

# Capítulo 20

## Análisis factorial: El procedimiento *Análisis factorial*

### Introducción

El análisis factorial es una técnica de reducción de datos que sirve para encontrar grupos homogéneos de variables a partir de un conjunto numeroso de variables. Esos grupos homogéneos se forman con las variables que correlacionan mucho entre sí y procurando, inicialmente, que unos grupos sean independientes de otros.

Cuando recogemos un gran número de variables de forma simultánea, como por ejemplo en un cuestionario de satisfacción laboral, podemos estar interesados en averiguar si las preguntas del cuestionario se agrupan de alguna forma característica. Aplicando un análisis factorial a las respuestas de los sujetos podemos encontrar grupos de variables con significado común y conseguir de esta manera reducir el número de dimensiones necesarias para explicar las respuestas de los sujetos.

El análisis factorial es, por tanto, una técnica de reducción de la dimensionalidad de los datos. Su propósito último consiste en buscar el número mínimo de dimensiones capaces de explicar el máximo de información contenida en los datos.

A diferencia de lo que ocurre en otras técnicas como el análisis de varianza o el de regresión, en el análisis factorial todas las variables del análisis cumplen el mismo papel: todas ellas son *independientes* en el sentido de que no existe *a priori* una dependencia conceptual de unas variables sobre otras.

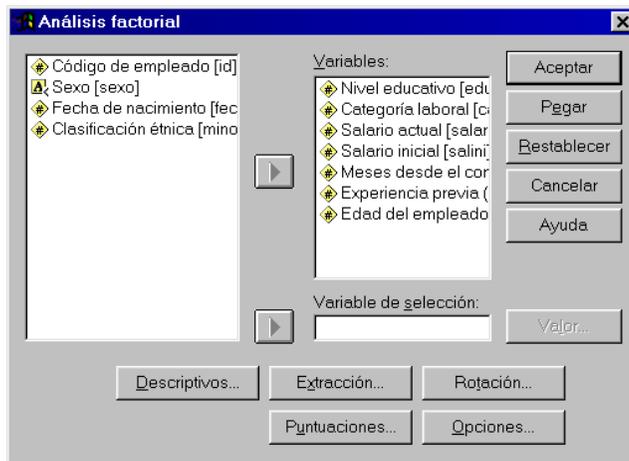
## Análisis factorial

El análisis factorial consta de cuatro fases características: el cálculo de una matriz capaz de expresar la variabilidad conjunta de todas las variables, la extracción del número óptimo de factores, la rotación de la solución para facilitar su interpretación y la estimación de las puntuaciones de los sujetos en las nuevas dimensiones. Para ejecutar correctamente un análisis factorial será necesario tomar algunas decisiones en cada una de estas fases. La estructura del procedimiento **Análisis factorial** del SPSS se ajusta a las cuatro fases mencionadas. Este capítulo explica cuáles son las especificaciones mínimas para obtener una solución inicial y cuáles son las opciones disponibles para personalizar la ejecución del procedimiento.

Para llevar a cabo un análisis factorial:

- ▶ Seleccionar la opción **Reducción de datos > Análisis factorial...** del menú **Analizar** para acceder al cuadro de diálogo *Análisis factorial* que se muestra en la figura 20.1.

**Figura 20.1.** Cuadro de diálogo *Análisis factorial*.



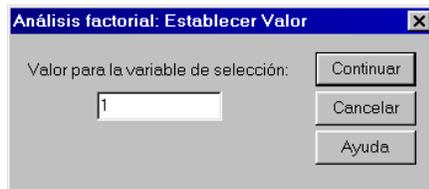
La lista de variables del archivo de datos contiene un listado de todas las variables del archivo, incluidas las variables de cadena (aunque éstas sólo pueden utilizarse como variables de *selección*). Para llevar a cabo un análisis factorial:

- ▶ Seleccionar el conjunto de variables que se desea analizar y trasladarlas a la lista **Va-riables**.

**Variable de selección.** Este cuadro permite seleccionar una de las variables del archivo de datos como variable de *filtro*: para definir una sub-muestra de sujetos que cumplan una determinada condición. Esta opción es especialmente útil cuando se ha reservado un porcentaje de los sujetos de la muestra para llevar a cabo una *validación cruzada* del modelo final. Para utilizar una variable de selección:

- ▶ Trasladar la variable al cuadro **Variable de selección** y pulsar en el botón **Valor...** para acceder al subcuadro de diálogo que muestra la figura 20.2.

**Figura 20.2.** Subcuadro de diálogo *Análisis factorial: Establecer valor*.



- ▶ Introducir en el cuadro de texto el valor de la variable de selección que identifica a los casos que se desea incluir en el análisis. (Para aprovechar al máximo las posibilidades de esta opción, podemos utilizar el proceso *Ejecutar casos no seleccionados*, el cual actúa sobre la tabla de notas del análisis conmutando los casos seleccionados por los no seleccionados y volviendo a ejecutar el procedimiento **Análisis factorial** con las mismas especificaciones establecidas en el análisis precedente).

## Ejemplo (Análisis factorial)

Este ejemplo muestra cómo ejecutar el procedimiento **Análisis factorial** con las especificaciones que el programa tiene establecidas por defecto. Vamos a comprobar si es posible resumir, mediante un número reducido de dimensiones o factores, la información disponible sobre las características laborales de un conjunto de empleados de banca (archivo de datos: *Datos de empleados.sav*). Para ello:

- ▶ En el cuadro de diálogo *Análisis factorial* (ver figura 20.1), seleccionar las variables *educ*, *catlab*, *salario*, *salini*, *tiempemp*, *expprev* y *edad* y trasladarlas a la lista **Variab-les**. (La variable *edad* se ha creado a partir de la variable *fechnac* mediante la expresión «*edad* = CTIME.DAYS(DATE.DMY(31,12,1997) – *fechnac*)/365», obteniendo así la edad en años a fecha 31/12/1997).

Aceptando estas selecciones, el *Visor* ofrece los resultados que muestran las tablas 20.1 a la 20.3.

La tabla 20.1 contiene las *comunalidades* asignadas inicialmente a las variables (*inicial*) y las comunalidades reproducidas por la solución factorial (*extracción*). La comunalidad de una variable es la proporción de su varianza que puede ser explicada por el modelo factorial obtenido. Estudiando las comunalidades de la extracción podemos valorar cuáles de las variables son peor explicadas por el modelo. En nuestro ejemplo, la variable *nivel educativo* es la peor explicada: el modelo sólo es capaz de reproducir el 68,2% de su variabilidad original.

En una nota a pie de tabla se indica que, para llegar a esta solución factorial, se ha utilizado un método de extracción denominado *componentes principales*. Dicho método de extracción, que es el que actúa por defecto, asume que es posible explicar el 100% de la varianza observada y, por ello, todas las comunalidades iniciales son iguales a la unidad (que es justamente la varianza de una variable en puntuaciones típicas).

A partir de esta tabla podemos empezar a plantearnos si el número de factores obtenidos (enseguida veremos cuáles son esos factores) es suficiente para explicar todas y cada una de las variables incluidas en el análisis. También podemos empezar a plantearnos en este momento si, dando por bueno el número de factores extraído, alguna de las variables incluidas podría quedar fuera del análisis.

**Tabla 20.1.** Comunalidades.

	Inicial	Extracción
Nivel educativo	1.000	.682
Categoría laboral	1.000	.782
Salario actual	1.000	.901
Salario inicial	1.000	.887
Meses desde el contrato	1.000	.997
Experiencia previa (meses)	1.000	.894
Edad (años)	1.000	.889

Método de extracción: Análisis de Componentes principales.

En la tabla de *porcentajes de varianza explicada* (tabla 20.2) se ofrece un listado de los *autovalores* de la matriz de varianzas-covarianzas y del porcentaje de varianza que representa cada uno de ellos. Los autovalores expresan la cantidad de la varianza total que está explicada por cada factor; y los porcentajes de varianza explicada asociados a cada factor se obtienen dividiendo su correspondiente autovalor por la suma de los autovalores (la cual coincide con el número de variables). Por defecto, se extraen tantos factores como autovalores mayores que 1 tiene la matriz analizada. En nuestro ejemplo hay 3 autovalores mayores que 1, por lo que el procedimiento extrae 3 factores que consiguen explicar un 86,17 % de la varianza de los datos originales. La tabla muestra también, para cada factor con autovalor mayor que 1, la suma de las saturaciones (ver tabla 20.3) al cuadrado. Las sumas de cuadrados de la columna *Total* (que coinciden con los autovalores cuando se utiliza el método *componentes principales*, pero no cuando se utilizan otros métodos de extracción), pueden ayudarnos, según veremos, a determinar el número idóneo de factores.

La información de esta tabla puede utilizarse para tomar una decisión sobre el número idóneo de factores que deben extraerse. Si quisiéramos explicar, por ejemplo, un mínimo del 90% de la variabilidad contenida en los datos, sería necesario extraer cuatro factores.

La matriz de varianzas-covarianzas analizada por defecto es la matriz de correlaciones entre las 7 variables incluidas en el análisis. Puesto que esta matriz es de dimensiones 7x7, es posible extraer hasta 7 factores independientes. Tal como muestra la columna de porcentajes acumulados (*% acumulado*), con los 7 factores que es posible extraer se consigue explicar el 100% de la varianza total, pero con ello no se consigue el objetivo de reducir el número de dimensiones necesarias para explicar los datos.

**Tabla 20.2.** Porcentajes de varianza explicada.

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3.167	45.242	45.242	3.167	45.242	45.242
2	1.857	26.528	71.769	1.857	26.528	71.769
3	1.008	14.405	86.174	1.008	14.405	86.174
4	.429	6.125	92.299			
5	.247	3.523	95.822			
6	.194	2.776	98.598			
7	.098	1.402	100.000			

Método de extracción: Análisis de Componentes principales.

En la tabla 20.3 se encuentra la *solución factorial* propiamente dicha. Contiene las correlaciones entre las variables originales (o *saturaciones*) y cada uno de los factores. Conviene señalar que esta matriz cambia de denominación dependiendo del método de extracción elegido. En este caso se denomina matriz de *componentes* porque en nuestro ejemplo hemos utilizado el método de *componentes principales* como método de extracción (es el método que actúa por defecto). Más adelante veremos que también recibe el nombre de matriz de *estructura factorial*.

Comparando las saturaciones relativas de cada variable en cada uno de los tres factores podemos apreciar que el primer factor está constituido por las variables *nivel educativo*, *categoría laboral*, *salario actual* y *salario inicial*. Todas estas variables saturan en un único factor porque constituyen un grupo diferenciado de variables dentro de la matriz de correlaciones. Este factor parece reflejar la dimensión de "promoción" dentro de la empresa. El segundo factor recoge el grupo de las variables *experiencia previa* y *edad*, por lo que podría representar la "veteranía laboral". Por último, el tercer factor está formado por una única variable, los *meses desde el contrato*, o lo que es lo mismo, la "antigüedad en el puesto", que es independiente de la "promoción" y de la "veteranía laboral" (puesto que los factores son independientes entre sí y la variable no satura en los otros dos factores).

**Tabla 20.3.** Matriz de componentes (matriz de la estructura factorial).

	Componente		
	1	2	3
Nivel educativo	.806	-.172	.047
Categoría laboral	.843	.260	-.061
Salario actual	.944	.089	.041
Salario inicial	.910	.232	-.077
Meses desde el contrato	.043	.053	.996
Experiencia previa (meses)	-.179	.927	-.041
Edad (años)	-.232	.914	.026

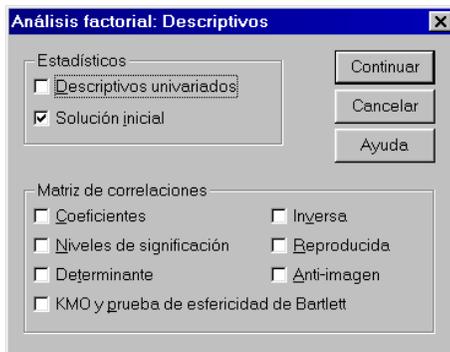
Método de extracción: Análisis de componentes principales.

