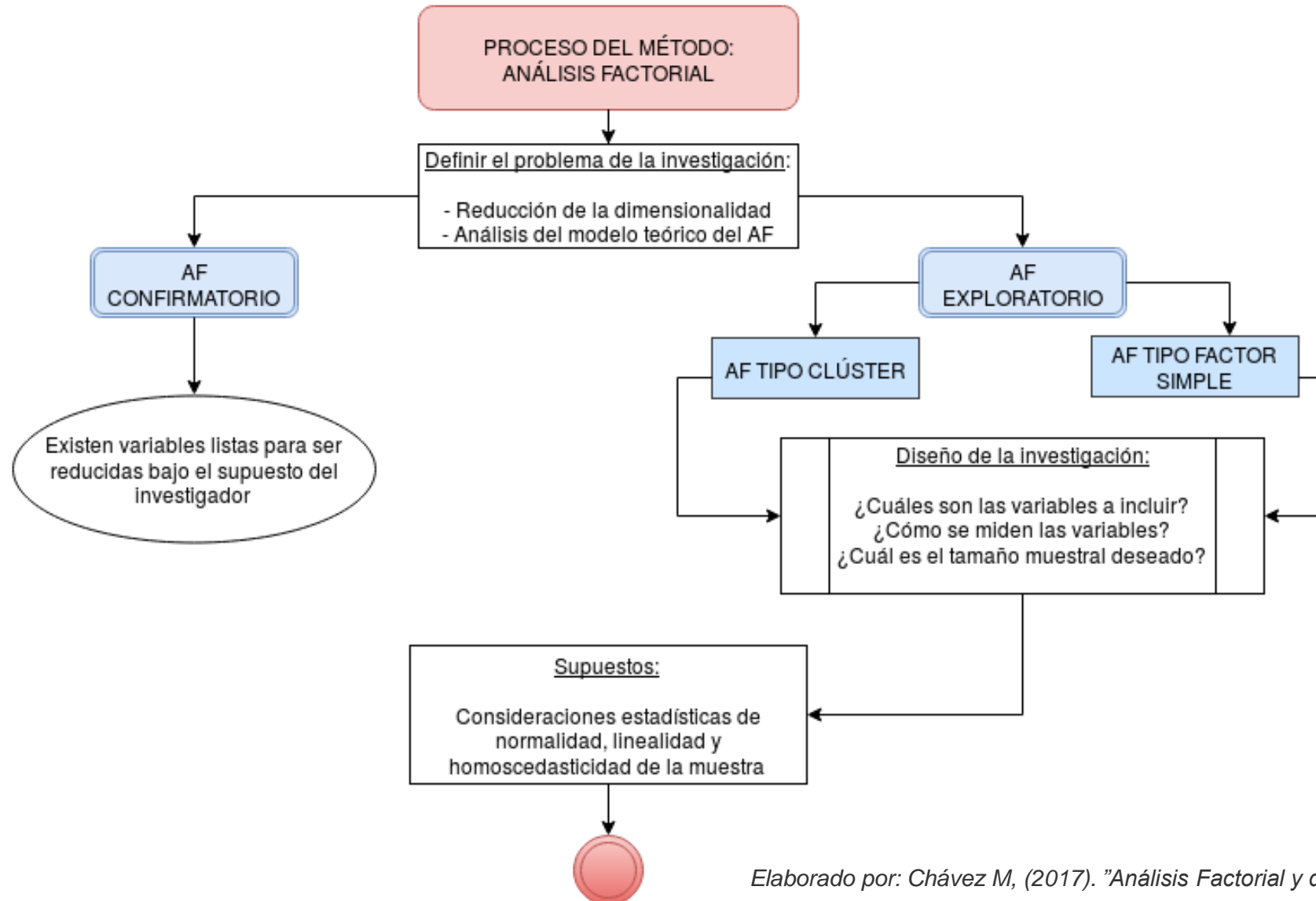




Análisis Factorial

Dr. Misael Erikson Maguiña Palma

Diagrama de la decisión del análisis factorial



Análisis Factorial Exploratoria

Una técnica de resumen

- Se caracteriza porque no se conocen a priori el numero de factores y es la aplicación empírica donde se determina este numero.
- Estima y agrupa un grupo de ítems en variables latentes.

Variables latentes, son las variables que no se observan directamente sino que son inferidas a partir de otras variables que se observan.

Etapas del análisis

- 1) Planteamiento de problema (incluir un numero de variables relacionas teóricamente con el tema de investigación)

Análisis Factorial Exploratoria

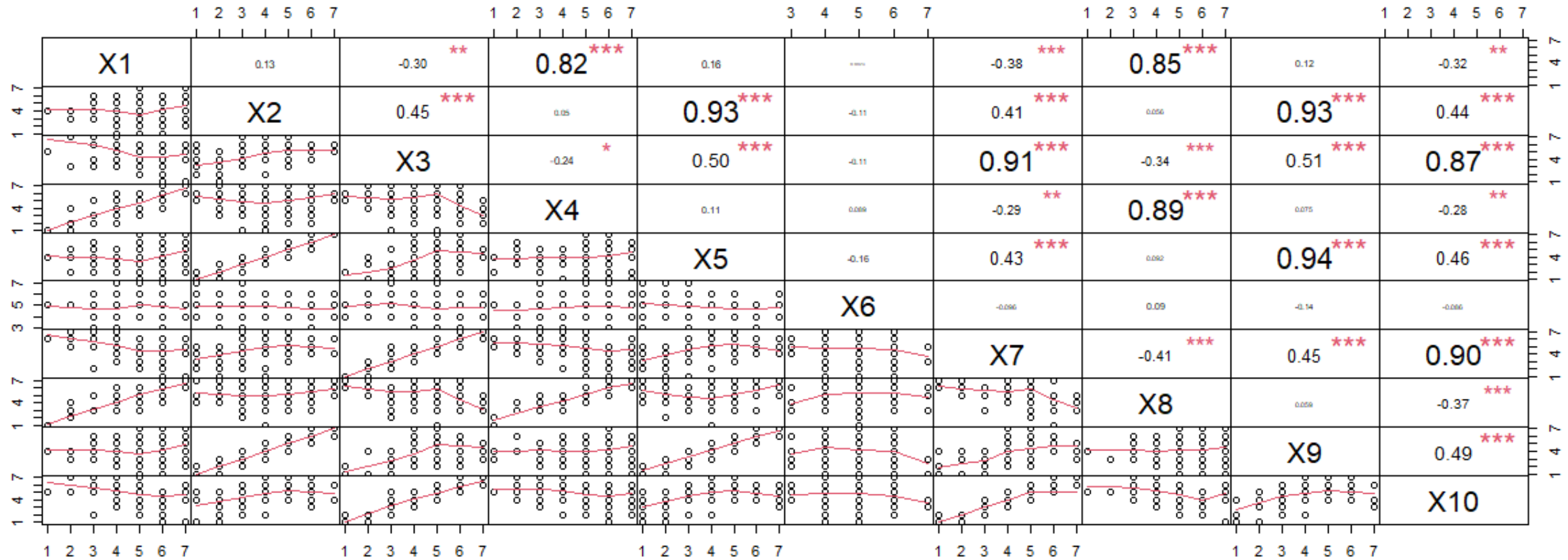
Etapa del análisis (Continuación....)

2) Estudio de condiciones previas

Antes de iniciar con el AF, debe ser realizado un estudio previo de:

2.1 Análisis de la matriz de correlaciones

Uno de los requisitos que debe cumplir el AF, es que las variables estén altamente intercorrelacionadas. Esto implica que las **correlaciones tiene que ser mayor de 0.30** en su mayoría para poder aplicar un AF.



Análisis Factorial Exploratoria

Etapas del análisis (Continuación....)

2.2 Test de Esfericidad de Bartlett

Contraste la hipótesis nula y concluir que las variables de la muestra están suficientemente correlacionadas entre sí para realizar el análisis factorial.

2.3 La prueba de adecuación de Kaiser-Meyer Olkin (KMO)

Permite comprobar el grado en que cada una de las variables es predecible a partir de las demás. Este estadístico se distribuye en valores entre 0 y 1.

