & \dots & \vdots \\ \frac{m_{I1}}{m_{I\cdot}} & \dots & \frac{m_{IT}}{m_{I\cdot}} \end{bmatrix}, \quad N = N_{(T,I)} = \begin{bmatrix} \frac{m_{11}}{m_{\cdot 1}} & \dots & \frac{m_{I1}}{m_{\cdot 1}} \\ \vdots & \dots & \vdots \\ \frac{m_{1T}}{m_{\cdot T}} & \dots & \frac{m_{IT}}{m_{\cdot T}} \end{bmatrix}. \tag{6}$$

Denote $P_0 = MN$ the product of these two matrices M and N, and P_0^∞ its stationary matrix

$$P_0^\infty = \lim_{n\to\infty} P_0^n . \tag{7}$$

We assume that the matrix P_0 is regular. One can see that rows of the matrix P_0^∞ look like the left eigenvector λ^T of matrix P_0, where

$$\lambda = \left(\frac{m_{1\cdot}}{m_{\cdot\cdot}}, \dots, \frac{m_{I\cdot}}{m_{\cdot\cdot}} \right)^T .$$

Define the vector of the group corrections as

$$G = \sum_{i=0}^{\infty} P_0^i (ML - P_0 K) . \tag{8}$$

Proposition 1. *The mathematical expectations of the introduced vectors are given by*

$$\mathbb{E}L = b + Na, \quad \mathbb{E}K = a + Mb , \quad \mathbb{E}G = Mb .$$

Proof. Using the expressions (3),(4) and (5), we get

$$\mathbb{E}x_{\dots} = \mu, \qquad \mathbb{E}x_{i\cdot t} = \mu + \alpha_i + \beta_t + \gamma_{it},$$

$$\mathbb{E}x_{i\cdot\cdot} = \frac{1}{m_{i\cdot}} \sum_{t=1}^{T} m_{it} \mathbb{E}x_{i\cdot t} = \mu + \alpha_i + \frac{1}{m_{i\cdot}} \sum_{t=1}^{T} m_{it}\beta_t,$$

$$\mathbb{E}(x_{i\cdot\cdot} - x_{\dots}) = \alpha_i + \frac{1}{m_{i\cdot}} \sum_{t=1}^{T} m_{it}\beta_t ,$$

which results in $\mathbb{E}K = a + Mb$. Analogically we get

$$x_{\cdot\cdot t} = \frac{1}{m_{\cdot t}} \sum_{i=1}^{I} m_{it} \mathbb{E} x_{i\cdot t} = \mu + \frac{1}{m_{\cdot t}} \sum_{i=1}^{I} m_{it} \alpha_i + \beta_t,$$

$$\mathbb{E}(x_{i\cdot t} - x_{\cdot\cdot\cdot}) = \frac{1}{m_{\cdot t}} \sum_{i=1}^{I} m_{it} \alpha_i + \beta_t,$$

which results in $\mathbb{E}L = b + Na$. Substitute values $\mathbb{E}L$ and $\mathbb{E}K$ in $\mathbb{E}G$,

$$\mathbb{E}G = \sum_{i=0}^{\infty} \mathbb{E} P_0^i (ML - P_0 K) = M\mathbb{E}L - P_0 \mathbb{E}K + P_0 M \mathbb{E}L - P_0^2 \mathbb{E}K + \ldots =$$

$$= M(b + Na) - P_0(a + Mb) + P_0 M(b + Na) - P_0^2(a + Mb) + \ldots - P_0^\infty Mb =$$

$$= Mb - P_0^\infty Mb = Mb, \quad \text{since} \quad \text{according} \quad \text{to} \quad (3)$$

$$Mb = \{\beta.(i)\}_{i=1}^{I} = \left\{ \frac{1}{m_{i\cdot}} \sum_{t=1}^{T} m_{it} \beta_t \right\}_{i=1}^{I},$$

$$m_{\cdot\cdot} P_0^\infty Mb = \sum_{i=1}^{I} m_{i\cdot} \beta.(i) = \sum_{t=1}^{T} \sum_{i=1}^{I} \beta_t = \sum_{t=1}^{T} m_{\cdot t} \beta_t = 0\,.$$

□

Proposition 2. *Let L and K be differences between the means (4), matrices M, N are defined in (6),*

$$Q_0 = (\mathbb{I} - P_0 + P_0^\infty)^{-1}, \quad A = P_0^\infty + Q_0 - \mathbb{I}, \quad B = Q_0 M\,, \tag{9}$$

where P_0 is defined in (7), $\mathbb{I}$ is the identity matrix. Then the vector of group corrections $G = \sum_{i=0}^{\infty} P_0^i (ML - P_0 K)$ is given by $G = -AK + BL$, and $A = Q_0 P_0 = BN$.

