
和歌山県における健康寿命の延伸 「健康長寿日本一わかやま」を目指して

和歌山県データ利活用推進センター 滋賀大学データサイエンス教育研究
センター
<https://www.pref.wakayama.lg.jp/prefg/020100/data/center.html>
www.ds.shiga-u.ac.jp

目次

第1章 1 はじめに	2
第2章 2 方法 (Method)	3
2.1 データ	3
第3章 結果	4
3.1 分析に用いる説明変数	4
3.2 線形回帰モデルからの探索	4
3.3 因子分析	4
3.4 線形回帰分析	7
3.5 一般化線形モデル: Gamma dist	7
3.6 一般化線形モデル: logit model.	7
3.7 ベイズ推定	7
3.8 aaa	8
第4章 4 考察 (Discussion)	11
第5章 about glm.	12

第 1 章 1 はじめに

- 研究背景
- 現状（平均寿命と健康寿命の格差の問題等）
- 健康寿命に関しての先行研究の紹介等

第 2 章 2 方法 (Method)

2.1 データ

解析に用いるデータは公的データ：全て 168 変数、その中、平均寿命、健康寿命は目的変数、県単位で集計された公的データを利用する。

変数中、一部の変数は性別の区別のない変数がある。

例えば、??? 変数は男性の??? と女性の??? であるが、??? 変数は県のデータとなる。

データソースの説明

滋賀県の研究を基に、更新されたデータを和歌山県が収集して滋賀大学に提出した旨 変数の説明「詳は添付資料」

分析分析手法等

R version 4.0.4

第 3 章 結果

3.1 分析に用いる説明変数

今回の分析では平均寿命及び健康寿命を目的変数とし、目的変数と統計的な関連性が示した以下の変数を説明変数として採用する。男性の説明変数は 18 個、女性は 10 個の説明変数である。

f_var	m_var
1 受療率_外来_脳血管疾患_2017	受療率_入院_心疾患_2017
2 人口・世帯_老年人口割合 2020	自然環境_年平均気温
3 人口・世帯_生産年齢人口割合 2020	健康・医療_保健師数 (人口 10 万人当たり)
4 自然環境_年平均気温	家計_貯蓄現在高
5 労働_完全失業率	人口・世帯_高齢単身世帯の割合
6 居住_都市公園数 (可住地面積 100km ² 当たり)	悪性新生物 (大腸)_年齢調整死亡率 2015
7 高血圧疾患_外来 2014 年	自己啓発・訓練_パソコンなどの情報処理
8 悪性新生物 (大腸)_年齢調整死亡率 2015	一定のバリアフリー化率_2018
9 ボランティア総行動率 - 総数	自己啓発・訓練 - 芸術・文化
10 受療率_外来_心疾患_2017	自己啓発・訓練 - 英語以外の外国語
11 居住_一戸建住宅比率	
12 75 歳未満調整死亡率_悪政新生物_2019	
13 診療所数_2019	
14 バリアフリー_手すりがある 2018	
15 循環器専門医数_2020	
16 家計_スマートフォン所有数量 (千世帯当たり)	
17 ボランティア総行動率 - 高齢者を対象とした活動	

Table. 3.1 $\beta_0 X_1 + \beta_0 X_2$ 寿命

3.2 線形回帰モデルからの探索

var_name_Jpn	estimate	statistic	p.value
1 受療率_外来_脳血管疾患_2017	111.81	17.79	0.00
2 人口・世帯_老年人口割合 2020	-0.00	-2.80	0.01
3 人口・世帯_生産年齢人口割合 2020	-0.16	-2.40	0.02
4 自然環境_年平均気温	-0.25	-3.26	0.00
5 労働_完全失業率	-0.05	-2.00	0.05
6 居住_都市公園数 (可住地面積 100km ² 当たり)	-0.25	-2.97	0.01
7 高血圧疾患_外来 2014 年	0.00	1.79	0.08
8 悪性新生物 (大腸)_年齢調整死亡率 2015	-0.02	-1.39	0.18
9 ボランティア総行動率 - 総数	-0.09	-3.22	0.00
10 受療率_外来_心疾患_2017	-0.04	-1.71	0.10
11 居住_一戸建住宅比率	-0.00	-1.86	0.07
12 75 歳未満調整死亡率_悪政新生物_2019	-0.03	-3.01	0.01
13 診療所数_2019	-0.00	-0.11	0.91
14 バリアフリー_手すりがある 2018	-0.00	-0.22	0.82
15 循環器専門医数_2020	0.00	0.20	0.84
16 家計_スマートフォン所有数量 (千世帯当たり)	-0.00	-0.20	0.84
17 ボランティア総行動率 - 高齢者を対象とした活動	0.00	1.16	0.25
	0.14	1.95	0.06