Si el valor de KMO es mayor o igual a 0.75 la idea de realizar un AF es buena, si se encuentra entre 0.5 y 0.75 la idea es aceptable y KMO es inferior a 0.5 es inaceptable realizar el AF exploratorio.

Análisis Factorial Exploratoria

Test de Esfericidad de barlett

```
Bartlett test of homogeneity of variances  
data: .  
Bartlett's K-squared = 49.595, df = 9, p-value  
= 1.284e-07
```

Se verifica que existe correlación

Kaiser-Meyer-Olkin factor adequacy

```
Overall MSA = 0.81  
MSA for each item =  
X1 X2 X3 X4 X5 X6 X7 X8 X9 X10  
0.85 0.81 0.85 0.73 0.82 0.59 0.78 0.77 0.84 0.87
```

$$|R| = 8.021167e - 06$$

El determinante de la matriz de correlaciones es cercano a cero. La variable esta correlacionada. Tiene sentido realizar el análisis factorial exploratorio.

Análisis Factorial Exploratoria

Etapa del análisis (Continuación....)

3) Extracción de factores

Una vez que se ha determinado que el AF es una técnica apropiada para analizar los datos, debe seleccionarse el método adecuado para la extracción de los factores. Existen diversos métodos cada uno de ellos con sus ventajas e inconvenientes. Los más conocidos con el método de los componentes principales y el de **Máxima verosimilitud**.

Algo que debe resaltar es que el software R la función que realiza el análisis factorial por defecto trabaja con el método MLE (Maximum-Likelihood Estimation) es decir, Estimación de máxima verosimilitud.

Análisis Factorial Exploratoria

El método de estimación

Factoring method:

“minres”,

“uls”,

“wls”,

“gls”,

“pa”,

“ml”,

“minchi”,

“minrank”,

“old.min”,

“Alpha”

¿con cuál nos quedamos?

Aquel modelo que tenga la **proporción Var más alta** y el **RMSR mas pequeño**

La matriz desde la que estimamos

Correlaciones:

“cor” es Pearson,

“cov” es covarianza,

“tet”, es tetrachoric,

“poly”, es polychoric,

“mixed”, utiliza cor mixto para una mezcla de tetrachórico, policórico, pearson, biserial y polyserials.

Análisis Factorial Exploratoria

	item	ML1	ML2	ML3	h2	u2	com
X7	7	0.98			0.995	0.005	1.1
X3	3	0.95			0.913	0.087	1.0
X10	10	0.95			0.911	0.089	1.0
X6	6				0.027	0.973	2.5
X5	5	0.54	0.81		0.967	0.033	1.8
X2	2	0.54	0.79		0.957	0.043	1.9
X9	9	0.58	0.78		0.976	0.024	1.9
X1	1	-0.37	0.61	0.58	0.849	0.151	2.6
X4	4		0.51	0.76	0.927	0.073	2.1
X8	8	-0.41	0.56	0.68	0.955	0.045	2.6
			ML1	ML2	ML3		
SS loadings			4.08	2.89	1.51		
Proportion Var			0.41	0.29	0.15		
Cumulative Var			0.41	0.70	0.85		
Proportion Explained			0.48	0.34	0.18		
Cumulative Proportion			0.48	0.82	1.00		

Análisis Factorial Exploratoria

Etapa del análisis (Continuación....)

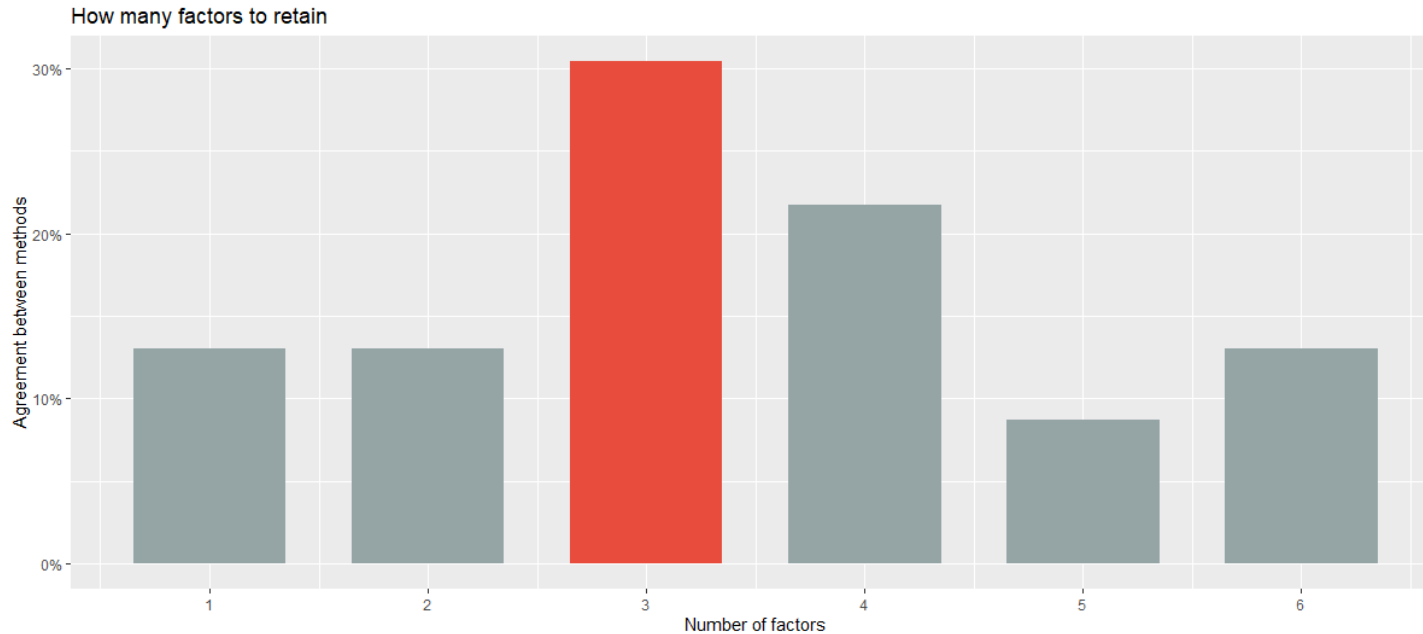
4) Determinación del número de factores

En esta parte, lo que se busca es determinar el número óptimo de factores a conservar. **Generalmente, hay un conjunto reducido de factores, los primeros, que contienen casi toda la información.** Los otros factores suelen contribuir relativamente poco, El objetivo principal es cumplir el principio de parsimonia

Numero de factores a extraer

1	Bentler	12	R2
2	t	13	VSS complexity 2
3	p	14	Velicer's MAP
4	Acceleration factor	15	BIC
5	VSS complexity 1	16	BIC
6	TLI	17	CNG
7	RMSEA	18	beta
8	Optimal coordinates	19	BIC (adjusted)
9	Parallel analysis	20	CRMS
10	Kaiser criterion	21	Bartlett
11	SE Scree	22	Anderson
		23	Lawley

Numero de factores a extraer



n_Factors	Method	Family
1	t	Multiple_regression
2	p	Multiple_regression
3	RMSEA	Fit
4	Optimal coordinates	Scree
5	Acceleration factor	Scree
6	VSS complexity 1	VSS
7	Bentler	Bentler
8	CNG	CNG
9	Parallel analysis	Scree
10	Kaiser criterion	Scree
11	BIC	BIC
12	BIC (adjusted)	BIC
13	BIC	Fit
14	beta	Multiple_regression
15	SE	Scree
16	R2	Scree_SE
17	VSS complexity 2	VSS

```
eigen() decomposition
$values
[1] 4.40007626 3.03714062 1.04959788 0.90892902
[5] 0.17675658 0.12843317 0.10230114 0.08123850
[9] 0.05912470 0.05640212
```

n_Factors	n_Methods
1	3
2	3
3	7
4	5
5	2
6	3

Análisis Factorial Exploratoria

Etapa del análisis (Continuación....)

