

睡眠見守りセンサーデータの因果分析 その2 －潜在成長モデルによるアプローチ－

大石 信弘^{1,*} 石田 明男²

A Causal Analysis of Sleep Monitoring Sensor Data Part 2 - Latent Growth Model Approach -

Nobuhiro Oishi^{1,*}, Akio Ishida²

In this paper, we applied structural equation modeling (SEM) and latent growth model (LGM) to evaluate the effects of vital signs and environmental data, which were obtained during 90 minutes after falling asleep, on sleep quality. The data used for the analysis is obtained by a care support device used in an elderly care facility. Applying SEM analysis to the data 90 minutes after falling asleep revealed factors that deepen sleep. Then applying LGM analysis, we found that each sleep level factors have a different sleep-deepening effect. Statistical analysis environment R and the lavaan package were used for both SEM and LGM analysis in this paper. The results show that SEM and LGM can be used to build a rational model of the effects of vital signs and environmental conditions on deepening sleep.

キーワード：因果分析、構造方程式モデリング、潜在成長モデル、睡眠レベル、R

Keywords : causal analysis, structural equation modeling, latent growth model, sleep level, R language

1. はじめに

近年、大量のデータが容易に手に入るビッグデータの時代になり⁽¹⁾、それとともに、大量のデータをどのように処理するかが重要になってきた。

例えば、観測される情報間の関係性を統計的に分析して、今まで曖昧だった観測変数間の関係性をはっきりと示すことは、ビッグデータから重要な情報のみを抽出することに役立つ。それを可能にする分析法としては、重回帰分析、さらには潜在変数を組み込んでモデル化した探索的因子分析 (EFA ; Exploratory Factor Analysis) や確認的因子分析 (CFA ; Confirmatory Factor Analysis) など⁽²⁾や、観測変数間の因果関係を明らかにする、グラフィカルモデリング (GM; Graphical Modeling)⁽³⁾や構造方程式モデリング (SEM; Structural Equation Modeling)⁽³⁻⁶⁾がある。さらには、縦断的なデータに対して、変化を促す効果を明らかにする分析法として、潜在成長モデル (LGM; Latent Growth Model)⁽⁶⁾が

ある。

翻ってこれからの日本が直面する社会的な問題を考えると、高齢化社会と出生率低下が引き起こす労働力人口の減少問題がある。この問題はすでに老人介護施設において顕著に現れており、人手不足による介護福祉士への過度な労働負担により入所者へのサービス低下などの問題が既に起きている⁽⁷⁾。

前回の論文では、この問題の解決のために、介護福祉士の労働負担を減らすべく、入所者が夜間眠れずにベッドを離床しないでよりよい睡眠をとってもらうことで、夜間の見回り回数を減らし介護福祉士の労働負担を軽減する案を提案した。そのために、老人介護施設で使用されている介護サポート機器から得られる環境データと睡眠の深さとの因果関係 SEM を用いて分析し、よりよい睡眠を得るための知見を得ることができた^(8,9)。

だが、睡眠レベルが深くなっていく過程については明らかにすることができなかった。また、睡眠の質を決めるのは、入眠してから 90 分間の時間帯にどれだけ深い眠りを得られているかで決まると言われているので、分析に用いるデータを入眠から 90 分間に限った分析を行うことで、眠りの質をよくするための因果分析も行えると考えた。そこで、本論文の目的として、入眠から 90 分のデータを用いて、睡眠の質を決める条件を見つけること、および睡眠の深さを深くする要因を特定することとした。これらのこと

¹ 電子情報システム工学系
〒861-1102 熊本県合志市須屋 2659-2
Faculty of Electronics and Information Systems Engineering
2659-2 Suya, Koshi-shi, Kumamoto, Japan 861-1102

² リベラルアーツ系
〒861-1102 熊本県合志市須屋 2659-2
Faculty of Liberal Arts,
2659-2 Suya, Koshi-shi, Kumamoto, Japan 861-1102

* Corresponding author:
E-mail address: oishi@kumamoto-nct.ac.jp (N. Oishi).

