

Investigation of the Dependence between Insurance Premiums and Socio- Economic Development Indices with Canonical Correlation Analysis

Betül Zehra KARAGÜL, Ismail GÜR, Hüseyin TATLIDIL

Hacettepe University, Ankara, TURKEY

3rd European Actuarial Journal Conference

Lyon, France, Sept 5-8 2016

Outline

- Introduction
- Canonical Correlation Analysis
- Assumptions
- Linear Canonical Correlation
- Partial Canonical Correlation
 - Income Statement of Insurance Types for Turkey
 - Life Insurance
 - Nonlife Insurance
- Cultural and Economic Development Indices
 - Application
 1. Canonical correlation analysis
 2. Partial canonical correlation analysis
- Conclusions
- Reference

Introduction

It is important to guess the demand for the products which are created by insurance companies. The premium increase in different lines of business can be a good clue to forecast this demand.

In this study we use the Canonical Correlation Analysis to reveal the multidimensional relationship between life insurance premiums and non-life insurance premiums. We use the Turkey insurance data. The effect of the some cultural and economic development indices which are created by OECD and TURKSTAT.

Canonical Correlation Analysis

We can measure the linear relationship between two sets of multidimensional variables. Canonical correlation is appropriate in the same situations where multiple regression would be, but where there are multiple intercorrelated outcome variables. Canonical correlation analysis determines a set of canonical variates, orthogonal linear combinations of the variables within each set that best explain the variability both within and between sets.

