

# Modificaciones del Análisis Paralelo para la Determinación del Número de Factores:

Un Estudio de Simulación

**Luis Garrido**

**Francisco José Abad**

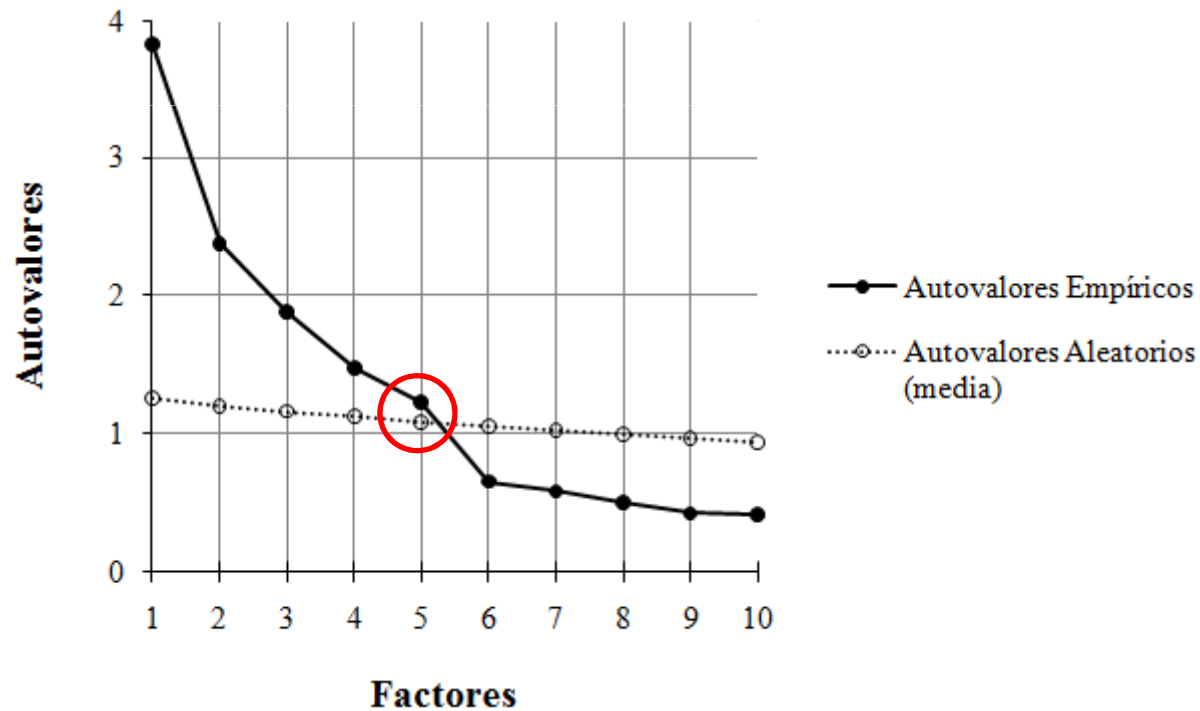
**Vicente Ponsoda**



# INTRODUCCIÓN

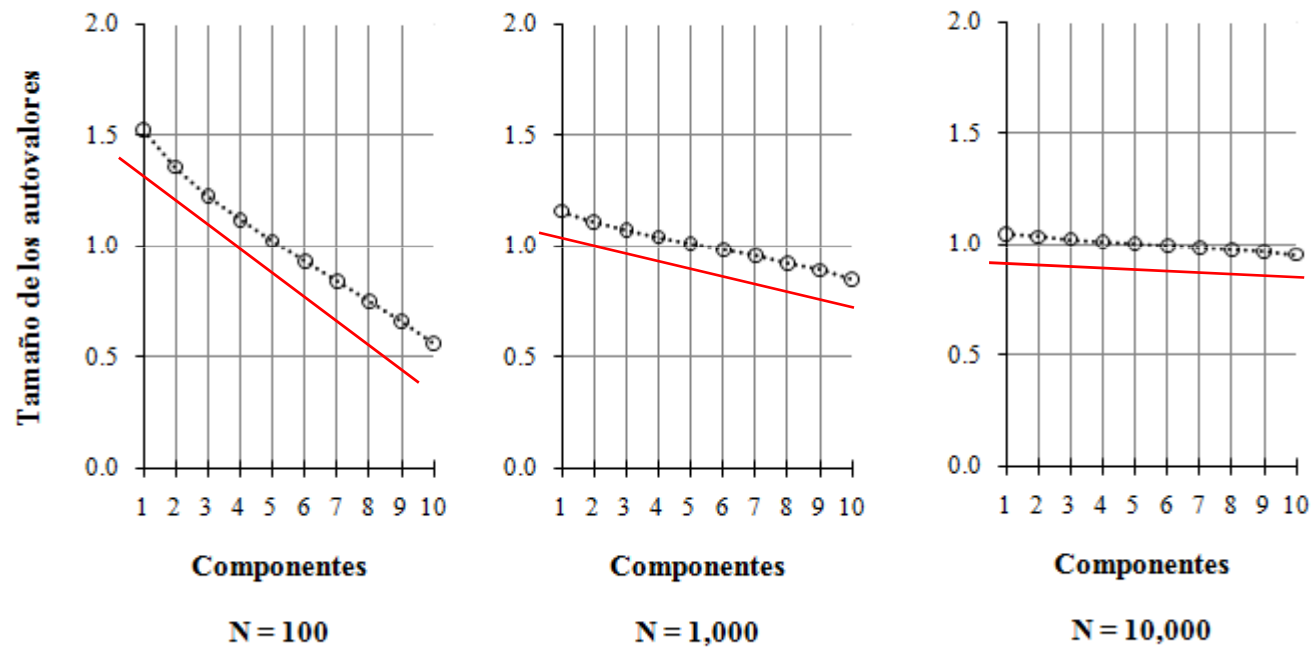
## EL ANÁLISIS PARALELO

- El análisis paralelo determina el número de factores al comparar los autovalores empíricos con los autovalores de datos aleatorios.



# INTRODUCCIÓN

## EL ANÁLISIS PARALELO



Notas. Número de variables = 10; Réplicas muestrales = 1,000.



# INTRODUCCIÓN

## ESTUDIOS DE SIMULACIÓN PREVIOS

- El análisis paralelo ha emergido como el método más eficaz de retención de factores para variables continuas<sup>1</sup>.
- Se han sugerido y estudiado diversas variaciones del análisis paralelo, algunas para todo tipo de variables<sup>2</sup> y otras para variables categóricas<sup>3</sup>.
- Estas variaciones aún no se han estudiado de forma conjunta mediante un diseño multifactorial y en el contexto de los demás métodos de retención de factores.

1. Mumford, Ferron, Hines, Hogarty & Kromrey (2003); Velicer, Eaton & Fava, (2000); Zwick & Velicer (1986).