Proof. In the series G from (8) we add and subtract K and then regroup elements:

$$G = (ML - P_0 K) + P_0(ML - P_0 K) + P_0^2(ML - P_0 K) \ldots =$$

$$= K - (K - ML) - P_0(K - ML) - P_0^2(K - ML) - \ldots + P_0^\infty K =$$

$$= (\mathbb{I} - P_0^\infty)K - (\mathbb{I} + P_0 + P_0^2 + \ldots)(K - ML)\,.$$

Remark that $P_0^\infty (K - ML) = 0$, because

$$\frac{1}{m_{\cdot\cdot}} \sum_{i=1}^{I} m_{i\cdot} (x_{i\cdot\cdot} - x_{\cdot\cdot\cdot} - \frac{1}{m_{i\cdot}} \sum_{t=1}^{T} m_{it} x_{\cdot\cdot t} + x_{\cdot\cdot\cdot}) = x_{\cdot\cdot\cdot} - \frac{1}{m_{\cdot\cdot}} \sum_{i=1}^{I} \sum_{t=1}^{T} m_{it} x_{\cdot\cdot t} = 0\,.$$

Using this fact and (9) in the following expression

$$\sum_{k=0}^{\infty} P_0^k (K-ML) = \sum_{k=0}^{\infty} (P_0 - P_0^{\infty})^k (K-ML) = (\mathbb{I} - P_0 + P_0^{\infty})^{-1}(K-ML)\,,$$

we get

$$G = (\mathbb{I} - P_0^{\infty})K + Q_0(ML-K) = -AK + BL\,.$$

Expression $Q_0 P_0 = A$ is also valid, since

$$Q_0^{-1} A = (\mathbb{I} - P_0 + P_0^{\infty})(P_0^{\infty} + Q_0 - \mathbb{I}) =$$

$$= P_0^{\infty} + Q_0 - \mathbb{I} - P_0(P_0^{\infty} + Q_0 - \mathbb{I}) + P_0^{\infty}(P_0^{\infty} + Q_0 - \mathbb{I})) =$$

$$= P_0^{\infty} + Q_0 - \mathbb{I} - P_0^{\infty} - P_0 Q_0 + P_0^{\infty} + P_0 + P_0^{\infty} Q_0 - P_0^{\infty} =$$

$$= Q_0 - \mathbb{I} + P_0 - P_0 Q_0 + P_0^{\infty} Q_0 = -\mathbb{I} + P_0 + Q(\mathbb{I} - P_0 + P_0^{\infty}) = P_0\,.$$

Therefore, $A = Q_0 P_0 = Q_0 MN = BN$.

□

Proposition 3. *Let $A = \{a_{i\iota}\}_{i,\iota=1}^{I}$ and $B = \{b_{it}\}_{i=1}^{I}\,{}_{t=1}^{T}$ be matrices from (9),*

$$F_0(i,k) = \sum_{t=1}^{T} \frac{b_{it} b_{kt}}{m_{\cdot t}} - \sum_{\iota=1}^{I} \frac{a_{i\iota} a_{k\iota}}{m_{\iota\cdot}}\,,$$

$$F_1(\iota, j, i) = -\frac{a_{i\iota}}{m_{\iota\cdot}} + \sum_{t \in N_{\iota j}} \frac{b_{it}}{n_{\iota j} m_{\cdot t}}\,,$$

$$F_2(\iota, t, i) = -\frac{a_{i\iota}}{m_{\iota\cdot}} + \frac{b_{it}}{m_{\cdot t}}\,.$$

Then a random component of i-group corrections $\varepsilon_i = G_i - \mathbb{E}G_i$ does not depend on individual errors e^1_{ij} and is given by

$$\varepsilon_i = \sum_{k=1}^{I} \sum_{t=1}^{T} m_{kt} \left(\frac{b_{it}}{m_{t\cdot}} - \frac{a_{ik}}{m_{k\cdot}} \right) e_{k\cdot t}, \tag{10}$$

and has the following properties:

$$\mathbb{E}\varepsilon_i e_{\iota jt} = \sigma^2 I_{[\iota=i]} \left(\frac{b_{it}}{m_{\cdot t}} - \frac{a_{i\iota}}{m_{\iota\cdot}} \right)\;;\quad \mathbb{E}\varepsilon_i \varepsilon_k = \sigma^2 F_0(i,k),$$

$$\mathbb{E}e_{\iota\cdot}\varepsilon_i = \sigma^2 F_1(\iota, j, i), \qquad \mathbb{E}e_{\iota\cdot t}\varepsilon_i = \sigma^2 F_2(\iota, t, i)\,.$$

Proof. Let us find a random component of vector G using an expression $G = -AK + BL = -AK + Q_0 ML$ from Proposition 2. Denote the random components of vectors ML and AK by $E = ML - \mathbb{E}ML$ and $J = AK - \mathbb{E}AK$ with elements E_i and J_i respectively, $i = 1, \ldots, I$, and $c = e^1_{\cdot\cdot} + e_{\cdots}$. According to (4) and (6) we can get ith element of vector ML, wherefrom element E_i:

$$(ML)_i = \sum_{t=1}^{T} \frac{m_{it}}{m_{i\cdot}} (x_{\cdot\cdot t} - x_{\cdots}) = \sum_{t=1}^{T} \frac{m_{it}}{m_{i\cdot}} (\alpha(t) + \beta_t + e^1_{\cdot\cdot}(t) + e_{\cdot\cdot t} - e^1_{\cdot\cdot} - e_{\cdots}) \,,$$

$$E_i = \sum_{t=1}^{T} \frac{m_{it}}{m_{i\cdot}} (e^1_{\cdot\cdot}(t) + e_{\cdot\cdot t}) - c \,.$$

