

Auxiliar 7

Otoño 2012-IN3401



Profesor: Marcelo Henríquez

Auxiliares: Paulina Céspedes, Sebastián Rojas A.

P1) Los siguientes cuadros son obtenidos al aplicar un ACP a un conjunto de datos asociados a diez países, principalmente europeos, publicados en 2006. Los datos corresponden a consumo de licor, vino y cerveza (en litros anuales por habitante), expectativa de vida (en años) y tasa de enfermedades cardíacas (en casos cada 100 mil habitantes, por año).

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
Licor	1,750	,9132	10
Vino	20,990	24,9270	10
Cerveza	68,260	38,6718	10
Expectativa_vida	75,90	3,213	10
Tasa_cardiaca	153,870	110,8182	10

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,331
Bartlett's Test of Sphericity	Approx. Chi-Square	16,516
	df	10
	Sig.	,086

Correlation Matrix

	Licor	Vino	Cerveza	Expectativa_vida	Tasa_cardiaca
Sig. (1-tailed)					
Licor		,452	,081	,144	,108
Vino			,133	,060	,129
Cerveza				,396	,243
Expectativa_vida					,012
Tasa_cardiaca					

Component Matrix^a

	Component				
	1	2	3	4	5
Licor	-,525	,720	,164	,413	,097
Vino	,675	,479	,472	-,291	,081
Cerveza	-,112	-,918	,325	,134	,146
Expectativa_vida	,887	-,110	,206	,369	-,149
Tasa_cardiaca	-,878	-,055	,432	-,115	-,164

Extraction Method: Principal Component Analysis.

a. 5 components extracted.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2,301	46,028	46,028	2,301	46,028	46,028
2	1,606	32,115	78,142	1,606	32,115	78,142
3	,584	11,685	89,827	,584	11,685	89,827
4	,422	8,444	98,271	,422	8,444	98,271
5	,086	1,729	100,000	,086	1,729	100,000

Extraction Method: Principal Component Analysis.

Rotated Component Matrix^a

	Component	
	1	2
Licor	-,438	,776
Vino	,726	,398
Cerveza	-,217	-,899
Expectativa_vida	,869	-,211
Tasa_cardiaca	-,878	,047

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Justificando adecuadamente, responde:

- ¿Cuáles bebidas se asocian a creciente expectativa de vida? ¿Cuáles a decreciente?
- ¿Se adecuan los datos al análisis factorial?
- ¿Qué representan los coeficientes de la Matriz de Componentes?
- Explicite la relación entre la Matriz de Componentes y el cuadro de Varianza Total explicada.
- ¿Cuántas componentes sugiere considerar en modelo de reducción de datos?
- Indique las variables que saturan en cada una de las componentes escogidas en (e) (estructura no rotada)
- Grafique las saturaciones factoriales (estructura rotada) en el primer plano factorial.
- De acuerdo al modelo escogido en (e), ¿el consumo de vino tiene algo que ver con expectativa de vida y con la tasa de enfermedades cardíacas? ¿Le parece esto realista?
- Beber cerveza y whiskey, ¿es una combinación favorable en los países analizados?

P2) Transantiago cuenta con la información de la evasión mensual promedio en términos porcentuales en los acercadores de distintas zonas del Gran Santiago, así como en distintos recorridos troncales para un año calendario (Octubre 2009 -Septiembre 2010). Se realiza un Análisis Factorial con el fin de reducir la información mensual y poder graficar en un mapa perceptual las distintas zonas/recorridos. La información relevante se despliega al final del enunciado.

- (1 pts.) ¿Se adecúa el problema a Análisis Factorial? Argumente cuantitativamente.
- (1 pts.) Describa la solución encontrada en términos de reducción de información con respecto al conjunto de datos original, tanto a nivel global como a nivel de variables. ¿Es una buena solución?
- (0,5 pts.) Bajo el criterio de extracción a partir de los valores propios. ¿Cuántos factores deberían extraerse y por qué?
- (1 pts.) Para este problema en particular, ¿qué ventaja tiene la rotación de ejes?
- (2,5 pts.) Metodológicamente hablando, ¿Cómo se construye el mapa perceptual presentado? Interprete cada eje a partir de la información disponible.

Hint: Identifique las zonas en el mapa perceptual y relaciónelas con los estadísticos descriptivos presentados.