Table. 3.2 女性の線形回帰 (平均寿命)

var_name_Jpn	estimate	statistic	p.value
1 受療率_入院_心疾患_2017	79.56	58.01	0.00
2 自然環境_年平均気温	0.00	1.40	0.17
3 健康・医療_保健師数 (人口 10 万人当たり)	0.06	3.17	0.00
4 家計_貯蓄現在高	0.01	2.77	0.01
5 人口・世帯_高齢単身世帯の割合	0.00	2.37	0.02
6 悪性新生物 (大腸)_年齢調整死亡率 2015	-0.10	-3.83	0.00
7 自己啓発・訓練_パソコンなどの情報処理	-0.10	-5.43	0.00
8 一定のバリアフリー化率_2018	0.08	2.66	0.01
9 自己啓発・訓練 - 芸術・文化	0.01	0.80	0.43
10 自己啓発・訓練 - 英語以外の外国語	0.01	0.15	0.88
11	0.10	1.70	0.10

Table. 3.3 男性の線形回帰 (平均寿命)

var_name_Jpn	estimate	statistic	p.value
1 受療率_外来_脳血管疾患_2017	52.26	5.02	0.00
2 人口・世帯_老年人口割合 2020	-0.01	-2.76	0.01
3 人口・世帯_生産年齢人口割合 2020	0.25	2.25	0.03
4 自然環境_年平均気温	0.35	2.75	0.01
5 労働_完全失業率	0.29	7.03	0.00
6 居住_都市公園数 (可住地面積 100km ² 当たり)	-0.29	-2.12	0.04
7 高血圧疾患_外来 2014 年	-0.00	-0.08	0.94
8 悪性新生物 (大腸)_年齢調整死亡率 2015	0.14	4.94	0.00
9 ボランティア総行動率 - 総数	0.17	3.61	0.00
10 受療率_外来_心疾患_2017	0.02	0.44	0.66
11 居住_一戸建住宅比率	-0.00	-2.04	0.05
12 75 歳未満調整死亡率_悪政新生物_2019	0.07	4.51	0.00
13 診療所数_2019	-0.10	-4.41	0.00
14 バリアフリー_手すりがある 2018	-0.03	-4.60	0.00
15 循環器専門医数_2020	-0.00	-4.62	0.00
16 家計_スマートフォン所有数量 (千世帯当たり)	0.00	4.37	0.00
17 ボランティア総行動率 - 高齢者を対象とした活動	-0.01	-6.02	0.00
18	-0.19	-1.63	0.11

Table. 3.4 女性の線形回帰 (健康寿命)

3.3 因子分析

$$X_1 = l_{11}F_1 + l_{12}F_2 + \varepsilon \quad (3.1)$$

$$\vdots \quad (3.2)$$

$$X_p = l_{p1}F_1 + l_{p2}F_2 + \varepsilon$$

$$x = Lf + \varepsilon \quad (3.3)$$

$$X_1 = l_{11}F_1 + l_{12}F_2 + \varepsilon$$

$$X_1 = l_{11}F_1 + l_{12}F_2 + \varepsilon$$

因子分析とは因子分析結果の読み方因子分析とは

因子分析とは、多変量データに潜む共通因子を探り出すための手法として、消費者を理解するためによく使われる多変量解析手法です。因子分析は「知能」という潜在的な概念を研究する中

var_name_Jpn	estimate	statistic	p.value
1	75.96	28.80	0.00
2 受療率_入院_心疾患_2017	0.00	0.04	0.98
3 自然環境_年平均気温	0.00	0.08	0.93
4 健康_医療_保健師数(人口 10 万人当たり)	0.00	0.62	0.54
5 家計_貯蓄現在高	0.00	1.59	0.12
6 人口_世帯_高齢単身世帯の割合	-0.16	-3.08	0.00
7 悪性新生物(大腸)_年齢調整死亡率_2015	-0.06	-1.71	0.09
8 自己啓発_訓練_パソコンなどの情報処理	-0.02	-0.26	0.80
9 一定のバリアフリー化率_2018	-0.03	-0.80	0.43
10 自己啓発_訓練_芸術・文化	-0.12	-1.47	0.15
11 自己啓発_訓練_英語以外の外国語	0.20	1.82	0.07

Table. 3.5 男性の線形回帰(健康寿命)