Changing the order of summation, we get

$$E_i + c = \sum_{t=1}^{T} \frac{m_{it}}{m_{i\cdot}} \left(\frac{1}{m_{\cdot t}} \sum_{k=1}^{I} \sum_{j \in M_{kt}} (e^1_{kj} + e_{kjt}) \right) = \sum_{k=1}^{I} \sum_{j=1}^{\nu_k} (e^1_{kj} + e_{kjt}) \sum_{t \in N_{kj}} \frac{m_{it}}{m_{\cdot t} m_{i\cdot}} \,.$$

Since the matrix $Q_0 = \{q_{i\iota}\}_{i,\iota=1}^{I}$ is stochastic, then the ith random-component of vector $Q_0 ML$ is given by

$$(Q_0 ML)_i = \sum_{\iota=1}^{I} q_{i\iota} E_\iota + c = \sum_{\iota=1}^{I} q_{i\iota} \sum_{k=1}^{I} \sum_{j=1}^{\nu_k} (e^1_{kj} + e_{kjt}) \sum_{t \in N_{kj}} \frac{m_{\iota t}}{m_{\cdot t} m_{i\cdot}} =$$

$$= \sum_{k=1}^{I} \sum_{t=1}^{T} (e^1_{k\cdot} + e_{k\cdot t}) m_{kt} \sum_{\iota=1}^{I} \frac{q_{i\iota} m_{\iota t}}{m_{\iota\cdot} m_{t\cdot}} \,.$$

Also select the random component $J = AK - \mathbb{E}AK$. The $(AK)_\iota$ is given by

$$(AK)_\iota = \sum_{i=1}^{I} a_{\iota i} (x_{\cdot\cdot} - x_{\cdots}) = \sum_{i=1}^{I} a_{\iota i} (\alpha_i + \beta_{\cdot}(i) + e^1_{i\cdot} + e_{i\cdot\cdot} - e^1_{\cdot\cdot} - e_{\cdots}),$$

hence, using $c = e^1_{\cdot\cdot} + e_{\cdots}$, we get

$$J_\iota = \sum_{k=1}^{I} a_{\iota k} (e^1_{k\cdot} + e_{k\cdot\cdot}) - c \,,$$

$$J_\iota + c = \sum_{k=1}^{I} \frac{a_{\iota k}}{m_{k\cdot}} \sum_{t=1}^{T} \sum_{j \in M_{kt}} (e^1_{kj} + e_{kjt}) = \sum_{k=1}^{I} \frac{a_{\iota k}}{m_{k\cdot}} \sum_{t=1}^{T} m_{kt} (e^1_{k\cdot} + e_{k\cdot t}) \,.$$

Thus ε_ι is given by

$$\varepsilon_\iota = \sum_{i=1}^{I} q_{\iota i} E_i - J_\iota = \sum_{k=1}^{I} \sum_{t=1}^{T} m_{kt} \left(\frac{1}{m_{\cdot t}} \sum_{i=1}^{I} \frac{q_{\iota i} m_{it}}{m_{i\cdot}} - \frac{a_{\iota k}}{m_{k\cdot}} \right) (e^1_{k\cdot} + e_{k\cdot t}) \,.$$

Remark that $b_{it}=\sum_{\iota=1}^{I} q_{i\iota} m_{\iota t}/m_{\iota\cdot}$ is an element of matrix $B=Q_0M$. Then one can get rid of the components, depending on the e^1_{ij}. Since $BN=A$, then

$$\sum_{k=1}^{I}\sum_{t=1}^{T} m_{kt}\left(\frac{b_{\iota t}}{m_{\cdot t}}-\frac{a_{\iota k}}{m_{k\cdot}}\right)e^1_{k\cdot}=\sum_{k=1}^{I} e^1_{k\cdot}\left(\sum_{t=1}^{T}\frac{b_{\iota t}m_{kt}}{m_{\cdot t}}-a_{\iota k}\right)=0\,.$$

Covariance can be calculated directly, for example,

$$\mathbb{E}\varepsilon_\iota e_{ij\tau}=\sum_{k=1}^{I}\sum_{t=1}^{T} m_{kt}\left(\frac{b_{\iota t}}{m_{\cdot t}}-\frac{a_{\iota k}}{m_{k\cdot}}\right)\mathbb{E}e_{k\cdot t}e_{ij\tau}=\sigma^2\left(\frac{b_{\iota\tau}}{m_{\cdot\tau}}-\frac{a_{\iota i}}{m_{i\cdot}}\right)\,.$$

□

Proposition 4. *In the case of the same proportion of complete data in groups*

$$\frac{m_{i\cdot}}{m_{\cdot\cdot}}=\frac{m_{it}}{m_{\cdot t}} \tag{11}$$

the group correctionis not needed because $G=\boldsymbol{0},\quad Q_0=\mathbb{I},\quad A=P_0^\infty,\quad \varepsilon_\iota=0.$