2. Buja & Eyuboglu (1992); Hayton, Allen & Scarpello (2004). Crawford, Green, Levy, Lo, Scott, Svetina & Thompson (2008).

3. Weng & Cheng (2005); Liu & Rijmen (2008); Cho, Li & Bandalos (2009).



# INTRODUCCIÓN

## VARIACIONES SUGERIDAS PARA EL ANÁLISIS PARALELO

1. Media vs. percentil 95.
2. Componentes principales vs. análisis factorial.
3. Distribución continua vs. categórica.
4. Correlación Pearson vs. policórica.



# INTRODUCCIÓN

## MÉTODOS DE RETENCIÓN DE FACTORES A ESTUDIAR

1. Análisis paralelo (AP)<sup>1</sup>.
2. Regla de Kaiser (KAISER)<sup>2</sup>.
3. Standard error scree (SESCREE)<sup>3</sup>.
4. Minimum average partial (MAP)<sup>4</sup>.

1. Horn (1965).
2. Kaiser (1960).
3. Zoski & Jurs (1996).
4. Velicer (1976).



# INTRODUCCIÓN

## CRITERIOS DE EFICACIA UTILIZADOS

1. Proporción de aciertos.
2. Media de la diferencia algebraica entre el valor estimado y el real.

$$\frac{\sum_{i=1}^n \# \text{ factores estimados} - \# \text{ factores reales}}{N}$$

3. Desviación típica respecto al valor real

$$\sqrt{\frac{\sum_{i=1}^n (\# \text{ factores estimados} - \# \text{ factores reales})^2}{N}}$$

- Estos son los criterios utilizados por Velicer, Eaton y Fava (2000).



