The Factor Analysis of Earthquake Disaster in India by using Statistical Techniques

Shubham Kumar*[*1] , Rajan Singh [2]

[1] Research Scholar, Department of Mathematics, IFTM University, Mooradabad 244001, India

[2] Assistant Professor, Department of Mathematics, IFTM University, Mooradabad 244001, India

Abstract-

Factor analysis is an important tool of research works which helps us to understand the nature of problem, interpretation of this problem in terms of variables. By factor analysis we can interpret our result in terms of some very needful variables. The main objective of this study was to develop a mathematical model for studying the different factors of earthquake, using Statistical techniques. Earthquake is a sudden movement of earth's crust due to the result of release of stress by volcanic activity. Earthquake damage depends upon what area it hit. India is one of the most earthquake disaster prone countries of the world. It is very difficult to predict earthquake disaster.

Keywords- Factor, Factor loadings, Communality, Eigen value (or latent root), Rotation, Factor score.

Introduction-

This paper included many factors of Earthquake which are responsible for lethalness of earthquake and as well the factors which can play vital role in disaster management of earthquake related casualties. Earthquake is a sudden movement of earth's crust due to the result of release of stress by volcanic activity. Earthquake damage depends upon what area it hit. India is one of the most earthquake disaster prone countries of the world. It is very difficult to predict earthquake disaster. So, it is very much needed to have a very strong reinforcement system. India has had a number of the world's greatest earthquakes in the last century. In fact, more than 58% area in the country is considered prone to damaging earthquakes. Factor analysis is the very convenient method to represent a set of observed variables X_1, X_2, X_3……Xn in terms of a number of 'common' factor plus a factor which is unique to each variable. The common factors (sometimes called latent variables) are hypothetical variable which explains why a number of variables correlated with each other. It is because they have one or more factor in common.

Research Methodology-

There are several method of factor analysis, but they do not necessarily give same results. As such factor analysis is not a single unique method but a set of techniques. The centroid method is an important method of factor analysis.
The following research methodologies are adopted for the proposed research paper:

● Identification of the problem and defining the parameters for study

● Collection and study of available related literature

● Mathematical formulation of the problem by using statistical techniques.

● Numerical solution of the problem

● Interpretation of results.

● Conclusion

Factors related to earthquake

There are several factors that determine just how destructive an earthquake can be.

Location- Location is most important factor of earthquake. An earthquake that hits in a populated area is more likely to do damage that one that hits an unpopulated area or the middle of the ocean.

Magnitude- Earthquake magnitude is a measure of the size of the earthquake reflecting the elastic energy released by the earthquake. It is referred by a certain real number on the Richter scale (*e.g.*, magnitude 6.5 earthquake).

Depth- Earthquake can happen anywhere from at the surface to 700 kilometers below. In general, deeper earthquakes are less damaging because their energy dissipated before it reaches the surface.

Distance from the epicenter – The epicenter is the point at the surface right above where the earthquake originated and is usually the place where the earthquake's intensity is the greatest.

Local geologic condition- The nature of the ground at the surface of an earthquake can have a profound influence on the level of damage. Loose, sandy, soggy soil, can liquefy if the shaking is strong and long enough secondary effect.

Architecture- Even the strong buildings may not survive a bad earthquake, but architecture plays a important role in what and who survives a quake.

Moon effect- There is a popular belief that earthquakes are more frequent when the moon is close to full. The explanation is that a full moon has the strongest tidal pull. The study concluded that there is a very less amount of moon tides affects the occurrence of earthquakes,

Relationship between different factors of earthquake

Relation between magnitude and distance of an earthquake - Earthquake magnitude, energy release, and shaking intensity are all related variable if an earthquake. Their relationship with other variable is complicated but can be determined by using suitable data. The location and magnitude of an earthquake can be determined by the data recorded by seismometer.