## Descriptivos

La opción **Descriptivos** ofrece algunos estadísticos descriptivos, además de la matriz de correlaciones y otras matrices y estadísticos relacionados con ella. Para obtener estos estadísticos:

- ▶ Pulsar en el botón **Descriptivos...** del cuadro de diálogo *Análisis factorial* (ver figura 20.1) para acceder al subcuadro de diálogo *Análisis factorial: Descriptivos* que muestra la figura 20.3.

**Figura 20.3.** Subcuadro de diálogo *Análisis factorial: Descriptivos*.



**Estadísticos.** Este apartado contiene las opciones que permiten seleccionar los estadísticos descriptivos del análisis:

- Descriptivos univariados.** Muestra, para cada variable, el número de casos válidos, la media y la desviación típica.
- Solución inicial.** Permite obtener las communalidades iniciales, los autovalores de la matriz analizada y los porcentajes de varianza asociados a cada autovalor. Esta opción actúa por defecto y la información que ofrece aparece en las tablas 20.1 y 20.2.

---

**Matriz de correlaciones.** En este apartado se encuentran las opciones necesarias para obtener información sobre la matriz de correlaciones y algunos estadísticos asociados a ella.

- Coefficientes.** Muestra la matriz con los coeficientes de correlación entre las variables utilizadas en el análisis.
- Niveles de significación.** Incluye en la matriz de correlaciones los niveles críticos unilaterales asociados a cada coeficiente.
- Determinante.** Muestra el determinante de la matriz de correlaciones. El valor del determinante aparece en una nota a pie de tabla. Los determinantes próximos a cero están indicando que la variables utilizadas están linealmente relacionadas, lo que significa que el análisis factorial es una técnica pertinente para analizar esas variables.
- Inversa.** Muestra la inversa de la matriz de correlaciones. Esta matriz es la base para el cálculo de las comunalidades iniciales en algunos métodos de extracción y para el cálculo de la matriz anti-imagen (ver más abajo).
- Reproducida.** Muestra la matriz reproducida. La matriz reproducida es la matriz de correlaciones que se obtiene a partir de la solución factorial hallada. Si el modelo es bueno y el número de factores el adecuado, la estructura factorial debe ser capaz de reproducir la matriz de correlaciones. En la diagonal de la matriz reproducida se encuentran las comunalidades finales. Junto con la matriz de correlaciones reproducidas se muestra la *matriz de correlaciones residuales*, la cual contiene los residuos, es decir, las diferencias entre las correlaciones observadas y las correlaciones reproducidas. Si el modelo es el correcto, el número de residuos con valores elevados debe ser mínimo.

- **Anti-imagen.** Muestra la matriz de covarianzas anti-imagen y la matriz de correlaciones anti-imagen. La matriz de *covarianzas anti-imagen* contiene los negativos de las covarianzas parciales y la matriz de *correlaciones anti-imagen* contiene los coeficientes de correlación parcial cambiados de signo (la correlación entre dos variables se parcializa teniendo en cuenta el resto de variables incluidas en el análisis). En la diagonal de la matriz de correlaciones anti-imagen se encuentran las medidas de *adecuación muestral* para cada variable. Si el modelo factorial elegido es adecuado para explicar los datos, los elementos de la diagonal de la matriz de correlaciones anti-imagen deben tener un valor próximo a 1 y el resto de elementos deben ser pequeños.
  
- **KMO y prueba de esfericidad de Bartlett.** La medida de *adecuación muestral KMO* (Kaiser-Meyer-Olkin) contrasta si las correlaciones parciales entre las variables son suficientemente pequeñas. Permite comparar la magnitud de los coeficientes de correlación observados con la magnitud de los coeficientes de correlación parcial. El estadístico KMO varía entre 0 y 1. Los valores pequeños indican que el análisis factorial puede no ser una buena idea, dado que las correlaciones entre los pares de variables no pueden ser explicadas por otras variables. Los menores que 0,5 indican que no debe utilizarse el análisis factorial con los datos muestrales que se están analizando.

La *prueba de esfericidad de Bartlett* contrasta la hipótesis nula de que la matriz de correlaciones es una matriz identidad, en cuyo caso no existirían correlaciones significativas entre las variables y el modelo factorial no sería pertinente.

## Ejemplo (Análisis factorial > Descriptivos)

Este ejemplo muestra cómo obtener estadísticos descriptivos adicionales a la solución ofrecida por defecto. Además de los estadísticos descriptivos, también veremos que es posible obtener estadísticos inferenciales para contrastar algunas hipótesis relevantes en el contexto del análisis factorial. Seguimos utilizando las mismas siete variables que en el ejemplo anterior. Para obtener estos estadísticos:

- ▣ En el subcuadro de diálogo *Análisis factorial: Descriptivos* (ver figura 20.3), seleccionar todas las opciones de los distintos apartados.

Aceptando estas selecciones, el *Visor* ofrece los resultados que muestran las tablas 20.4 a 20.8. La tabla 20.4 muestra, para cada una de las variables incluidas en el análisis, algunos *estadísticos descriptivos* univariados: la media, la desviación típica y el número de casos válidos para el análisis (que puede diferir del número de casos del archivo de datos).

Si se mantienen las especificaciones que el programa tiene establecidas por defecto y el análisis se basa en la matriz de correlaciones, las diferencias de escala y de variabilidad entre las variables carecen de relevancia. Sin embargo, si se decide que el análisis se base en la matriz de varianzas-covarianzas, las variables con mayor variabilidad tendrán mayor importancia en la solución final.

**Tabla 20.4.** Estadísticos descriptivos.

	Media	Desviación típica	N del análisis
Nivel educativo	13.49	2.89	473
Categoría laboral	1.41	.77	473
Salario actual	\$34,418.45	\$17,093.72	473
Salario inicial	\$17,009.25	\$7,877.56	473
Meses desde el contrato	81.14	10.05	473
Experiencia previa (meses)	95.95	104.68	473
Edad (años)	43.3961	11.7872	473

La tabla 20.5 ofrece la matriz de correlaciones, es decir, los coeficientes de correlación de Pearson entre cada par de variables. Si no se especifica lo contrario, ésta es, según hemos señalado ya, la matriz de la cual parte el análisis. Con el método de extracción *componentes principales* (método que actúa por defecto), la matriz de correlaciones se auto-descompone en sus autovalores y autovectores para alcanzar la solución factorial. El resto de los métodos de extracción se basan en una transformación de la matriz de correlaciones.

**Tabla 20.5.** Matriz de correlaciones.

		Nivel educativo	Categ. laboral	Salario actual	Salario inicial	Meses desde el contrato	Experien. previa (meses)	Edad (años)
Correlación <sup>a</sup>	Nivel educativo	1.000	.515	.661	.633	.050	-.252	-.281
	Categoría laboral	.515	1.000	.780	.755	.004	.062	.010
	Salario actual	.661	.780	1.000	.880	.084	-.097	-.144
	Salario inicial	.633	.755	.880	1.000	-.018	.045	-.009
	Meses desde el contrato	.050	.004	.084	-.018	1.000	.002	.053
	Experiencia previa (meses)	-.252	.062	-.097	.045	.002	1.000	.802
	Edad (años)	-.281	.010	-.144	-.009	.053	.802	1.000
Sig. (Unilateral)	Nivel educativo		.000	.000	.000	.137	.000	.000
	Categoría laboral	.000		.000	.000	.468	.088	.414
	Salario actual	.000	.000		.000	.033	.017	.001
	Salario inicial	.000	.000	.000		.344	.162	.423
	Meses desde el contrato	.137	.468	.033	.344		.485	.127
	Experiencia previa (meses)	.000	.088	.017	.162	.485		.000
	Edad (años)	.000	.414	.001	.423	.127	.000	

a. Determinante = 1.196E-02

Para que el análisis sea fructífero es conveniente que la matriz contenga grupos de variables que correlacionen fuertemente entre sí. Una matriz de correlaciones próxima a una matriz identidad indica que el análisis factorial conducirá a una solución deficiente. Para formarnos una idea sobre el grado de relación existente entre las variables, la tabla 20.4 ofrece, además de la matriz de correlaciones, el nivel crítico unilateral (*Sig. unilateral*) asociado a cada coeficiente de correlación (el nivel crítico bilateral se obtiene multiplicando por dos el unilateral). Un nivel crítico menor que 0,05 indica que la correlación poblacional entre el correspondiente par de variables puede ser considerada significativamente distinta de cero. Lo deseable, por tanto, es encontrar muchos niveles críticos pequeños.

Por último, en una nota a pie de tabla aparece el valor del *determinante* de la matriz de correlaciones. Si las variables de la matriz están linealmente relacionadas, el valor del determinante se aproxima a cero, lo cual es un buen síntoma de cara a la idoneidad del análisis.

La tabla 20.6 recoge la *inversa de la matriz de correlaciones*. Esta matriz se encuentra estrechamente relacionada con la *matriz anti-imagen* que se muestra más abajo (ver tabla 20.8). Si el determinante de la matriz de correlaciones vale exactamente cero, el programa emite una advertencia indicando que no es posible calcular la matriz inversa, en cuyo caso tampoco será posible utilizar algunos de los métodos de extracción (por ejemplo, *ejes principales* o *máxima verosimilitud*).

**Tabla 20.6.** Inversa de la matriz de correlaciones.

	Nivel educativo	Categoría laboral	Salario actual	Salario inicial	Meses desde el contrato	Experien. previa (meses)	Edad (años)
Nivel educativo	2.029	-.057	-.534	-.783	-.085	.291	.258
Categoría laboral	-.057	2.800	-1.676	-.591	.129	-.244	-.101
Salario actual	-.534	-1.676	6.333	-3.997	-.599	.423	.433
Salario inicial	-.783	-.591	-3.997	5.490	.499	-.524	-.345
Meses desde el contrato	-.085	.129	-.599	.499	1.074	.052	-.205
Experiencia previa (meses)	.291	-.244	.423	-.524	.052	2.927	-2.211
Edad (años)	.258	-.101	.433	-.345	-.205	-2.211	2.917

La tabla 20.7 contiene dos estadísticos que permiten valorar la bondad de ajuste o adecuación de los datos analizados a un modelo factorial: la medida de adecuación muestral KMO y la prueba de esfericidad de Bartlett.

La *medida de adecuación muestral de Kaiser-Meyer-Olkin* (KMO) es un índice que compara la magnitud de los coeficientes de correlación observados con la magnitud de los coeficientes de correlación parcial:

$$KMO = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} r_{ij,m}^2}$$

donde  $r_{ij}$  representa el coeficiente de correlación simple entre las variables  $i$  y  $j$  y  $r_{ij,m}$  representa la correlación parcial entre las variables  $i$  y  $j$  eliminado el efecto de las restantes  $m$  variables incluidas en el análisis. Puesto que la correlación parcial entre dos variables debe ser pequeña cuando el modelo factorial es adecuado (véase más adelante), el denominador debe aumentar poco si los datos corresponden a una estructura factorial, en cuyo caso KMO tomará un valor próximo a 1. Si el valor de la medida de adecuación muestral es reducido (los valores por debajo de 0,6 se consideran mediocres) puede que no sea pertinente utilizar el análisis factorial con esos datos. (La diagonal de la matriz de correlaciones anti-imagen incluye los coeficientes de adecuación muestral para cada variable individualmente considerada).