Canonical Correlation Analysis

Assumptions

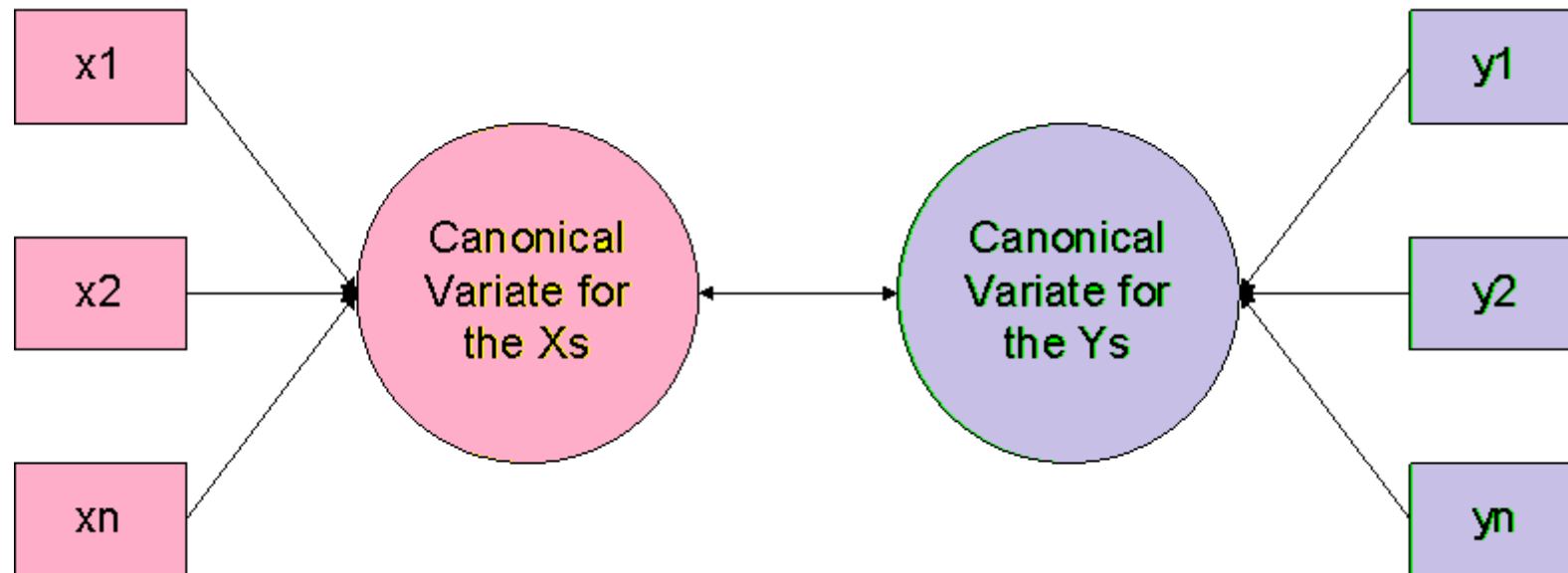
○ Normality

- Univariate - univariate normality is not explicitly required for multivariate multiple regression
- Multivariate - multivariate normality is required and there is not way to test for except establishing univariate normality on all variables, even though this is still no guarantee.

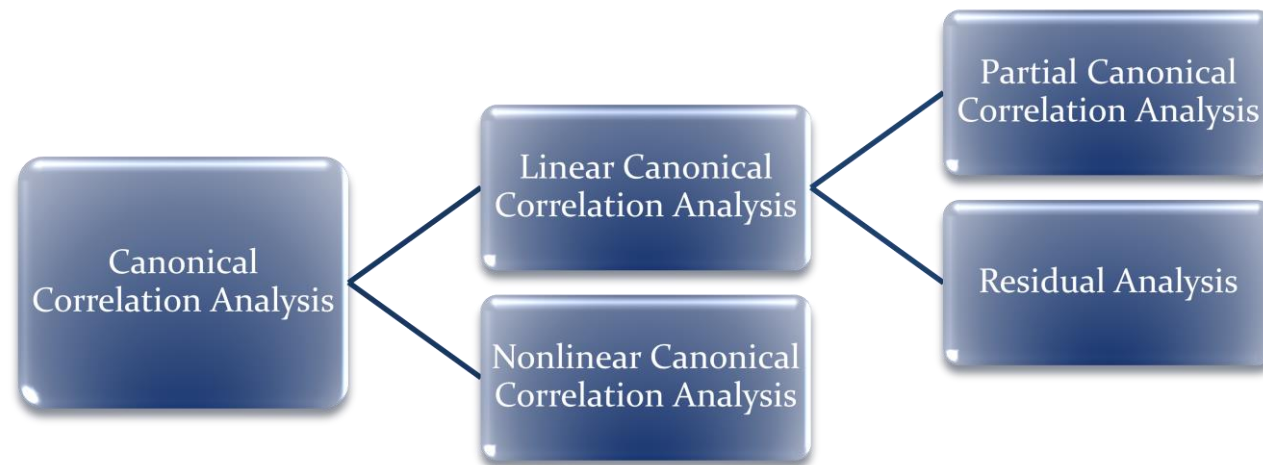
○ Linearity - linear relationship assumed for all variables in each set and also between sets

○ Homoskedasticity - needs to be checked for all pairs of variables within and between sets.

Canonical Correlation Analysis



Canonical Correlation Analysis



Linear Canonical Correlation Analysis

$$X' = [X_1 \ X_2 \ \dots \ X_p]$$
$$Y' = [Y_1 \ Y_2 \ \dots \ Y_p]$$

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$
$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