が分かることで、入所者各人がよりよい睡眠を得られるようになり、健康な生活を送ることができるとともに、入床後の巡回タイミングを適切にでき、一人当たりの介護福祉士にかかる労働負担を軽減することが可能になると期待できる。

本論文におけるデータの分析には、統計解析環境 R^(10,11)を用いた。R はフリーウェアであり、教育的にもこの言語の使用が有益であると考えられる⁽¹²⁾。因果関係の分析手法としては前回と同じ SEM を引き続き採用するとともに、睡眠の深さを深める要因を分析するために、LGM を新たに採用する。どちらの手法も R の lavaan パッケージ⁽⁶⁾を用いて実装した。このパッケージは現在も開発が続いているパッケージではあるが、分析結果は厳密で正確であると分析性能については定評がある⁽⁴⁾。また、このパッケージで採用されているモデル記述文法も簡単であるとともに、様々なモデルを表現できる汎用性がある。

2. 分析に用いたデータ

本論文で分析に用いたデータは、前回の論文と同じものを使用した。そのため、データの詳細については、前回の論文を参照^(8,13)してほしい。

今回の分析のために用意したデータは前回と同じ1人分に関する、5分間隔の測定モードで得られたデータであるが、この後の議論のために、改めて7個の観測変数を再掲する；室温 (*tm*)、湿度 (*hm*)、気圧 (*at*)、照度 (*il*)、1分前の呼吸数 (*rr1*)、1分前の心拍数 (*hr1*) および睡眠レベル (*ss*)。分析に用いた時系列データは欠損値を除外した2926時点(約10日分)である。LGMに用いたデータはこのデータをもとに加工したものであるが、これについては後述する。

3. SEMによる因果分析

前回の論文では、この2926時点のデータを用いて、変数選択や潜在変数の割り当てを行った後、(1)式に示す測定方程式及び構造方程式で表されるモデルを想定し、SEM分析を行いパス係数を得た。後に、入眠後90分でのSEM分析のパス係数と比較するために、前回の論文のパス図を

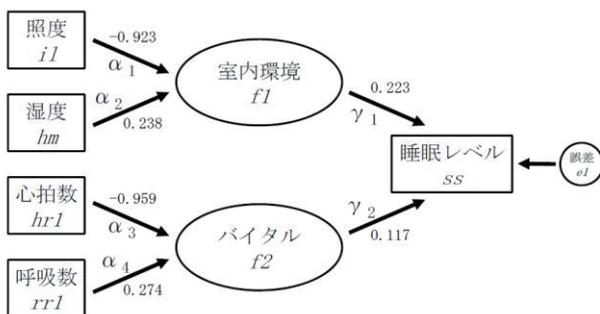


図1 全時刻におけるデータを用いたSEMのパス図

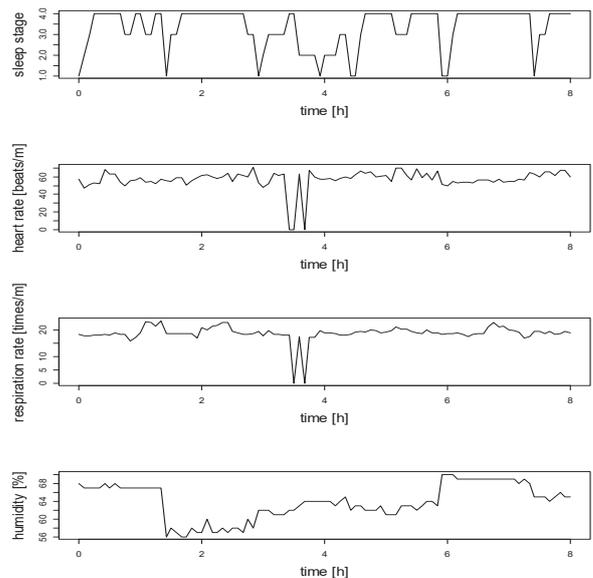


図2 入眠から90分間の時系列データ(8時間分)

1に再掲する。

$$\begin{cases} f1 = \alpha_1 \cdot il + \alpha_2 \cdot hm \\ f2 = \alpha_3 \cdot hr1 + \alpha_4 \cdot rr1 \\ ss = \gamma_1 \cdot f1 + \gamma_2 \cdot f2 + e_1 \end{cases} \dots\dots\dots (1)$$

これは、部屋を暗くしたうえで、心拍数を小さくするように落ち着いてゆったりとした雰囲気になれば睡眠の深さをより深くできることを意味する。

前述のとおり、睡眠の質を決めるのは入眠から90分間でどれだけ深い睡眠レベルにあるかなので、2926時点の全データの分析では睡眠の質をよくするための分析はできない。そこで、入床データの入眠イベント発生時刻をもとに、その時点から90分間の時系列データを抽出してSEM