Table for data

S.No.	Magnitude range	Distance range(Degree)
01	4-6.5	0-10
01	5-5.5	1-90
02	5.5-7.0	20-180
04	7.00-8.5	25-185
05	8.5-9.00	30-190

Correlation coefficient between magnitude and distance of an earthquake-

X	***Y***	***Mid -value of X***	***Mid- value of Y***	***X^2***	***Y^2***	***XY***
4-6.5	0-10	4.75	5	22.5625	25	23.75
5-5.5	10-90	5.25	50	27.5625	2500	262.5
5.5-7.0	20-180	6.25	100	39.0625	10000	625
7.00-8.5	25-185	7.75	105	60.0625	11025	813.15
8.5-9.00	30-190	7.75	110	60.0625	12100	852.5
		$\sum X = 31.75$	$\sum Y = 370$	$\sum X^2 = 209.3125$	$\sum Y^2 = 35650$	$\sum XY = 2476.90$

By using Karl Pearson method of Correlation-

$$r(X,Y) = \frac{\frac{1}{n}\sum XY - \bar{X}\bar{Y}}{\sqrt{\left(\frac{1}{n}\sum X^2 - \bar{X}^2\right)\ \left(\frac{1}{n}\sum y^2 - \bar{Y}^2\right)}}$$

$$\bar{X} = \frac{1}{n}\sum X = \frac{31.75}{5} = 6.35$$

$$\bar{Y} = \frac{1}{n}Y = \frac{370}{5} = 74$$

$$\text{So } r(X,Y) = \frac{\frac{1}{5}X\ 2476.90 - \ 469.90}{\sqrt{\left(\frac{1}{5}X\ 209.3125 - 40.3225\right)\ \left(\frac{1}{5}\sum 35650 - 5476\right)}}$$

$$= \frac{495.38 - 469.90}{50.4693}$$

***r(X,Y)* = 0.50486**

Relation between magnitude and depth of earthquake - The depth of the earthquake give us important information about the Earth's structure and tectonic setting where the earthquake is occurring.

S. No.	*Magnitude (X)*	*Depth (Cm) (Y)*	*(X^2)*	*(Y^2)*	*(XY)*
01	5.1	120	26.01	14400	612
02	5.2	142	27.04	20164	738.4
03	5.3	150	28.09	22500	795
04	5.4	155	29.16	24025	837
05	5.5	162	30.25	26244	891
06	5.6	173	31.36	29929	968.8
07	5.7	181	32.49	32761	1031.7
08	5.8	195	33.64	38025	1131
09	5.9	205	34.81	42025	1209.5
10	6.0	217	36	47089	1302
	$\sum X = 55.5$	$\sum Y = 1700$	$\sum X^2 = 3058.85$	$\sum Y^2 = 297162$	$\sum XY = 9516.4$

By using Karl Pearson method of Correlation-

$$r(X,Y) = \frac{\frac{1}{n}\sum XY - \bar{X}\bar{Y}}{\sqrt{\left(\frac{1}{n}\sum X^2 - \bar{X}^2\right)\left(\frac{1}{n}\sum y^2 - \bar{Y}^2\right)}}$$

$$\bar{X} = \frac{1}{n}\sum X = \frac{55.5}{10} = 5.55$$

$$\bar{Y} = \frac{1}{n} Y = \frac{1700}{10} = 170$$

$$\text{So } r(X,Y) = \frac{\frac{1}{10} X\ 9516.4 - \frac{1}{10} X\ 943.5}{\sqrt{\left(\frac{1}{10} X\ 3058.85 - 30.8025\right)\left(\frac{1}{10}\sum 297162 - 28900\right)}}$$

$$= \frac{951.64 - 94.35}{2850.1051}$$

r(X,Y) **= 0.3007**

Relation between magnitude and local geological conditions of an earthquake- Of the various earthquake-causing activities, the filling of large reservoirs is among the most important.

Zone V- Zone 5 covers the areas with the highest risks zone that suffers earthquakes of intensity MSK IX or greater. The IS code assigns zone factor of 0.36 for Zone 5.

Zone IV- This zone is called the High Damage Risk Zone and covers areas liable to MSK VIII. The IS code assigns zone factor of 0.24 for Zone 4.

Zone III- This zone is classified as Moderate Damage Risk Zone which is liable to MSK VII and also 7.8 The IS code assigns zone factor of 0.16 for Zone 3.

Zone II-This region is liable to MSK VI or less and is classified as the Low Damage Risk Zone. The IS code assigns zone factor of 0.10.

Zone I- Since the current division of India into earthquake hazard zones does not use Zone 1, no area of India is classed as Zone 1.