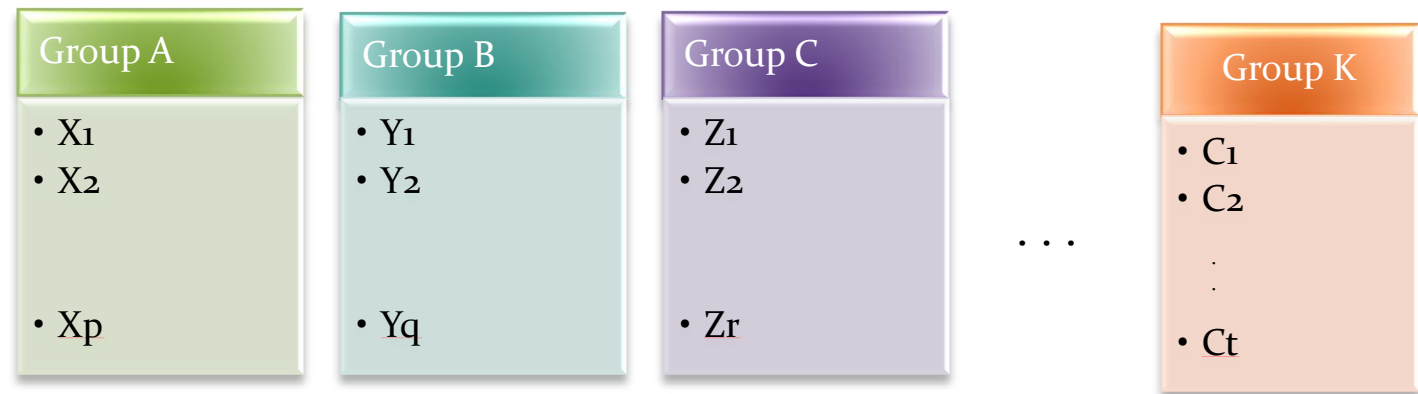
$$U = \alpha'X$$
$$V = \gamma'Y$$

$$\sigma_U^2 = \alpha' \Sigma_{11} \alpha \qquad E(U) = E(\alpha'X) = \alpha'E(X) = 0$$
$$\sigma_V^2 = \gamma' \Sigma_{22} \gamma \qquad E(V) = E(\gamma'Y) = \gamma'E(Y) = 0$$

$$r_{UV} = \alpha' \Sigma_{12} \gamma = \rho$$

Partial Canonical Correlation Analysis

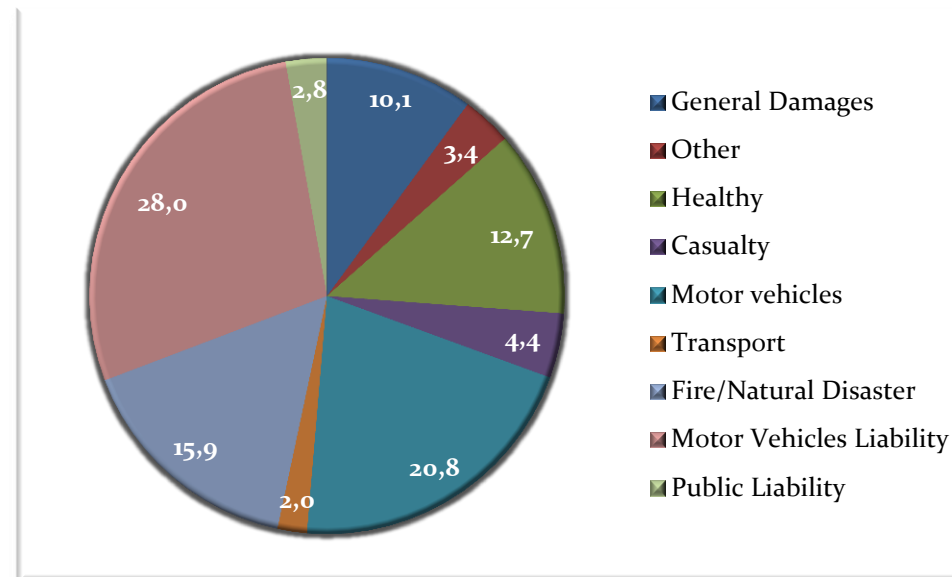
- Tree or more data sets in canonical correlation



- We chose one of the 3 data sets which each of them have respectively p , q , r variables and isolate the effect of it on the others then do canonical correlation analysis between these two data sets.

Income Statement of Insurance Types for Turkey

In line with the global insurance market, insurers may operate on life or non-life business according to Turkish Insurance Regulation. Since 2008, 18 non-life insurance and 8 life insurance lines of business are regulated.