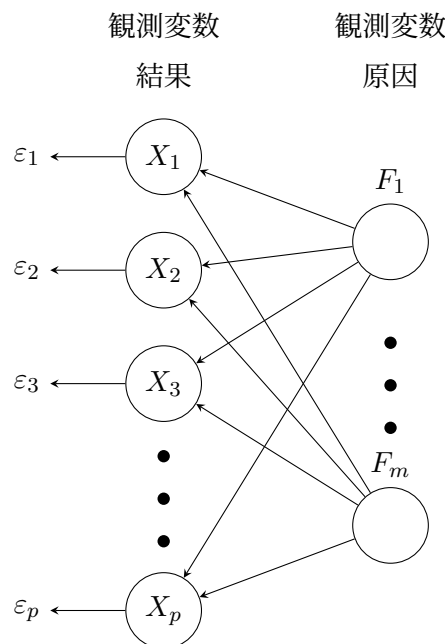


Fig. 3.1 因子分析の概念図

から生まれた分析手法です (Spearman, 1904)。これは目には見えず、直接測ることができない「知能」というものが存在し、それが具体的な知能テストや試験などの結果として現れる（観測できる）という考え方を元としています。様々な事象（観測変数）を手がかりにして、潜在的に存在する概念（潜在変数）を推定するしかない、ということになります。このような考え方の視点を広げてみると、消費者の意識、態度は全て潜在的な概念と考えることができ、具体的にモデルを描くと以下の図 1 のようになります。楕円が「悲しい気持ち」という潜在変数で、共通因子といいます。そして四角の $X_1 \sim X_3$ までの観測変数が、私達が見ることができる実際の現象とすることができます。もちろん他の観測変数

でも、「悲しい」という心が出する可能性はあるでしょう。消費者の意識、態度を潜在的な概念と考へたモデル図図 1 また、この図にある $e_1 \sim e_3$ は独自因子といい、それぞれの観測変数に固有の情報を表しています。つまり、それぞれの観測変数が以下のような形で分解されることになり

ります。観測変数 = 共通因子 + 独自因子

$$\text{観測変数} = \text{共通因子} + \text{独自因子}$$
 の統計的な概念を図示すると、図 2 のようになります。観測変数 z を共通因子空間 $S(F)$ で説明するとして、 z のなかで $S(F)$ で説明できる部分は h となり、できなかった部分が e となります。ここで、 e と h は直行するので、 z, h, e で直角三角形ができます。このことが、 z を h と e に分解する、という意味になります。通常、分析をするにあたって関心を持つのは共通因子の方になりますので、単純に「因子」といったときは独自因子ではなく、共通因子を指します。ここで着目していただきたいのが、観測変数の全てを説明するモデルを作るのではなく、他の観測変数の中から共通する部分だけを共通因子空間によって説明し、それだけでは説明できない固有の要素を独自変数として残す、という点です。つまり、より小さな変数で人々の意識を理解する試みが行われており、「枝葉を捨てて、エッセンスを見る」というのが根本的な考え方となっています。この時、それぞれの観測変数に固有の枝が独自因子 e と言うことができます。（朝野熙彦「マーケティング・リサーチ」講談社の第 2 章から引用）観測変数 = 共通因子 + 独自因子の統計的な概念図 2 因子分析結果の読み方

因子分析で得られる指標

因子負荷量

直交解を求めた場合に限りませんが各変数と各因子の相関を表します。その場合は因子負荷量は、相関係数なので 0 から ± 1 の値をとります。バリマックス回転が直交解の方法としてよく利

用されます。しかし、次の項の共通性の推定により、独自因子の情報は共通因子空間から除かれていることに注意してください。通常、この因子負荷量が高い変数を考慮して、因子の名前をつけます。共通性

各変数が因子空間で表される分散を表しています。0 から 1 の値をとります。これも直交解を求めた場合に限りませんが共通性は、各因子負荷量の 2 乗和となります。寄与率

ある因子がどの程度の説明力を持っているかを割合を表します。独自因子の割合＝独自性です。Uniqueness と言います。因子分析で得られる指標図 3 因子得点

因子得点は、各因子ごとの各個体（対象者）のスコアを表します。因子得点が高い人は、その因子に影響されている度合いが高いといえます。下記の表は、適性検査の成績を因子分析した結果の一部です。「計算能力」、「図形処理能力」、「言語能力」、「記憶能力」という 4 つの因子が抽出され、対象者ごとの因子得点を求めたものです。因子得点から、対象者を 3 つのグループに分けることができました。因子得点図 4

バリマックス回転 因子分析における直交回転法のひとつで、もっともよく利用されてきた。回転の目的は因子の解釈を容易にすることであり、バリマックス回転 (varimax rotation) が解釈しやすい結果を与えることが多かったために、研究者や実務家に非常に頻繁に利用されてきた。実は回転方法は無数にある。素朴な疑問として「回転」とは何なのか、なぜ「回転」するのか、ということを理解したいが、それに先立って回転前（初期解）と回転後（回転解）の実例を示す。これはブランド戦略サーベイの企業イメージ 25 変数の因子分析である。どのような変化が回転前後で生じているであろうか。