Table for factor load

S. NO.	Zone No.	Factor load	Average factor loading
01	V	0.36	0.1720
02	IV	0.24	
03	III	0.16	
04	II	0.10	
05	I	0	

Relation between magnitude and architecture of a place for earthquake- A large part of India is prone to strong earthquake-induced shaking, and with highly vulnerable constructions, there is huge risk of death and destruction. Areas in seismic zones IV and V of the Indian zone map encompass some of the most populated regions of India, including Delhi, the world's second most populous city. Only 10 % of total structures of India are liable to resist the earthquake. So we can say that this factor is highly correlated with the location of any place where then earthquake occur. The factor loading for this factor is 0.80.

Relation between magnitude and moon effect of a place for earthquake- An analysis of Earth deformation, earthquakes and tides has been undertaken using Earth tide, GPS data from stations in Southern California and the Parkfield area and interferometric SAR data from ESA satellites. We can say that that there is a very weak relationship in earthquake and tides of a particular place. So we can take a factor loading 0.005 for these two factors.

Centroid Method of Factor Analysis

This method of factor analysis, developed by L.L. Thurstone. The centroid method tends to maximizing the sum of loadings, disregarding signs; it is the method which extracts the largest sum of absolute loadings for each factor in turn.

Correlation matrix-

	Magnitude	Distance	Depth	Local Geo Condition	Architecture	Moon effect
Magnitude	1	0.5084	0.3007	0.1720	0.8000	0.0005
Distance	0.5084	1	-0.5610	0.9690	0.8670	0.0004
Depth	0.3007	-0.5610	1	0.8810	0.7240	0.0002
Local Geo Conditions	0.1720	0.9690	0.8810	1	0.9150	0.0006
Architecture	0.8000	0.8670	0.7240	0.9150	1	0.0001
Moon effect	0.0005	0.0004	0.0002	0.0006	0.0001	1

Let us denote the above variables as:

Variable No.	Description
1	Magnitude
2	Distance
3	Depth
4	Local Geo Conditions
5	Architecture
6	Moon effect

So, now we can tabulate our data and conduct the calculations for factor analysis

Variables

		1	2	3	4	5	6
Variables	**1**	1	0.5084	0.3007	0.172	0.8	0.0005
	2	0.5084	1	-0.5610	0.9690	0.8670	0.0004
	3	0.3007	-0.5610	1	0.881	0.7240	0.0002
	4	0.1720	0.9690	0.8810	1	0.9150	0.0006
	5	0.8	0.8670	0.7240	0.9150	1	0.0001
	6	0.0005	0.0004	0.0002	0.0006	0.0001	1
Column Sum		2.7816	2.7838	2.3449	3.9376	4.3061	1.0018

Sum of Column sums (T) = 16.4318 so $\sqrt{T}$ = 4.0536

First centroid factor **A**	$2.7816/\sqrt{T}$	$2.7838/\sqrt{T}$	$2.3449/\sqrt{T}$	$3.9376/\sqrt{T}$	$3.5821/\sqrt{T}$	$1.0018/\sqrt{T}$
	0.686	**0.687**	**0.578**	**0.971**	**0.884**	**0.247**

We can also state this information as under:

Variables	Factor loadings concerning first centroid factor A
1	**0.686**
2	**0.687**
3	**0.578**
4	**0.971**
5	**0.884**
6	**0.247**

To obtain the second centroid factor B, we first of all develop the first matrix of factor cross product Q_1

First Matrix of Factor Cross Product (Q_1)

	0.686	**0.687**	**0.578**	**0.971**	**0.884**	**0.247**
0.686	0.471	0.471	0.397	0.667	0.606	0.170
0.687	0.471	0.472	0.397	0.667	0.607	0.170
0.578	0.397	0.397	0.335	0.562	0.511	0.143
0.971	0.667	0.667	0.562	0.944	0.858	0.240
0.884	0.606	0.607	0.511	0.858	0.781	0.218
0.247	0.169	0.170	0.143	0.240	0.218	0.061

Now we obtain first matrix of residual coefficient (R_1) by subtracting Q_1 from R as shown below

First Matrix of Residual Coefficient (R_1)

