

災害被害統計における次元削減に関する研究 —インドネシアを事例として—

佐々木 大輔¹・森山 佳奈²・小野 裕一³

¹正会員 東北大学助教 災害科学国際研究所 (〒980-0845 仙台市青葉区荒巻字青葉468-1 S302)

E-mail: dsasaki@irides.tohoku.ac.jp

²非会員 パシフィックコンサルタンツ株式会社 (〒101-8462 東京都千代田区神田錦町三丁目22番地)

³非会員 東北大学教授 災害科学国際研究所 (〒980-0845 仙台市青葉区荒巻字青葉468-1 S302)

2015年に仙台で開催された第3回国連防災世界会議において、仙台防災枠組 (Sendai Framework for Disaster Risk Reduction 2015-2030) が策定された。当該枠組では、災害による人的・物的損失の削減に向けた結果目標を含む7つの目標が設定され、今後、各国政府には達成状況のモニタリングの実施が求められている。一方で、モニタリングのベースとなる災害被害統計について、特に集計項目の妥当性に関し、これまで十分な学術的研究がなされてきたとはいえない。そこで本研究では、災害被害統計の特徴・解釈方法を明確にするための手法として、多変量解析の1つである主成分分析を用いた次元削減 (Dimensionality Reduction) を提示する。なお、具体的な事例として、インドネシアにおける既存の災害被害統計データベースであるDIBI (Data Information Bencana Indonesia) を取り上げ、分析を実施した。その結果、「facilities (施設)」「agricultural losses (農業被害)」「Mortality (死亡)」に関連する主成分の存在が確認された。

Key Words : Sendai Framework, DIBI, Dimensionality Reduction, Principal Component Analysis

1. はじめに

2015年3月に宮城県仙台市で開催された第3回国連防災世界会議において、災害による人的・物的損失の削減に向けた結果目標を含む7つの目標を設定した仙台防災枠組 (Sendai Framework for Disaster Risk Reduction 2015-2030) が187カ国により採択された¹⁾。7つの目標とは、災害による死亡者数の大幅な削減 (目標(a))、災害による被災者数の大幅な削減 (目標(b))、災害による直接的経済損失の削減 (目標(c)) 医療・教育施設を含めた重要インフラへの損害や基本サービスの途絶の大幅な削減 (目標(d)) の4つの結果目標 (Outcome Target) と、国家・地方の防災戦略を有する国家数の大幅な増加 (目標(e))、開発途上国の施策を補完する適切で持続可能な支援、開発途上国への国際協力の大幅な強化 (目標(f))、マルチハザードに対応した早期警戒システムと災害リスク情報・評価へのアクセスの大幅な向上 (目標(g)) の3つの行動目標 (Output Target) からなる²⁾。今後、各国政府には達成状況のモニタリングの実施が求められており、その実施要領 (Technical Guidance) についても、2017年12月に国連国際防災戦略事務局 (UNISDR) から提供されている³⁾。このように、仙台防災枠組に基づくより客観的

なアプローチが導入されれば、従前にも増して災害科学における定量分析の必要性が高まるものと思われる。

一方、モニタリングのベースとなる災害被害統計について、特に集計項目の妥当性に関し、これまで十分な学術的研究がなされてきたとはいえない。こうした状況の下、東北大学災害科学国際研究所 (IRIDeS) では2015年4月に、仙台防災枠組のモニタリング支援や災害統計グローバルデータベースの構築等を目的として、災害統計グローバルセンター (Global Centre for Disaster Statistics) を設立した⁴⁾。当該センターには、国連開発計画 (UNDP) や民間企業も参画しており、産官学の幅広い視点から科学的根拠に基づく政策提言等を行うことが期待されている。

これらの背景を踏まえ、本研究では、災害被害統計の特徴・解釈方法を明確にするための手法を明らかにすることを目的としている。既存の災害被害統計においては、国ごとに集計項目等が異なり、国際比較を行うことが困難であるのが実情である。そこで、多変量解析の手法を用いて次元削減 (Dimensionality Reduction) を行うことにより、災害被害統計の本質的な特徴を明確化し、より一般的な形での解釈を可能にすることは、一定の意義があるものと考えられる。そこで、本研究では、インドネシアにおける既存の災害被害統計データベースであるDIBI

(Data Information Bencana Indonesia) を事例として、試行的な分析を実施した。

2. 方法論

(1) 主成分分析 (Principal Component Analysis)