La *prueba de esfericidad de Bartlett* contrasta la hipótesis nula de que la matriz de correlaciones observada es en realidad una matriz identidad. Asumiendo que los datos provienen de una distribución normal multivariante, el estadístico de Bartlett se distribuye aproximadamente según el modelo de probabilidad *chi-cuadrado* y es una transformación del determinante de la matriz de correlaciones. Si el nivel crítico (*Sig.*) es mayor que 0,05, no podremos rechazar la hipótesis nula de esfericidad y, consecuentemente, no podremos asegurar que el modelo factorial sea adecuado para explicar los datos.

**Tabla 20.7.** KMO y prueba de Bartlett.

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		.724
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	2075.310
	gl	21
	Sig.	.000

La tabla 20.8 ofrece la matriz de varianzas-covarianzas anti-imagen y la matriz de correlaciones anti-imagen. La matriz de correlaciones anti-imagen se encuentra relacionada con la matriz analizada por el método de extracción *Análisis Imagen* y se utiliza como diagnóstico de la adecuación de los datos a un modelo factorial.

**Tabla 20.8.** Matrices anti-imagen.

		Nivel educativo	Categ. laboral	Salario actual	Salario inicial	Meses desde el contrato	Experien. previa (meses)	Edad (años)
Covarianza anti-imagen	Nivel educativo	.493	-.010	-.042	-.070	-.039	.049	.044
	Categoría laboral	-.010	.357	-.095	-.038	.043	-.030	-.012
	Salario actual	-.042	-.095	.158	-.115	-.088	.023	.023
	Salario inicial	-.070	-.038	-.115	.182	.085	-.033	-.022
	Meses desde el contrato	-.039	.043	-.088	.085	.931	.017	-.065
	Experiencia previa (meses)	.049	-.030	.023	-.033	.017	.342	-.259
	Edad (años)	.044	-.012	.023	-.022	-.065	-.259	.343
Correlación anti-imagen	Nivel educativo	.921 <sup>a</sup>	-.024	-.149	-.235	-.058	.119	.106
	Categoría laboral	-.024	.881 <sup>a</sup>	-.398	-.151	.074	-.085	-.035
	Salario actual	-.149	-.398	.723 <sup>a</sup>	-.678	-.230	.098	.101
	Salario inicial	-.235	-.151	-.678	.743 <sup>a</sup>	.206	-.131	-.086
	Meses desde el contrato	-.058	.074	-.230	.206	.098 <sup>a</sup>	.029	-.116
	Experiencia previa (meses)	.119	-.085	.098	-.131	.029	.538 <sup>a</sup>	-.757
	Edad (años)	.106	-.035	.101	-.086	-.116	-.757	.548 <sup>a</sup>

a. Medida de adecuación muestral

En este contexto, un coeficiente de correlación parcial expresa el grado de relación existente entre dos variables tras eliminar el efecto de las restantes variables incluidas en el análisis. Cuando las variables incluidas en el análisis comparten gran cantidad de información debido a la presencia de factores comunes, la correlación parcial entre cualquier par de variables debe ser reducida. Por el contrario, cuando dos variables comparten gran cantidad de información entre ellas, pero no la comparten con las restantes variables (ni, consecuentemente, con los factores comunes), la correlación parcial entre ellas será elevada, siendo esto un mal síntoma de cara a la idoneidad del análisis.

Por otro lado, las correlaciones parciales son también estimaciones de las correlaciones entre los factores únicos (existe un factor único para cada variable del modelo). Y puesto que los factores únicos son independientes entre sí, las correlaciones parciales deben ser próximas a cero.

La correlación anti-imagen es el negativo de la correlación parcial entre dos variables. Si la matriz de correlaciones anti-imagen contiene una gran proporción de coeficientes elevados, el modelo factorial puede no ser adecuado para analizar los datos.

La diagonal de la matriz de correlaciones anti-imagen contiene una medida de *adecuación muestral* para cada variable. Esta medida es similar a la medida KMO, pero para cada variable individualmente considerada.

Los valores de la diagonal de la matriz de covarianza anti-imagen se obtienen restando a 1 la correlación múltiple al cuadrado entre cada variable y las restantes variables del análisis. Representan, por tanto, una estimación de la *unicidad* de cada variable, o lo que es lo mismo, una estimación de lo que cada variable tiene de propio o de no compartido con las demás. Habitualmente, los valores de estas dos matrices se muestran en notación científica (en formato exponencial). Si se desea reformar la tabla para que los valores no se muestren en notación científica, sino en notación decimal, se puede ejecutar el proceso de SPSS *Deshacer notación científica.sbs*. Para ello:

- ▶ Seleccionar la tabla en el *Visor de resultados*.
- ▶ Seleccionar la opción **Ejecutar proceso...** del menú **Utilidades**.
- ▶ En la carpeta *Scripts* (que cuelga de la carpeta en la que está instalado el SPSS), seleccionar el archivo *Deshacer notación científica.sbs*.

La tabla 20.9 muestra la matriz de *correlaciones reproducidas*. El *Visor* ofrece esta tabla al final de los resultados de la extracción y no junto al resto de estadísticos descriptivos.

La matriz de correlaciones reproducidas contiene las correlaciones que es posible reproducir utilizando tan sólo la información contenida en la solución factorial. Es decir, utilizando la matriz de la tabla 20.3. En concreto, la matriz reproducida se obtiene post-multiplicando la matriz factorial por su traspuesta.

Además de la matriz de correlaciones reproducidas, la tabla 20.9 también incluye la *matriz residual*, la cual contiene los residuos del análisis factorial. Cada residuo expresa la diferencia existente entre la correlación observada entre dos variables (ver tabla 20.5) y la correlación reproducida por la estructura factorial para esas dos variables. Si el análisis ha sido fructífero, la mayoría de las correlaciones reproducidas se parecerán a las correlaciones observadas y los residuos serán muy pequeños. De hecho, como orientación, la tabla incluye una nota a pie de tabla que contabiliza el número de residuos mayores que 0,05 (un valor arbitrariamente pequeño) y el porcentaje que ese número representa sobre el total de correlaciones no redundantes de la matriz.

**Tabla 20.9.** Matriz de correlaciones reproducidas y matriz residual.

		Nivel educativo	Categoría laboral	Salario actual	Salario inicial	Meses desde el contrato	Experien. previa (meses)	Edad (años)
Correlación reproducida	Nivel educativo	.682 <sup>b</sup>	.632	.748	.690	.073	-.306	-.344
	Categoría laboral	.632	.782 <sup>b</sup>	.816	.832	-.010	.093	.041
	Salario actual	.748	.816	.901 <sup>b</sup>	.876	-.087	-.088	-.137
	Salario inicial	.690	.832	.876	.887 <sup>b</sup>	-.025	.056	-.001
	Meses desde el contrato	.073	-.010	.087	-.025	.997 <sup>b</sup>	.000	.064
	Experiencia previa (meses)	-.306	.093	-.088	.056	.000	.894 <sup>b</sup>	.888
	Edad (años)	-.344	.041	-.137	-.001	.064	.888	.889 <sup>b</sup>
Residual <sup>a</sup>	Nivel educativo		-.116	-.087	-.057	-.022	.054	.062
	Categoría laboral	-.116		-.036	-.076	.014	-.031	-.031
	Salario actual	-.087	-.036		.004	-.002	-.010	-.007
	Salario inicial	-.057	-.076	.004		.006	-.010	-.008
	Meses desde el contrato	-.022	.014	-.002	.006		.002	-.012
	Experiencia previa (meses)	.054	-.031	-.010	-.010	.002		-.085
	Edad (años)	.062	-.031	-.007	-.008	-.012	-.085	

Método de extracción: Análisis de Componentes principales.

a. Los residuos se calculan entre las correlaciones observadas y reproducidas. Hay 7 (33.0%) residuos no redundantes con valores absolutos > 0.05.

b. Comunalidades reproducidas

Existen varias razones por las que el análisis podría desembocar en una matriz residual con un gran número de residuos altos (en valor absoluto). En primer lugar, podría ocurrir que se hubiera extraído un número insuficiente de factores y que, consecuentemente, la estructura factorial no fuera capaz de reproducir adecuadamente la matriz de correlaciones observada. En segundo lugar, podría ocurrir que las correlaciones observadas estuvieran mal estimadas, bien por la presencia de sesgos en la medida de las variables, bien porque el coeficiente de correlación de Pearson no fuera el apropiado para cuantificar la relación por causa de la escala utilizada para medir las variables. Por último, podría ocurrir que el modelo factorial no fuera pertinente para analizar los datos (porque las variables no están linealmente relacionadas, porque en los datos analizados no existe ningún tipo de estructura factorial, etc.).

## Extracción

La opción **Extracción** permite controlar varios aspectos relacionados con la fase de extracción de los factores. Entre otras cosas, permite decidir qué modelo factorial se desea utilizar, en qué matriz de datos basar el análisis y cuántos factores deben extraerse. Para controlar los aspectos relacionados con el proceso de extracción de factores:

- ▶ Pulsar en el botón **Extracción...** del cuadro de diálogo *Análisis factorial* (ver figura 20.1) para acceder al subcuadro de diálogo *Análisis factorial: Extracción* que muestra la figura 20.4.

**Figura 20.4.** Subcuadro de diálogo *Análisis factorial: Extracción*.

The image shows a dialog box titled "Análisis factorial: Extracción". It contains several settings for the extraction process:

- Método:** A dropdown menu set to "Componentes principales".
- Analizar:** Two radio buttons: "Matriz de correlaciones" (selected) and "Matriz de covarianza".
- Mostrar:** Two checkboxes: "Solución factorial sin rotar" (checked) and "Gráfico de sedimentación" (unchecked).
- Extraer:** Two radio buttons: "Autovalores mayores que:" (selected) with a text input field containing "1", and "Número de factores:" (unchecked) with an empty text input field.
- Nº máximo de iteraciones para convergencia:** A text input field containing "25".
- Buttons:** "Continuar", "Cancelar", and "Ayuda" are located on the right side of the dialog.

**Método.** En esta lista desplegable se puede seleccionar el modelo factorial que será utilizado para estimar las saturaciones de las variables en los factores. Los distintos métodos difieren tanto en el algoritmo de cálculo como en la matriz que será analizada (se asume que la matriz seleccionada es la matriz de correlaciones). Los distintos métodos disponibles son:

- **Componentes principales.** Método de extracción en el que los factores obtenidos son los autovectores de la matriz de correlaciones re-escalados.
- **Mínimos cuadrados no ponderados.** Método de extracción que minimiza la suma de los cuadrados de las diferencias entre las matrices de correlaciones observada y reproducida, ignorando los elementos de la diagonal.
- **Mínimos cuadrados generalizados.** Método de extracción que minimiza la suma de los cuadrados de las diferencias entre las matrices de correlaciones observada y reproducida. Las correlaciones se ponderan por el inverso de su unicidad, de manera que las variables cuya unicidad es alta reciben un peso menor que aquellas cuyo valor es bajo. Este método genera un estadístico de bondad de ajuste *chi-cuadrado* que permite contrastar la hipótesis nula de que la matriz residual es una matriz nula.
- **Máxima verosimilitud.** Método de extracción que proporciona las estimaciones de los parámetros que con mayor probabilidad han producido la matriz de correlaciones observada, asumiendo que la muestra procede de una distribución normal multivariada. Las correlaciones se ponderan por el inverso de la unicidad de las variables y se emplea un algoritmo iterativo. Este método genera un estadístico de bondad de ajuste *chi-cuadrado* que permite contrastar la bondad del modelo para explicar la matriz de correlaciones.
- **Ejes principales.** Método de estimación iterativo en el que, como estimación inicial de la comunalidad, la matriz de correlaciones original se reduce sustituyendo los unos de su diagonal por las estimaciones de la correlación múltiple al cuadrado entre cada variable y todas las demás. La matriz reducida se auto-descompone y se corrigen las estimaciones iniciales de la comunalidad por las nuevas estimaciones resultantes. El proceso continúa hasta que no existe diferencia entre las estimaciones de las comunalidades entre dos pasos sucesivos o se alcanza alguno de los criterios de parada.

