IOSR Journal of Nursing and Health Science (IOSR-JNHS)
e-ISSN: 2320–1959.p- ISSN: 2320–1940 Volume 5, Issue 6 Ver. IV (Nov. - Dec. 2016), PP 10-13
www.iosrjournals.org

Risk Factor of Heart Attack in Sulaimany City

Sara Noori Mohammad Ali
Assistant Lecture, Department of Statistics. Administration and Economic School, University of Sulaimany.

Abstract: *This paper shows the main risk factor of heart attack (or the science name is Myocardial Infarctio) in sulaimany city;because heart attack is one of most sick published in sulaimany city with different age and gender , at application I take 300 patients have heart attack with different year, also take twenty three variables of each patients after that using Principal component analysis of sample data reduce the large number of data set to new set of factor and each factor contain one variable or more. And the number of factor is ten factors less than number of variables, list of factor not randamely but respectively one by one by more effect or risk factor of heart attack.The main of this paper is to show which variables in each factor have more effect of heart attack, respectively for each factor and variable.*

I. Introduction

The main idea of principal component analysis (PCA) is to reduce the dimensionality of a data set consisting of a large number of some variables. This is achieved by transforming to a new set of factors, the principal components (PCs), and which are ordered so that the first few retain most of the variation present in all of the original variables.also each factor contain one or more variables frome more effect to few effect at each factor also it is true about factors .

Principal component analysis has been widely adapted in climate research. The intention is to define coherent patterns of variability of a data set in either time or space. The technique of PCA has been introduced by PEARSON(1901). HOTELLING (1933) improved the PCA, since about 30 years it has frequently been used in climate research. PREISENDORFER (1988) gives a detailed overview over PCA techniques. The PCA in this study was implemented using a freely available software package from MURTAGH (2007).**[5]**

Principal components analysis is one of a family of techniques for high-dimensional data, and using the dependencies between the variables to represent it in a more tractable, lower-dimensional form, without losing too much information.

II. Goals and disadvantages of PCA

Principal component analyses (PCA) like some thingth of statistics have advantages and disadvantages.here I show little of it. The goals of PCA are to:

(1) Extract the most important information from the data table;
(2) Compress the size of the data set by keeping only this important information;
(3) Simplify the description of the data set; and
(4) Analyze the structure of the observations and the variables.[1]

But The disadvantages of PCA are:

(1) It is only able to find a linear subspace and thus cannot deal properly with data lying on nonlinear manifolds.
(2) One does not know how many principal components to keep, although some thumb rules are applied in practice. For example, eliminate components whose eigenvalues are smaller than a fraction of the mean eigenvalue, or keep as many as necessary to explain a certain fraction of the total variance.[6]

III. Definition and Derivation of Principal Component:

Suppose that $\mathbf{x}$ is a vector of p random variables, and that the variances of the p random variables and the structure of the covariances or correlations between the p variables are of interest. Unless p is small, or the structure is very simple, it will often not be very helpful to simply look at the p variances and all of the $1/2p(p-1)$ correlations or covariances. An alternative approach is to look for a few ($<< p$) derived variables that preserve most of the information given by these variances and correlations or covariances. Although PCA does not ignore covariances and correlations, it concentrates on variances.**[2]**

IV. Principal Component Analysis:

PCA is a way to recognize variables or we can say frome PCA few number of variables to factors but the variables not out to the analysis but some of them addition in one factor. But must not be all factor conain more than one variables may be contain one variable or more variables.

PCA will provide a mechanism to recognizethis geometric similarity through algebraic means [4].Principal components are equivalent to major axis regressions. As such, principal components analysis is subject to the same restrictions as regression, in particular multivariate normality, the distributions of each variable should be checked for normality and transforms used where necessary to correct high degrees of skewness in particular. Outliers should be removed from the data set as they can dominate the results of a principal components analysis.**[7]**

Principal component analysis (PCA) is a mainstay of modern data analysis a black box that is widely used but poorly understood. This tutorial does not shy away from explaining the ideas informally, nor does it shy away from the mathematics**.[4]**

V. Rotation

After the number of components has been determined, some time analysis need involves a rotation of the components that were retained. There are two main types of rotation are used:

Orthogonal: when the new axes are also orthogonal to each other, and

Oblique: when the new axes are not required to be orthogonal.

When and Why Using Rotations? The main reason for using rotation is to facilitate the interpretation. When the data follow a model stipulating that each variable load on only one factor and that there is a clear difference in intensity between the relevant factors and the noise, then the rotation is likely to provide a solution that is more reliable than the original solution. However, if this model does not accurately represent the data, then rotation will make the solution less replicable and potentially harder to interpret because the mathematical properties of PCA have been lost.**[1]**

VI. Application

From application principle component must be : the number of variable at least two variable and dependent of subject to variable ratio (STV) rule must be the number of observation or subject fifth of variables or we can say at least we have 100 observation or subject.