5) Interpretación de los factores

En esta fase se sugieren los dos pasos siguientes:

- Identificar las variables cuya correlaciones con el factor son las más elevadas en valor absoluto
- Intentar dar un nombre a los factores. El nombre debe asignarse de acuerdo con la estructura de sus correlaciones con las variables. Analizando con qué variables tiene una relación fuerte es posible, en muchos casos, hacerse una idea más o menos clara de cual es el significado de un factor.

Análisis Factorial Exploratoria

	item	ML1	ML2	ML3	h2	u2	com
X7	7	0.98			0.995	0.005	1.1
X3	3	0.95			0.913	0.087	1.0
X10	10	0.95			0.911	0.089	1.0
X6	6				0.027	0.973	2.5
X5	5	0.54	0.81		0.967	0.033	1.8
X2	2	0.54	0.79		0.957	0.043	1.9
X9	9	0.58	0.78		0.976	0.024	1.9
X1	1	-0.37	0.61	0.58	0.849	0.151	2.6
X4	4		0.51	0.76	0.927	0.073	2.1
X8	8	-0.41	0.56	0.68	0.955	0.045	2.6
		ML1	ML2	ML3			
SS loadings		4.08	2.89	1.51			
Proportion Var		0.41	0.29	0.15			
Cumulative Var		0.41	0.70	0.85			
Proportion Explained		0.48	0.34	0.18			
Cumulative Proportion		0.48	0.82	1.00			

Análisis Factorial Exploratoria

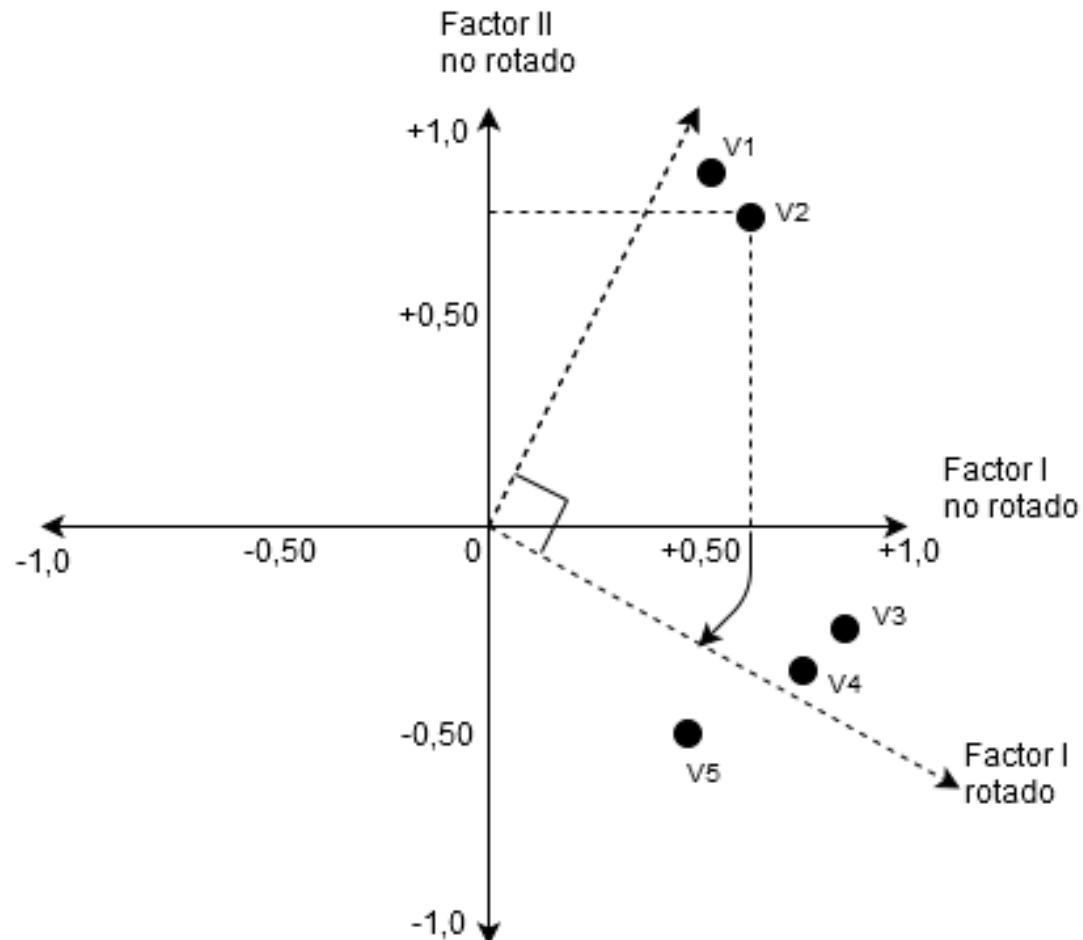
Etapa del análisis (Continuación....)

6) Rotación de factores

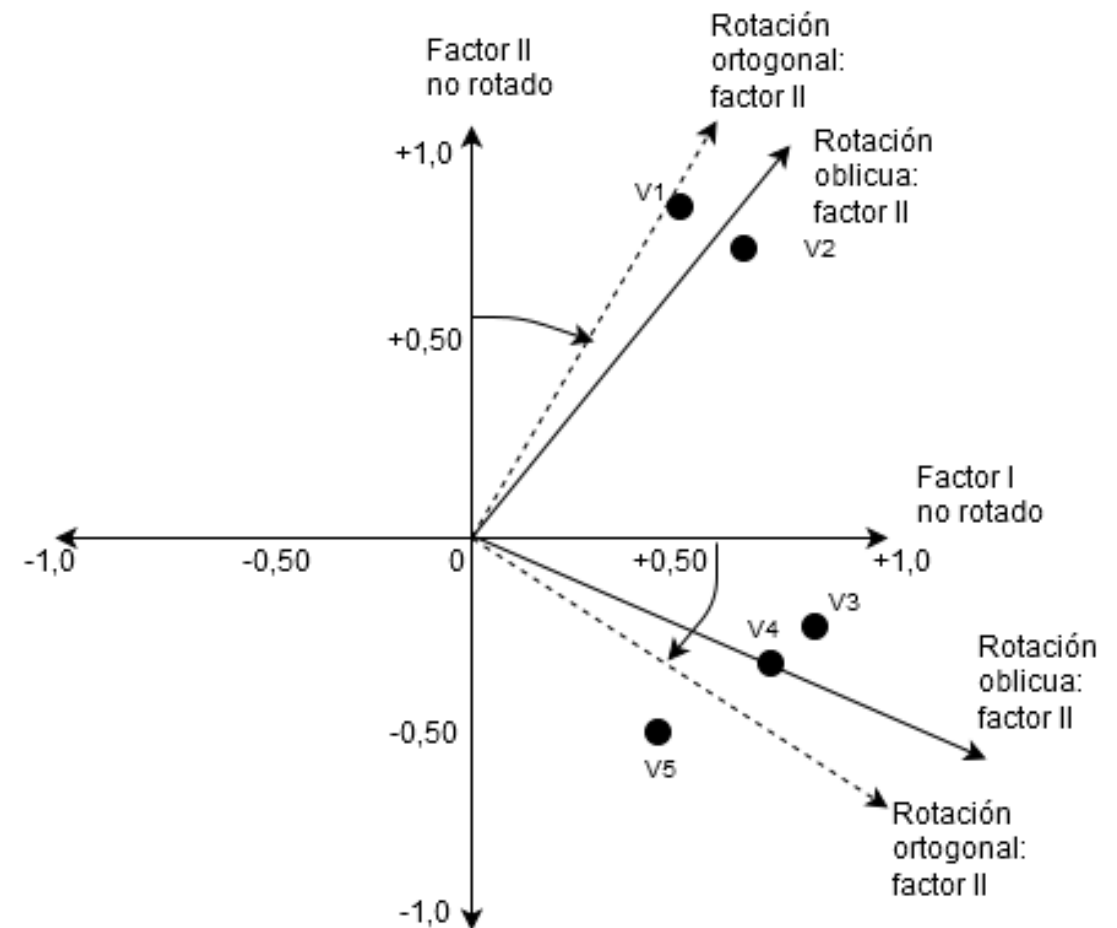
Cuando no se logra interpretar adecuadamente la matriz de cargas factoriales se realiza el procedimiento de rotación de factores que, busca obtener, a partir de la solución inicial, uno factores cuya matriz de cargas factoriales los haga más fácilmente interpretables, los más populares usando son el varimax para rotación ortogonal u promax para rotación oblicua