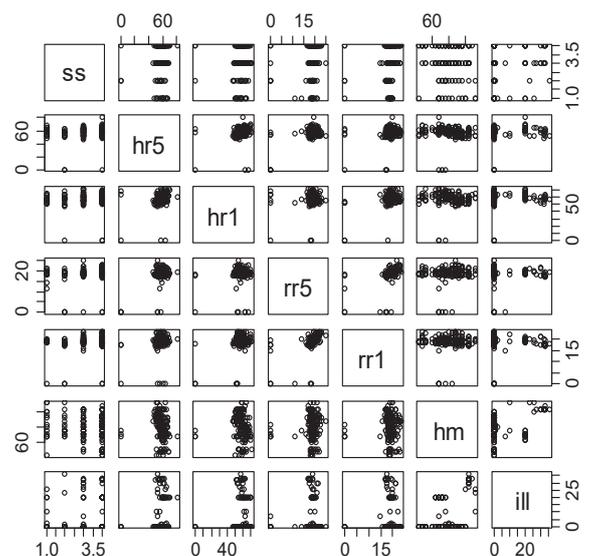


図3 入眠から90分のデータの散布図行列

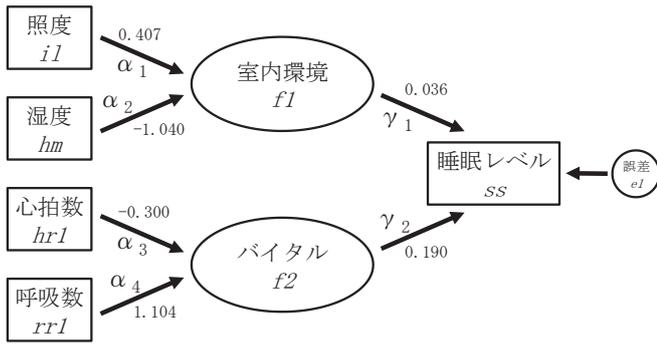


図4 入眠から90分のパス図 (図1と同じモデル)

による分析を行った。図2に抽出した時系列データの8時間分を、図3にその時系列データの散布図行列を示す。照度 (*il*) はほとんどの時点で0または小さな値で変化がなかったため、図2から割愛した。

この入眠から90分間を抽出した時系列データから得られるパス係数と図1の全時刻におけるパス係数を比較するために、図1と同じモデルである(1)式を適用してSEM分析を行った。得られたパス図を図4に示す。潜在変数への影響は、全時刻での観測と違い、特に湿度 (*hm*) と呼吸数 (*rrl*) が大きく貢献している。このことは、睡眠の質をよくするには入眠当初に呼吸数が増えるべく多い睡眠にするという質の良い睡眠が得られるということを示している。

4. LGMによる睡眠を深める効果の分析

では、各観測変数は睡眠の深さを深くするのにどの程度効果があるのだろうか。この疑問に答えるために、潜在成長モデル (LGM) を入眠後90分の抽出データを用いて分析を試みた。

LGMは、例えば、子供の成長(身長伸びや体重の増加など)の時系列変化といったように、同じ対象に複数回の測定を行って得られるデータ(縦断的データ)を分析する手法である。経時的な変化の場合、時間に不変な潜在変数を「切片」とし、時間の変化によって影響が変化する潜在変数を「傾き」として、構造方程式モデリングを行う方法である。これにより継時的な変化に観測変数を与える影響を明らかにすることができる。また、個体差を比較することもできる特徴を持っている。

今回の分析では、睡眠レベルを深める効果を明らかにしたいので、経時的なデータとしてとらえることはできない。時系列データにおいては時間を独立変数として取り扱うが、今回の場合は、睡眠レベルを独立変数とする必要がある。そこで、入眠から90分間の時系列データを4つの睡眠レベルごとに分け、各観測変数の平均値をその各観測変数のスコアとした。また、各睡眠レベルには90分間のうちに占めた時間の長さをスコアとして観測データを加工した。表1に各睡眠ごとの観測変数のスコアを示す。このデータを用いて、睡眠レベルを深めるための切片と傾きに対する効果の程度をLGM分析によって求めた。