From above rule at application we can say we don't have any problem about sample size (number of observation) and number of variables.

By using spss program to find what is the risk factor of heart attack in sulaimany city I take 300 patients all of them have heart attack befor ,I took some variables (Did you take medication for your Disease regularly?, Did you visit physician regularly?, Diabetes mellitus, Are you doing Exercise?, Do you have regular exercise?, Rheumatoid arthritis, Dyslipidaemia, Hypertension, Hypertension Family History, Previous Heart attack, Cardiac Catheterization, Smoking, Waist - Hip ratio, Do you drink alcohol?, Do you drink alcohol?, Have you been exposed to critical events that affected your life through the previous years?, Do you agree about your body shape?, BMI, Sudden death Family History, Heart attack Family History, Diabetes mellitus Family History, Dyslipidaemia Family History, Do you agree with your dietary pattern?, angina) for this research, after that at statistics I use PCA to find risk factor and the result of statistics as follow:

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.452	10.661	10.661	2.452	10.661	10.661	1.980	8.608	8.608
2	2.043	8.880	19.542	2.043	8.880	19.542	1.804	7.842	16.450
3	1.712	7.443	26.985	1.712	7.443	26.985	1.768	7.687	24.136
4	1.601	6.960	33.946	1.601	6.960	33.946	1.654	7.193	31.329
5	1.367	5.944	39.890	1.367	5.944	39.890	1.637	7.115	38.444
6	1.300	5.651	45.541	1.300	5.651	45.541	1.334	5.801	44.245
7	1.238	5.383	50.924	1.238	5.383	50.924	1.222	5.311	49.557
8	1.100	4.781	55.705	1.100	4.781	55.705	1.212	5.271	54.828
9	1.095	4.759	60.464	1.095	4.759	60.464	1.177	5.116	59.945
10	1.046	4.549	65.014	1.046	4.549	65.014	1.166	5.069	65.014
11	.961	4.177	69.190						
12	.914	3.974	73.164						
13	.845	3.673	76.837						
14	.800	3.478	80.315						
15	.767	3.334	83.649						
16	.643	2.794	86.443						
17	.616	2.678	89.121						
18	.512	2.224	91.346						
19	.503	2.187	93.532						
20	.455	1.980	95.513						
21	.399	1.733	97.246						

22	.363	1.577	98.823						
23	.271	1.177	100.000						

Table above show the number of component (ten component) that is calculated or found from (twenty three) variables or frome sample.also show the initial eigenvalues (with total, percent of variance and percent of cumulative), Extraction Sums of Squared Loadings (with total, percent of variance and percent of cumulative) and Rotation Sums of Squared Loadings (with total, percent of variance and percent of cumulative).

Also frome this table we can see the cumulative percent variance for ten factors is (65.014) it means (65.014) of heart attack is that ten factor from twenty variables. That component has different value of variance, important for select variables that poignant at heart atack. That ten component explanations each of them respectively (8.608, 7.842, 7.687, 7.193, 7.115, 5.801, 5.311, 5.271, 5.116 and 5.069) from total variance after rotation.

Rotated Component Matrixa

	Component									
	1	2	3	4	5	6	7	8	9	10
Did you take medication for your Disease regularly?	.770	.159	.160	.103	-.092	.107	.042	-.089	-.037	-.088
Did you visit physician regularly?	.752	-.051	.072	.242	.050	-.030	.023	-.056	-.053	.195
Diabetes mellitus	.607	-.081	-.013	-.128	-.144	.111	.118	.373	-.102	-.195
Are you doing Exercise?	-.035	.864	-.029	.072	.006	.138	.042	-.071	-.067	.043
Do you have regular exercise	.053	.851	-.065	-.018	.052	-.156	-.074	.097	-.002	-.053
Rheumatoid arthritis	.251	.384	-.229	-.074	.098	-.039	-.364	-.070	.197	.364
Dyslipidaemia	.000	-.062	.683	.196	-.042	-.113	.012	-.043	.100	.148
Hypertension	.184	-.103	.669	.027	-.076	.103	-.106	-.184	-.089	.050
Hypertension Family History	.051	.098	.660	-.307	.236	-.076	.107	.178	.225	-.178
Previous Heart attack	.072	-.050	.103	.804	.106	-.102	.015	.018	.014	.166
Cardiac Catheterization	.193	.115	-.041	.752	-.218	.113	-.025	.040	.001	-.186
Smoking	.184	-.017	.013	-.037	-.723	-.016	.223	-.009	.059	.115
Waist - Hip ratio	.149	-.005	-.004	-.183	.672	.099	.086	-.170	-.007	.168
Do you drink alcohol?	-.204	.241	.153	.220	.448	-.058	.240	.249	-.268	-.199
Have you been exposed to critical events that affected your life through the previous years?	-.148	.074	-.285	.009	.362	-.002	-.351	-.059	.237	-.112
Do you agree about your body shape?	.141	-.044	.023	-.011	.013	.745	-.172	.124	-.085	.110
BMI	.003	-.058	.142	.010	-.139	-.688	-.331	.067	-.246	.098
Sudden death Family History	.101	-.027	-.094	-.006	-.062	-.019	.725	-.179	.078	.204
Heart attack Family History	.090	-.071	.057	-.142	.134	-.096	.092	-.750	-.046	-.085
Diabetes mellitus Family History	.305	-.129	-.129	-.307	.120	-.092	-.170	.492	-.025	.106
Dyslipidaemia Family History	-.158	-.074	.156	.040	-.105	.067	.011	.021	.822	.003
Do you agree with your dietary pattern?	-.169	-.119	.343	.114	-.309	.380	-.230	-.077	-.405	.084
Angina	-.042	-.003	.145	.025	-.048	.066	.220	.130	-.040	.802