Rotación de factores

Ortogonal: Los ejes de rotación forman un ángulo 90 grados



Oblicua: Los ejes de rotación forman distintos ángulos



Tipos de rotación a emplear

Ortogonal

“none”,
“varimax”,
“quartimax”,
“bentlerT”,
“equamax”,
“varimin”,
“geominT”,
“bifactor”

No Ortogonal

“Promax”,
“promax”,
“oblimin”,
“simplimax”,
“bentlerQ”,
“geominQ”,
“biquartimin”,
“cluster”

Package ‘psych’

January 9, 2020

Version 1.9.12.31

Date 2019-12-31

Title Procedures for Psychological, Psychometric, and Personality
Research

Análisis Factorial Exploratoria

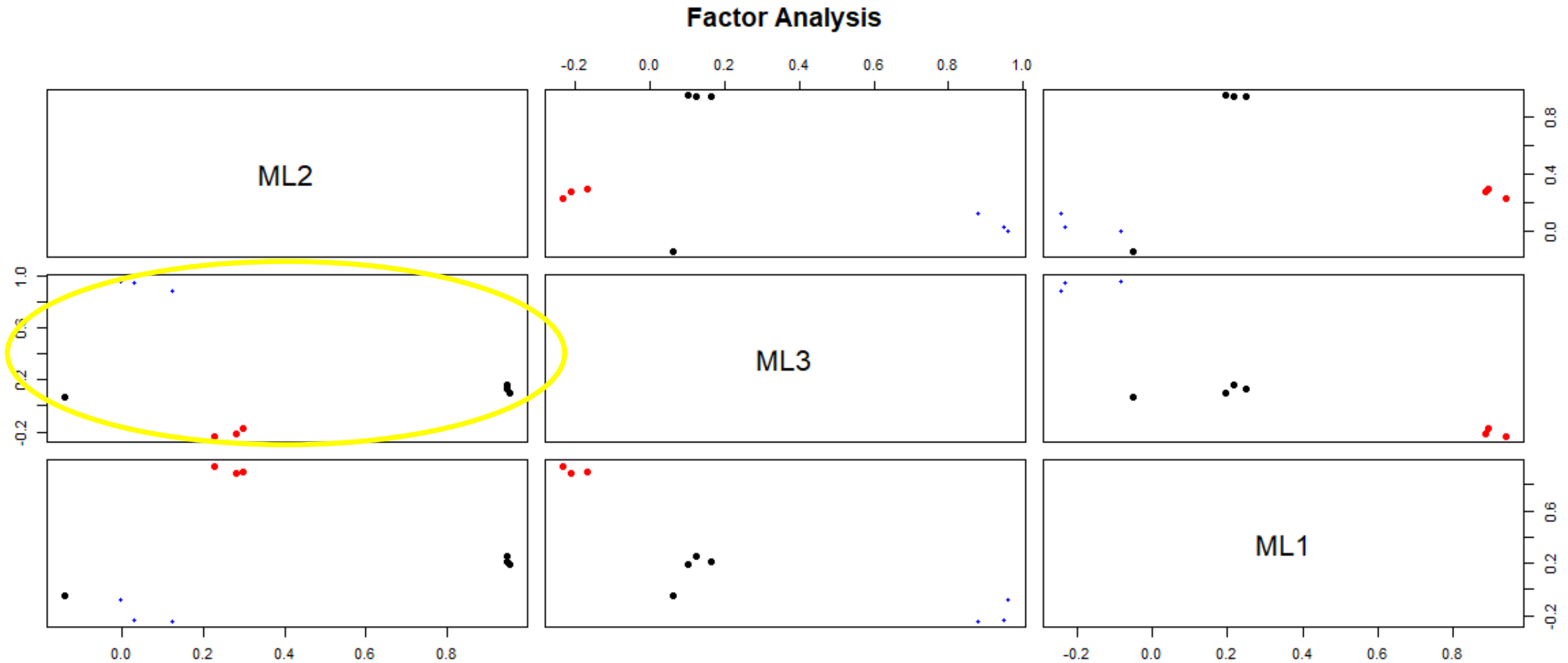
	item	ML2	ML3	ML1	h2	u2	com
X2	2	0.95			0.957	0.043	1.1
X9	9	0.95			0.976	0.024	1.2
X5	5	0.95			0.967	0.033	1.2
X6	6				0.027	0.973	1.7
X4	4		0.96		0.927	0.073	1.0
X8	8		0.95		0.955	0.045	1.1
X1	1		0.88		0.849	0.151	1.2
X7	7			0.94	0.995	0.005	1.2
X3	3			0.89	0.913	0.087	1.3
X10	10			0.89	0.911	0.089	1.3

X6 muy mal representado
tiene una muy baja
comunalidad (h2).

	ML2	ML3	ML1
SS loadings	2.96	2.78	2.74
Proportion Var	0.30	0.28	0.27
Cumulative Var	0.30	0.57	0.85
Proportion Explained	0.35	0.33	0.32
Cumulative Proportion	0.35	0.68	1.00

Contiene el 85% de todos
los datos de las 10
variables

Análisis Factorial Exploratoria

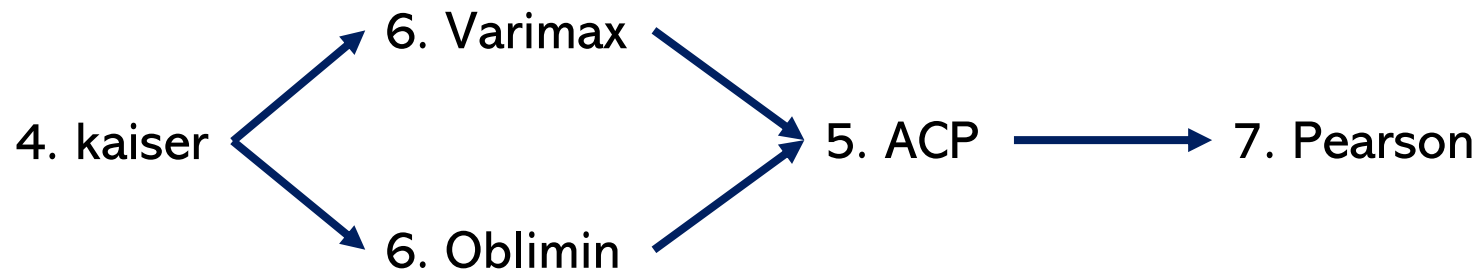


Análisis Factorial Exploratoria

Etapa del análisis (Continuación....)

7) Cálculo de puntuaciones factoriales.

Aquí se calculan los nuevos valores que tomarán los individuos en cada uno de los factores. Estos valores se pueden calcular a través de diferentes métodos como: método de regresión, el método de barlett, Método de Anderson-Rubin.



Puntuaciones factoriales.

	ML2	ML3	ML1
[1,]	-0.48782198	-1.3339874150	-0.149221425
[2,]	-1.12454665	-0.6006835035	-0.516881292
[3,]	0.51268812	-0.6225988984	0.642662522
[4,]	0.43338882	-0.3723605870	0.798196078
[5,]	1.39649405	0.9114274367	0.063624936
[6,]	-0.96784665	0.9415180339	0.740834499
[7,]	-0.21670854	1.2041357943	0.300865695
[8,]	-1.40747002	0.3602716421	0.943165564
[9,]	0.08599275	-1.1081117993	1.796773930
[10,]	0.88998723	-0.5131266743	1.042509620
[11,]	0.77265670	1.1018376670	-0.237858523
[12,]	0.26226568	1.1807634050	0.231798514
[13,]	-1.61891197	0.3715450495	-1.057088938
[14,]	-1.03780569	-1.0824980970	-0.726873588

Análisis Factorial Exploratoria

Etapa del análisis (Continuación....)