多変量解析の1手法である主成分分析 (Principal Component Analysis) については、豊富な既往研究が存在している (例えば, Hotelling⁹⁾ や Wold et al.¹⁰⁾ など)。また昨今では、災害科学及びその関連領域においても、主成分分析を方法論に用いる研究が公開されている (例えば, Lou et al.⁷⁾ など)。ここでは、詳細な定式化は割愛し、当該分析手法の特徴について述べる。

相関行列を C とすると、主成分分析とは数学的に次の式を解くことを意味している。

$$C\mathbf{x} = \lambda\mathbf{x} \quad (1a)$$

(\mathbf{x} : C の固有ベクトル, λ : C の固有値)

ここで得られる固有値 λ が、それぞれ各主成分の分散に対応しており、値が大きいほど情報を多く有していると考えることができる。すなわち、固有値の大きい順に、第1主成分、第2主成分、第3主成分と対応させることで、情報の散逸を抑えながら変数を減らすことを可能にするのである。なお、採用する主成分の数については、通常、一定の基準を定めることになるが、本研究では、固有値基準 (対応する固有値が1を上回っている主成分を採用するもの) 及びスクリープロット (Scree Plot) 基準 (固有値をグラフにして傾きが緩やかになるまでの主成分を採用するもの) を基に、総合的に検討し、決定することにした。

(2) 分析対象データ

本研究では、インドネシアにおける既存の災害被害統計データベースである DIBI (Data Information Bencana Indonesia) に格納されている全データを抽出し、欠損値を1つでも含むレコードを削除 (Listwise Deletion) したものを分析対象とした。なお、削除後のレコード数は、14,295件である。

(3) ソフトウェアパッケージ

本研究では、主成分分析を行うにあたり、IBM社が提供している統計解析ソフトウェアのパッケージである SPSS Statistics 25を使用した。

3. 結果と考察

説明された分散の合計については、表-1に示したとおりである。固有値基準に基づけば、対応する固有値が1を上回っているか否かで主成分の採否を判断するので、今回の結果では、8個の主成分が抽出されることになる。ただし、第4主成分以降は、対応する固有値が1に近い値となっており、相対的に見て多くの情報を有しているとはいえない。一方で、図-1に示したスクリープロットを見ると、第4主成分以降、明らかに傾きが緩やかになっている。したがって、これらの結果を総合的に検討し、本研究では第3主成分までを採用することにする。

次に、表-2には成分行列を示した。これを基に、第1主成分、第2主成分、第3主成分が、各々どのような特性を有しているか検討する。まずは第1主成分であるが、injured, heavily damage houses, lightly damage houses, moderate damage houses, offices, health facilities, education facilities, worship facilitiesの8個の変数について、大きな値を示している。このことから、第1主成分は「facilities (施設)」に関連していると捉えることが可能である (なお、injuredについては、facilitiesが倒壊する際等に発生すると考えることができる)。また、第2主成分については、damage in crops, plantation, submerged housesの3個の変数に係る値が大きい。したがって、第2主成分は「agricultural losses (農業被害)」に関連していると捉えることができよう。そして第3主成分では、明らかにdeath, disappearの2個の変数に係る値が大きいことから、第3主成分は「Mortality (死亡)」に関連していると考えられる。

4. 結論

以上のように、固有値基準及びスクリープロット基準を用いて主成分の採否を総合的に判断した後、採用した主成分について成分行列を基にその解釈を与えることが可能である。本研究では、当該災害被害統計の特徴・解釈方法を明確にすることが可能である。本研究では、インドネシアの災害被害統計において、「facilities (施設)」、「agricultural losses (農業被害)」、及び「Mortality (死亡)」に関連する主成分が存在していることが確認された。

本研究では、次元削減により災害被害統計の特徴・解釈方法を明確にするための手法を明らかにすることを目的としたため、方法論として主成分分析を採用したが、潜在変数を想定して因果構造の解明を図る場合は、共分散構造分析・因子分析といった手法も存在する。今後は、こうした方法論も活用し、災害被害統計の因果構造を解

表-1 主成分分析の結果 (説明された分散の合計)

