

5カ国におけるうつ病症状数の要因分析と予測モデル提案 Proposal for the Analysis and Prediction Model to Increasing Depressive Disorders at 5 Countries

1741087 中野 優真

Yuma NAKANO

指導教員 秋葉 知昭

Recently, depressive disorders are increasing in the world. In this study, I proposed results of factor analysis and prediction model to increase depressive disorders at 5 countries. From these results, I evaluated and compared the accuracy general methods and prediction model.

1. 緒言

近年の日本の精神疾患において、うつ病患者数は年々増加している。平成8年の43.3万人に対し、平成20年では104.1万人と12年間で2.4倍に増加した[1]。世界的に見ても、うつ病を含めた精神疾患の患者が増加傾向である。WHOによる2015年の世界のうつ病患者率データでは、世界の人口の4.4%となっており、世界で3億2200万人の人々がうつ病で苦しんでいる状況である。また、うつ病予備軍と言われるうつ病の一步手前である人の数は世界で2640万人となっている[2]。患者自身や周囲の人々にかかる負担は心身ともに計り知れない。

精神的な疾患によって失業や医療費、国の制度等で多額な費用がかかる。従って、うつ病などの精神疾患の増加は経済も逼迫することになる。自殺・うつ病がなくなった場合の経済的便益の推定では、2009年日本の例では、合計が2兆6782億円にも上る。うつ病だけで見れば、7754億円となっている[1]。そこで、考えられるうつ病の要因を解析し、予測式を作れば、考え付いた改善案でうつ病患者数がどう変化するか正確な予測が行える。それが、今後多くの国での精神疾患対策や医療業界での病床率の予測等に役立てられると考える。

本研究では、1990年から2010年のうつ病症状数の分析を行い、その結果から、予測式を提案する。2011年から2015年の症状数の予測を行い、実際の結果と比較する。また、うつ病症状数を相対誤差や他の一般的な手法による予測との精度を比較し、うつ病症状数予測モデルを評価する。

2. 分析対象と使用する変数

本研究では、うつ病症状数を分析するために、our world in data[3]から1990年から2015年のうつ病症状数のデータを対象とする。うつ病症状数は、うつ病、躁うつ病、抗うつ状態などの症状が出た実態に近い数である。対象国は日本、アメリカ、イギリス、中国、インドの5カ国である。

最終的に使用する変数は以下の通りである。

目的変数：うつ病症状数

説明変数：一人当たりCO₂排出量、災害死亡率、都市生活率、自殺率、人口増加率

3. 分析結果

変数間の相関を調べるため、相関分析を予備解析として行った。結果を抜粋して表1に示す。

表1 日本の相関分析の結果

変数名	GDPpercapit	internet	disasters de	healthcare	total tax re	mental or su	urban areas	suicide sha	population	burden of w
1 GDPpercapit	1.000	0.323	0.046	-0.065	-0.084	-0.262	0.076	0.222	-0.268	-0.268
2 internet	0.553	1.000	0.225	0.054	0.192	-0.411	-0.531	0.091	0.588	-0.258
3 GDPpercapit	0.323	0.225	1.000	-0.272	0.274	-0.080	0.188	0.000	0.190	-0.838
4 internet	0.046	0.225	-0.272	1.000	-0.222	-0.029	0.163	-0.212	-0.272	0.198
5 disasters de	-0.065	0.054	-0.272	-0.222	1.000	-0.198	0.205	0.000	0.183	-0.268
6 healthcare	0.262	0.192	0.054	-0.029	-0.198	1.000	-0.447	0.000	0.183	-0.268
7 total tax re	-0.262	-0.411	-0.080	-0.029	-0.198	-0.447	1.000	-0.212	-0.272	-0.268
8 mental or su	0.076	-0.531	0.188	0.163	0.205	0.000	0.447	1.000	-0.212	-0.272
9 urban areas	0.222	0.091	0.188	-0.212	0.205	0.000	-0.212	-0.272	1.000	-0.268
10 suicide sha	0.222	0.091	0.188	-0.212	0.205	0.000	-0.212	-0.272	0.222	1.000
11 population	-0.268	0.588	-0.838	-0.268	-0.268	-0.268	-0.268	-0.268	-0.268	-0.268
12 burden of w	-0.268	-0.258	-0.838	-0.268	-0.268	-0.268	-0.268	-0.268	-0.268	-0.268
13 working	-0.268	-0.258	-0.838	-0.268	-0.268	-0.268	-0.268	-0.268	-0.268	-0.268