Graph 1. Gross Premium Production per LoB in nonlife (2015)

Income Statement of Insurance Types for Turkey

Nonlife

In 2015,

- the business accounted for 26% of total claim payments and with a 20% premium share in total non-life premium.
- Motor vehicles liability insurance accounted for 28% of direct premium volume in non-life insurance.
- Fire and natural disasters insurance accounts for 16% of non-life direct premium volume.

Income Statement of Insurance Types for Turkey

Life

2011	2012	2013	2014	2015	
886.126	1.008.956	1.214.761	1.375.975	1.333.958	Term Life Ins. (Yearly)
1.256.670	1.210.537	1.573.065	1.476.614	1.968.900	Term Life Ins. (Long Term)
319.040	273.827	383.891	240.380	246.681	Whole Life Insurance
87.414	91.891	83.288	66.545	61.523	Pure Endowment Insurance
22.025	49.013	47.691	50.470	95.925	Endowment Life Insurance
113.131	75.887	91.984	69.613	53.592	Annuity
177	125	83	20	190	Others
2.684.583	2.710.237	3.394.763	3.279.617	3.760.768	Total

Table 1. Gross Premium Production per LoB in life

- ✓ Term Life Yearly
- ✓ Term Life Long term

Cultural and Economic Development Indices

Total household spending (TURKSTAT): A per person breakdown of general living expenses. It includes the amount paid for lodging, food consumed within the home, utilities paid and other expenses. The sum of all the expenses is then divided by the number of family members residing in the house in order to find each member's part of the total expense.

Unemployment rate (OECD): Unemployment rate is the number of unemployed people as a percentage of the labour force, where the latter consists of the unemployed plus those in paid or self-employment.

Passenger transport(OECD):Passenger transport refers to the total movement of passengers using inland transport on a given network.

Alcohol consumption (TURKSTAT): Alcohol consumption is defined as annual sales of pure alcohol in liters per person aged 15 years and older.

Application

1) Canonical Correlation Analysis

Correlations between Non-Life insurance branches

Correlation Analysis

Pearson Correlation Coefficients

	Y (1)	Y(2)	Y(3)
Y(1): fire and natural disaster	1.00000	0.95171	0.99870
Y(2): Land vehicles liability	0.75171	1.00000	0.74178
Y(3): land vehicles liability (TPL)	0.99870	0.94178	1.00000

Fire and natural disaster and land vehicle liability (TPL) are highly correlated

Application

Correlations between Life insurance branches

Pearson Correlation Coefficients

	X (1)	X(2)
X(1): Term life ins(yearly)	1.00000	0.96326
X(2): Term life ins.(long term)	0.96326	1.00000

Life Insurance
 X_1 : Yearly term insurance
 X_2 : Long term insurance



Nonlife Insurance
 Y_1 : fire and natural disaster
 Y_2 : land vehicles liability

Application

Correlations between Life insurance and Non-Life insurance branches

Pearson Correlation Coefficients		
	Y (1)	Y(2)
X(1): Term life ins(yearly)	0.96827	0.94639
X(2): Term life ins.(long term)	0.89801	0.93579

Canonical Correlation Analysis					
	Canonical Correlation	Eigenvalues	F Value	D F	P Value
1	0.976982	20.9750	13.87	6	0.0004
2	0.698066	0.9504	5.70	2	0.0542

As a result of linear canonical correlation analysis, two canonical correlation coefficients are obtained and one of these is statistically significant in %95 confidence level. Thus, the relationship between total premium productions of nonlife and life insurance is positively 0.976982.