成分	初期の固有値			抽出後の負荷量平方和		
	合計	分散の%	累積%	合計	分散の%	累積%
1	3.462	15.737	15.737	3.462	15.737	15.737
2	2.347	10.668	26.405	2.347	10.668	26.405
3	1.530	6.953	33.358	1.530	6.953	33.358
4	1.177	5.351	38.709	1.177	5.351	38.709
5	1.156	5.254	43.963	1.156	5.254	43.963
6	1.059	4.812	48.775	1.059	4.812	48.775
7	1.034	4.701	53.476	1.034	4.701	53.476
8	1.002	4.554	58.029	1.002	4.554	58.029
9	1.000	4.545	62.575			
10	.987	4.485	67.059			
11	.975	4.433	71.492			
12	.942	4.281	75.773			
13	.871	3.957	79.730			
14	.828	3.764	83.494			
15	.787	3.578	87.072			
16	.671	3.048	90.120			
17	.567	2.577	92.697			
18	.501	2.275	94.972			
19	.394	1.790	96.762			
20	.372	1.692	98.454			
21	.340	1.546	100.000			
22	7.730E-10	3.514E-9	100.000			

因子抽出法: 主成分分析

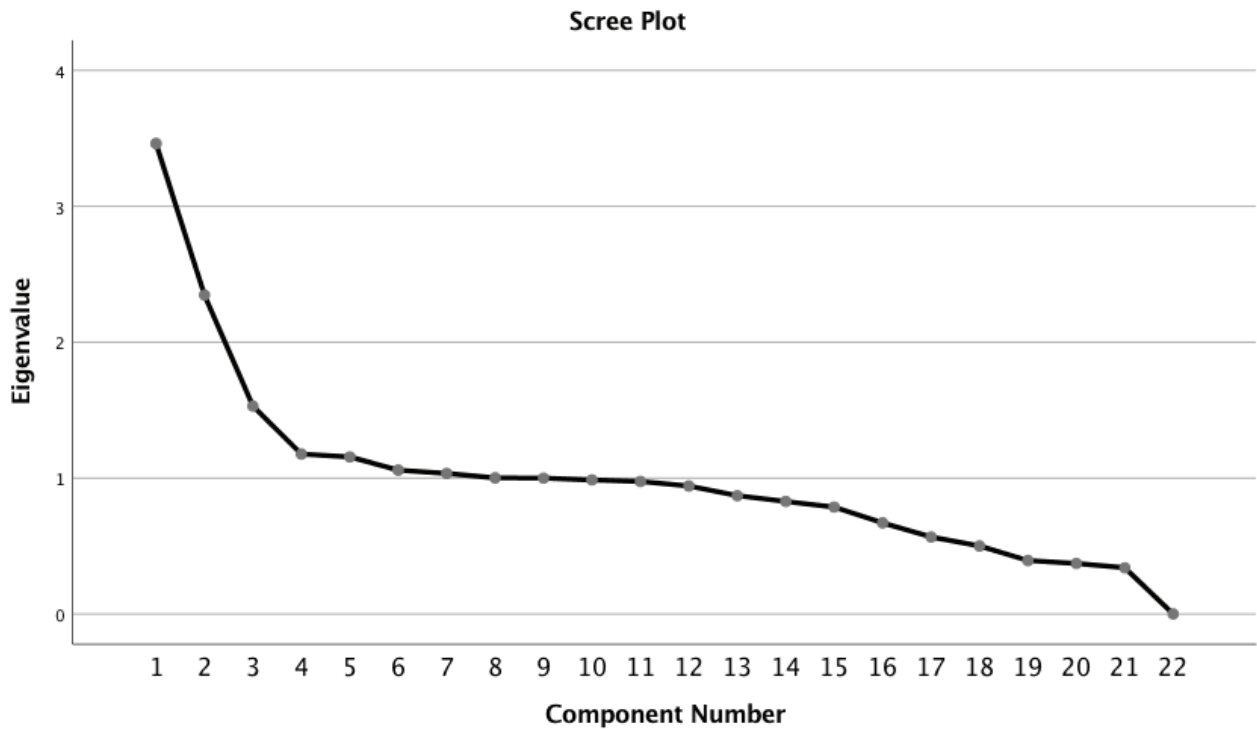


図-1 主成分分析の結果 (スクリープロット)

表-2 主成分分析の結果 (成分行列)