初期解（回転前）の因子負荷行列

バリマックス回転後（回転解）の因子負荷行列

＜回転とは何か＞ 回転は幾何学的概念である。一方、因子負荷量は解析的概念である。データ解析では、しばしば幾何学的表現と解析的表現が、同じ文脈で混在するので、慣れていないと混乱するであろう。因子負荷行列を図的に表現すれば下図のようになる。因子は 2 個としてあるので、因子を縦軸と横軸にすれば平面を描くことができる（3 因子による空間表示でもかまわない）。変数は 6 個にして色分けしてある。●は因子 1 と因子 2 の因子負荷量の値を座標値とした位置にあるが、見やすいように原点からのベクトルで表現してある。これが因子負荷行列の図的表現である。回転とは、この平面つまり座標空間で因子（軸）を「回転させる」という幾何学的イメージに準拠している。6 変数の相対的位置は変わっていない。下図の回転前後は因子（軸）ではなく、変数が回転しているように見えるが、因子の方を回転しているのである。そして、この回転は 360 度、どのように回転することもできる。無数の回転解が存在する。因子分析は座標空間だけを定めたのである。ちなみに、これをネガティブに「因子の不定性」とか、ポジティブに「回転の自由度」などという。

因子負荷行列の図的表現

＜なぜ回転するのか？＞ 解釈しやすい解を得るためである。解釈しやすいとは、どのような状態であろうか。それは単純構造の時である。単純構造という概念は Thurstone が提案したのだが、これを解析的に実現したのが Kaiser で、1958 年に "The varimax criterion for analytic rotation in factor analysis" という論文として Psychometrika に発表した。この時、バリマックス回転が成立した。

因子負荷行列の図的表現をみると、回転前は 6 変数のすべてが因子 1 と関係している。回転後では最初の 3 変数は縦軸と、後の 3 変数は横軸と強い関係を持つように分離している。別の見方をすると、6 変数が 3 変数ごとにグループ化

された。関係の強さは幾何学的には因子の軸と各変数ベクトルとの角度の小ささである。回転後は、3 変数はある因子と強く関係し、他の因子とは弱い関係になった。回転によって単純構造に接近したのである。Kaiser は単純構造を得るには、因子負荷行列の要素の分散を（標準化したうえで）最大化すれば実現できる、というアイデアを得た。大きい負荷はより大きく、小さい負荷はより小さくなるような規準に向かって回転させるので、分散（variance）の最大化（max）、すなわち varimax という名前にしたのである。

最初に示した「ブランド戦略サーベイ」の初期解（回転前）は、因子 1 にほぼすべての変数は高い負荷を持つ。しかし、バリマックス回転後は単純構造に向かって、因子と変数とのコントラストが強化されていることが分かる。これで 4 因子についての解釈は容易な方向に改善されたのである。＜因子負荷量の計算＞ 因子分析の数理的な目標は、因子負荷量の推定である。しかし因子分析モデルは強い制約条件をもつ統計モデルである。そのため、まず計算しやすいような解を最初に求めている。それが「初期解」という名前の意味である。初期解は第 1 因子の分散が最大になるように計算し、次に第 2 因子の分散、という順番に解を求めているので、因子 1 の負荷量がすべて大きかったのである。因子分析モデルの制約条件が多い理由は、因子が観測されていない潜在変数であるためである。方程式の本数よりも未知数の方が多いと、一意に解を求めることができないので制約条件を設定して計算している。回転解とはそのような制約のあとに「有用な」解を求めていくことである。なお、回転しても因子分析モデルの共通性や独自性、モデルの適合度などは変化しない。

	var_name_Jpn	F1	F2
1	受療率_外来_脳血管疾患_2017	0.28	0.64
2	人口・世帯_老年人口割合_2020	0.01	0.86
3	人口・世帯_生産年齢人口割合_2020	0.13	-0.92
4	自然環境_年平均気温	-0.25	-0.34
5	労働_完全失業率	0.58	-0.01
6	居住_都市公園数(可住地面積 100km ² 当たり)	0.17	-0.85
7	高血圧疾患_外来_2014 年	0.36	-0.84
8	悪性新生物(大腸)_年齢調整死亡率_2015	0.56	-0.07
9	ボランティア総行動率-総数	-0.80	0.08
10	受療率_外来_心疾患_2017	-0.02	0.47
11	居住_一戸建住宅比率	-0.31	0.82
12	75 歳未満調整死亡率_悪政新生物_2019	0.80	0.33
13	診療所数_2019	-0.19	0.08
14	バリアフリー_手すりがある_2018	0.33	-0.86
15	循環器専門医数_2020	0.27	-0.84
16	家計_スマートフォン所有数量(千世帯当たり)	-0.32	-0.73
17	ボランティア総行動率-高齢者を対象とした活動	-0.77	0.24