8) Validación del modelo.

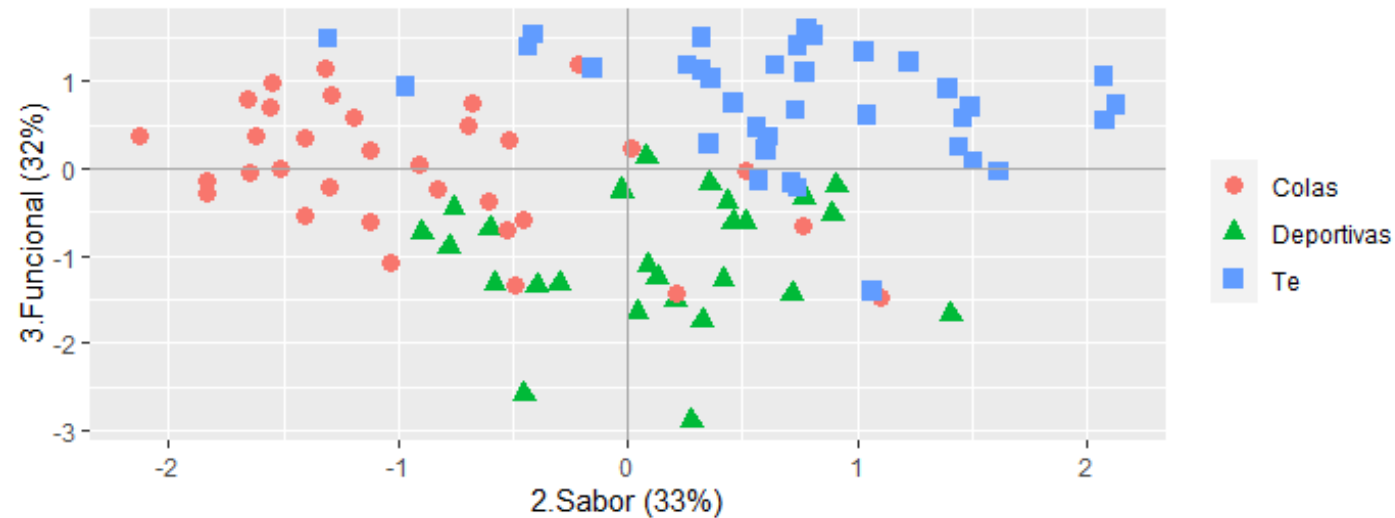
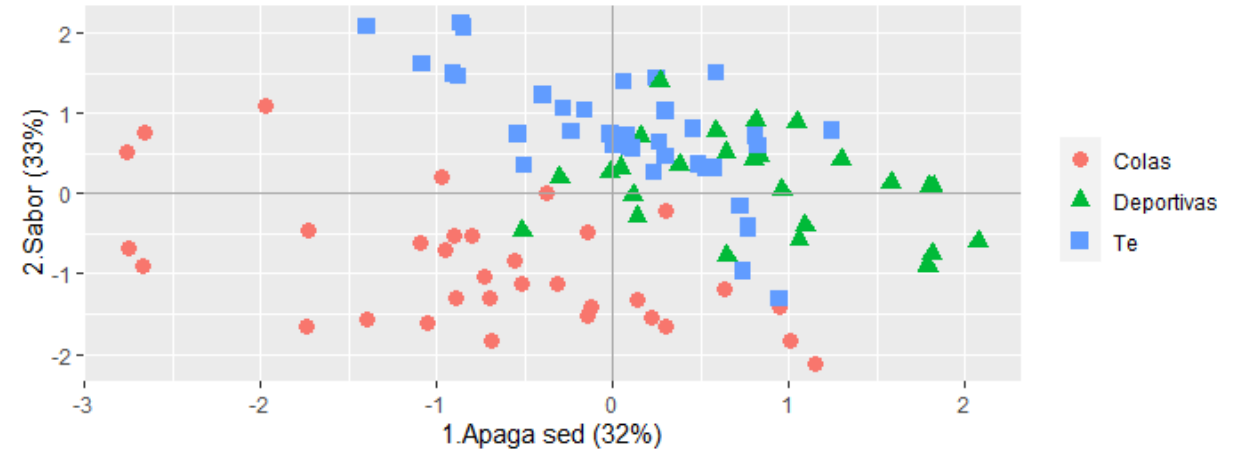
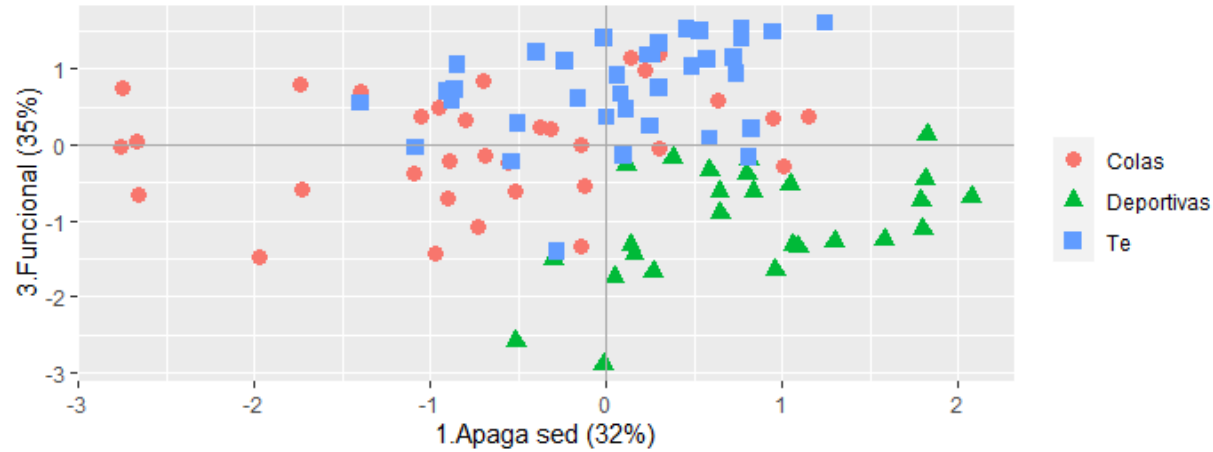
El ultimo paso en el AF es estudiar la validez del modelo. Dicha validación debe hacerse en dos direcciones: analizando la bondad de ajuste del mismo y la generalidad de sus conclusiones.

The root mean square of the residuals (RMSR) is 0.01
The df corrected root mean square of the residuals is 0.02

Teóricamente un modelo presenta una solución adecuada cuando el RSMR es menor o igual a 0.08.

