

Ejercicio Tema 7: Análisis Factorial

Problema y objetivos de investigación

Identificar si en la medida de la inteligencia de los sujetos, en lugar de tener que trabajar con las notas de las seis asignaturas, estas pueden estar contenidas en un número inferior de factores que las integren.

Con la información del enunciado y los datos de un fichero SPSS con las características siguientes:

- Casos: 200
- Variables: v1 (Castellano), v2 (Inglés), v3 (Historia), v4 (Aritmética), v5 (Álgebra) y v6 (Física)

Se pide:

1. Analizar la adecuación de la técnica al objetivo propuesto
2. Analizar la aplicabilidad del análisis factorial a los datos existentes
3. Determinar si es razonable el número de factores obtenidos
4. Interpretar los factores obtenidos de acuerdo con el objetivo de la investigación

```
FACTOR
/MATRIX=IN(COR=*)
/VARIABLES v1 v2 v3 v4 v5 v6 /MISSING LISTWISE /ANALYSIS v1 v2 v3 v4 v5 v6
/PRINT UNIVARIATE INITIAL CORRELATION SIG KMO AIC EXTRACTION ROTATION
/CRITERIA MINEIGEN(1) ITERATE(25)
/EXTRACTION PC
/CRITERIA ITERATE(25)
/ROTATION VARIMAX
/METHOD=CORRELATION .
```

Adecuación del análisis factorial al objetivo

Como técnica de reducción de datos, en principio la técnica es adecuada al objetivo ya que permitiría reducir el número de variables explicativas y encontrar alguna estructura común entre algunas de ellas.

Aplicabilidad del análisis

Las variables a analizar son continuas, requisito para la aplicación de esta técnica

El tamaño de la muestra es de 220 casos, superior al requerimiento de 100 para la aplicabilidad.

La relación observaciones a variables es de 36 a 1, muy superior a la de 10 a 1 requerida.

Correlation Matrix

		V1	V2	V3	V4	V5	V6
Correlation	V1	1,000	,439	,410	,288	,329	,248
	V2	,439	1,000	,351	,354	,320	,329
	V3	,410	,351	1,000	,164	,190	,181
	V4	,288	,354	,164	1,000	,595	,470
	V5	,329	,320	,190	,595	1,000	,464
	V6	,248	,329	,181	,470	,464	1,000
Sig. (1-tailed)	V1		,000	,000	,000	,000	,000
	V2	,000		,000	,000	,000	,000
	V3	,000	,000		,007	,002	,004
	V4	,000	,000	,007		,000	,000
	V5	,000	,000	,002	,000		,000
	V6	,000	,000	,004	,000	,000	

En la matriz de correlaciones se observan 10 de las 15 como mayores a 0,3 y observando la significatividad solo 2 no serían significativas (v4 con v3 y v6 con v3), por lo que en principio la cantidad de correlaciones parece ser suficiente como para que por lo menos algunas de ellas se agrupen significativamente en factores representativos.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,775
Bartlett's Test of Sphericity	Approx. Chi-Square	310,841
	df	15
	Sig.	,000

Tanto el Test de Esfericidad de Bartlett que resulta significativo ($p=0,000$) como la medida de adecuación muestral KMO cuyo índice (0,775) supera el valor mínimo requerido de 0,5 dan prueba adicional de la aplicabilidad de la técnica.

Observando la medida de adecuación muestral en la matriz anti-imagen se observa que todas las variables tienen valores superiores a 0,5, con lo que todas ellas son igualmente susceptibles de ser incluidas en el análisis.

Análisis

Aplicando el método de Componentes Principales para lograr la máxima reducción de datos se obtienen dos factores.

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	2,733	45,548	45,548	2,733	45,548	45,548	2,091	34,853	34,853
2	1,130	18,830	64,378	1,130	18,830	64,378	1,771	29,525	64,378
3	,615	10,253	74,630						
4	,601	10,020	84,651						
5	,525	8,747	93,397						
6	,396	6,603	100,000						

Extraction Method: Principal Component Analysis.

Observando los autovalores se ve que los componentes 1 y 2 tienen valores mayores a 1, aunque este procedimiento puede resultar en la extracción de pocos factores para análisis con menos de 20 variables como es este caso.

En el % de varianza explicada se ve que los factores 1 y 2 acumulan el 64.378% de la misma, superando el 60% e indicando que explicarían razonablemente las variables agrupadas.

Observando el gráfico de sedimentación también se ve que los factores 1 y 2 están antes del “valle”.

Obtención de los factores

Rotated Component Matrix^a

	Component	
	1	2
V1	,225	,764
V2	,349	,660
V3	-2,589E-03	,821
V4	,833	,147
V5	,814	,180
V6	,750	,154

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

a. Rotation converged in 3 iterations.

De la observación de las cargas factoriales rotadas se puede ver que el factor 1 está compuesto por las variables v4 (Aritmética), v5 (Álgebra) y v6 (Física) y el factor 2 está compuesto por las variables v1 (Castellano), v2 (Inglés) y v3 (Historia). Todas las variables contribuyen a los mencionados factores con cargas superiores a 0,40, valor indicado para tamaños muestrales de 200 casos.

Tanto en el factor 1 como en el 2 las cargas de las variables son todas positivas, lo que indica que además estas contribuyen en igual sentido.

Communalities

	Initial	Extraction
V1	1,000	,634
V2	1,000	,558
V3	1,000	,674
V4	1,000	,716
V5	1,000	,694
V6	1,000	,586

Extraction Method: Principal Component Analysis.

Se observa también que la varianza de cada variable recogida por los factores es representativa en todos los casos.

Resultados

1. La técnica del análisis factorial es adecuada para los objetivos planteados.
2. La misma también resulta aplicable una vez analizadas las condiciones de la muestra y los datos.

3. La cantidad de factores obtenidos (2) es razonable tanto según indicadores estadísticos como desde la interpretación de los mismos

Interpretación

El factor 1 agrupa las 3 variables que miden inteligencia en matemáticas (aritmética, álgebra y Física), se ve que todas se presentan y comportan en un mismo sentido y que son adecuadamente representadas por el factor. Por ello, podemos interpretar a este factor como de inteligencia analítica o matemática.

El factor 2 agrupa a las 3 variables que miden inteligencia en lenguaje (castellano, inglés e historia) y también se ve que se comportan similarmente y están representadas adecuadamente. Así, podemos interpretar a este factor como de inteligencia lingüística o verbal.

Validación del modelo

Al contarse con la matriz de correlaciones y no con la muestra original no se puede realizar una validación del modelo repitiendo el análisis en mitades de la muestra.