Proof. Firstly prove that at correct expression (11) $ML=\mathbf{0}$, $NK=\mathbf{0}$. In fact

$$(ML)_i=\sum_{t=1}^{T}\frac{m_{it}}{m_{i\cdot}}(x_{\cdot\cdot t}-x_{\cdot\cdot\cdot})\overset{(11)}{=}\sum_{t=1}^{T}\frac{m_{\cdot t}}{m_{\cdot\cdot}}(x_{\cdot\cdot t}-x_{\cdot\cdot\cdot})=0\,.$$

$$(NK)_i=\sum_{i=1}^{I}\frac{m_{it}}{m_{\cdot t}}(x_{i\cdot\cdot}-x_{\cdot\cdot\cdot})\overset{(11)}{=}\sum_{i=1}^{I}\frac{m_{i\cdot}}{m_{\cdot\cdot}}(x_{i\cdot\cdot}-x_{\cdot\cdot\cdot})=0\,.$$

Hence $G=\sum_{k=0}^{\infty}P_0^k(ML-MNK)=\mathbf{0}$. Further in cases if (11) correct we have $N=P_0^\infty$, $P_0=MN=N=P_0^\infty$. Therefore $Q_0=(\mathbb{I}-P_0+P_0^\infty)^{-1}=\mathbb{I}$, $A=P_0+Q_0-\mathbb{I}=P^\infty$, $B=M$ and

$$\varepsilon_\iota=\sum_{k=1}^{I}\sum_{t=1}^{T} m_{kt}\left(\frac{b_{\iota t}}{m_{\cdot t}}-\frac{a_{\iota k}}{m_{k\cdot}}\right)e_{k\cdot t}=\sum_{k=1}^{I}\sum_{t=1}^{T} m_{kt}\left(\frac{1}{m_{\cdot\cdot}}-\frac{1}{m_{\cdot\cdot}}\right)e_{k\cdot t}=0\,.$$

□

2.2 Individual corrections

Consider the incidence matrix of missing data J^i with ν_i rows and T columns in ith group. Denote: the diagonal matrix Λ_{ν_i} of dimension ν_i with elements $\frac{1}{n_{ij}}$; the diagonal matrix Λ_{iT} of dimension T with elements $\frac{1}{m_{it}}$; the matrix $R_i=\Lambda_{\nu_i}J^i$ and the stochastic matrix $P_i=R_i\Lambda_{iT}(J^i)^T$ with left eigenvector

$$\pi(i)=\left(\frac{n_{i1}}{m_{i\cdot}},\ldots,\frac{n_{i\nu_i}}{m_{i\cdot}}\right)\,.$$

The stochastic matrix P_i is regular when at least at one point t there are complete data. Denote $U_i = \{x_{i\cdot t}\}_{t=1}^{T}$, $V_i = \{x_{ij\cdot}\}_{j=1}^{\nu_i}$ and describe the recurrent sequence

$$A_i(k) = P_i A_i(k-1), \tag{12}$$

with the initial vector $A_i(0) = R_i U_i - P_i V_i$. Obviously, $A_i(k) = P_i^k A_i(0)$. Define the vector of individual corrections in ith group as

$$H_i = \sum_{k=1}^{\infty} A_i(k)\,.$$

Proposition 5. *1. Let $\beta.(i)$ be defined in (3) and $\{\beta.(i)\}_1^{\nu_i}$ as the vector of dimension ν_i with identical elements $\beta.(i)$, where ν_i is equal the number of individuals in ith group. Then $\mathbb{E}H_i = R_i(b+g_i) - \{\beta.(i)\}_1^{\nu_i}$ is the vector with elements*

$$\mathbb{E}H_{ij} = \frac{1}{n_{ij}} \sum_{\tau \in N_{ij}} (\beta_\tau + \gamma_{i\tau}) - \frac{1}{m_{i\cdot}} \sum_{t=1}^{T} m_{it}\beta_t\,, \quad j = 1, 2, \ldots, \nu_i\,. \tag{13}$$

2. Denote $P_i^\infty = \lim_{n\to\infty} P_i^n$, $Q_i = (\mathbb{I} - P_i + P_i^\infty)^{-1}$, $C_i = P_i^\infty + Q_i$, $D_i = Q_i R_i$. Then $H_i = (\mathbb{I} - C_i)V_i + D_i U_i = \mathbb{E}H_i + \mathcal{E}_i$, where $\mathcal{E}_i = \{\mathcal{E}_{ij}\}_{j=1}^{\nu_i}$ is the vector with elements

$$\mathcal{E}_{ij} = e_{ij\cdot} - \sum_{l=1}^{\nu_i} c_{jl}^i e_{il\cdot} + \sum_{\tau=1}^{T} d_{j\tau}^i e_{i\cdot\tau}\,. \tag{14}$$

Proof. Consider jth element of initial vector

$$A_{ij}(0) = \frac{1}{n_{ij}} \sum_{\tau \in N_{ij}} \frac{1}{m_{i\tau}} \sum_{l \in M_{i\tau}} (x_{il\tau} - x_{il\cdot})\,.$$