- **Alfa.** Método de extracción que considera las variables incluidas en el análisis como una muestra del universo de las variables posibles. Este método maximiza la generalizabilidad de los factores calculada como el *alfa* de Cronbach.
- **Imagen.** Método de extracción en el que se auto-descompone la matriz de correlaciones imagen. Se asume que la comunalidad es igual al cuadrado de la correlación múltiple entre una variable y todas las demás. Al solicitar este método de extracción, los resultados incluyen una tabla con la matriz de covarianza imagen.

**Analizar.** Las opciones de este apartado sólo están disponibles cuando se seleccionan los métodos de *componentes principales*, *ejes principales* y *análisis imagen*. Permite seleccionar el tipo de matriz que será analizada.

- **Matriz de correlaciones.** El análisis se basa en la matriz de correlaciones, en la matriz de correlaciones reducida, o en la matriz de correlaciones anti-imagen, según el método seleccionado.
- **Matriz de covarianza.** El análisis se basa en la matriz de varianzas-covarianzas, en la matriz de varianzas-covarianzas reducida, o la matriz de covarianzas anti-imagen, según el método seleccionado.

**Extraer.** Este apartado contiene opciones que permiten determinar el número de factores que se extraerán en la solución factorial, bien a partir de una regla heurística, bien especificando un número concreto:

- **Autovalores mayores que.** Si la matriz analizada es la de *correlaciones*, esta opción permite utilizar el tamaño de los autovalores como criterio para decidir el número de factores que estarán presentes en la solución factorial. Por defecto se extraen los factores cuyos autovalores son mayores que la unidad (a este criterio se le denomina *regla K1*). Si la matriz analizada es la de *varianzas-covarianzas*, la regla expresa el número de veces que un autovalor debe ser mayor que el autovalor promedio de la matriz para que el correspondiente factor sea retenido en la solución.

El valor que actúa por defecto es 1, pero este valor puede cambiarse introduciendo otro distinto (entre cero y el número de variables) en el correspondiente cuadro de texto.

- **Número de factores.** Permite especificar el número exacto de factores que se desea incluir en la solución. Se debe introducir dicho número en el recuadro de texto.

**Mostrar.** Estas opciones permiten seleccionar los resultados de la extracción que aparecerán en el visor de resultados.

- **Solución factorial sin rotar.** Muestra las saturaciones factoriales sin rotar (la matriz de componentes o factorial), las comunalidades y los autovalores de la solución factorial.
- **Gráfico de sedimentación** (también llamado *prueba de sedimentación de Cattell*). Muestra una representación gráfica de la magnitud de los autovalores. El corte en la tendencia descendente sirve de regla para la determinación del número óptimo de factores que deben estar presentes en la solución. Siempre se muestra la representación de los autovalores de la matriz de correlaciones (o de covarianzas) originales, independientemente del método de extracción seleccionado.

**Nº de iteraciones para convergencia.** Este cuadro de texto permite establecer el número máximo de iteraciones que los algoritmos pueden realizar para encontrar la solución factorial final. El valor por defecto es 25, habitualmente suficiente para obtener una solución. Este valor puede cambiarse introduciendo un entero positivo.

## Ejemplo (Análisis factorial > Extracción)

Este ejemplo muestra cómo seleccionar un método de extracción (*ejes principales*) distinto del método que actúa por defecto (*componentes principales*) y describe algunas opciones relacionadas con la extracción de factores. Seguimos utilizando las mismas siete variables que en los ejemplos anteriores. Para seleccionar un método de extracción:

- ▶ Pulsar en el botón **Extracción...** del cuadro de diálogo *Análisis factorial* (ver figura 20.1) para acceder al subcuadro de diálogo *Análisis factorial: Extracción* (ver figura 20.4).
- ▶ Desplegar la lista **Método** y seleccionar la opción *Ejes principales*.
- ▶ Seleccionar la opción **Gráfico de sedimentación** del apartado **Mostrar**.

Aceptando estas selecciones, el *Visor* ofrece los resultados que muestran las tablas 20.10 a la 20.13 y la figura 20.6.

La tabla de *comunalidades* (tabla 20.10) recoge ahora una estimación inicial de las comunalidades de las variables. Esta estimación se realiza calculando, para cada variable, la correlación múltiple al cuadrado entre esa variable y las restantes variables incluidas en el análisis. Se asume que si una variable está muy relacionada con las restantes variables del análisis, tenderá a compartir su información en un factor común. Sin embargo, los supuestos de la regresión múltiple (en la que se basa el coeficiente de correlación múltiple) no son los mismos que los del análisis factorial y por ello, la estimación inicial de la comunalidad y la cantidad de varianza que cada variable comparte con las demás a través de los factores comunes, rara vez coincide. De hecho, en la tabla 20.10 es posible apreciar que los valores de las comunalidades iniciales y los valores de las comunalidades de la extracción son diferentes.

Las comunalidades de la extracción se calculan, para una variable dada, elevando al cuadrado las saturaciones de esa variable en cada uno de los factores y sumando después esos cuadrados.

El método *ejes principales* es un método iterativo, es decir, un método que se ejecuta repetitivamente hasta alcanzar la solución idónea. La manera de proceder es la siguiente. Se comienza estimando la comunalidad inicial de cada variable mediante el coeficiente de correlación múltiple entre esa variable y todas las demás. Esas comunalidades estimadas sustituyen a los valores originales de la diagonal de la matriz de correlaciones, dando lugar a la *matriz de correlaciones reducida* (denominada así porque los nuevos valores son normalmente menores que los unos originales de la matriz de correlaciones). A continuación, la matriz de correlacio-

nes reducida se auto-descompone en sus autovalores y autovectores. Los autovectores resultantes (en número igual al de autovalores mayores que 1 en la matriz original, si no se ha variado este criterio), se re-escalan y pasan a formar la matriz factorial de la primera iteración. A partir de la matriz factorial se calculan las estimaciones de las comunalidades de la extracción.

Las comunalidades estimadas mediante la correlación múltiple al cuadrado y las estimadas a partir de la matriz factorial rara vez coinciden. Además, las estimaciones de las comunalidades obtenidas a partir de la matriz factorial dependen del número de factores contenidos en la solución factorial y podría ocurrir que el número de factores de la solución obtenida no fuera el óptimo. En cualquier caso, el objetivo de la extracción con el método *ejes principales* consiste en insertar en la diagonal de la matriz de correlaciones la mejor estimación posible de las comunalidades de las variables a partir del número de factores seleccionado. Por ello, en una segunda iteración, las comunalidades inicialmente estimadas se sustituyen por las comunalidades estimadas a partir de la matriz factorial y se repite la auto-descomposición de la nueva matriz reducida. Con ello se obtiene una nueva matriz factorial y un nuevo conjunto de estimaciones de las comunalidades. Las nuevas estimaciones se comparan con las últimas insertadas en la diagonal de la matriz de correlaciones al iniciar la iteración. Si los valores de las estimaciones no difieren sensiblemente entre sí la extracción concluye (por defecto, se admite una diferencia menor de 0,001); en caso contrario, se repiten las iteraciones hasta alcanzar el criterio de convergencia o el máximo de 25 iteraciones.

En nuestro ejemplo, podemos apreciar que la variable *meses desde el contrato* apenas correlaciona con las restantes variables, por lo que la estimación inicial de su comunalidad es bastante reducida (0,069). Tras finalizar la extracción, la comunalidad para esta variable no ha mejorado demasiado (0,233).

**Tabla 20.10.** Comunalidades.

	Inicial	Extracción
Nivel educativo	.507	.529
Categoría laboral	.643	.661
Salario actual	.842	.939
Salario inicial	.818	.895
Meses desde el contrato	.069	.233
Experiencia previa (meses)	.658	.803
Edad (años)	.657	.818

Método de extracción: Factorización de Ejes principales.

La tabla de *porcentajes de varianza explicada* (tabla 20.11) también ha cambiado con el nuevo método de extracción. La parte izquierda de la tabla sigue ofreciendo los autovalores iniciales de la matriz de correlaciones original. Sin embargo, las sumas de las saturaciones al cuadrado de cada factor ahora no coinciden con los autovalores iniciales (pueden compararse estos resultados con los de la tabla 20.2). Esto es debido a que la matriz de estructura factorial ha cambiado. Ahora, el tercer factor apenas explica un 3,8% de la varianza total y su nuevo valor es muy inferior a 1. Este dato nos debe hacer pensar que seguramente el tercer factor de la extracción no sea relevante desde el punto de vista de la proporción de varianza que consigue explicar. Más adelante podremos valorar mejor la pertinencia de mantener este tercer factor en la solución factorial final.

**Tabla 20.11.** Porcentajes de varianza explicada.

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3.167	45.242	45.242	2.953	42.190	42.190
2	1.857	26.528	71.769	1.661	23.728	65.919
3	1.008	14.405	86.174	.263	3.761	69.680
4	.429	6.125	92.299			
5	.247	3.523	95.822			
6	.194	2.776	98.598			
7	9.813E-02	1.402	100.000			

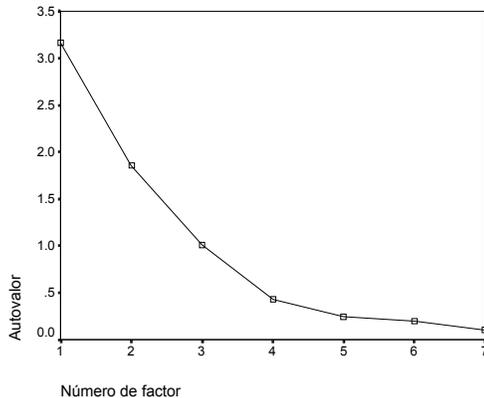
Método de extracción: Factorización de Ejes principales.

El *gráfico de sedimentación* (ver figura 20.5) sirve para determinar el número óptimo de factores. Originalmente propuesto por Cattell (1966), consiste simplemente en una representación gráfica del tamaño de los autovalores.

Según hemos señalado ya, un autovalor indica la cantidad de varianza explicada por una componente principal. Tanto la tabla de porcentajes de varianza explicada (tabla 20.11) como el gráfico de sedimentación (figura 20.5) ofrecen los autovalores ordenados de mayor a menor: el primer autovalor es el mayor de los posibles, el segundo autovalor es el segundo mayor, y así sucesivamente. Si un autovalor se aproxima a cero, esto significa que el factor correspondiente a ese autovalor es incapaz de explicar una cantidad relevante de la varianza total. Por tanto, un factor al que corresponde un autovalor próximo a cero se considera un factor residual y carente de sentido en el análisis.

Al representar todos los autovalores según su tamaño, es posible formarse muy rápidamente una idea sobre si la cantidad de varianza asociada a cada uno de ellos es relevante para el análisis o si por el contrario se trata sólo de varianza residual. Los autovalores residuales se encuentran en la parte derecha del gráfico, formando una planicie de poca inclinación, frente a la fuerte pendiente formada por los autovalores que explican la mayor parte de la varianza disponible. Por ello, es conveniente inspeccionar el gráfico de sedimentación de izquierda a derecha, buscando el punto de inflexión en el que los autovalores dejan de formar una pendiente significativa y comienzan a describir una caída de poca inclinación.