StatWorks ver.5.82 を使い、上記の変数を日本、アメリカ、イギリス、中国、インドの5カ国で多変量解析をした。その結果の例が、表2である

表2 日本の重回帰分析の結果

変数名	重相関係数	寄与率R ²	Re ²	R** ²	残差標準偏差	分散比	P値 (上側)	偏回帰係数	標準偏回帰	トレランス	VIF
number	0.990	0.981	0.974	0.968	16967.652	118.3776	0.000	3115057.370			
2 rate	0.1327	0.013 +				9.1327	0.013 +				
3 GDPpercapit	7.7180	0.015 +				7.7180	0.015 +				
4 CO2percapit	13.6671	0.002	59976.095	0.172	0.600	13.6671	0.002	59976.095	0.172	0.600	1.668
5 population	46.5640	0.000 +				46.5640	0.000 +				
6 internet (%)	0.3027	0.591 -				0.3027	0.591 -				
7 disasters de	0.5274	0.479	24163.922	0.030	0.777	0.5274	0.479	24163.922	0.030	0.777	1.287
8 healthcare	0.4690	0.505 -				0.4690	0.505 -				
9 total tax re	9.9615	0.010 -				9.9615	0.010 -				
10 mental or su	0.0818	0.779 -				0.0818	0.779 -				
11 schooling me	7.9074	0.014 +				7.9074	0.014 +				
12 urban areas	19.2647	0.001	9994.656	0.455	0.120	19.2647	0.001	9994.656	0.455	0.120	8.307
13 suicide shar	10.6448	0.005	-93262.070	-0.173	0.459	10.6448	0.005	-93262.070	-0.173	0.459	2.181
14 population g	26.0847	0.000	-519228.827	-0.531	0.119	26.0847	0.000	-519228.827	-0.531	0.119	8.376
15 burden of de	1.0254	0.328 -				1.0254	0.328 -				
17 working	6.2789	0.025 -				6.2789	0.025 -				

寄与率が高く、P値の異常値もないため、結果として、採択した。以下より、解析結果で目立つ変数を抜粋して考察する。

まず、5カ国で変数の種類が異なるため、最終的な変数の種類は少し異なる。多くの国で標準偏回帰係数の値が大きかったのは、一人当たりCO₂排出量であった。中位所得先進国である中国、インドでは、その影響が大きく、今後の環境問題へ

の対策が必要である。一方、先進国では、その影響は2カ国と比べて小さい。企業の海外進出や、環境対策によるものだと考える。

次に、都市生活率が先進国では影響が強かった。特に日本、アメリカでの標準偏回帰係数高く、都市化による利便性の向上の反面、森林不足や、排気ガス、過密等によるストレスが大ききうつ病症状数が増加すると考察する。また、自殺率や災害死亡率は少子化や高齢化などの人口の増減によるせい、影響度は低い。

4. 予測結果と比較からの考察

多変量解析から得られた偏回帰係数により、以下の予測式を作成した。日本の例を示す。

$$S_i = a(1) + b(1)m_1 + c(1)n_1 + d(1)o_1 + e(1)p_1 + f(1)q_1 \quad (1)$$

i : 予測する国 j : 予測する年

S_1 : 5年分のうつ病症状数ベクトル

m_1 : 5年分の一人当たりCO₂排出量ベクトル

n_1 : 5年分の災害死亡率ベクトル

o_1 : 5年分の都市生活率ベクトル

p_1 : 5年分の自殺率ベクトル

q_1 : 5年分の人口増加率ベクトル

$a(1), b(1), c(1), d(1), e(1), f(1)$ はそれぞれに対する偏回帰係数の値である。 $a(1)$ は定数項を表す。

算出した予測式の結果の一部を抜粋して表3に示す。

表3 日本の予測式の精度

年	予測値	実測値	残差	相対誤差	絶対値	MAD
2011	4331998	4342819	10821	0.002498	0.002498	0.006279
2012	4385619	4368761	-16858	-0.003844	0.003844	
2013	4407492	4400455	-7037	-0.001597	0.001597	
2014	4406817	4437730	30913	0.007015	0.007015	
2015	4406945	4479393	72448	0.016439	0.016439	

MAD値で見ると高い精度だが、2015年の相対誤差の値が大きいため、今後の予測をする際は不安定な数値になると考える。また、日本だけでなく、他の国においても精度が高いが、不安定な動きであった。

得られた結果の精度をMAD値で他手法と表4にて比較する。

表4 他手法との比較

国 \ 手法	日本	アメリカ	イギリス	中国	インド
移動平均法3年	0.016907	0.018194	0.014847	0.027798	0.039407
移動平均法5年	0.020601	0.023737	0.023737	0.044883	0.048737
指数平滑法	0.008502	0.011599	0.008917	0.016443	0.026395
本研究	0.006279	0.003804	0.013684	0.012488	0.031776

他手法と比較すると、本研究では、説明変数の多い予測式ほど精度が高いことがわかる。日本、アメリカの精度がかなり良く、イギリス、中国、インドは低くなった。説明変数の目的変数に与えている影響が異なるため、国によって大きな差が出るといえる。また、対象国の5カ国全ての指数平滑法の数値が高い理由として、現在の数値の重要性が非常に高いことがわかった。また、5カ国とも高い相対誤差であった理由は目的変数が単調増加に近いため、予測がしやすかったからだと考える。説明変数の相関を考慮し、より多くの変数のデータを用意すれば、いずれの予測式も精度が向上すると考える。5カ国ともMAD値のみで見ると、高い精度で予測することができた。一方で、本研究では、相対誤差の乱高下する値から、より長期的な予測の安定性は低いと考える。

5. 結 言

うつ病症状数の要因分析を行った結果、都市生活率と人口増加率の影響が大きかった。これは都市化に関する因子であることから都市化がうつ病症状数増加に強く関係していることがわかる。また、一人当たりCO₂排出量は、先進国では、環境保全や企業の海外進出などの動きから、影響が小さく、中位所得先進国では、工業化に伴う影響が大きかった。今後、うつ病症状数増加対策として、企業の拠点分散や在宅ワークにより人々の都市流用に歯止めをかけ、ストレスを減少することが重要である。

文 献

- [1] 厚生労働省：自殺・うつ病対策プロジェクトチームのまとめについて、自殺・うつ対策の経済的便益（2010年）
- [2] WHO：Global Health Estimates（2017年）
- [3] Our World in Data：https://ourworldindata.org/（2021年1月22日現在）