# MÉTODO

## DISEÑO FACTORIAL PARA LOS DATOS GENERADOS

Factores manipulados		Nivel 1	Nivel 2	Nivel 3	Nivel 4
Tamaño muestral	N	150	300	500	
Saturación	Sat	.40	.55	.70	
Variables por factor	VarFac	4	8	12	
Número de factores	Fac	2	5	8	
Correlación factorial	CorrFac	.00	.20	.40	
Categorías de respuesta*	CatRes	2	3	4	5

*Nota.* \* Sólo para datos categóricos.

I) Diseño para datos continuos:  $3 \times 3 \times 3 \times 3 \times 3 = 243$  combinaciones.

Réplicas muestrales = 100.

Total de casos = 24,300.

II) Diseño para datos categóricos:  $3 \times 3 \times 3 \times 3 \times 3 \times 4 = 972$  combinaciones.

Réplicas muestrales = 25.

Total de casos = 24,300.





# MÉTODO

## DISEÑO FACTORIAL PARA LOS MÉTODOS ESTUDIADOS

Método	Factores Manipulados			
	Corr	F2	F3	F4
AP1	P	DCont	CP	media
AP2	P	DCont	CP	95%
AP3	P	DCont	AF	media
AP4	P	DCont	AF	95%
AP5	P	DCat	CP	media
AP6	P	DCat	CP	95%
AP7	P	DCat	AF	media
AP8	P	DCat	AF	95%
AP9	PO	DCat	CP	media
AP10	PO	DCat	CP	95%
AP11	PO	DCat	AF	media
AP12	PO	DCat	AF	95%
KAISER1	P			
KAISER2	PO			
SESCREE1	P			
SESCREE2	PO			
MAP1	P	P2		
MAP2	P	P4		
MAP3	PO	P2		
MAP4	PO	P4		

Leyenda	
Corr	= Correlación
F	= Factor
P	= Pearson
PO	= Policórica
P2	= 2da Potencia
P4	= 4ta Potencia
DCont	= Distribución Continua
DCat	= Distribución Categórica
CP	= Componentes Principales
AF	= Análisis Factorial



# RESULTADOS

## EFFECTIVIDAD PARA VARIABLES CONTINUAS

Método	Factores Manipulados		Criterios de Eficacia		
	F1	F2	Aciertos	Diferencia	Desviación
AP1	CP	media	→ .88	-.11	.77
AP2	CP	95%	.87	-.27	1.04
AP3	AF	media	.78	.35	1.17
AP4	AF	95%	.86	-.03	.84
KAISER			.38	5.28	8.92
SESCREE			.46	7.83	15.90
MAP1	P2		→ .65	-1.32	2.63
MAP2	P4		.55	-1.35	2.46

Leyenda
F = Factor
P2 = 2da Potencia
P4 = 4ta Potencia
CP = Componentes Principales
AF = Análisis Factorial



# RESULTADOS

## EFFECTIVIDAD PARA VARIABLES CATEGÓRICAS

Método	Factores Manipulados				Criterios de Eficacia		
	Corr	F2	F3	F4	Aciertos	Diferencia	Desviación
AP1	P	DCont	CP	media	→ .78	-.10	1.17
AP2	P	DCont	CP	95%	.77	-.48	1.50
AP3	P	DCont	AF	media	.63	.71	2.00
AP4	P	DCont	AF	95%	.74	-.10	1.38
AP5	P	DCat	CP	media	.78	-.10	1.17
AP6	P	DCat	CP	95%	.78	-.48	1.50
AP7	P	DCat	AF	media	.63	.71	2.02
AP8	P	DCat	AF	95%	.75	-.10	1.37
AP9	PO	DCat	CP	media	.77	-.10	1.18
AP10	PO	DCat	CP	95%	.77	-.49	1.52
AP11	PO	DCat	AF	media	.57	2.97	11.00
AP12	PO	DCat	AF	95%	.68	-.48	2.72
KAISER1	P				→ .21	7.28	10.83
KAISER2	PO				.22	7.24	10.81
SESCREE1	P				→ .38	8.35	16.80
SESCREE2	PO				.35	12.92	24.53
MAP1	P	P2			→ .50	-1.98	3.24
MAP2	P	P4			.40	-2.00	3.09
MAP3	PO	P2			.50	2.45	13.27
MAP4	PO	P4			.50	-.85	7.84