**Figura 20.5.** Gráfico de sedimentación.



En nuestro ejemplo, la pendiente pierde inclinación a partir del cuarto autovalor (hacia su derecha). O lo que es lo mismo, el cuarto autovalor no provoca pendiente respecto del quinto (es decir, el cuarto valor se encuentra en la misma planicie sin pendiente que el quinto), por lo que deberemos considerar que sólo deben extraerse los tres primeros factores y desechar del cuarto en adelante.

Es importante resaltar que el gráfico de sedimentación no varía con el número de factores seleccionado. Por otra parte, el gráfico siempre muestra todos los posibles autovalores de la matriz de correlaciones original y no los autovalores de la matriz analizada, que puede ser distinta de la de correlaciones, según el método de extracción utilizado.

La tabla 20.12 muestra la *matriz de estructura factorial* obtenida con el método de extracción *ejes principales*. En una nota a pie de la tabla se advierte de que no se ha conseguido alcanzar el criterio de convergencia. Puesto que se han sobrepasado las 25 iteraciones establecidas por defecto y dado que alguna de las comunalidades (posiblemente la correspondiente a la variable *meses desde el contrato*) todavía difiere de su estimación en la iteración anterior en más de 0,001 (en concreto, difiere 0,0029), la extracción no se considera satisfactoria. Muy posiblemente, aumentando el número de iteraciones admisibles se podría llegar a una estimación estable, sin embargo, antes de tomar esta decisión, es mejor valorar la solución obtenida hasta el momento.

Comparando la solución actual con la obtenida en el primer ejemplo mediante el método de *componentes principales*, vemos que no existen diferencias muy notables. El primer factor sigue agrupando las variables correspondientes a la “promoción dentro de la empresa” y el segundo factor agrupa a las variables correspondientes a la “veterana laboral”. En cuanto al tercer factor, sólo la variable *meses desde el contrato* satura en él y dicha variable no satura en ningún otro factor. Asumiendo la filosofía de encontrar factores comunes que agrupen variables que correlacionen entre sí, el tercer factor carece de sentido y también carece de sentido incluir en el análisis la variable *meses desde el contrato*, puesto que no correlaciona con ninguna otra variable de las utilizadas (véase la matriz de correlaciones de la tabla 20.5). Por todo ello, parece más razonable excluir dicha variable del análisis que aumentar el número de iteraciones (como veremos enseguida).

**Tabla 20.12.** Matriz de la estructura factorial.

	Factor <sup>a</sup>		
	1	2	3
Nivel educativo	.711	-.146	.050
Categoría laboral	.783	.214	-.050
Salario actual	.961	.077	.098
Salario inicial	.914	.220	-.109
Meses desde el contrato	.033	.030	.481
Experiencia previa (meses)	-.167	.879	-.042
Edad (años)	-.220	.875	.065

Método de extracción: Factorización del eje principal.

- a. Se han intentado extraer 3 factores. Requidas más de 25 iteraciones. (Convergencia=2.911E-03). Se ha terminado la extracción.

De momento, antes de excluir del análisis la variable *meses desde el contrato*, vamos a solicitar la extracción de sólo dos factores. Manteniendo esa variable en el análisis podremos valorar su influencia en el método de extracción. Para forzar una solución factorial de sólo 2 factores (seguimos utilizando las mismas siete variables):

- ▶ Pulsar en el botón **Extracción...** del cuadro de diálogo *Análisis factorial* (ver figura 20.1) para acceder al subcuadro de diálogo *Análisis factorial: Extracción* (ver figura 20.4).
- ▶ Mantener *Ejes principales* como método de extracción y, tras seleccionar la opción **Número de factores** en el apartado **Extraer**, introducir un 2 en el cuadro de texto.

Aceptando estas selecciones, el *Visor* ofrece los resultados que muestran las tablas 20.13 a 20.15 y la figura 20.5.

La tabla de comunalidades (tabla 20.13) muestra que, en esta nueva situación, todas las comunalidades estimadas consiguen mejorar tras la extracción, excepto la de la variable *meses desde el contrato*, que empeora.

Puesto que no se ha extraído el factor en el que podría saturar de manera única la variable *meses desde el contrato*, la estimación de su comunalidad no puede ser actualizada (como ocurría anteriormente) y la solución factorial alcanzada no la contempla como una variable común a las demás. Visto este resultado, la decisión más acertada sería excluir dicha variable del análisis, sin variar ya el número factores.

**Tabla 20.13.** Comunalidades.

	Inicial	Extracción
Nivel educativo	.507	.528
Categoría laboral	.643	.662
Salario actual	.842	.922
Salario inicial	.818	.874
Meses desde el contrato	.069	.002
Experiencia previa (meses)	.658	.813
Edad (años)	.657	.796

Método de extracción: Factorización de Ejes principales.

Ahora, la tabla de *porcentajes de varianza explicada* (tabla 20.14) ofrece las sumas de las saturaciones al cuadrado (es decir, las estimaciones de los autovalores tras la extracción) sólo para los dos factores solicitados. Ambas sumas de cuadrados son mayores que 1, lo que es un buen indicador desde el punto de vista del número idóneo de factores. Sin embargo, en esta tabla todavía se encuentra incluida la información correspondiente a la variable *meses desde el contrato*, lo cual representa un inconveniente pues los cálculos asumen que se desea explicar la variabilidad de 7 variables cuando en realidad sabemos que sólo 6 de ellas son pertinentes (recordemos que la proporción de varianza explicada por un factor se obtiene dividiendo ese autovalor por el número de variables). En nuestro ejemplo, se están realizando esos cocientes sobre 7 variables cuando deberían estar calculados sobre 6. Si hiciéramos correctamente los cálculos, los dos factores extraídos estarían explicando el 76,59% de la varianza y no el 65,66% (ver la tabla 20.18).

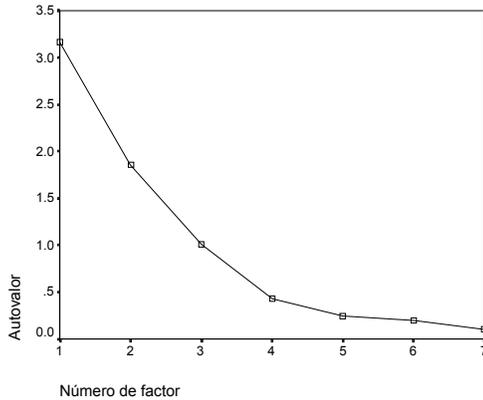
**Tabla 20.14.** Porcentajes de varianza explicada.

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3.167	45.242	45.242	2.942	42.024	42.024
2	1.857	26.528	71.769	1.655	23.639	65.664
3	1.008	14.405	86.174			
4	.429	6.125	92.299			
5	.247	3.523	95.822			
6	.194	2.776	98.598			
7	.098	1.402	100.000			

Método de extracción: Factorización de Ejes principales.

Puesto que no hemos alterado el número de variables, el *gráfico de sedimentación* (figura 20.6) no ha cambiado (recordemos que en este gráfico están representados los autovalores iniciales de la matriz de correlaciones original). Para que el gráfico de sedimentación sea de utilidad debemos excluir del análisis la variable que no forma parte de la solución factorial.

**Figura 20.6.** Gráfico de sedimentación de Cattell.



La *matriz de estructura* (tabla 20.15) muestra ahora claramente que la variable *meses desde el contrato* no comparte información con las restantes variables: no satura en ninguno de los dos factores presentes en la solución.

La extracción ha alcanzado el criterio de convergencia en 10 iteraciones, a pesar de la presencia de la variable de *ruido* que hemos dejado permanecer en el análisis. Puesto que el número de factores estaba bien determinado, la variable de ruido no ha afectado sensiblemente en la estimación (aunque sí ha afectado a los cálculos de la varianza explicada).

**Tabla 20.15.** Matriz de la estructura factorial.

	Factor <sup>a</sup>	
	1	2
Nivel educativo	.712	-.146
Categoría laboral	.784	.216
Salario actual	.957	.077
Salario inicial	.909	.219
Meses desde el contrato	.031	.026
Experiencia previa (meses)	-.168	.886
Edad (años)	-.220	.865

Método de extracción: Factorización del eje principal.

a. 2 factores extraídos. Requeridas 10 iteraciones.

Como paso final, vamos a excluir del análisis la variable *meses desde el contrato* y a repetir la extracción forzando una solución de dos factores. Para ello,

- ▶ En el cuadro de diálogo *Análisis factorial* (ver figura 20.1), seleccionar la variable *meses desde el contrato* y eliminarla de la lista **Variables**.
- ▶ Repetir el análisis utilizando *ejes principales* como método de extracción y limitando a 2 el número de factores.

En la tabla de *comunalidades* (tabla 20.16) puede apreciarse que las estimaciones iniciales, ahora que nos hemos deshecho de la variable *meses desde el contrato*, son, todas ellas, bastante buenas y bastante próximas a las estimaciones obtenidas tras la extracción.

**Tabla 20.16.** Comunalidades.

	Inicial	Extracción
Nivel educativo	.505	.527
Categoría laboral	.641	.662
Salario actual	.833	.919
Salario inicial	.810	.878
Experiencia previa (meses)	.658	.820
Edad (años)	.653	.788

Método de extracción: Factorización de Ejes principales.

La tabla de *porcentajes de varianza explicada* (tabla 20.17) muestra que, dado que ahora hemos incluido 6 variables en el análisis, sólo es posible extraer 6 factores como máximo. Y sólo dos de los autovalores contienen información sustancial sobre la varianza común disponible; los restantes autovalores son marcadamente menores que 1. Sin embargo, con sólo dos factores es posible explicar el 83,7% de la varianza total de la matriz de correlaciones. Tras la extracción, los autovalores de la matriz de correlaciones reducida suponen el 76,6% de la varianza total. (Este dato debe interpretarse de manera aproximada dado que, como se ha señalado, el porcentaje de varianza explicada está calculado a partir del número de variables de la matriz de correlaciones y no a partir de la traza de la matriz de correlaciones reducida, que sería, posiblemente, más adecuado).

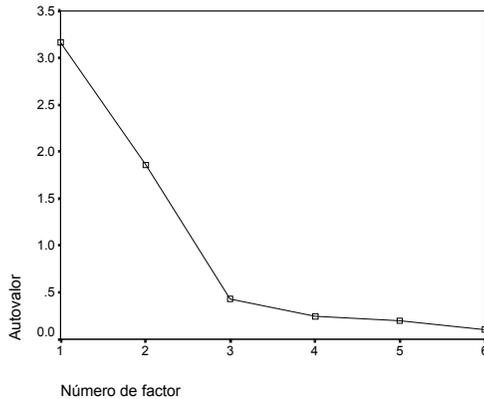
**Tabla 20.17.** Porcentajes de varianza explicada.

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3.166	52.760	52.760	2.941	49.013	49.013
2	1.856	30.927	83.688	1.654	27.572	76.585
3	.431	7.185	90.873			
4	.247	4.110	94.983			
5	.196	3.270	98.253			
6	.105	1.747	100.000			

Método de extracción: Factorización de Ejes principales.

El *gráfico de sedimentación* (figura 20.7) resulta más claro ahora que hemos eliminado la variable *meses desde el contrato*. Se aprecia una clara inflexión a partir del tercer autovalor, por lo que debemos concluir que es pertinente extraer sólo dos factores. Los restantes factores deben considerarse residuales y si su valor no es exactamente cero es sólo debido al error muestral inherente a las estimaciones de las correlaciones.