Variables

		1	2	3	4	5	6
	1	0.529	0.037	-0.096	-0.495	0.194	0.001
	2	0.037	0.528	-0.958	0.302	0.260	0.000
Variables	**3**	-0.096	-0.958	0.665	0.319	0.213	0.000
	4	-0.495	0.302	0.319	0.056	0.057	0.001
	5	0.194	0.260	0.213	0.057	0.219	0.000
	6	-0.169	-0.169	-0.143	-0.239	-0.218	1.000

Reflected Matrix of Residual Coefficients (R_1')

Variables

		1	2	3	4	5	6
	1	0.529	0.037	0.096	0.495	0.194	0.001
	2	0.037	0.528	0.958	0.302	0.26	0
Variables	**3**	0.096	0.958	0.665	0.319	0.213	0
	4	0.495	0.302	0.319	0.056	0.057	0.001
	5	0.194	0.26	0.213	0.057	0.219	0
	6	0.169	0.169	0.143	0.239	0.218	1
Colum Sum		1.52	2.254	2.394	1.468	1.161	1.002

Sum of Column sums (T) = 9.799 so$\sqrt{T}$ = 3.1303

Second centroid factor **B**	1.520/$\sqrt{T}$	2.254/$\sqrt{T}$	2.394/$\sqrt{T}$	1.468/$\sqrt{T}$	1.161/$\sqrt{T}$	1.002/$\sqrt{T}$
	0.486	**0.720**	**0.765**	**0.469**	**0.371**	**0.320**

We can also state this information as under:

Variables	Factor loadings concerning second centroid factor B
1	**0.486**
2	**0.720**
3	**0.765**
4	**0.469**
5	**0.371**
6	**0.320**

We can find out communality for our data as under:

Variables	Factor Loadings		Communality(h^2)
	Centroid Factor A	Centroid Factor B	
1	0.686	0.486	$(0.686)^2 + (0.486)^2$ = **0.707**
2	0.687	0.720	$(0.687)^2 + (0.720)^2$ =**0.990**
3	0.578	0.765	$(0.578)^2 + (0.765)^2$ =**0.757**
4	0.971	0.469	$(0.971)^2 + (0.469)^2$ =**1.164**
5	0.884	0.371	$(0.884)^2 + (0.371)^2$ =**0.919**
6	0.247	0.320	$(0.247)^2 + (0.320)^2$ =**0.163**

Conclusion-

The main merit of this method is that it is relatively simple, can be easily understood and involve simpler calculations. Here Communality (h^2),shows how much of each variable is accounted for by the underlying factor taken together. A high value of communality means that not much the variable is left over after whatever the factors represent is taken into consideration. Here we have concluded that the variable 4 (Local Geo Conditions) have maximum communality i.e.1.164 and the variable no. 6 (Moon effect) has minimum communality i.e. 0.163. The relative communality of different variables shows the importance of each factor taken into consideration.

References

[1] **Andrews, J.F.:** "A Mathematical model for the continuous culture of microorganism utilizing inhibitory substrates", Biotech. Bioengg. 1968; 10,707-23

[2] **Bernardelli, H.** : :Population waves", Bermuda Research Soc., 1941;31,1-15

[3] **Bulm, J.S.**: Concentration profiles in and around capillaries, Am. Journ. Phys. 1960;198,991-98

[4] **Borsellino, A.** and **Torre, V.**: "Limits to growth from Volterra theory, Kybernetik, 1974, 16,113-18

[5] **C.A. Bender** "An introduction to Mathematical Modelling" Wiley Inter science, New York.

[6] **C.C. Linand** L.A. Segal" Mathematics applied to Deterministic problem in Natural Sciences, McMillan, New York, 1974.

[7] **Dabes, J.N.**, **Finn, R.K.**, and **Wilkie, C.R**.: "Equations for substrate- limited growth: The case for Blackman Kinetics", Biotech Bioengg; 1973;15, 1159-77

[8] **Doubleday, W.G.** : Harvesting in matrix population models, Biometrika , 1975;31,189-200

[9] **Dr. Priyanka Banerji**, **Ms. Nidhi Singh**, "Comparative Analysis of disaster management between Japan & India" IOSR13(6) 2013; 62:74

[10] **Gilpin, M.E. and Justice. K.E.**: Reinterpretation of invalidation of the principle of competitive exclusion, Nature, 1972, 236,273-301