Application

Standardized Canonical Coefficients for the Nonlife insurance

	V1*	V2
Fire and natural disaster	1.0736	-3.0754
Land vehicles liability	-0.0776	3.2565

Standardized Canonical Coefficients for the Life insurance

	U1*	U2
Term life ins(yearly)	1.5235	-3.3976
Term life ins.(long term)	-0.5551	3.6819

$$U_1^* = 1.5235X_{(1)} - 0.5551X_{(2)}$$

$$U_2 = -3.3976X_{(1)} + 3.6819X_{(2)}$$

$$V_1^* = 1.0736Y_{(1)} - 0.0776Y_{(2)}$$

$$V_2 = -3.0754Y_{(1)} + 3.2565Y_{(2)}$$

Application

	U_1^*	V_1^*
term life ins(yearly)	0.9888	0.9661
term life ins.(long term)	0.9125	0.8915

	U_1^*	V_1^*
fire and natural disaster	0.9997	0.9767
Land vehicles liability	0.9441	0.9224

Canonical loadings(U_1^*) are correlations between variables and the canonical variates. Both life and nonlife insurance branches has higher loadings for U_1^* and V_1^*

Variance and Redundancy Analysis for Nonlife and Life insurance premium productions

NonLife

Canonical Variable Number	Their Own Canonical Variables		Canonical R-Square	The Opposite Canonical Variables	
	Proportion	Cumulative Proportion		Proportion	Cumulative Proportion
1	0.9454	0.9454	0.9545	0.9024	0.9024
2	0.0546	1.0000	0.4873	0.0266	0.9290

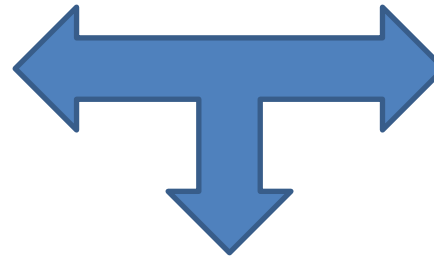
Life

Canonical Variable Number	Their Own Canonical Variables		Canonical R-Square	The Opposite Canonical Variables	
	Proportion	Cumulative Proportion		Proportion	Cumulative Proportion
1	0.9052	0.9052	0.9545	0.8640	0.8640
2	0.0948	1.0000	0.4873	0.0462	0.9102

Application

2) Partial Canonical Correlation Analysis / Economic Indices

Life Insurance
 X_1 : Yearly term insurance
 X_2 : Long term insurance



Nonlife Insurance
 Y_1 : fire and natural disaster
 Y_2 : land vehicles liability

Economic Indices
 Z_1 : Total household expense
 Z_2 : Unemployment Rate

Application

2) Partial Canonical Correlation Analysis / Economic Indices

Correlation Coefficients between Nonlife and Life insurance productions, when economic indices are partially stable.

Pearson Partial Correlation Coefficients		
	X (1)	X(2)
X(1): Term life ins(yearly)	1.0000	0.79673
X(2): Term life ins.(long term)	0.79673	1.0000

Pearson Partial Correlation Coefficients		
	Y(1)	Y(2)
Y(1): Fire and natural disaster	1.0000	0.85429
Y(2): Land vehicles liability	0.85429	1.0000

Application

Pearson Partial Correlation Coefficients		
	X(1): Term life ins(yearly)	X(2): Term life ins.(long term)
Y(1): Fire and natural disaster	0.97363	0.76777
Y(2): Land vehicles liability	0.89414	0.89794

	Canonical Correlation	Eigen Values	F Value	D F	P Value
1	0.984152	30.808	9.15	6	0.0100
2	0.607358	0.5845	2.34	2	0.2010

Standardized Canonical Coefficients for the Nonlife Insurance

	V1*	V2
fire and natural disaster	0.6794	-1.7998
Land vehicles liability	0.3551	1.8907

Standardized Canonical Coefficients for the Life insurance

	U1*	U2
term life ins(yearly)	0.8607	-1.4132
term life ins.(long term)	0.1683	1.6461