Descriptive Statistics			
	Mean	Std. Deviation	Analysis N
Oct2010	,1793	,08983	14
Nov2010	,1800	,08143	14
Dic2010	,1836	,09009	14
Ene2010	,1807	,07800	14
Feb2010	,1829	,06900	14
Mar2010	,1786	,06803	14
Abr2010	,1750	,06892	14
May2010	,1714	,07941	14
Jun2010	,1843	,07035	14
Jul2010	,1979	,07329	14
Ago2010	,1800	,06668	14
Sep2010	,1850	,05735	14

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,766
Bartlett's Test of Sphericity	Approx. Chi-Square	297,598
	df	66
	Sig.	,000

Communalities		
	Initial	Extraction
Oct2010	1,000	,897
Nov2010	1,000	,922
Dic2010	1,000	,823
Ene2010	1,000	,936
Feb2010	1,000	,983
Mar2010	1,000	,935
Abr2010	1,000	,785
May2010	1,000	,888
Jun2010	1,000	,813
Jul2010	1,000	,827
Ago2010	1,000	,868
Sep2010	1,000	,680

Extraction Method: Principal Component Analysis.

Component Matrix ^a		
	Component	
	1	2
Oct2010	,800	,508
Nov2010	,936	,215
Dic2010	,895	-,148
Ene2010	,963	-,094
Feb2010	,991	-,033
Mar2010	,966	-,054
Abr2010	,882	-,080
May2010	,820	,465
Jun2010	,817	-,381
Jul2010	,899	,135
Ago2010	,917	-,164
Sep2010	,740	-,364

Extraction Method: Principal Component Analysis.

a. 2 components extracted

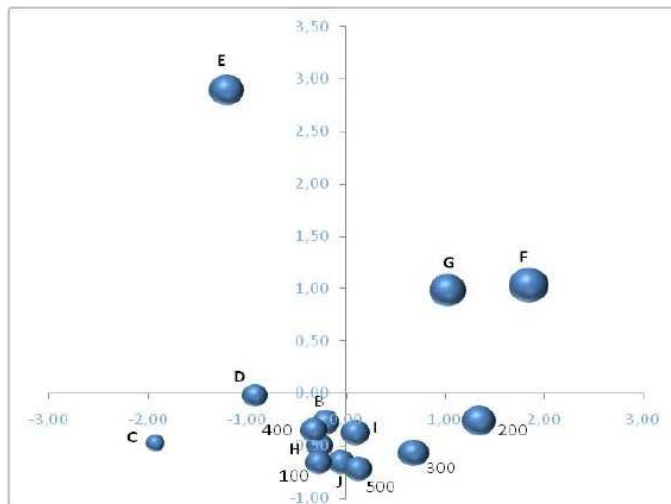
Rotated Component Matrix ^a		
	Component	
	1	2
Oct2010	,272	,908
Nov2010	,566	,775
Dic2010	,773	,474
Ene2010	,789	,560
Feb2010	,771	,624
Mar2010	,765	,592
Abr2010	,719	,518
May2010	,315	,888
Jun2010	,867	,248
Jul2010	,591	,691
Ago2010	,801	,477
Sep2010	,798	,210

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

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	9,474	78,948	78,948	9,474	78,948	78,948	5,787	48,225	48,225
2	,884	7,364	86,312	,884	7,364	86,312	4,571	38,088	86,312
3	,606	5,052	91,364						
4	,482	4,017	95,381						
5	,306	2,551	97,932						
6	,128	1,064	98,996						
7	,066	,546	99,543						
8	,041	,338	99,881						
9	,013	,104	99,985						
10	,001	,012	99,997						
11	,000	,002	99,999						
12	,000	,001	100,000						

Extraction Method: Principal Component Analysis.



ZONA	Prom	DesvEst	CoefVar
B	0,159	0,034	0,2115
C	0,067	0,011	0,1655
D	0,137	0,026	0,1889
E	0,250	0,070	0,2814
F	0,318	0,054	0,1698
G	0,273	0,029	0,1067
H	0,145	0,030	0,2090
I	0,169	0,016	0,0939
J	0,152	0,041	0,2693
100	0,140	0,032	0,2257
200	0,235	0,027	0,1133
300	0,190	0,035	0,1846
400	0,149	0,032	0,2159
500	0,156	0,031	0,2002