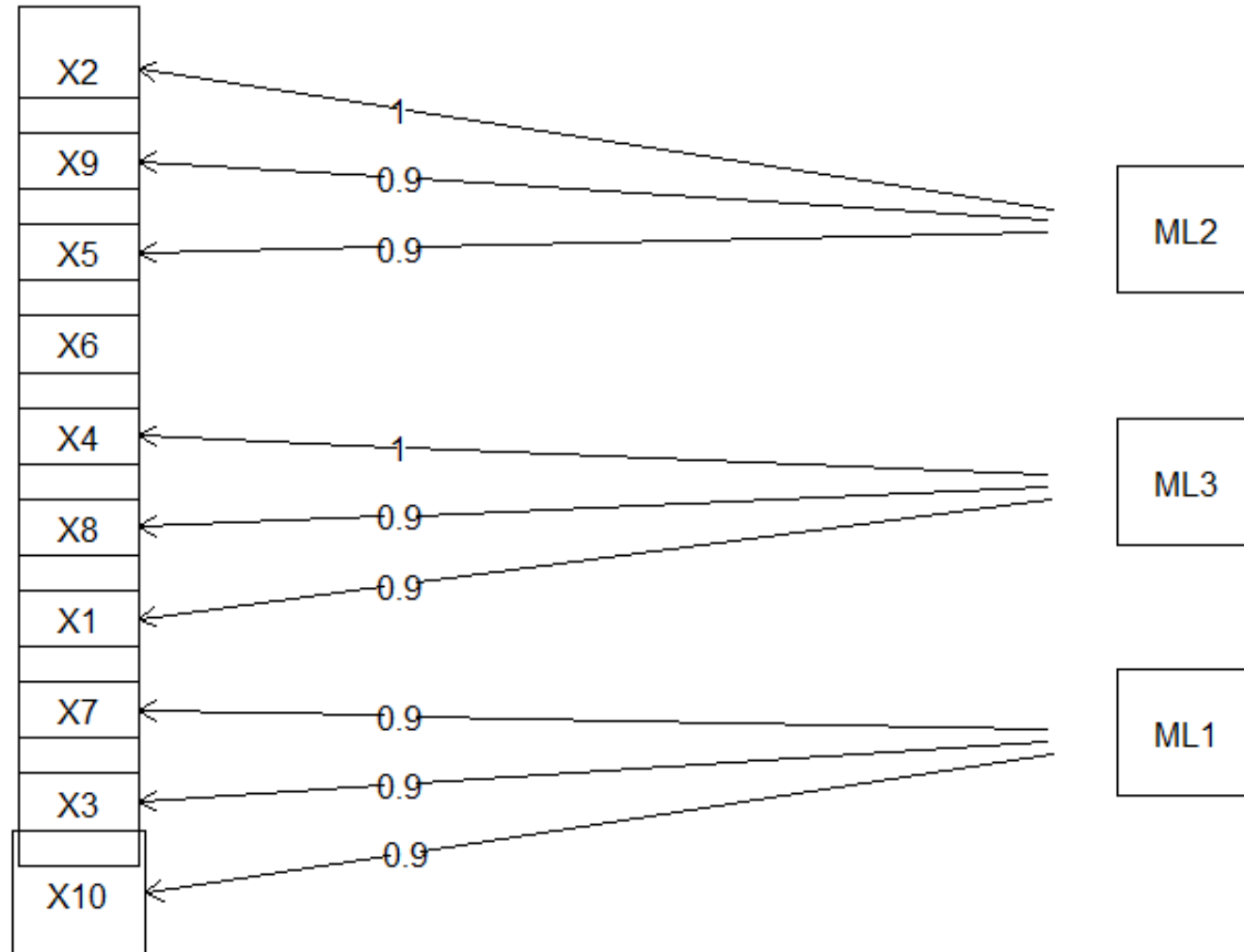
¿con cuál nos quedamos?
Aquel modelo que tenga la proporción Var más alta y el RMSR mas pequeño

Análisis Factorial Exploratoria



Graficando los factores

Factor Analysis



Percepciones de los consumidores sobre marcas de refrescos



Una empresa de estudios de mercado ha realizado una encuesta a 95 consumidores para determinar sus percepciones sobre seis marcas de refrescos que compiten entre sí. Las marcas son las siguientes⁴: (1) Pepsi; (2) Coca Cola; (3) Gatorade; (4) Allsport; (5) Lipton; (6) Nestea. Para ello los entrevistados respondieron en una escala donde 1=Totalmente en desacuerdo hasta 7=Totalmente de acuerdo a las preguntas que aparecen en el cuadro 12.16. El objetivo es evaluar los factores que subyacen en la configuración de la imagen de las marcas y, a la vez, obtener el mapa perceptual de las seis marcas.

Después de lo expuesto durante el tema, optaremos por una extracción de factores por ejes principales con una rotación varimax. Pero antes hemos de ser capaces de determinar (a) si los datos son factorizables, es decir, si la matriz de correlaciones es distinta de la identidad y (b) si la muestra se adecua global

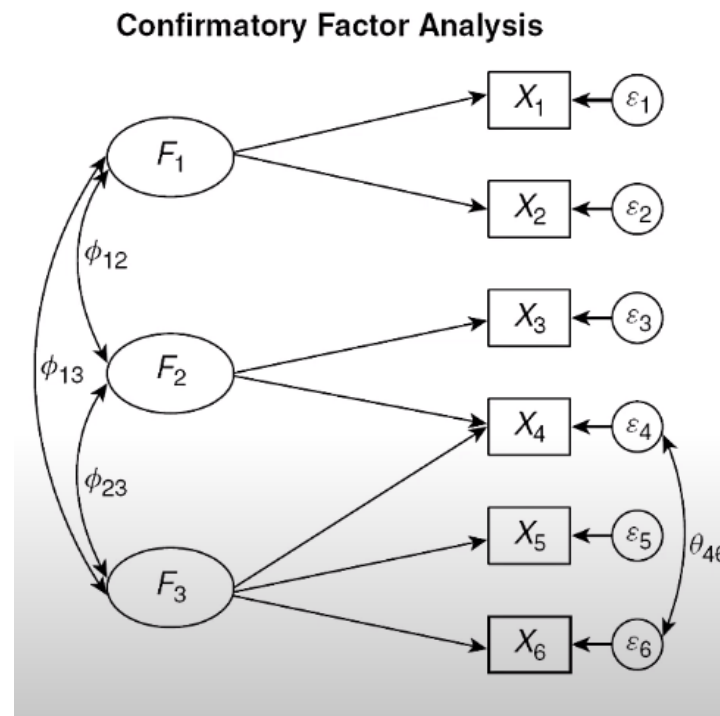
Percepciones de los consumidores sobre marcas de refrescos



Etiqueta	Enunciado / Descripción
X1	La marca X tiene un sabor refrescante
X2	Prefiero X porque tiene menos calorías que otras marcas
X3	La marca X apaga mi sed inmediatamente
X4	Me gusta el sabor dulce de la marca X
X5	Prefiero X después del ejercicio porque me da energía
X6	Prefiero X porque viene en un envase respetuoso con el medioambiente
X7	X tiene minerales y vitaminas y me recupera además de quitar la sed
X8	X tiene un sabor muy distinto a las demás
X9	X tiene la combinación justa de minerales y vitaminas para ser saludable
X10	Prefiero tomar la marca X cuando tengo sed
Marca	1=Pepsi; 2=Coca Cola; 3=Gatorade; 4=Allsport; 5=Lipton; 6=Nestea
ID	Etiqueta para identificar al entrevistado

Análisis Factorial Confirmatorio

Los factores son conocidos a priori, generalmente descritos en la teoría, y el objetivo descrito en la teoría, y el objetivo es comprobar si dicha estructura teórica previa se ajusta a los datos a través de contrastes de hipótesis.

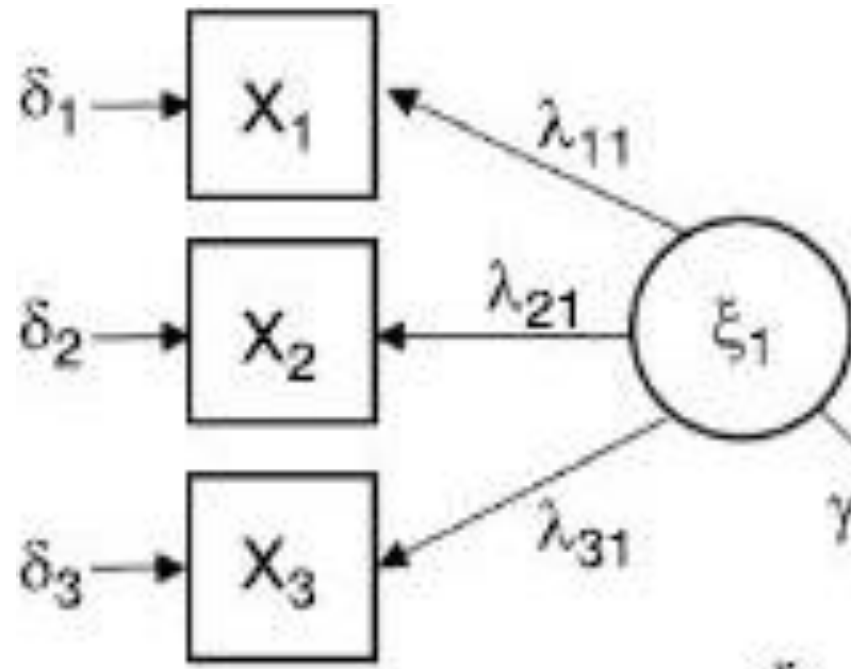


Introducción

- SEM es una combinación de análisis factorial con regresión lineal múltiple.