According to model (1) we have the expressions

$$x_{ijt} - x_{ij\cdot} = \beta_t + \gamma_{it} + e_{ijt} - (\beta.(i,j) + \gamma_{i\cdot}(j) + e_{ij\cdot}), \text{ where}$$

$$\beta.(i) = \frac{1}{m_{i\cdot}} \sum_{t=1}^{T} m_{it}\beta_t, \quad \beta.(i,j) = \frac{1}{n_{ij}} \sum_{t\in N_{ij}} \beta_t, \quad \text{and} \quad \gamma_{i\cdot}(j) = \frac{1}{n_{ij}} \sum_{t \in N_{ij}} \gamma_{it}.$$

For short in i-group denote $A_i = A, H_i = H, J^i = J, P_i = P, V_i = V, U_i = U, \nu_i = \nu, R_i = R, \Lambda_{iT} = \Lambda_T$. Then

$$A(0) = RL - PK \quad \text{and} \quad H = \sum_{k=0}^{\infty} P^k (RL - PK)\,, \tag{15}$$

where $L = \{\beta_t + \gamma_{it} + e_{i\cdot t}\}_{t=1}^{T}$, $K = \{\beta.(i,j) + \gamma_{i\cdot}(j) + e_{ij\cdot}\}_{j=1}^{\nu}$ with mathematical expectations $\mathbb{E}L = b + g$, $\mathbb{E}K = R(b+g)$. Hence

$$\mathbb{E}H = R(b+g) - PR(b+g) + PR(b+g) - P^2R(b+g) + \ldots = R(b+g) - P^\infty R(b+g)\,.$$

Considering the expression $P^\infty Rb + P^\infty Rg$, we get firstly $P^\infty Rg = \mathbf{0}$, because for each from $\nu = \nu_i$ elements of the vector $P^\infty Rg$, according to (2), we have in ith group

$$\frac{1}{m_{i\cdot}}\sum_{j=1}^{\nu} n_{ij}\frac{1}{n_{ij}}\sum_{t\in N_{ij}}\gamma_{it} = \frac{1}{m_{i\cdot}}\sum_{t\in N_{ij}} m_{it}\gamma_{it} = 0\,.$$

Other expression $P^\infty Rb$ is equal 0 in the case of complete data or when it is correct (11),

$$\frac{1}{m_{i\cdot}}\sum_{j=1}^{\nu} n_{ij}\frac{1}{n_{ij}}\sum_{t\in N_{ij}}\beta_t = \frac{1}{m_{i\cdot}}\sum_{t\in N_{ij}} m_{it}\beta_t = \frac{1}{m_{\cdot\cdot}}\sum_{t\in N_{ij}} m_{\cdot t}\beta_t = 0\,.$$

Therefore (13) is correct. To obtain the random component of individual corrections, it is necessary to perform the following lineal transformations

$$\begin{gathered} H = RL - PK + P(RL-PK) + P^2(RL-PK) + \ldots = \\ = K - (K-RL) - P(K-RL) - P^2(K-RL) - P^3K + \ldots = \\ = K - (\mathbb{I} + P + P^2 + \ldots)(K-RL) - P^\infty K = \\ = (\mathbb{I} - P^\infty)K - (\mathbb{I} - P + P^\infty)^{-1}(K-RL) = \\ = (\mathbb{I} - P^\infty - Q)K + QRL = K - CK + DL, \quad \text{where} \\ Q = (\mathbb{I} - P + P^\infty)^{-1}, \quad C = P^\infty + Q\,, \quad D = QR\,. \end{gathered}$$

So random components of individual corrections look like

$$\mathcal{E}_{ij} = e_{ij\cdot} - \sum_{l=1}^{\nu_i} c^i_{jl} e_{il\cdot} + \sum_{\tau=1}^{T} d^i_{j\tau} e_{i\cdot\tau}\,.$$

If U, V is considered instead of L, K in (15) then one can get similarly $H = (\mathbb{I}-C)V+DU$ which is used for calculating. □

2.3 United corrections

The following statement can be directly obtained from Proposition 5.

Proposition 6. *Let $\Delta_{ij} = e_{ij\cdot} - \mathcal{E}_{ij} - \varepsilon_i$ be the combined error, where ε_i $\mathcal{E}_{ij}$ are random components of G_i and H_{ij} from (10) and (14) respectively, matrices $C_i = \{c^i_{jl}\}_{j,l=1}^{\nu_i}$, $D_i = \{d^i_{jt}\}_{j,t=1}^{\nu_i,T}$ are defined in Proposition 5. Then*

$$\Delta_{ij} = \sum_{l=1}^{\nu_i} c^i_{jl} e_{il\cdot} - \sum_{t=1}^{T} d^i_{jt} e_{i\cdot t} - \varepsilon_i\,.$$