Table. 3.6 女性の FA

	var_name_Jpn	F1	F2
1	受療率_入院_心疾患_2017	0.02	-0.61
2	自然環境_年平均気温	0.50	0.16
3	健康・医療_保健師数(人口 10 万人当たり)	-0.36	-0.76
4	家計_貯蓄現在高	-0.43	0.65
5	人口・世帯_高齢単身世帯の割合	0.20	-0.51
6	悪性新生物(大腸)_年齢調整死亡率_2015	0.65	-0.11
7	自己啓発・訓練-パソコンなどの情報処理	0.08	0.90
8	一定のバリアフリー化率_2018	-0.93	0.07
9	自己啓発・訓練-芸術・文化	-0.11	0.87
10	自己啓発・訓練-英語以外の外国語	0.17	0.82

Table. 3.7 男性の FA

3.4 線形回帰分析

3.5 一般化線形モデル: Gamma dist

3.6 一般化線形モデル: logit model

3.7 ベイズ推定

bayesian multilevel モデル

<https://www.stata.com/features/overview/bayesian-multilevel-models/>

何のことか？

multilevel モデルは、グループ固有の効果を組み込んだ回帰モデル。

グループは、病院、病院内にネストされた医師、病院内にネストされた医師内にネストされた患者など、グループ固有の効果は、いくつかの事前分布、(通常は正規分布に従って)、グループ間でランダムに変化すると想定。

さまざまなレベルの階層を表す場合があ。この仮定により、multilevel モデルはベイズ分析の自然な候補にな。bayes multilevel モデルはさらに、回帰係数や分散成分（グループ固有の効果の分散）などの他のモデルパラメーターもラン

term	estimate	std.error	statistic	p.value	term	estimate	std.error	statistic	p.value
1 (Intercept)	87.02	0.04	1988.52	0.00	1 (Intercept)	74.94	0.09	792.01	0.00
2 d.f._FA\$OBS.rotate1	-0.26	0.04	-5.91	0.00	2 d.f._FA\$OBS.rotate1	-0.13	0.10	-1.35	0.18
3 d.f._FA\$OBS.rotate2	-0.09	0.04	-2.13	0.04	3 d.f._FA\$OBS.rotate2	0.07	0.10	0.73	0.47

Table. 3.8 女性の回帰 withFA(平均寿命)

Table. 3.10 女性の回帰 withFA(健康寿命)

term	estimate	std.error	statistic	p.value	term	estimate	std.error	statistic	p.value
1 (Intercept)	80.65	0.06	1393.83	0.00	1 (Intercept)	72.06	0.07	1000.10	0.00
2 d_m_FA\$OBS.rotate1	-0.25	0.06	-4.29	0.00	2 d_m_FA\$OBS.rotate1	-0.10	0.07	-1.41	0.17
3 d_m_FA\$OBS.rotate2	0.34	0.06	5.89	0.00	3 d_m_FA\$OBS.rotate2	0.15	0.07	2.11	0.04

Table. 3.9 男性の回帰 withFA(平均寿命)

Table. 3.11 男性の回帰 withFA(健康寿命)

ダムであると想定。

bayesmultilevel モデルを使用する理由 bayes 分析の標準的な理由に加えて、bayesmultilevel モデリングは、グループの数が少ない場合、または多くの階層レベルが存在する場合によく使用されます。

逸脱度情報量基準 (DIC) などのベイズ情報量基準も、multilevel モデルの比較によく使用されます。グループの比較が主な関心事である場合、bayesmultilevel モデリングは、グループ固有の効果の分布全体を提供できます。

multilevel コマンドの前にベイズを付けるだけ。

bayesmultilevel モデルを Stata に適合させることができ、これを簡単に行うことができます。ベイズ：混合 y x1 x2 — id：もちろん、「簡単に」と言うときは、モデルの定式化ではなく、モデルの仕様を指します。他のモデリングタスクと同様に、bayesmultilevel モデリングでは慎重に検討する必要があります。