**Figura 20.7.** Gráfico de sedimentación de Cattell.



A partir de las saturaciones de la matriz de la *estructura factorial* (tabla 20.18) es fácil inferir la correspondencia existente entre cada variable y cada uno de los dos factores extraídos. Factor 1 (“promoción”): *nivel educativo, categoría laboral, salario actual y salario inicial*. Factor 2 (“veteranía laboral”): *experiencia previa y edad*. Todas las variables del análisis saturan mayoritariamente en uno de los dos factores y ninguna de ellas presenta indicios que hagan sospechar de su falta de adecuación a la solución.

**Tabla 20.18.** Matriz de la estructura factorial.

	Factor <sup>a</sup>	
	1	2
Nivel educativo	.711	-.146
Categoría laboral	.785	.217
Salario actual	.956	.076
Salario inicial	.910	.220
Experiencia previa (meses)	-.169	.890
Edad (años)	-.220	.860

Método de extracción: Factorización del eje principal.

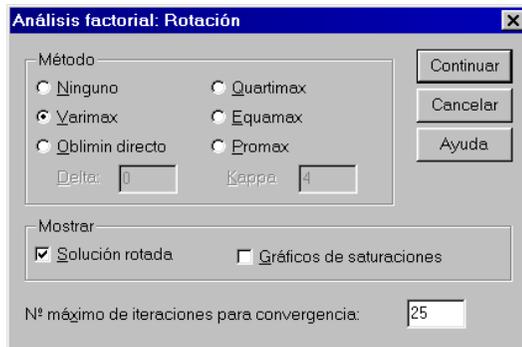
a. 2 factores extraídos. Requeridas 12 iteraciones.

## Rotación

La opción **Rotación** permite controlar la fase de rotación del análisis. Con esta opción podemos definir el método de rotación que deseamos utilizar para facilitar la interpretación de la solución factorial y solicitar la representación gráfica de las saturaciones. Por defecto, no se encuentra seleccionado ningún método de rotación. Para seleccionar el método de rotación:

- ▶ Pulsar en el botón **Rotación...** del cuadro de diálogo *Análisis factorial* (ver figura 20.1) para acceder al subcuadro de diálogo *Análisis factorial: Rotación* que muestra la figura 20.8.

**Figura 20.8.** Subcuadro de diálogo Análisis factorial: Rotación.



**Método.** En este apartado se puede seleccionar el método de rotación de la solución factorial. Se encuentran disponibles tres procedimientos de rotación *ortogonal*, mediante los cuales se respeta la independencia entre factores de la solución inicial, y dos procedimientos de rotación *oblicua*, mediante los cuales pueden obtenerse factores relacionados entre sí. Los métodos disponibles son:

- **Ninguno.** No se aplica ningún método de rotación. Es la opción que actúa por defecto. Cuando la solución consta de un único factor y no se ha marcado esta opción, el *Visor de resultados* muestra un mensaje de advertencia.
- **Varimax.** Método de rotación ortogonal que minimiza el número de variables que tienen saturaciones altas en cada factor. Simplifica la interpretación de los factores optimizando la solución por columna.
- **Quartimax.** Método de rotación ortogonal que minimiza el número de factores necesarios para explicar cada variable. Simplifica la interpretación de las variables observadas optimizando la interpretación por filas.
- **Equamax.** Método de rotación que es combinación del método Varimax, que simplifica los factores, y el método Quartimax, que simplifica las variables. Se minimiza tanto el número de variables que saturan alto en un factor como el número de factores necesarios para explicar una variable.
- **Oblimin directo.** Método para la rotación oblicua (no ortogonal). Cuando *delta* es igual a cero (el valor por defecto), las soluciones son las más oblicuas. A medida que delta se va haciendo más negativo, los factores son menos oblicuos. Para anular el valor por defecto de delta, puede introducirse un número menor o igual que 0,8.
  - Delta.** El valor de delta permite controlar el grado de oblicuidad que pueden llegar a alcanzar los factores de la solución.
- **Promax.** Rotación oblicua que permite que los factores estén correlacionados. Puede calcularse más rápidamente que una rotación Oblimin directa, por lo que es útil para grandes conjuntos de datos.
  - Kappa.** Parámetro que controla el cálculo de la rotación Promax. El valor por defecto es 4. Este valor es adecuado para la mayoría de los análisis.

**Mostrar.** Este cuadro se permite seleccionar los resultados de la rotación que se mostrarán en el *Visor de resultados*. Por defecto se muestra la solución rotada cuando se selecciona alguno de los métodos de rotación. Si se encuentra seleccionada la opción *Ninguna* del recuadro *Método* no será posible seleccionar ninguna de las opciones de este recuadro.

- Solución rotada.** Permite obtener una o más tablas con los resultados del proceso de rotación. Al seleccionar una rotación ortogonal, esta opción permite obtener la matriz de estructura factorial rotada y la matriz de transformación necesaria para rotar los factores a partir de la solución inicial. Además, en la tabla de *porcentajes de varianza explicada* aparecen columnas adicionales que contienen la varianza total explicada por los factores rotados. Al seleccionar una rotación oblicua, esta opción permite obtener la matriz de configuración rotada, que contiene las saturaciones de las variables en los factores, y la matriz de estructura, que contiene las correlaciones entre las variables observadas y los factores (cuando la rotación es ortogonal, ambas matrices son idénticas). Además, ofrece la matriz de correlaciones entre los factores y desecha la matriz de transformación para la rotación. En la tabla de *porcentajes de varianza explicada* sólo se incluyen los autovalores de los factores rotados (ya que no tiene sentido hablar de porcentajes de varianza independientes).
- Gráficos de saturaciones.** Esta opción genera un gráfico de dispersión que refleja la ubicación de las variables en el espacio definido por los factores. Se trata de un gráfico de las saturaciones. El gráfico muestra, por defecto, los tres primeros factores de la solución factorial en un gráfico tridimensional. Si se desea representar otros factores, es necesario editar el gráfico y elegir esos otros factores. Los ejes factoriales se representan siempre en formato ortogonal, aunque exista correlación entre los factores.

**Nº máximo de iteraciones para convergencia.** Permite determinar el número máximo de iteraciones que puede recorrer el algoritmo para la estimación de la solución rotada. Por defecto se efectúan un máximo de 25 iteraciones, lo que es suficiente para la mayoría de las situaciones.

### Ejemplo (Análisis factorial > Rotación ortogonal)

Este ejemplo muestra cómo rotar la estructura factorial original mediante un método ortogonal. La estructura que vamos a rotar es la obtenida en el ejemplo anterior (2 factores) con el método de extracción *ejes principales*. Primero vamos a representar gráficamente la solución no rotada y posteriormente la compararemos con la solución rotada.

Para representar gráficamente la solución factorial no rotada del ejemplo anterior (6 variables, 2 factores, método de extracción *ejes principales*):

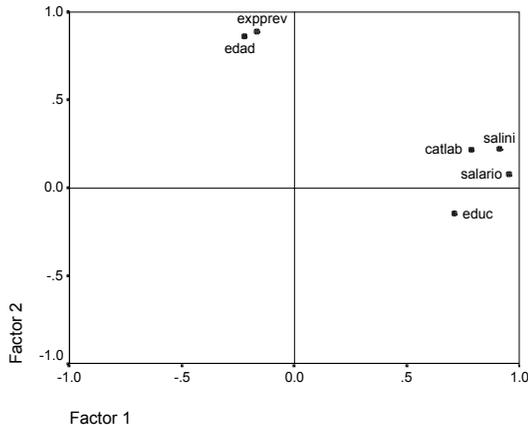
- ▶ Pulsar en el botón **Rotación...** del cuadro de diálogo *Análisis factorial* (ver figura 20.1) para acceder al subcuadro de diálogo *Análisis factorial: Rotación* que muestra la figura 20.8.
- ▶ En el apartado **Mostrar**, seleccionar la opción **Gráficos de saturaciones**.

Aceptando estas selecciones obtenemos, además de los resultados del ejemplo anterior, el gráfico de las saturaciones factoriales que muestra la figura 20.9.

Un *gráfico de factores* o de *saturaciones factoriales* representa el espacio factorial definido por los factores contenidos en la solución factorial. Si la solución contiene un único factor, el gráfico no se genera y aparece una advertencia indicando tal circunstancia; si la solución contiene dos factores se genera un diagrama de dispersión simple; si la solución contiene tres o más factores se genera un gráfico de dispersión tridimensional en el que sólo se representan los tres primeros factores. Cuando la solución contiene más de tres factores, el gráfico tridimensional representa los tres primeros factores, pero almacena también la información correspondiente a los restantes factores. Para representar factores distintos de los tres primeros pueden hacerse dos cosas, ambas desde el *Editor de gráficos* (al cual se accede pinchando dos veces sobre un gráfico): 1) seleccionar en el menú **Series** los tres factores que se desea representar, o 2) solicitar en el menú **Galería** un diagrama de dispersión matricial para representar simultáneamente todos los factores dos a dos.

Un *gráfico de saturaciones factoriales* es un diagrama de dispersión en el que los factores definen los ejes del espacio y las variables constituyen los puntos del diagrama. Las coordenadas de una variable en cada factor se corresponden con las saturaciones de la variable en dichos factores, es decir, con los valores de la matriz factorial (ver tabla 20.18).

**Figura 20.9.** Gráfico de las saturaciones factoriales.



La rotación de la solución original se realiza con el objetivo de mejorar la interpretación de la estructura factorial. Las restricciones de la auto-descomposición de la matriz de correlaciones imponen que el primer factor explique el máximo de la varianza común disponible en los datos, que el segundo factor explique el máximo de la varianza común restante (e independiente de la explicada por el primer factor), y así sucesivamente hasta el último de los factores. Estas restricciones se imponen para deshacer la indeterminación intrínseca a la solución del sistema homogéneo de ecuaciones que da lugar a los autovectores. Un efecto indeseable de estas restricciones es que los primeros factores tienden a capitalizar la información de covariación contenida en la matriz de correlaciones, acumulando más información de la que posiblemente les corresponda. Este hecho se aprecia en que las saturaciones de las variables en los primeros factores (y en especial en el primer factor) suelen encontrarse infladas, llevando esto a conceder excesiva importancia a los primeros factores. Cuando la estructura factorial es clara y cada variable del análisis se encuentra inequívocamente asignada a un único factor, el efecto “contaminante” de las restricciones no suele apreciarse. Sin embargo, cuando las variables saturan en más de un factor o existe un factor general que domina la solución, la rotación puede ser de gran utilidad para interpretar los resultados.

Otro de los motivos que justifican la rotación es que la solución factorial original es siempre ortogonal (los factores no rotados son siempre independientes entre sí). Sin embargo, existe un gran número de situaciones (y en especial en las ciencias sociales) en las que los factores pueden estar relacionados entre sí. En estos casos, si se desea estimar el grado de relación existente entre los factores, debe recurrirse a una rotación oblicua.

En el gráfico de la figura 20.10 pueden apreciarse dos grupos diferenciados de variables. El primer grupo se encuentra próximo al extremo positivo del factor 1, formado por las variables *salario actual* (salario), *salario inicial* (salini), *categoría laboral* (catlab) y *nivel educativo* (educ). El segundo grupo se encuentra próximo al extremo positivo del factor 2 y está formado por las variables *experiencia previa* (expprev) y *edad*. Un panorama similar obtenemos si estudiamos con detenimiento la matriz factorial de la tabla 20.18. En ella podemos apreciar que el grupo de variables pertenecientes al segundo factor también saturan, aunque poco, en el primer factor. Además, la variable *nivel educativo*, aunque se encuentra relativamente distante del primer factor, satura bastante en él (tal vez de manera excesiva). Si efectuamos una rotación ortogonal, posiblemente se aclare un poco más la estructura de las variables.