Proposition 7. *Let J_{ijt} denote the indicator of jth individual from ith group in time-moment t, $I_{[A]}$ denote the indicator of set A, then*

$$1.\ \mathbb{E}e_{ij\cdot}e_{i_1j_1\cdot} = \frac{\sigma^2 I_{[i=i_1]} I_{[j=j_1]}}{n_{ij}}\,;\quad 2.\ \mathbb{E}e_{ij\cdot}e_{k\cdot t} = \frac{I_{[i=k]}\sigma^2}{n_{ij}m_{it}} J_{ijt}\,.$$

$$3.\ \mathbb{E}e_{ij\cdot}e_{\cdots} = \mathbb{E}e_{i\cdot t}e_{\cdots} = \mathbb{E}e_{i\cdot\cdot}e_{\cdots} = \mathbb{E}e_{\cdot\cdot t}e_{\cdots} = \frac{\sigma^2}{m_{\cdot\cdot}}\,.\ 4.\ \mathbb{E}e_{ij\cdot}e_{k\cdot\cdot} = \frac{I_{[i=k]}\sigma^2}{m_{i\cdot}}\,.$$

$$5.\ \mathbb{E}e_{ij\cdot}e_{\cdot\cdot t} = \frac{I_{[i=k]}\sigma^2}{n_{ij}m_{\cdot t}} J_{ijt}\,.\ \ 6.\ \mathbb{E}e_{i\cdot t}e_{k\cdot\cdot} = \frac{\sigma^2 I_{[i=k]}}{m_{i\cdot}}\,.\ \ 7.\ \mathbb{E}e_{i\cdot\cdot}e_{\cdot\cdot t} = \frac{m_{it}\sigma^2}{m_{i\cdot}m_{\cdot t}}\,.$$

Proposition 8. *Denote $C_i = C$, $D_i = D$, $Q_i = Q$. Let $F_1(i,\cdot,k) = \{F_1(i,\cdot,k)\}_{j=1}^{\nu_i}$, $F_2(i,\cdot,k) = \{F_2(i,\cdot,k)\}_{j=1}^{\nu_i}$ be from Proposition 7. Then*

$$1)\ P_i^\infty F_1(i,\cdot,k) = \boldsymbol{0}.\quad 2)\ DF_2(i,\cdot,k) = CF_1(i,\cdot,k).$$

Proof. 1) All rows of the stochastic matrix P_i^∞ are the same, so consider only one row

$$\sum_{j=1}^{\nu_i} \frac{n_{ij}}{m_{i\cdot}} \left(-\frac{a_{ki}}{m_{i\cdot}} + \sum_{t\in N_{ij}} \frac{b_{kt}}{n_{ij}m_{\cdot t}} \right) = \frac{1}{m_{i\cdot}} \left(-a_{ki} + \sum_{t=1}^{T} \frac{b_{kt}m_{it}}{m_{\cdot t}} \right) = 0,$$

because, according Proposition 2,

$$A = BN = \left\{ \sum_{t=1}^{T} \frac{b_{kt}m_{it}}{m_{\cdot t}} \right\}_{k,i=1}^{I}.$$

2) On this basis one can see that

$$\frac{1}{\sigma^2} CF_1(i,\cdot,k) = (P_i^\infty + Q)F_1(i,\cdot,k) = Q \left\{ -\frac{a_{ki}}{m_{i\cdot}} + \sum_{t\in N_{ij}} \frac{b_{kt}}{n_{ij}m_{\cdot t}} \right\}_{j=1}^{\nu_i} =$$

$$= -\frac{a_{ki}}{m_{i\cdot}} + Q_i\Lambda_{\nu_i}J^i \left\{ \frac{b_{kt}}{m_{\cdot t}} \right\}_{t=1}^{T}.$$

On the other hand

$$\frac{1}{\sigma^2} D_iF_2(i,\cdot,k) = Q_i\Lambda_{\nu_i}J^i \left\{ -\frac{a_{ki}}{m_{i\cdot}} + \frac{b_{kt}}{m_{\cdot t}} \right\}_{t=1}^{T} = -\frac{a_{ki}}{m_{i\cdot}} + Q_i\Lambda_{\nu_i}J^i \left\{ \frac{b_{kt}}{m_{\cdot t}} \right\}_{t=1}^{T}.$$

□

The individual correction is needed in cases when the missing data change the individual means over time. For example, a trend is increasing but the data of individuals have less missing data at the beginning of trial than at the end. In this case, a time mean is less than it should be. The group correction is needed in cases when the data of one group have less missing data than that of the other.

Theorem 1. *Let H_{ij} and G_i be the individual and the group corrections respectively, $z_{ij} = x_{ij\cdot} - (H_{ij} + G_i)$ and $y_{ijt} = x_{ijt} - z_{ij}$, the combined error Δ_{ij} is defined in Proposition 6,*

$$M_{ij,kl\tau} = \frac{1}{\sigma^2}\mathbb{E}\Delta_{ij}e_{kl\tau}, \quad \tilde{D}_{ij,kl} = \frac{1}{\sigma^2}\mathbb{E}\Delta_{ij}\Delta_{kl}\,.$$