連続、打ち切り、バイナリ、序数、カウント、GLM、および生存の結果がサポートされています。

サポートされている multilevel コマンドの完全なリストを参照してください。

複数レベルの階層、ネストおよびクロスされたランダム効果、ランダム切片と係数、ランダム効果共分散構造など、

すべての multilevel 機能を利用できます。

[BAYES] bayesmh コマンドによって提供され

るすべてのベイズ機能は、

multilevel コマンドでベイズプレフィックスを使用する場合にサポートされます。

bayesmultilevel モデリングの新機能もご覧ください。

$$p(\theta|y) \propto p(\theta)p(y|\theta) \quad (3.4)$$

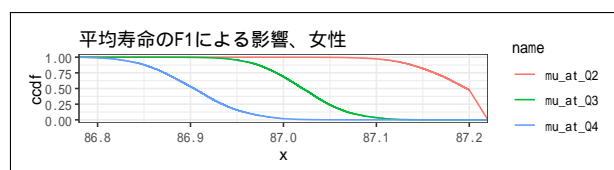


Fig. 3.2 Bayes_LE.f.ccdf.F1

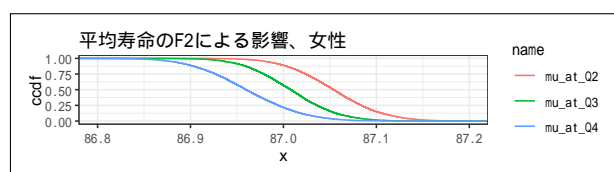


Fig. 3.3 Bayes_LE.f.ccdf.F2

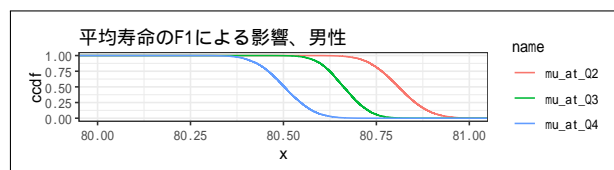


Fig. 3.4 Bayes_LE.m.ccdf.F1

3.8 aaa

	term	estimate	std.error	statistic	p.value
1	(Intercept)	4.466	0.001	8876.243	0.000
2	F1	-0.003	0.001	-5.913	0.000
3	F2	-0.001	0.001	-2.138	0.038

Table. 3.12 女性の一般化線形モデル
withFA(平均寿命)

	term	estimate	std.error	statistic	p.value
1	(Intercept)	4.390	0.001	6101.247	0.000
2	F1	-0.003	0.001	-4.279	0.000
3	F2	0.004	0.001	5.881	0.000

Table. 3.13 男性の一般化線形モデル
withFA(平均寿命)

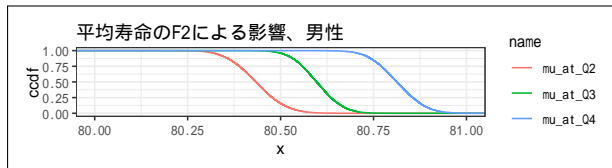


Fig. 3.5 Bayes.LE_m.ccdf.F2

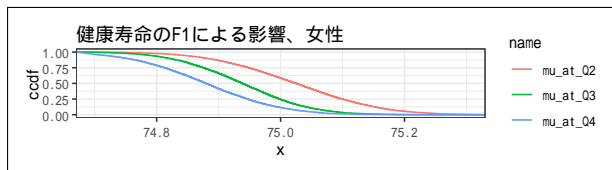


Fig. 3.6 Bayes.HLE_f.ccdf.F1

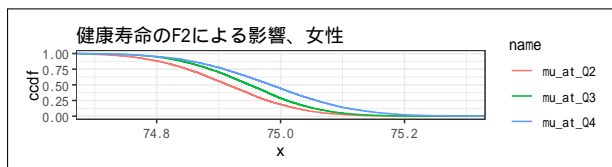


Fig. 3.7 Bayes.HLE_f.ccdf.F2

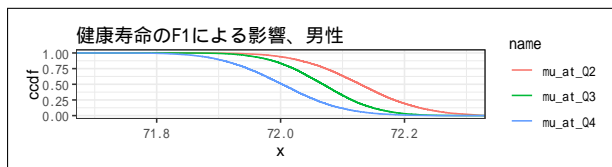


Fig. 3.8 Bayes.HLE_m.ccdf.F1

	term	estimate	std.error	statistic	p.value
1	(Intercept)	4.317	0.001	3418.628	0.000
2	F1	-0.002	0.001	-1.354	0.183
3	F2	0.001	0.001	0.731	0.469

Table. 3.14 女性の一般化線形モデル
withFA(健康寿命)

	term	estimate	std.error	statistic	p.value
1	(Intercept)	4.278	0.001	4278.460	0.000
2	F1	-0.001	0.001	-1.411	0.165
3	F2	0.002	0.001	2.110	0.041