Para solicitar un método de rotación (manteniendo las mismas 6 variables, forzando una solución de 2 factores y utilizando *ejes principales* como método de extracción):

- ▶ Pulsar en el botón **Rotación...** del cuadro de diálogo *Análisis factorial* (ver figura 20.1) para acceder al subcuadro de diálogo *Análisis factorial: Rotación* que se muestra en la figura 20.8.
- ▶ En el apartado **Método**, seleccionar la opción **Varimax**.

Aceptando estas selecciones, el *Visor* ofrece los resultados que muestran las tablas 20.19 a la 20.22 y la figura 20.11.

La tabla de comunalidades es la misma que la obtenida en la extracción no rotada (ver tabla 20.16). Es importante resaltar este aspecto, pues el proceso de rotación busca clarificar la interpretación de la estructura factorial sin alterar la situación relativa de unas variables respecto a las otras, y sin alterar tampoco el porcentaje de la varianza de cada variable que es capaz de explicar cada factor. Para que la comunalidad de las variables cambie es necesario variar el número de factores de la solución.

Lo que sí cambia en el proceso de rotación es el porcentaje de varianza total explicada por cada factor (y cambia tanto más cuanto más éxito tiene la rotación). Ahora, La tabla de porcentajes de varianza explicada (tabla 20.19) incorpora información adicional referente a la suma de las saturaciones tras la rotación de los factores. En nuestro ejemplo, podemos comprobar que las sumas de los cuadrados de las saturaciones no coinciden con las de la extracción no rotada, aunque difieren poco (por lo que podemos pensar que la rotación no mejora demasiado la interpretación de la solución factorial y que la extracción inicial ofrece ya una solución suficientemente clara).

**Tabla 20.19.** Porcentajes de varianza explicada antes y después de la rotación.

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3.166	52.760	52.760	2.941	49.013	49.013	2.902	48.359	48.359
2	1.856	30.927	83.688	1.654	27.572	76.585	1.694	28.226	76.585
3	.431	7.185	90.873						
4	.247	4.110	94.983						
5	.196	3.270	98.253						
6	.105	1.747	100.000						

Método de extracción: Factorización de Ejes principales.

La matriz de la *estructura factorial no rotada* (ver tabla 20.20) contiene la solución factorial antes de la rotación (es decir, contiene las saturaciones de las variables en los factores no rotados). Esta matriz es idéntica a la obtenida en la solución no rotada (ver tabla 20.18). El *Visor* ofrece esta tabla (que es el punto de partida del proceso de rotación) justo antes de presentar la matriz de la *estructura factorial rotada*.

**Tabla 20.20.** Matriz factorial (matriz de la estructura factorial no rotada).

	Factor <sup>a</sup>	
	1	2
Nivel educativo	.711	-.146
Categoría laboral	.785	.217
Salario actual	.956	.076
Salario inicial	.910	.220
Experiencia previa (meses)	-.169	.890
Edad (años)	-.220	.860

Método de extracción: Factorización del eje principal.

a. 2 factores extraídos. Requeridas 12 iteraciones.

La matriz de la *estructura factorial rotada* aparece en la tabla denominada *matriz de factores rotados* (ver tabla 20.21). Comparándola con la matriz no rotada de la tabla 20.20, podemos comprobar que ha mejorado algo la saturación de las dos variables agrupadas en el segundo factor: se han incrementado sus saturaciones en dicho factor y han disminuido en el primero (recordemos que si la saturación de una variable aumenta en un factor, su saturación en los restantes factores debe disminuir para que se mantenga inalterado el valor de su comunalidad). La variable *categoría laboral* también se ha desplazado hacia su factor, el primero, disminuyendo su saturación en el segundo factor. Sin embargo, la variable *nivel educativo*, que antes saturaba fundamentalmente en el primer factor, a perdido parte de su correlación con él en beneficio del segundo factor, con el que ahora comparte más información.

En definitiva, el proceso de rotación busca lo que Thurstone (1947) denominó una *estructura simple*: variables que saturan, a ser posible, en un único factor, y factores que contengan un número reducido de variables que saturan inequívoca y exclusivamente en ellos. Con todo, las variables que compartan información con varios factores, si existen, entorpecerán el proceso de rotación y, en lugar de una única saturación elevada en un único factor, tenderán a mostrar saturaciones moderadas en varios factores (como ocurre en nuestro ejemplo, en parte, con la de la variable *nivel educativo*).

**Tabla 20.21.** Matriz de factores rotados (matriz de la estructura factorial rotada).

	Factor <sup>a</sup>	
	1	2
Nivel educativo	.675	-.268
Categoría laboral	.810	.076
Salario actual	.954	-.092
Salario inicial	.935	.058
Experiencia previa (meses)	-.011	.906
Edad (años)	-.066	.885

Método de extracción: Factorización del eje principal.

Método de rotación: Varimax con normalización de Kaiser.

a. La rotación ha convergido en 3 iteraciones.

La tabla 20.22 muestra la *matriz de transformación de los factores*, que es la matriz utilizada para rotar la solución inicial. Esta matriz adopta la forma:

$$\Lambda^* = \Lambda T$$

$$\text{con } \begin{cases} \mathbf{T} = \begin{pmatrix} \cos \phi & \text{sen } \phi \\ -\text{sen } \phi & \cos \phi \end{pmatrix} \text{ si la rotación se hace en el sentido de las agujas del reloj} \\ \mathbf{T} = \begin{pmatrix} \cos \phi & -\text{sen } \phi \\ \text{sen } \phi & \cos \phi \end{pmatrix} \text{ si la rotación se hace en el sentido contrario al de las agujas del reloj} \end{cases}$$

Donde  $\Lambda$  es la matriz de estructura factorial antes de la rotación,  $\mathbf{T}$  es la matriz de transformación y  $\Lambda^*$  es la matriz de estructura factorial después de la rotación. Igualando los términos de la matriz anterior con los de la tabla 20.22 y despejando, obtenemos que el ángulo de rotación es de aproximadamente  $10^\circ$  en el sentido contrario al de las agujas del reloj. Como el método de rotación utilizado consiste en una rotación ortogonal, los ejes rotados seguirán siendo ortogonales e independientes entre sí, es decir, seguirán formando un ángulo de  $90^\circ$  entre ellos.

**Tabla 20.22.** Matriz de transformación de los factores.

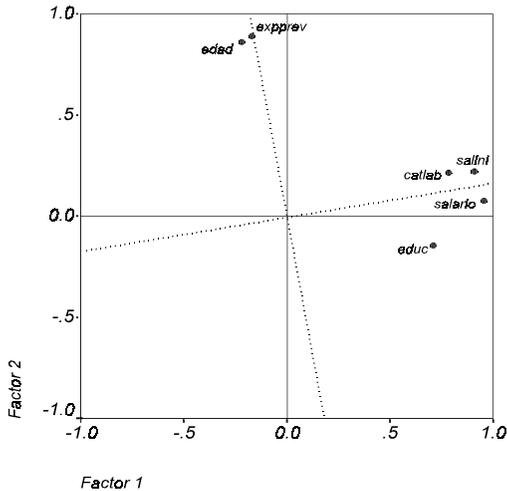
Factor	1	2
1	.985	-.175
2	.175	.985

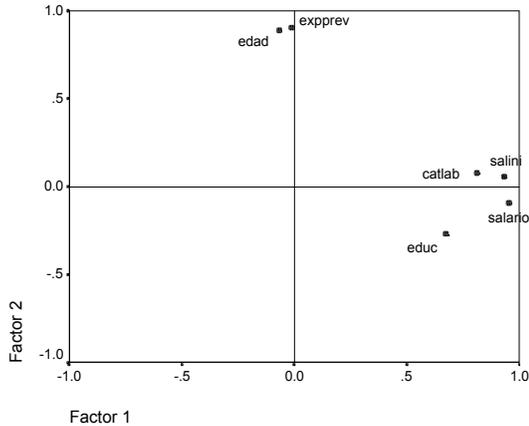
Método de extracción: Factorización del eje principal.

Método de rotación: Varimax con normalización de Kaiser.

Para ayudar a percibir con claridad el efecto de la rotación, la figura 20.10 muestra los ejes rotados superpuestos sobre el gráfico de la solución no rotada de la figura 20.9. No obstante, en la rotación que ofrece el *Visor* (ver figura 20.11), da la impresión de que son las variables las que se desplazan hacia los ejes en lugar de ser los ejes los que se desplazan hacia las variables.

**Figura 20.10.** Gráfico de las saturaciones factoriales (estructuras rotada y no rotada).



**Figura 20.11.** Gráfico de las saturaciones factoriales (estructura rotada).

El gráfico muestra con claridad cómo las variables pertenecientes al factor 2 se han aproximado más a él, cómo el grupo de variable pertenecientes al factor 1 ahora se encuentran atravesadas por el eje que representa dicho factor y cómo la variable *nivel educativo* se ha distanciado del factor 1, llevándonos a pensar que esta variable de hecho comparte información con el factor 2. A la vista del gráfico y de la matriz de estructura rotada, debemos pensar que la personas de mayor edad y experiencia tienden a presentar un menor nivel educativo en esta empresa y vice-versa, las personas de menor edad y experiencia laboral tienden a presentar un mayor nivel educativo.

### Ejemplo (Análisis factorial > Rotación oblicua)

Este ejemplo muestra cómo rotar la solución factorial original mediante un método oblicuo y compara los resultados de la rotación oblicua con los de la rotación ortogonal del ejemplo anterior. La estructura rotada es la ya obtenida con el método de extracción *ejes principales* (6 variables, 2 factores). Para seleccionar un método de rotación oblicuo:

- ▶ Pulsar el botón **Rotación...** del cuadro de diálogo *Análisis factorial* (ver figura 20.1) para acceder al subcuadro de diálogo *Análisis factorial: Rotación* que muestra la figura 20.8.
- ▶ En el apartado **Método**, seleccionar la opción **Oblimin directo**.

Aceptando estas selecciones, el *Visor* ofrece los resultados que muestran las tablas 20.23 a la 20.27.

La tabla de comunalidades es la misma que la ya obtenida en la extracción no rotada (ver tabla 20.16). Recordemos que el proceso de rotación busca clarificar la interpretación de la estructura factorial sin alterar la situación relativa de las variables ni el porcentaje de la varianza de cada variable que es capaz de explicar el conjunto de factores.

La tabla de porcentajes de varianza explicada (tabla 20.23) muestra información parcial sobre el resultado de la rotación: si bien la tabla recoge las sumas de cuadrados de las saturaciones de las variables en cada factor (última columna de la tabla), no es posible sumar esas sumas de cuadrados e interpretarlas como porcentajes de varianza explicada. Esto es debido a que los factores ya no tienen por qué cumplir la restricción de ser ortogonales entre sí y, consecuentemente, la varianza de la suma de los factores ya no es igual a la suma de las varianzas de los factores (ensanchando la última columna de la tabla puede verse su encabezamiento completo: *Suma de las saturaciones al cuadrado de la rotación*).

**Tabla 20.23.** Porcentajes de varianza explicada.

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total
1	3.166	52.760	52.760	2.941	49.013	49.013	2.922
2	1.856	30.927	83.688	1.654	27.572	76.585	1.718
3	.431	7.185	90.873				
4	.247	4.110	94.983				
5	.196	3.270	98.253				
6	.105	1.747	100.000				

Método de extracción: Factorización de Ejes principales.

- a. Cuando los factores están correlacionados, no se pueden sumar las sumas de los cuadrados de las saturaciones para obtener una varianza total.