Leyenda	
Corr	= Correlación
F	= Factor
P	= Pearson
PO	= Policórica
P2	= 2da Potencia
P4	= 4ta Potencia
DCont	= Distribución Continua
DCat	= Distribución Categórica
CP	= Componentes Principales
AF	= Análisis Factorial



# RESULTADOS

## REGRESIÓN LOGÍSTICA SOBRE LA PROPORCIÓN DE ACIERTOS

Factores manipulados	Métodos de Retención de Factores			
	AP	KAISER	SESCREE	MAP
Tamaño muestral	.08	.08	.05	.01
Saturación	.31	.31	.09	.42
Variables por factor	.04	.19 *	.08 *	.33
Número de factores	.11 *	.18 *	.33 *	.02 *
Correlación factorial	.02 *			.03 *
Categorías de respuesta	.04	.08	.01	.06
Total	.60	.84	.56	.87

*Nota.* El estadístico utilizado es el  $R^2$  de Nagelkerke.

\* La relación es negativa.



# RESULTADOS

## EJEMPLO EFECTIVIDAD PARA DIFERENTES COMBINACIONES

Método	Criterios de Eficacia		
	Aciertos	Diferencia	Desviación
<u>Caso 1: Sat = .55 &amp; VarFac = 4</u>			
MAP	.00	-4.48	5.02
SESCREE	.63	-.04	1.26
AP	.74	-.44	1.24
KAISER	.36	1.57	2.23
<u>Caso 2: Sat = .55 &amp; VarFac = 12</u>			
MAP	.94	-.07	.39
SESCREE	.25	19.28	27.27
AP	.96	.00	.26
KAISER	.01	13.63	16.06



# DISCUSIÓN

- El análisis paralelo clásico (Pearson+DistCont +CP+Media) tiene un funcionamiento igual o mejor que las demás versiones del método.
- El análisis paralelo es más efectivo según los 3 criterios utilizados que el MAP, KAISER y SESCREE.
- No se recomienda el uso generalizado del MAP, KAISER o SESCREE dada su gran variabilidad y/o poca efectividad.



# Modificaciones del Análisis Paralelo para la Determinación del Número de Factores:

Un Estudio de Simulación

# Muchas Gracias



# REFERENCIAS

- Buja, A. & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research*, 27, 509-540.
- Cho, S.-J. , Li, F., and Bandalos, D. L. (In press). Accuracy of the parallel analysis procedure using polychoric correlations. *Educational and Psychological Measurement*.
- Crawford, A. V., Green, S. B., Levy, R., Lo, W. J., Scott, L., Svetina, D., & Thompson, M. S. (2008). Evaluation of parallel analysis methods for determining the number of factors. Manuscript under review.
- Hayton, J. C., Allen, D. G. & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis. *Organizational Research Methods*, 7, 191-205.
- Horn, J. L. (1965). A Rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141-151.
- Liu, O. L. & Rijmen, F. (2008). A modified procedure for parallel analysis of ordered categorical data. *Behavior Research Methods*, 40, 556-562.
- Mumford, K. R., Ferron, J. M., Hines, C. V., Hogarty, K. Y., y Kromrey, J. D. (2003, abril). *Factor retention in exploratory factor analysis: A comparison of alternative methods*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.
- Velicer, W. F. (1976) . Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, 321-327.





# REFERENCIAS

- Velicer, W. F., Eaton, C. A., y Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. En Goffin, R. D., y Helmes, E. (Eds.), *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy* (pp. 41-71). Boston: Kluwer.
- Weng, L. J. & Cheng, Ch. P. (2005) Parallel Analysis with Unidimensional Binary Data. *Educational and Psychological Measurement*, 65, 697-716 .
- Zoski, K.,&Jurs, S. (1996). An objective counterpart to visual test for factor analysis: The standard error scree. *Educational and Psychological Measurement*, 56, 443-451.
- Zwick, W. R., y Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99, 432-442.