Application

2) Partial Canonical Correlation Analysis / Economic Indices

$$U_1^* = 0.8607X_{(1)} + 0.1683X_{(2)}$$

$$U_2 = -1.4132X_{(1)} + 1.6461X_{(2)}$$

$$V_1^* = 0.6794Y_{(1)} + 0.3551Y_{(2)}$$

$$V_2 = -1.7998Y_{(1)} + 1.8907Y_{(2)}$$

NonLife

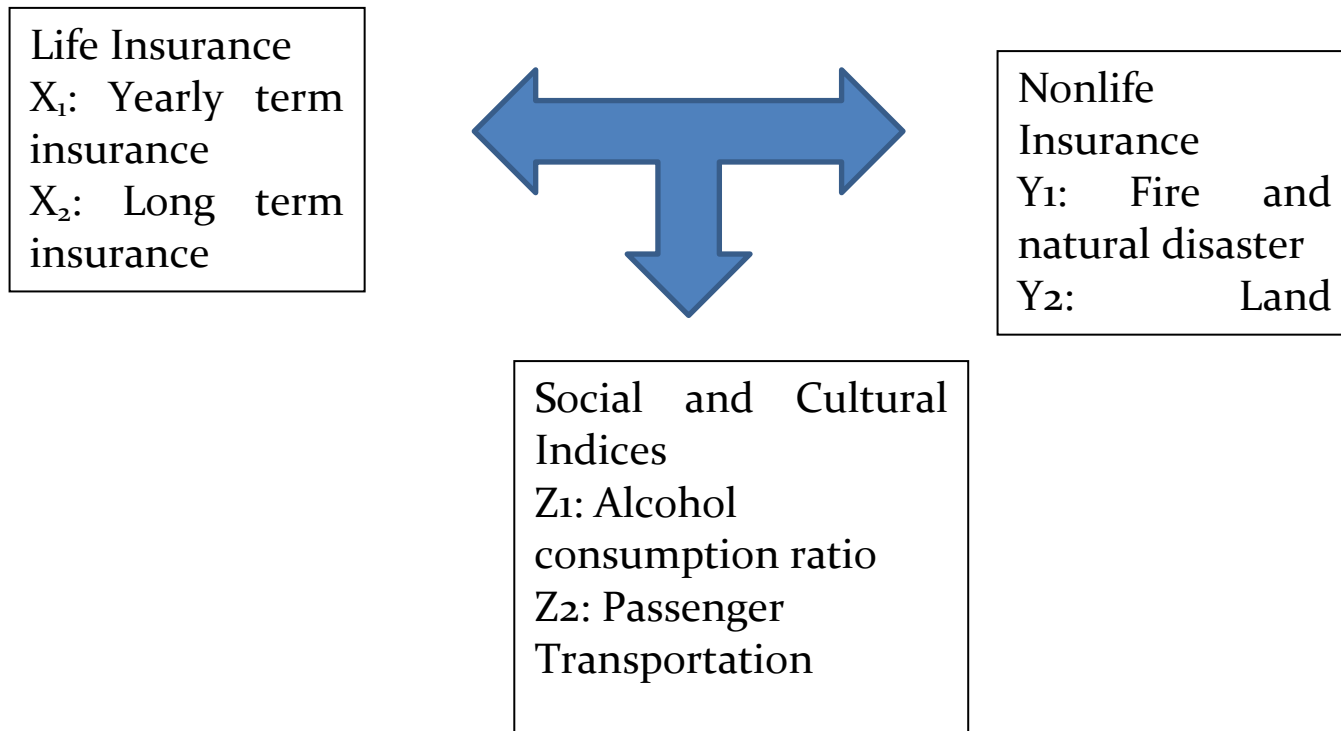
Canonical Variable Number	Their Own Canonical Variables		Canonical R-Square	The Opposite Canonical Variables	
	Proportion	Cumulative Proportion		Proportion	Cumulative Proportion
1	0.9206	0.9206	0.9686	0.8916	0.8916
2	0.0794	1.0000	0.3689	0.0293	0.9209

Life

Canonical Variable Number	Their Own Canonical Variables		Canonical R-Square	The Opposite Canonical Variables	
	Proportion	Cumulative Proportion		Proportion	Cumulative Proportion
1	0.8595	0.8595	0.9686	0.8325	0.8325
2	0.1405	1.0000	0.3689	0.0518	0.8843