LGMにおいては、切片と傾きを潜在変数とし、この潜在変数から各睡眠レベルに向けて1本の矢線が出ている。一方この潜在変数には、これらに影響を及ぼす変数から矢線が入ってくる。潜在変数に影響を及ぼす変数は各睡眠レベルにおける観測変数のスコアであるため、原因となる観測変数は1つの種類に絞ることが望ましい。そのため、各観測変数ごとにLGMによる分析を行ったが、データ点が10点と少ないため、収束しない観測変数が多かった。ここでは収束した例として、心拍数が睡眠レベルを深くする効果を取り上げる。独立変数である睡眠レベルが深くなる(増加する)と潜在変数である傾き (*s*) の分だけ睡眠レベル3 (*ss3*) のスコア(時間)から睡眠レベル4 (*ss4*) のスコア(時間)に増加させる。睡眠レベルの変化に関係なく睡眠レベルのスコアに影響を与えるのもう一つの潜在変数である切片 (*i*) である。これらの潜在変数に結び付いている観測変数は各睡眠レベルでの心拍数の平均値 (*hrss3* および *hrss4*) である。これらの関係を表した測定方程式と構造方程式を(2)式に示す。

$$\begin{cases} i = \alpha_1 \cdot hrss3 + \alpha_2 \cdot hrss4 \\ s = \alpha_3 \cdot hrss3 + \alpha_4 \cdot hrss4 \dots\dots\dots (2) \\ ss3 = 1 \times i + 0 \times s + e_1 \\ ss4 = 1 \times i + 1 \times s + e_1 \end{cases}$$

この方程式からlavaanパッケージのgrowth関数を用いてLGMのパス係数を求めた。その結果を図5に示す。また、分析に用いたRスクリプトを図6に示す。

図5のLGMの結果を見ると、*s*に向かうパス係数のうち *hrss3* からのものは負の値である。これは睡眠レベル3における平均心拍数が小さいほど睡眠を深くする効果が大きい

表1 抽出データの観測変数の睡眠レベルごとのスコア

day	ss1	ss2	ss3	ss4	hrss1	hrss2	hrss3	hrss4	rrss1	rrss2	rrss3	rrss4	hmss1	hmss2	hmss3	hmss4	ilss1	ilss2	ilss3	ilss4
1	1	1	5	10	60.6	52.1	53.3	57.6	18.4	17.7	19.4	18.7	68	67	67	67.2	0	0	0	0
2	1	0	4	13	51.4	51.4	56.7	61.4	18.9	18.9	18.5	19.6	56	56	58.3	57.4	0	0	0	0
3	2	8	6	2	59.7	44.4	58.3	28.5	16.4	14	18.7	9.3	63	63.5	62.3	62	0	0	0	0
4	2	0	4	12	66.4	67.6	63.4	62.5	18.9	18.5	19.7	19.2	62.5	63	62.5	62.5	20	20	20	20
5	2	0	1	15	50.9	48.5	54.8	55.1	19.3	18.6	18.1	19.3	70	70	70	68.9	0	0	0	0
6	1	1	4	12	65	58.5	61.9	59	19.1	19	19.2	19.5	65	66	64.5	64.7	2	1	0.5	0.8
7	1	2	4	11	57.5	53.1	58.1	53.8	19.1	18.4	18.7	19.5	69	68.5	68.3	68.1	0	0	0	0
8	2	2	7	7	57.8	52.1	57.1	57.9	19.1	20.2	18.8	18.3	73	72	70.9	71.1	18	29.5	30	32.3
9	2	2	10	4	58.7	54.5	60.7	56.2	16.3	9.9	17.7	19.7	62	63.5	63.5	65	0	0	0.7	0
10	1	0	6	11	60.4	60.4	55.1	59.4	21.1	21.1	20.2	18.4	66	66	66.5	66.3	0	0	0	0

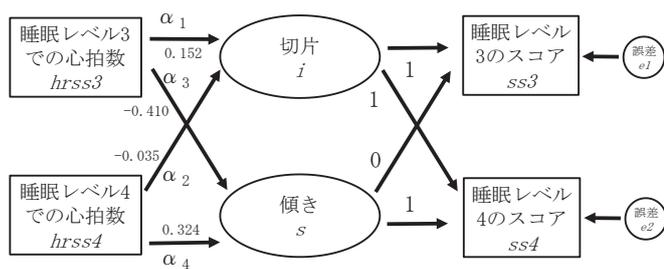


図5 入眠から90分のLGMのパス図

```

# lavaan パッケージの growth 関数を用いて、LGM 分析を行う
# 入眠から 90 分の時系列データを加工したものが
# "data.dat" に格納されている
# 睡眠レベルは 3 と 4 のみ、また観測変数は hr のみとする

library(lavaan)

# model : 図 5 の LGM モデルを lavaan の文法で記述
model <- '
# intercept and slope with fixed coefficients
i =~ 1*ss3 + 1*ss4
s =~ 0*ss3 + 1*ss4
# regressions
i =~ hrss3 + hrss4
s =~ hrss3 + hrss4

fit <- lavaan::growth(model, data=data.dat)
summary(fit, standardized=T)