Then two unbiased models look like $z_{ij} = \mu + \alpha_i + e^1_{ij} + \Delta_{ij}$, $\;y_{ijt} = \beta_t + \gamma_{it} + e_{ijt} - \Delta_{ij}$, besides

$$M_{ij,kl\tau} = I_{[i=k]}\left(\frac{c^i_{jl}}{n_{il}} - \frac{d^i_{j\tau}}{m_{i\tau}}\right) + \left(\frac{a_{ik}}{m_{k\cdot}} - \frac{b_{i\tau}}{m_{\cdot\tau}}\right),$$

$$\tilde{D}_{ij,kl} = F_0(i,k) + I_{[i=k]}\left(\sum_{\iota=1}^{\nu_i}\frac{c^i_{j\iota}c^k_{l\iota}}{n_{i\iota}} - \sum_{\iota=1}^{\nu_i}\sum_{\tau\in N_{i\iota}}\left(\frac{c^i_{j\iota}d^k_{l\tau}}{n_{i\iota}m_{k\tau}} + \frac{d^i_{j\tau}c^k_{l\iota}}{n_{k\iota}m_{i\tau}}\right) + \sum_{t=1}^{T}\frac{d^i_{jt}d^k_{lt}}{m_{it}}\right),$$

$$\mathbb{E}(e_{ijt} - \Delta_{ij})(e_{kl\tau} - \Delta_{kl}) = \sigma^2(I_{[i=k,j=l,t=\tau]} - M_{ij,kl\tau} - M_{kl,ijt} + \tilde{D}_{ij,kl})\,,$$

$$\mathbb{E}(e^1_{ij} + \Delta_{ij})(e^1_{kl} + \Delta_{kl}) = \sigma_1^2 I_{[i=k,j=l]} + \sigma^2\tilde{D}_{ij,kl}\,.$$

Proof. On the base of Proposition 7 the necessary covariance can be constructed:

$$\mathbb{E}e_{ijt}e_{kl\tau} = I_{[i=k,j=l,t=\tau]}\sigma^2,$$

$$M_{ij,kl\tau} = \mathbb{E}\Delta_{ij}e_{kl\tau} = \mathbb{E}e_{kl\tau}\left(\sum_{\iota=1}^{\nu_i}c^i_{j\iota}e_{i\iota\cdot} - \sum_{t=1}^{T}d^i_{jt}e_{i\cdot t} - \varepsilon_i\right) =$$

$$= \sigma^2 I_{[i=k]}\left(\frac{c_{jl}}{n_{il}} - \frac{d_{j\tau}}{m_{i\tau}}\right) + \sigma^2\left(\frac{a_{ik}}{m_{k\cdot}} - \frac{b_{i\tau}}{m_{\cdot\tau}}\right)$$

Calculate the covariance $\sigma^2\tilde{D}_{ijkl} = \mathbb{E}\Delta_{ij}\Delta_{kl} =$

$$= \mathbb{E}\left(\sum_{\iota=1}^{\nu_i}c^i_{j\iota}e_{i\iota\cdot} - \sum_{t=1}^{T}d^i_{jt}e_{i\cdot t} - \varepsilon_i\right)\left(\sum_{j_1=1}^{\nu_k}c^k_{lj_1}e_{kj_1\cdot} - \sum_{\tau=1}^{T}d^k_{l\tau}e_{k\cdot\tau} - \varepsilon_k\right) =$$

$$= \sum_{\iota=1}^{\nu_i}c^i_{j\iota}\sum_{j_1=1}^{\nu_k}c^k_{lj_1}\mathbf{E}e_{i\iota\cdot}e_{kj_1\cdot} - \sum_{\iota=1}^{\nu_i}c^i_{j\iota}\sum_{\tau=1}^{T}d^k_{l\tau}\mathbf{E}e_{i\iota\cdot}e_{k\cdot\tau} - \sum_{\iota=1}^{\nu_i}c^i_{j\iota}\mathbf{E}e_{i\iota\cdot}\varepsilon_k -$$

$$- \sum_{t=1}^{T}d^i_{jt}\sum_{j_1=1}^{\nu_k}c^k_{lj_1}\mathbf{E}e_{i\cdot t}e_{kj_1\cdot} + \sum_{t=1}^{T}d^i_{jt}\sum_{\tau=1}^{T}d^k_{l\tau}\mathbf{E}e_{i\cdot t}e_{k\cdot\tau} + \sum_{t=1}^{T}d^i_{jt}\mathbf{E}e_{i\cdot t}\varepsilon_k -$$

$$- \sum_{j_1=1}^{\nu_k}c^k_{lj_1}\mathbf{E}\varepsilon_i e_{kj_1\cdot} + \sum_{\tau=1}^{T}d^k_{l\tau}\mathbf{E}\varepsilon_i e_{k\cdot\tau} + \mathbf{E}\varepsilon_i\varepsilon_k =$$

$$= \sigma^2\left(I_{[i=k]}\left(e_{i\iota\cdot}\sum_{j_1=1}^{\nu_i}c_{lj_1}e_{kj_1\cdot}\sum_{\iota=1}^{\nu_i}\frac{c^i_{j\iota}c^k_{l\iota}}{n_{i\iota}} - \sum_{\iota=1}^{\nu_i}\sum_{\tau\in N_{i\iota}}\frac{c^i_{j\iota}d^k_{l\tau}}{n_{i\iota}m_{k\tau}}\right) - \sum_{\iota=1}^{\nu_i}c^i_{j\iota}F_1(i,\iota,k) -\right.$$

$$-\frac{I_{[i=k]}}{m_{it}}\left(\sum_{j_1=1}^{\nu_k}\sum_{t\in N_{ij_1}}\frac{d_{jt}^i c_{lj_1}^k}{n_{kj_1}}-\sum_{t=1}^{T}d_{jt}^i d_{lt}^k\right)+\sum_{t=1}^{T}d_{jt}^i F_2(i,t,k)-$$
$$\left.-\sum_{j_1=1}^{\nu_k}c_{lj_1}^k F_1(k,j_1,i)+\sum_{\tau=1}^{T}d_{l\tau}^k F_2(k,\tau,i)+F_0(i,k)\right)\,.$$