$$Y = \beta_0 + \beta_1 X + u$$

$$y = v + \Lambda \eta + \varepsilon$$

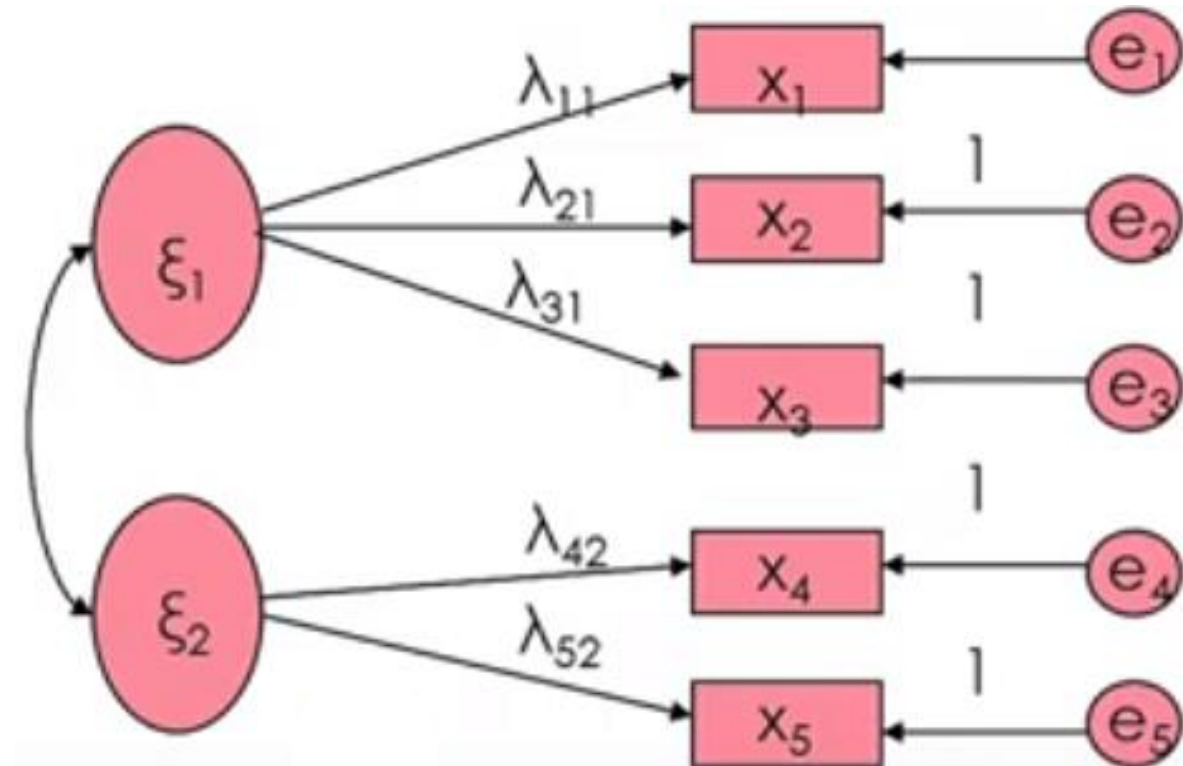


Pasos a desarrollar un SEM

- Especificidad del modelo Conceptual,
- Identificación del modelo matemático.
- Método de estimación
- Evaluación de la “Bondad de Ajuste”
- Re - especificación justificada.

Especificidad del modelo conceptual

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \end{bmatrix}$$



Identificación del modelo

- **Sub - identificado:** existe mas parámetros que datos (elementos de información) grados de libertad en Negativo.
- **Posiblemente identificado (saturado):** existe una correspondencia de 1 a 1 entre datos (elementos de información) y los parámetros . Tiene 0 grados de libertad.
- **Sobre – Identificado:** el numero de parámetros a estimarse es menor que el numero de datos (elementos de información). Resulta en: 0 grados de libertad en positivo.