Table. 3.15 男性の一般化線形モデル
withFA(健康寿命)

	term	estimate	std.error	statistic	p.value
1	(Intercept)	-0.185	0.350	-0.528	0.598
2	d.f._FA\$OBS.rotate1	-1.447	0.520	-2.784	0.005
3	d.f._FA\$OBS.rotate2	-0.575	0.371	-1.548	0.122

Table. 3.16 女性の一般化線形モデル
(logit)withFA(平均寿命)

	term	estimate	std.error	statistic	p.value
1	(Intercept)	-0.335	0.455	-0.737	0.461
2	d.m._FA\$OBS.rotate1	-1.778	0.658	-2.703	0.007
3	d.m._FA\$OBS.rotate2	2.566	0.798	3.213	0.001

Table. 3.17 男性の一般化線形モデル
(logit)withFA(平均寿命)

	term	estimate	std.error	statistic	p.value
1	(Intercept)	-0.098	0.322	-0.304	0.761
2	d.m._FA\$OBS.rotate1	-0.782	0.412	-1.896	0.058
3	d.m._FA\$OBS.rotate2	0.658	0.349	1.888	0.059

Table. 3.18 男性の一般化線形モデル
(logit)withFA(健康寿命)

	mean	sd	5.5%	94.5%
beta0	87.022	0.042	86.954	87.090
beta1	-0.261	0.043	-0.330	-0.193
beta2	-0.094	0.043	-0.163	-0.026
sigma	0.290	0.030	0.242	0.338

Table. 3.19 女性のBayes(平均寿命)

	mean	sd	5.5%	94.5%
beta0	80.652	0.056	80.563	80.742
beta1	-0.251	0.057	-0.342	-0.161
beta2	0.345	0.057	0.254	0.435
sigma	0.384	0.040	0.321	0.447

Table. 3.20 男性のBayes(平均寿命)

	mean	sd	5.5%	94.5%
beta0	74.940	0.092	74.794	75.086
beta1	-0.129	0.093	-0.277	0.018
beta2	0.070	0.093	-0.078	0.218
sigma	0.628	0.065	0.524	0.731

Table. 3.21 女性のBayes(健康寿命)

	mean	sd	5.5%	94.5%
beta0	72.064	0.070	71.953	72.175
beta1	-0.103	0.070	-0.215	0.010
beta2	0.153	0.070	0.041	0.266
sigma	0.478	0.049	0.399	0.557

Table. 3.22 男性のBayes(健康寿命)

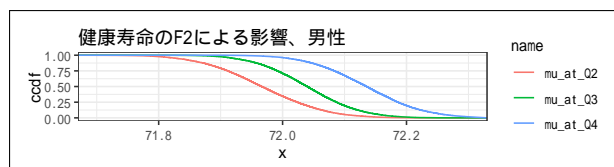
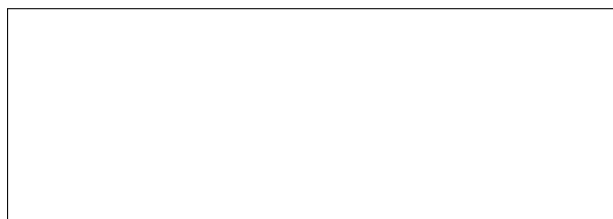


Fig. 3.9 Bayes_HLE_m.ccdf.F2

Fig. 3.10 DALY の算出概念図, 出典：
wikipedia.org, 障害調整生命年

第 4 章 4 考察 (Discussion)

諸々の結果を踏まえた考察先行研究と比較
今後の展望・提言
研究の限界ヘルスケア産業についても触れて
ください

5 参考資料 (Reference) 引用文献、参考文献

6 添付 (Appendix)

第 5 章 about glm

統計では、一般化線形モデル (GLM) は、通常の線形回帰を柔軟に一般化したものであり、応答変数に正規分布以外の誤差分布を持たせることができます。GLM は、線形モデルをリンク関数を介して応答変数に関連付けることを許可し、各測定値の分散の大きさをその予測値の関数にすることにより、線形回帰を一般化します。

一般化線形モデルは、線形回帰、ロジスティック回帰、ポアソン回帰など、他のさまざまな統計モデルを統合する方法として、ジョンネルダーとロバートウェダーバーンによって策定されました。[1] 彼らは、モデルパラメータの最尤推定のために繰り返し再重み付けされた最小二乗法を提案しました。最尤推定は依然として一般的であり、多くの統計計算パッケージのデフォルトの方法です。ベイジアンアプローチや分散安定化応答への最小二乗適合など、他のアプローチが開発されています。