	成分							
	1	2	3	4	5	6	7	8
losses	.005	.038	-.001	.111	.639	.056	-.364	-.033
death	.284	-.079	.841	.007	-.007	-.011	-.028	.012
injured	.700	-.207	.003	-.116	.010	.007	-.016	-.020
disappear	.113	-.041	.815	-.079	.004	.007	-.023	.004
evacuated	.299	.050	.045	-.252	.197	.046	.543	-.160
afected	.018	.056	-.002	.068	.290	-.019	.277	.130
heavily_damage_houses	.683	-.198	.009	-.355	.079	-.002	.004	.020
lightly_damage_houses	.551	-.190	-.164	-.449	.085	-.005	.010	.028
moderate_damage_houses	.611	-.176	-.199	-.218	-.005	.023	-.153	.079
submerged_houses	.241	.657	-.001	.105	.145	.026	.246	-.028
offices	.627	-.028	-.110	.519	-.055	-.002	.113	-.093
kiosk	.000	-.006	-.001	-.001	.030	-.152	-.133	-.043
infrastructure	-.021	-.013	.011	-.001	-.064	.726	-.001	-.032
plantation	.260	.914	.022	-.085	-.122	-.007	-.141	-.014
ponds	.036	.143	-.002	.090	.316	.008	.501	.092
irrigation	.004	-.004	.004	.028	-.033	.705	-.022	.098
other_buildings	.016	.071	.001	.082	.660	.039	-.289	-.015
health_facilities	.603	-.181	-.146	.454	-.072	-.001	.033	-.088
education_facilities	.490	-.123	.193	.426	-.072	-.029	-.052	.034
damage_in_crops	.260	.914	.022	-.085	-.122	-.007	-.141	-.014
damage_in_road	.022	.034	-.002	.058	-.007	-.043	.033	.950
worship_facilities	.673	.077	-.132	-.016	-.078	.001	-.184	.093

因子抽出法: 主成分分析

a. 8 個の成分が抽出されました

明する一般的な手法を確立していくことが求められる。

また併せて、仙台防災枠組で設定された7つの目標が既存の災害被害統計においてどのように位置付けられるか、詳細な分析が必要である。本研究の結果、インドネシアの災害被害統計において、「facilities (施設)」、 「agricultural losses (農業被害)」、及び「Mortality (死亡)」に関連する主成分の存在が確認されたが、これらは7つの目標とも密接に関連しているものと考えられる。すなわち、既存の災害被害統計の構造を解明することは、仙台防災枠組における7つの目標に係るモニタリングにも資するものであるといえよう。

謝辞：本研究は、富士通株式会社から東北大学災害科学国際研究所災害統計グローバルセンターへの寄附金により実施されたものである。

参考文献

- 1) United Nations Office for Disaster Risk Reduction (UNISDR): Sendai Framework for Disaster Risk Reduction 2015 - 2030, UNISDR, 2015.

- 2) 小野裕一：仙台防災枠組における目標設定までの道のり、用語・指標設定の現状、および災害統計グローバルセンターについて、学術の動向, Vol.21, No.3, pp.94-102, 2016.
- 3) United Nations Office for Disaster Risk Reduction (UNISDR): Technical guidance for monitoring and reporting on progress in achieving the global targets of the Sendai Framework for Disaster Risk Reduction, UNISDR, 2017.
- 4) Ito, T., Miyamoto, M., and Ono, Y.: Strengthening governance on disaster risk reduction through improved disaster damage statistics, *Journal of Disaster Research*, Vol.11, No.3, pp.470-475, 2016.
- 5) Hotelling, H.: Analysis of a complex of statistical variables into principal components, *Journal of Educational Psychology*, Vol.24, No.6, pp.417-441, 1933.
- 6) Wold, S., Esbensen, K., and Geladi, P.: Principal component analysis, *Chemometrics and Intelligent Laboratory Systems*, Vol.2, No.1-3, pp.37-52, 1987.
- 7) Lou, W., Chen, H., Qiu, X., Tang, Q. and Zheng, F.: Assessment of economic losses from tropical cyclone disasters based on PCA-BP, *Natural Hazards*, Vol.60, No.3, pp.819-829, 2011.

(???. ???. ?? 受付)

DIMENSIONALITY REDUCTION FOR DISASTER STATISTICS: A CASE STUDY IN INDONESIA DISASTER DATABASE

Daisuke SASAKI, Kana MORIYAMA and Yuichi ONO

At the Third UN World Conference on Disaster Risk Reduction held in March 2015 in Sendai City, Japan, the new framework called the Sendai Framework for Disaster Risk Reduction 2015-2030 was adopted by 187 countries including the seven global targets. Furthermore, monitoring and reporting on progress in achieving the seven global targets is mandatory for all countries in the very near future. Meanwhile, there exist few studies concerning disaster loss statistics that will be the basis of the monitoring and reporting process. Thus, this study aims at giving an idea of dimensionality reduction by utilizing principal component analysis in order to clarify the characteristics of disaster loss statistics. We conducted a case study of Indonesian disaster loss database called DIBI (Data Information Bencana Indonesia). As a result, we found the existence of three principal components corresponding to facilities, agricultural losses and mortality, respectively.