Application

2) Partial Canonical Correlation Analysis / Social and Cultural Indices



Application

Pearson Partial Correlation Coefficients		
	X(1): Term life ins(yearly)	X(2): Term life ins.(long term)
Y(1): Fire and natural disaster	0.95721	0.87601
Y(2): Land vehicles liability	0.38916	0.71987

	Canonical Correlation	Eigen Values	F Value	D F	P Value
1	0.987151	38.1641	11.71	4	0.0054
2	0.703691	0.9809	3.92	1	0.1187

Standardized Canonical Coefficients for the Nonlife Insurance

	V1	V2
fire and natural disaster	0.8879	-1.3237
Land vehicles liability	0.1390	1.5878

Standardized Canonical Coefficients for the Life insurance

	U1	U2
term life ins(yearly)	1.1170	-0.5651
term life ins(long term)	-0.2205	1.2322

Application

2) Partial Canonical Correlation Analysis / Social and Cultural Indices

$$U_1^* = 1.1170X_{(1)} - 0.2205X_{(2)}$$

$$U_2 = -0.5651X_{(1)} + 1.2322X_{(2)}$$

$$V_1^* = 0.8879Y_{(1)} + 0.1390Y_{(2)}$$

$$V_2 = -1.3237Y_{(1)} + 1.5878Y_{(2)}$$

NonLife

Canonical Variable Number	Their Own Canonical Variables		Canonical R-Square	The Opposite Canonical Variables	
	Proportion	Cumulative Proportion		Proportion	Cumulative Proportion
1	0.8410	0.8410	0.9745	0.8195	0.8194
2	0.1590	1.0000	0.4952	0.0787	0.8983

Life

Canonical Variable Number	Their Own Canonical Variables		Canonical R-Square	The Opposite Canonical Variables	
	Proportion	Cumulative Proportion		Proportion	Cumulative Proportion
1	0.5864	0.5864	0.9745	0.5714	0.5714
2	0.4136	1.0000	0.4952	0.2048	0.7762

Conclusions

- In this study first of all the dependency between the Gross Premium of lines of business in nonlife and in life is investigated with Canonical Correlation Analysis.
- Secondly, this dependency is investigated with the effect of some cultural and economic development indices which are thought can have important role on the gross premium by partial canonical correlation analysis.
- When we consider the effect of the development indices the result is more realistic because all the variables in this dynamic sector are interrelated and the ratio of explained variation decreases.
- Social and cultural development indices have more effect on the dependency between LoB in life and in nonlife than the economic development indices.

References

- Canonical Correlation Analysis, Uses and Interpretation(Quantitative Applications in the Social Sciences),Bruce Thompson
- Applied Multivariate Statistical Analysis,6th Edition,Richard A. Johnson and Dean W. Wichern
- Applied Multivariate Statistical Analysis,Wolfgang Karl Härdle,Léopold Simar
- <http://www.treasury.gov.tr/en-US/Reports-Page?mid=1202&cid=28&nm=714>
Insurance and Private Pension Reports -Undersecretariat of Treasury,Republic of Turkey Prime Ministry
- https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#cancorr_toc.htm- Tutorial for Canonical Correlation Analysis for SAS
- Bhatia, V.K. (2002) Canonical Correlation Analysis, IASRI,Library Avenue, Dew Delhi.
- Hotelling, H. (1935). The Most Predictable Criterion. Journal of Experimental Psychology. 26, 139-142. Johnson, R.A.,Wichern, D.W. (2007). Applied Multivariate Statistical Analysis, 6th ed. Pearson, Prentice Hall, Upper Saddle River, NJ.

- OECD (2011). <http://stats.oecd.org/Index.aspx?DataSetCode=INSINDs.105-127>.
- Tatlıdil, H. (2002). Uygulamalı Çok Değişkenli İstatistiksel Analiz. Ziraat Matbaacılık, Ankara.
- Tatlıdil, H., İçen D. Sigorta Sektörünün Çok Boyutlu İlişki Analizi Yöntemleri ile İncelenmesi, Bankacılar Dergisi, 2013