this table Rotated Component Matrixa show the variable or variables of each componet also This table shows the principal component analysis and statistics for all variables, also from this it shows the factors respectively have risk or effecting heart attack because the first factor have more effect or risk compare second factor on heart attack and second factor more effect compare third factor and third factor have more effect compare fourth factor ,so on for all factors in each factor the first variable have more effect of heart attack compare second variable and second variable have more effect compare third variable so on for all variable in factors.

Also at Component Score Covariance Matrix all value of this matrix is 0.000 expect orthogonal value all of them is (1).

VII. Results & Conclusions

By taken 300 patients in sulaimany city all of them have heart attack,I took some informations frome all patients (Did you take medication for your Disease regularly?, Did you visit physician regularly?, Diabetes mellitus, Are you doing Exercise?, Do you have regular exercise?, Rheumatoid arthritis, Dyslipidaemia, Hypertension, Hypertension Family History, Previous Heart attack, Cardiac Catheterization, Smoking, Waist - Hip ratio, Do you drink alcohol?, Do you drink alcohol?, Have you been exposed to critical events that affected your life through the previous years?, Do you agree about your body shape?, BMI, Sudden death Family History, Heart attack Family History, Diabetes mellitus Family History, Dyslipidaemia Family History, Do you agree with your dietary pattern?, angina) . All variables used to find the risk factor heart attack by using Principal component analysis we have this resealt:

1. We get ten factor frome twenty three variables.
2. The first risk factor of heart attack is (Did you take medication for your Disease regularly? , Did you visit physician regularly? and Diabetes mellitus) .

3. Second risk factor is (Are you doing Exercise? And do you have regular exercise, Rheumatoid arthritis)
4. Third factor contain variables (Dyslipidaemia, Hypertension and Hypertension Family History)
5. Fourth factor contain (Previous Heart attack and Cardiac Catheterization)
6. Also fifth risk factor has (Smoking, Waist - Hip ratio, Do you drink alcohol? and Have you been exposed to critical events that affected your life through the previous years?)
7. Sixth factor contain two variables are (Do you agree about your body shape?and BMI)
8. Seventh factor contain only (Sudden death Family History)
9. Eighth (Heart attack Family History and Diabetes mellitus Family History)
10. Ninth (Dyslipidaemia Family History and Do you agree with your dietary pattern?)
11. Tenth factor only contain Angina.

VIII. Sources

[1] Herve´ Abdi[1] and Lynne J. Williams[2]," Principal component analysis"
[2] Volume 2, July/August 2010.
[3] I.T. Jolliffe, "Principal Component Analysis", second edition, 2002.
[4] Jeff Jauregui, "Principal component analysis with linear algebra", August 31, 2012
[5] Jon Shlens, "A TUTORIAL ON PRINCIPAL COMPONENT ANALYSIS"
[6] Derivation, Discussion and Singular Value Decomposition, Version1, 25 March 2003.
[7] KAI BORN[1], ANDREAS H. FINK[1] and HEIKO PAETH[2], "Dry and wet periods in the northwestern Maghreb for present day and future climate conditions", [1]Institut f ¨ur Geophysik und Meteorologie der Universit¨at zu K¨oln, K¨oln, Germany, [2]Institut f ¨ur Geographie, W¨urzburg, Germany, (Manuscript received April 23, 2008; in revised form July 11, 2008; accepted July 11, 2008)
[8] Miguel A'. Carreira-Perpina'n, " A Review of Dimension Reduction Techniques_ " Technical Report CS-96-09, Dept. of Computer Science, University of Sheffeld
[9] January 27, 1997.
[10] Steven M. Holland,"PRINCIPAL COMPONENTS ANALYSIS (PCA)", Department of Geology, University of Georgia, Athens, GA 30602-2501, May 2008.