一般化線形モデルは、線形回帰、ロジスティック回帰、ポアソン回帰など、他のさまざまな統計モデルを統合する方法として、ジョンネルダーとロバートウェダーバーンによって策定されました。[1] 彼らは、モデルパラメータの最尤推定のために繰り返し再重み付けされた最小二乗法を提案しました。最尤推定は依然として一般的であり、多くの統計計算パッケージのデフォルトの方法です。ベイジアンアプローチや分散安定化応答への最小二乗適合など、他のアプローチが開発されています。

内容

1 直感 2 概要 3 モデルコンポーネント 3.1 確率分布 3.2 線形予測子 3.3 リンク機能 4 フィッティング 4.1 最尤法 4.2 ベイズ法 5 例 5.1 一般線形モ

デル 5.2 線形回帰 5.3 バイナリデータ 5.3.1 ロジックリンク機能 5.3.2 逆累積分布関数の一般的な選択肢としてのプロビットリンク関数 5.3.3 補完的な両対数 (cloglog) 5.3.4 ID リンク 5.3.5 分散関数 5.4 多項ロジット 5.4.1 注文された応答 5.4.2 順序付けられていない応答 5.5 カウントデータ 6 拡張機能 6.1 相関データまたはクラスター化されたデータ 6.2 一般化された加法モデル 7 関連項目 8 参考文献 8.1 引用 8.2 参考文献 9 さらに読む 10 外部リンク

通常の線形回帰は、特定の未知の量 (応答変数、確率変数) の期待値を、一連の観測値 (予測子) の線形結合として予測します。これは、予測子の一定の変化が応答変数の一定の変化につながることを意味します (つまり、線形応答モデル)。これは、応答変数がどちらかの方向に無期限に、またはより一般的には予測変数の変動と比較して比較的少量しか変動しない任意の量に対して、適切な近似で変動する可能性がある場合に適しています。人間の身長。

ただし、これらの仮定は、一部のタイプの応答変数には不適切です。たとえば、応答変数が常に正であり、広範囲にわたって変化すると予想される場合、一定の入力変化は、絶えず変化するのではなく、幾何学的に (つまり指数関数的に) 変化する出力変化につながります。例として、線形予測モデルがいくつかのデータ (おそらく主に大きなビーチから引き出されたもの) から、10 度の気温低下によりビーチを訪れる人が 1,000 人少なくなることを学習するとします。このモデルは、さまざまなサイズのビーチで一般化する可能性はほとんどありません。より具体的には、問題は、モデルを使用して、定期的に 50 人のビーチファンを受け入れるビーチの気温が 10 度低下する新しい出席を予測すると、不可能な出席値 -950 を予測することです。論理的には、より現実的なモデルでは、代わりに一定の割合でビーチへの出席が増えると予測されます

(たとえば、10 度の増加はビーチへの出席の倍増につながり、10 度の減少は出席の半減につながります)。このようなモデルは、指数応答モデル（または、応答の対数が線形に変化すると予測されるため、対数線形モデル）と呼ばれます。

同様に、はい/いいえの選択を行う確率を予測するモデル（ベルヌーイ変数）は、確率が両端で制限されているため（0 から 1 の間でなければならない）、線形応答モデルとしてはさらに適していません。たとえば、特定の人がビーチに行く可能性を温度の関数として予測するモデルを想像してみてください。合理的なモデルでは、たとえば、10 度の変化により、人がビーチに行く可能性が 2 倍になると予測される場合があります。しかし、確率の観点から「2 倍の可能性」とはどういう意味ですか？文字通り、確率値を 2 倍にすることを意味することはできません（たとえば、50 % が 100 % になる、75 % が 150 % になるなど）。むしろ、倍増しているのはオッズです。2:1 オッズ、4:1 オッズ、8:1 オッズなどです。このようなモデルは、対数オッズまたはロジスティックモデルです。

一般化線形モデルは、（単なる正規分布ではなく）任意の分布を持つ応答変数を許可し、応答変数の任意の関数（リンク関数）を（仮定するのではなく）予測子とともに線形に変化させることにより、これらすべての状況をカバーします。応答自体は線形に変化する必要があります。たとえば、上記の予測されたビーチ参加者数のケースは、通常、ポアソン分布とログリンクでモデル化されますが、ビーチ参加者の予測確率のケースは、通常、ベルヌーイ分布（または正確に当てはめて二項分布）でモデル化されます。問題の言い回し）と log-odds（または logit）リンク関数。

概要概要

一般化線形モデル（GLM）では、従属変数の各結果 Y は、指数型分布族の特定の分布、とりわけ正規分布、二項分布、ポアソン分布、ガン

マ分布を含む大きなクラスの確率分布から生成されると想定されます。分布の平均 μ は、次のように独立変数 X に依存します。