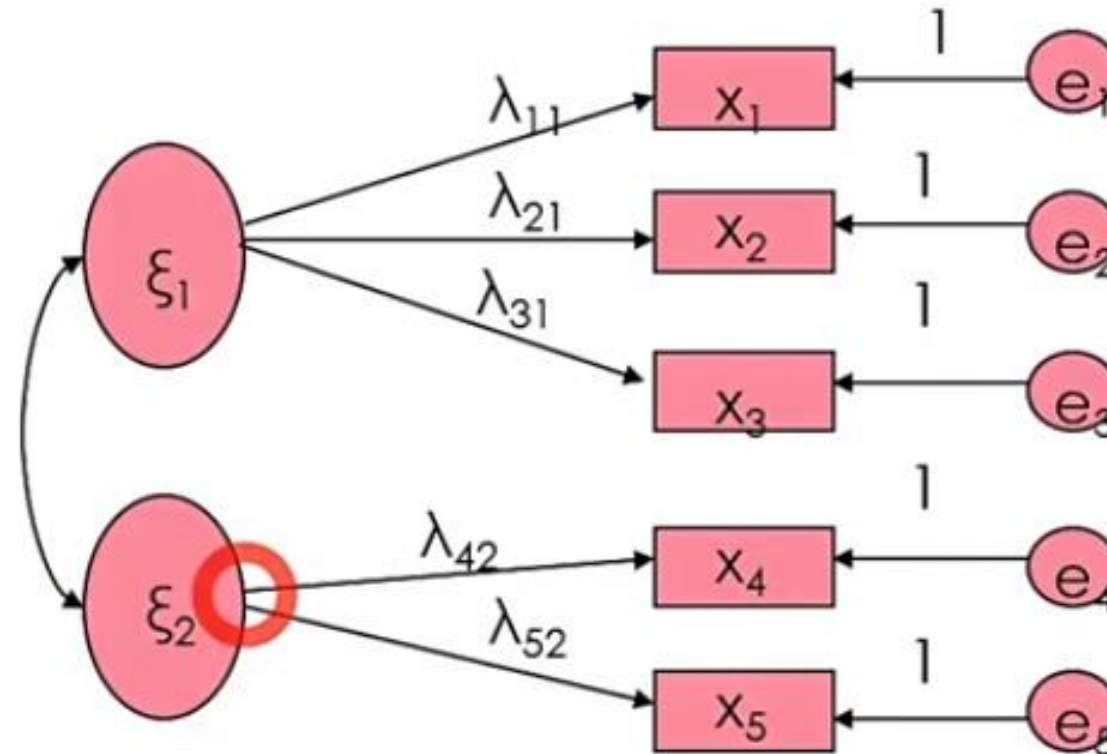
Identificación del modelo

$$Y = [p*(p+1)]/2$$

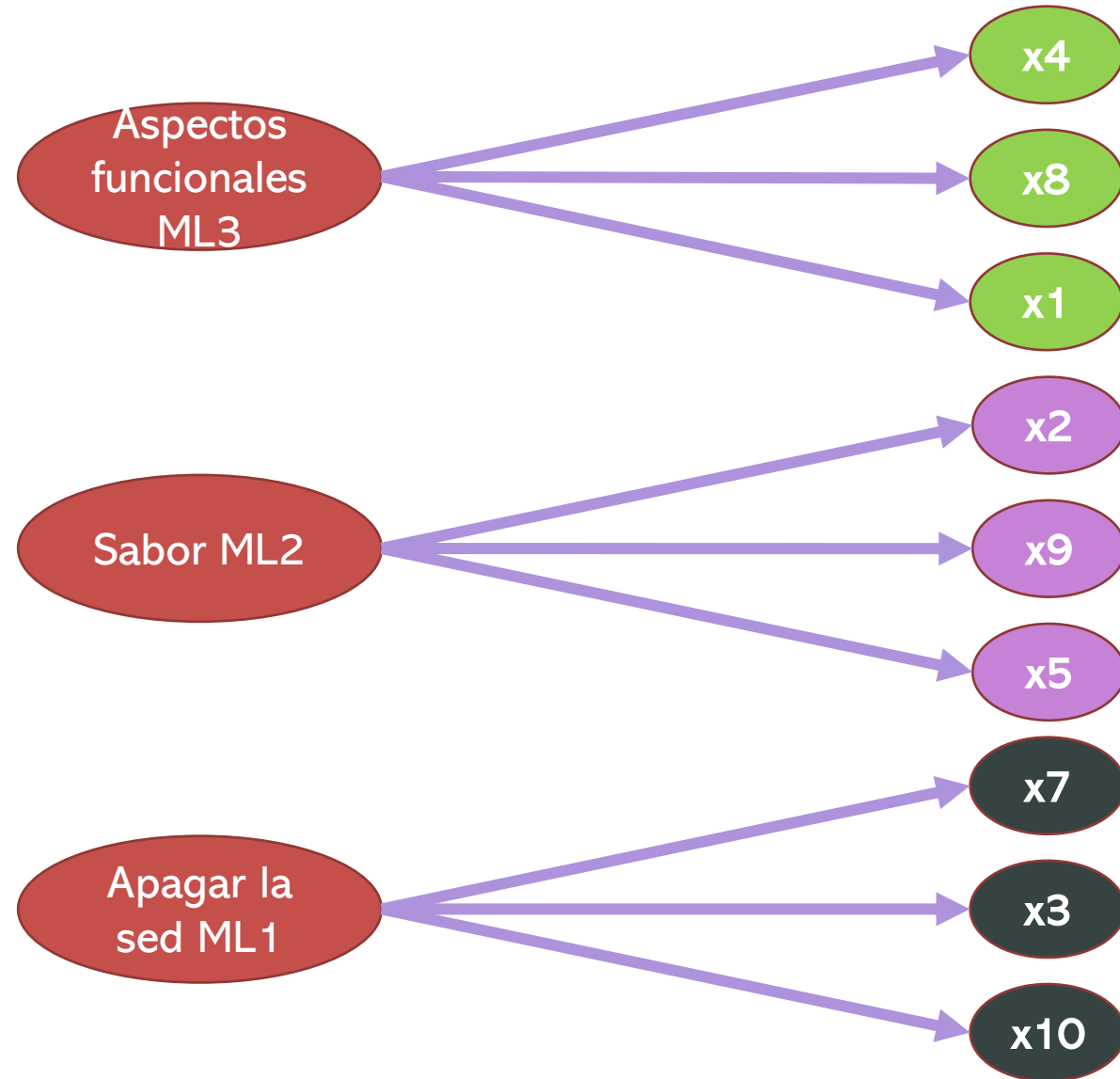
P=número de variables observadas

Z=variables a estimar. (no cuentan las covarianzas entre errores)

$$Y - Z = GI$$



Modelo FAC



Modelo FAC

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
ML_3 ≈				
X7	1.000			
X3	0.969	0.025	38.909	0.000
X10	0.969	0.021	46.099	0.000
ML_2 ≈				
X4	1.000			
X8	1.050	0.029	36.512	0.000
X1	0.973	0.029	33.765	0.000
ML_1 ≈				
X2	1.000			
X9	1.013	0.016	64.811	0.000
X5	1.006	0.014	74.276	0.000

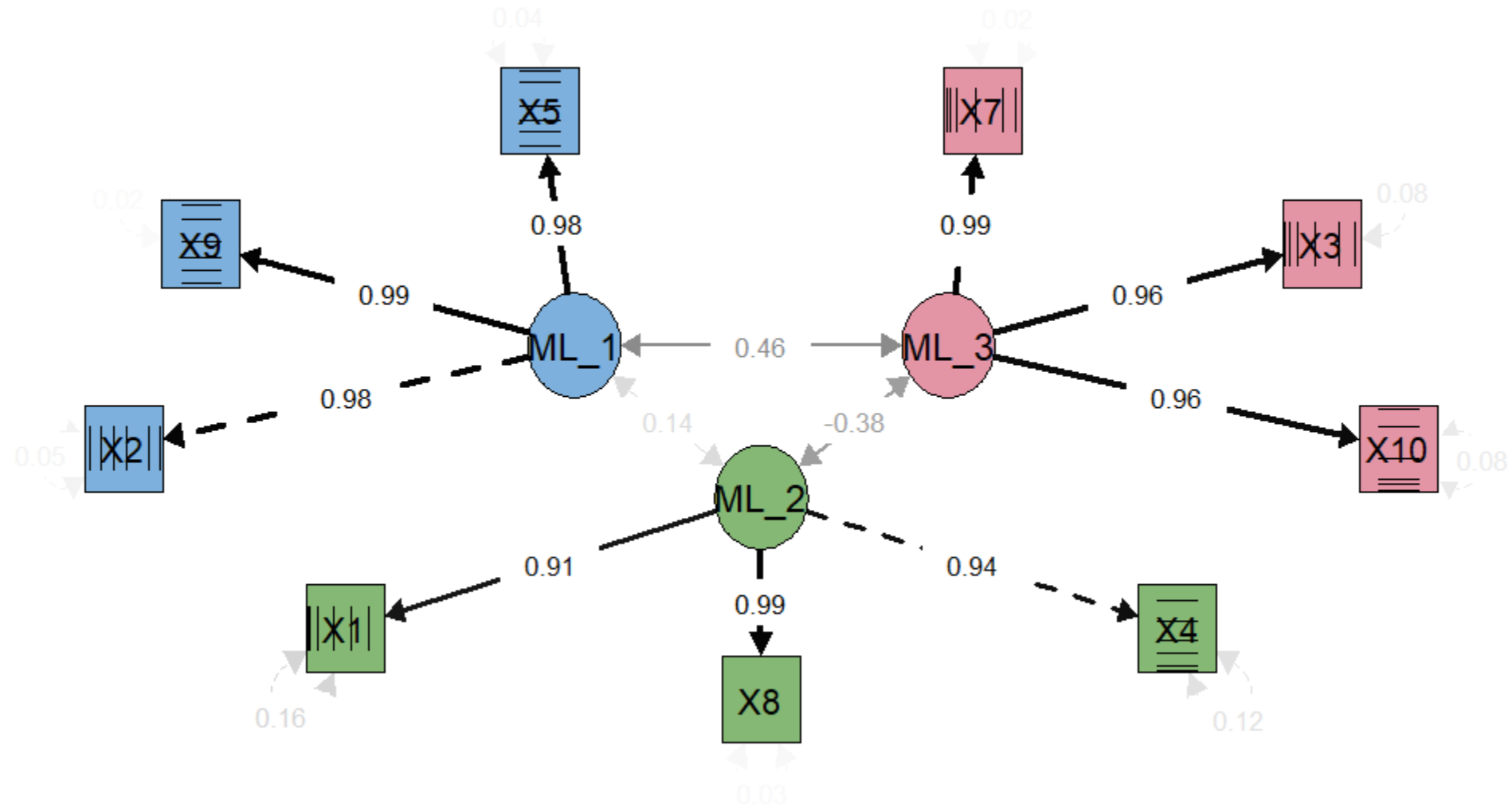
Notamos que las variables se relacionan con las variables latentes

Estamos corroborando si la matriz de varianzas y covarianzas muestrales y la teórica son iguales

	Standard	Robust
Test Statistic	7.612	20.489
Degrees of freedom	24	24
P-value (Chi-square)	0.999	0.669
Scaling correction factor		0.674
Shift parameter		9.197
simple second-order correction		

RMSEA	0.000	0.000
90 Percent confidence interval - lower	0.000	0.000
90 Percent confidence interval - upper	0.000	0.069
P-value RMSEA ≤ 0.05	1.000	0.870
Robust RMSEA		NA
90 Percent confidence interval - lower		0.000
90 Percent confidence interval - upper		NA
Standardized Root Mean Square Residual:		
SRMR	0.036	0.036

Modelo factorial



AFC.

Clima Organism.

Bueno
Malo
Regular

Dimensiones (AFE)

RRHH

Motivación

Comunicación.