The expectations of different combinations are calculated directly. For example,

$$\sum_{\iota=1}^{\nu_i}c_{j\iota}^i\sum_{\tau=1}^{T}d_{l\tau}^k \mathbf{E}e_{i\iota\cdot}e_{k\cdot\tau}=\sum_{\iota=1}^{\nu_i}c_{j\iota}^i\sum_{\tau=1}^{T}d_{l\tau}^k\frac{I_{[i=k]}\sigma^2}{n_{i\iota}m_{k\tau}}J_{i\iota\tau}=\sigma^2 I_{[i=k]}\sum_{\iota=1}^{\nu_i}\sum_{\tau\in N_{i\iota}}\frac{c_{j\iota}^i d_{l\tau}^k}{n_{i\iota}m_{k\tau}}\,.$$

On the basis of Proposition 8 one can make sure that

$$-\sum_{\iota=1}^{\nu_i}c_{j\iota}^i F_1(i,\iota,k)+\sum_{t=1}^{T}d_{jt}^i F_2(i,t,k)-\sum_{j_1=1}^{\nu_k}c_{lj_1}^k F_1(k,j_1,i)+\sum_{\tau=1}^{T}d_{l\tau}^k F_2(k,\tau,i)=0\,.$$

□

2.4 Biometrical example

So in the case of uniformity of missing values across the groups we have identical results with the group correction or without it. In the case of non-uniformity of missing values, the group correction can change the result, but maybe not. Variation of results may indicate specificity dropout.

For example in the clinical trials (Krupitsky E., Zvartau E., Verbitskaya E., Alexeyeva N. at al., 2012) described in the introduction, only from 6% to 32% patients in different groups remained until the very end of a trial. The patients from a placebo group dropped out more often than in any other group. At the same time, patients from a full drug group dropped out more rarely than in any other group.

One variable, which means the Beck Depression Inventory (BDI) was significantly higher in the full drug group than in any other group and it makes no difference whether it is calculated with the group correction or without it. It means that BDI would have remained the same, even if more patients dropped out. I.e. dependence does not exist between variable BDI and dropout.

Another variable Global Assessment of Functioning Scale (GAF-score) was significantly lower in the placebo group than in any other group without the group correction whereas once the group correction is done, the relevant difference in GAF disappears. It means that GAF would have increased if more patients from the placebo group remained. In fact, all the groups have the same GAF, but dependence exists between variable GAF and dropout. Patients in the placebo group with higher GAF in the absence of real medication dropped out more. For patients in the placebo group with lower GAF, was more effective the psychological care.

Thus comparison of results with and without the group correction can indicate the causes of dropout.

3 Summary

We calculate the group and individual corrections for the time mean which allow to balancing two biased model with off-diagonal covariance matrices errors. This makes it possible to verify the importance of the influence of various factors on the longitudinal data in the case of incomplete data.

4 Acknowledgment

The author would like to thank researcher Jaropolov A. and professor Bondarenko B.B. for thoughtful comments that led to a significantly improved presentation of the paper.

References

Afifi A., Azen S. (1972). *The Analysis of Variance.* Academic Press, New York.

Alexeyeva N. (2013). *Analysis of biomedical systems. Reciprocity. Ergodicity. Synonymy.* Publishing of the Saint-Petersburg State University, Saint-Petersburg.

Alexeyeva N.P., Tatarinova A.A., Bondarenko B.B. et al. (2011). Analysis of repeated cardiological incomplete data based on ergodic centralization of model. *Bulleten of Almazov Center*, 3(8):59–63.

Fitzmaurice G., Davidian M., Verbeke G., Molenberghs G., editor (2008). *Longitudinal Data Analysis*, chapter 5. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. Chapman and Hall/CRC, Boca Raton.

Hedden S.L., Woolson R.F., Carter R.E. et al. (2009). The impact of loss to follow-up hypothesis tests of the treatment effect for several statistical methods in substance abuse clinical trials. *Journal of Substance Abuse Treatment*, 37:54–63.

Krupitsky E., Zvartau E., Verbitskaya E., Alexeyeva N. at al. (2012). Double-blind, randomized, placebo-controlled studies of the use of implantable naltrexone forms retard (prodetoksona) for relapse prevention of opium addiction. *Addiction issues*, 6:3–27.

Myers W.R. (2000). Handling missing data in clinical trials: an overview. *Drug Information Journal*, 34:525–533.

Rubin D. B. (1976). Inference and missing data. *Biometrika*, 63:581–592.

Scheffe H. (1999). *The Analysis of Variance.* John Wiley and Sons, Canada.

Ufliand A., Alexeyeva N. (2014). The dependence of the ergodicity on the time effect in the repeated measures anova with missing data based on the unbiasedness recovery. In *Topics in Statistical Simulation, Springer Proceedings in Mathematics and Statistics 114*, pages 517–527. Springer Science + Business Media New York 2014.