```

図6 LGMのlavaanスクリプト

ことを表している。逆に $hrss4$ からのものは正の値であるため、睡眠レベル4における平均心拍数が大きいほど睡眠を深くする効果が大きいことを表している。つまり、睡眠レベル3において平均心拍数が小さくなることで睡眠レベル4の時間が長くなり、レベル4に移った後も平均心拍数が小さい場合そのままの深い睡眠を維持するということになる。

5. まとめ

睡眠の質を分析するために入眠から90分間の時系列データを抽出して、構造方程式モデリングによる因果関係および潜在成長モデルによる睡眠を深くする要因について分析を行うことができた。SEMの分析では、1日中観測した時系列データを用いた場合には影響が大きさと認められなかった観測変数である呼吸数と湿度が睡眠レベルに影響を及ぼすことが分かった。また、LGMの分析では、睡眠レベルがやや深い時の心拍数を小さくすると、より深い睡眠になることが分かった。

今回の分析では前回に引き続き lavaan パッケージを用いた。構造方程式を記述する文法が容易であり、潜在成長モデルも同じ文法で記述できるため、lavaan パッケージを用いたRが因果分析の場において広く普及することを期待している。

今回の潜在成長モデルでの分析に用いることができたデータ数はわずかに10点でしかなかったため、観測変数によっては解析の際に収束しないなどのエラーが出た。そのため、ベイズ統計を用いたマルコフ連鎖モンテカルロ法(MCMC)などを潜在成長モデルに組み込むことで、少ないデータ数でもパス係数を求められるようにし、睡眠レベルが深くなっていく過程を明らかにしたい。また、潜在成長モデルの特徴として個体差の比較ができることがあげられるので、それを睡眠レベルを深くする要因に適用し、深くしていく個人差についても明らかにしていきたい。

(令和2年9月25日受付)

(令和2年12月7日受理)

参考文献

- (1) 総務省:「平成29年版情報通信白書」, pp.52-62 (2018)
- (2) 水野欽司:「多変量データ解析講義」, pp.61-69, 朝倉書店 (1996)
- (3) 小島隆矢, 山本将史:「Excelで学ぶ 共分散構造分析とグラフィカルモデリング」, pp.89-179, オーム社 (2013).
- (4) 豊田秀樹:「共分散構造分析 [R 偏] —構造方程式モデリング—」, pp.18-195, 東京図書 (2014).
- (5) 豊田秀樹, 前田忠彦, 柳井晴夫:「原因をさぐる統計学 —共分散構造分析入門—」, pp.99-132, 講談社 (1992).
- (6) Yves Rosseel: “The lavaan tutorial”, pp.8-29, Ghent University(2019).
- (7) 厚生労働省:「平成30年版厚生労働白書—障害や病気などと向き合い、全ての人が活躍できる社会に—」, pp.369-398 (2019)
- (8) 大石信弘, 山本直樹, 石田明男, 村上純:「睡眠見守りセンサーデータの構造方程式モデリングによる因果分析」, 熊本高等専門学校研究紀要, 第11号, pp.83-86(2019).
- (9) N. Oishi, N. Yamamoto, A. Ishida, and J. Murakami, “A Causal Analysis by Structural Equation Modeling of Sleep Monitoring Sensor Data,” IJEEE, vol. 8, pp. 59-62, 2020.
- (10) “The R Project for Statistical Computing”, <https://www.r-project.org/>, Retrieved Sep. 25, 2020.
- (11) 山田剛史, 杉澤武俊, 村井純一郎:「Rによるやさしい統計学」, pp.309-319, オーム社 (2008).
- (12) 石田明男, 山本直樹, 大石信弘, 村上純:「多次元データ分解の手法を用いた立体パズルの解法(その6)」, 初等数学, 第88号, pp.28-32(2020).
- (13) 「まもるへの」, <https://mamoruno.miel.care/>, Retrieved Sep. 25, 2020.