La matriz de la *estructura factorial no rotada* (tabla 20.24) se ofrece en una tabla denominada *matriz factorial* y, al igual que ocurre con la tabla de comunalidades, también es idéntica a la obtenida anteriormente (ver tabla 20.18).

**Tabla 20.24.** Matriz factorial (matriz de la estructura factorial no rotada).

	Factor <sup>a</sup>	
	1	2
Nivel educativo	.711	-.146
Categoría laboral	.785	.217
Salario actual	.956	.076
Salario inicial	.910	.220
Experiencia previa (meses)	-.169	.890
Edad (años)	-.220	.860

Método de extracción: Factorización del eje principal.

a. 2 factores extraídos. Requeridas 12 iteraciones.

A diferencia de lo que ocurre en la rotación ortogonal, los resultados de la rotación oblicua no pueden representarse en una única matriz. Si los factores son ortogonales (independientes entre sí), la saturación de una variable en un factor, es decir, su proyección sobre el factor, es igual a la correlación de esa variable con el factor. Pero si los factores son oblicuos (correlacionan entre sí), la saturación y la correlación de una variable en un factor no coinciden. Por este motivo, al solicitar una rotación oblicua, el *Visor* muestra dos matrices para la estructura factorial rotada: una con las correlaciones (a la que llama *matriz factorial*) y otra con las saturaciones (a la que llama *matriz de configuración*).

La *matriz de configuración* ofrece las saturaciones de las variables en los factores de la solución rotada. Esas saturaciones, que son las que se representan en el gráfico del espacio factorial rotado, representan la contribución neta de cada variable en cada factor, por lo que constituyen la manera más fácil de interpretar la solución factorial.

**Tabla 20.25.** Matriz de configuración (saturaciones de la estructura rotada).

	Factor <sup>a</sup>	
	1	2
Nivel educativo	.665	-.231
Categoría laboral	.818	.122
Salario actual	.954	-.039
Salario inicial	.942	.110
Experiencia previa (meses)	.034	.909
Edad (años)	-.023	.885

Método de extracción: Factorización del eje principal.  
 Metodo de rotación: Oblimin con normalización de Kaiser.

a. La rotación ha convergido en 4 iteraciones.

La *matriz de estructura* contiene las correlaciones de las variables con los factores de la solución rotada. Estas correlaciones representan la contribución bruta de cada variable a cada factor. Cuando los factores correlacionan mucho entre sí (se encuentran muy próximos en el espacio), la matriz de estructura contiene correlaciones muy grandes entre todas las variables y todos los factores, lo cual hace muy difícil la interpretación por la imposibilidad de precisar a qué factor único hay asignar cada variable (si bien esto representa la situación real: las variables que correlacionan con un factor también lo harán con los factores relacionados con él).

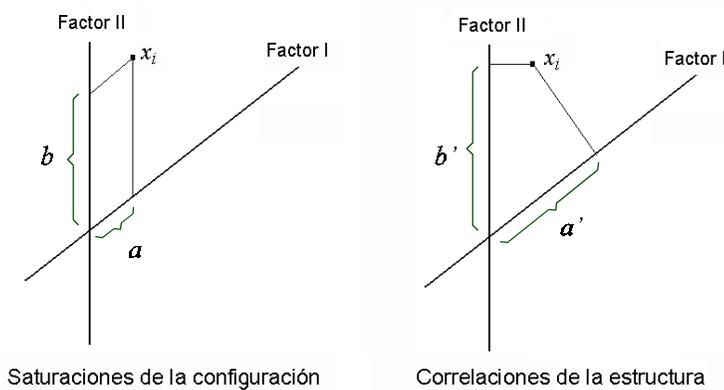
**Tabla 20.26.** Matriz de estructura (correlaciones de la estructura rotada).

	Factor	
	1	2
Nivel educativo	.689	-.301
Categoría laboral	.805	.036
Salario actual	.958	-.138
Salario inicial	.930	.012
Experiencia previa (meses)	-.061	.905
Edad (años)	-.115	.888

Método de extracción: Factorización del eje principal.  
 Metodo de rotación: Normalización Oblimin con Kaiser.

En la figura 20.12 están representadas las proyecciones espaciales que dan lugar a las saturaciones factoriales de la matriz de configuración y a las correlaciones variables-factores de la matriz de estructura. Todo ello dentro del espacio factorial definido por dos factores correlacionados entre sí (oblicuos). En la figura puede apreciarse que las saturaciones y las correlaciones más altas en un factor (proyecciones más largas:  $b$  y  $b'$ ) corresponden a las variables que más cerca se encuentran de él. Sin embargo, justamente por tratarse de factores oblicuos, mientras la saturación de una variable en el factor del que se encuentra más alejada es muy reducida (proyección  $a$ ), la correlación de esa misma variable con ese mismo factor (proyección  $a'$ ) es más alta que la saturación.

**Figura 20.12.** Representación espacial de las saturaciones y de las correlaciones.



La *matriz de correlaciones entre los factores* (tabla 20.27) permite apreciar el grado de proximidad existente entre los factores: cuanto mayor sea la correlación entre los factores (en valor absoluto), más próximos se encontrarán éstos en el espacio. Estas correlaciones están estrechamente relacionadas con el ángulo que forman los factores: equivalen al coseno del ángulo comprendido entre ellos. En nuestro ejemplo:  $\text{correlación}(\text{factor } 1, \text{factor } 2) = -0,104 = \cos(\theta)$  y  $\text{arccos}(-0,104) = \theta \cong 96^\circ$ . Los factores se encuentran abiertos un poco más de 90 grados; por tanto, no son ortogonales (si los factores fueran ortogonales o independientes, la correlación entre ambos valdría cero y el ángulo comprendido entre ellos sería exactamente de 90 grados).

**Tabla 20.27.** Matriz de correlaciones entre los factores.

Factor	1	2
1	1.000	-.104
2	-.104	1.000

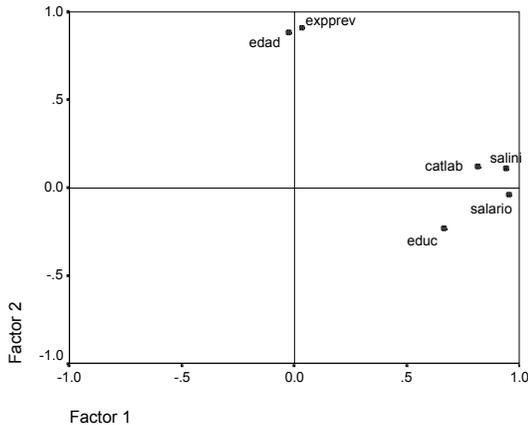
Método de extracción: Factorización del eje principal.

Método de rotación: Oblimin con normalización de Kaiser.

Puede darse el caso de que, aun solicitando una rotación oblicua, los factores permanezcan ortogonales. El algoritmo de rotación oblicua busca rotar de manera autónoma cada uno de los factores, pero eso no quiere decir que por ello los factores deban aproximarse entre sí cuando la solución ortogonal es la mejor de las posibles. En nuestro ejemplo, el ángulo entre factores ha permanecido prácticamente en 90 grados (casi ortogonales), razón por la cual la matriz de configuración y la matriz de estructura apenas difieren entre sí.

Cuando se realiza una rotación oblicua, el *gráfico de las saturaciones en el espacio factorial rotado* (figura 20.13) puede resultar engañoso. Aunque el gráfico representa la posición relativa de las variables en los factores, el ángulo entre los factores se mantiene en 90 grados, independientemente del ángulo real obtenido con la rotación. Para interpretar correctamente el gráfico debe tenerse en cuenta cuál es la posición de las variables respecto del factor en el que más saturan (que es el factor al que se encuentran más próximas). Sin embargo, la inclinación de los ejes debe intuirse a partir de los valores de la *matriz de correlaciones entre los factores*.

**Figura 20.13.** Gráfico de las saturaciones en el espacio factorial rotado.

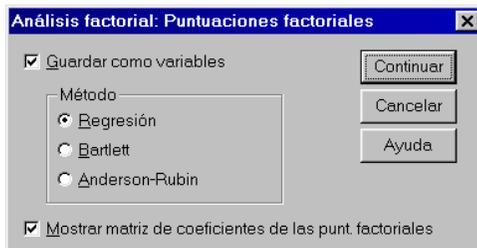


## Puntuaciones factoriales

Una vez alcanzada la solución factorial final, suele resultar interesante obtener una estimación de las puntuaciones de los sujetos en cada uno de los factores resultantes de la extracción a fin de valorar la situación relativa de cada sujeto en esas “dimensiones ocultas” capaces de resumir la información contenida en las variables originales. El cuadro de diálogo *Puntuaciones factoriales* contiene las opciones que permiten solicitar las estimaciones de las *puntuaciones factoriales* y seleccionar el método de estimación que se desea utilizar para obtener tales estimaciones. Para acceder a estas opciones:

- ▶ Pulsar en el botón **Puntuaciones...** del cuadro de diálogo *Análisis factorial* (ver figura 20.1) para acceder al subcuadro de diálogo *Análisis factorial: Puntuaciones factoriales* que muestra la figura 20.14.

**Figura 20.14.** Subcuadro de diálogo *Análisis factorial: Puntuaciones factoriales*.



- **Guardar como variables.** Activando esta opción se guardan automáticamente en el *Editor de datos* las puntuaciones factoriales estimadas para cada sujeto en cada uno de los factores obtenidos en la solución factorial. Para ello, el SPSS crea en el archivo de datos activo tantas variables nuevas como factores contenga la solución factorial. Si no se selecciona esta opción no es posible acceder a los métodos de estimación de las puntuaciones factoriales.

**Método.** Este apartado contiene varios métodos de estimación de las puntuaciones factoriales. Por defecto se encuentra seleccionado el método de *Regresión*, que es el de uso más generalizado. Es importante señalar que las opciones de este recuadro no tienen efecto alguno cuando se ha seleccionado *componentes principales* como método de extracción, ya que en ese modelo factorial las puntuaciones factoriales no son estimadas, sino calculadas directamente a partir de las variables originales.

- **Regresión.** Método de estimación de las puntuaciones factoriales en el que las estimaciones resultantes tienen una media de cero y una varianza igual al cuadrado de la correlación múltiple entre las puntuaciones factoriales estimadas y los valores factoriales verdaderos. Las puntuaciones factoriales estimadas con este método pueden estar correlacionadas incluso cuando los factores son ortogonales.
- **Bartlett.** Método de estimación de las puntuaciones factoriales en el que las estimaciones resultantes tienen una media de cero. Este método minimiza la suma de cuadrados de los factores únicos (es decir, minimiza la *unicidad* correspondiente a cada una de las variables incluidas en el análisis).
- **Anderson-Rubin.** Este método de estimación es una modificación del método de Bartlett que asegura la ortogonalidad de las puntuaciones factoriales estimadas. Las estimaciones resultantes tienen una media de cero, una desviación típica de uno y son independientes entre sí (incluso en el caso de que se haya solicitado una solución rotada oblicua).

**Mostrar matriz de coeficientes de las puntuaciones factoriales.** Esta opción permite obtener una tabla con los pesos o ponderaciones necesarios para calcular las puntuaciones factoriales a partir de las variables originales. Esta opción se encuentra desactivada por defecto. Por tanto, para obtener la matriz de coeficientes no basta con solicitar las puntuaciones factoriales.

## Ejemplo (Análisis factorial > Puntuaciones)

Este ejemplo muestra cómo obtener e interpretar las estimaciones de las puntuaciones factoriales. Se compararán varios métodos de estimación de las puntuaciones factoriales. Las puntuaciones factoriales de los sujetos dependerán del método de extracción utilizado, el método de rotación elegido y el método de estimación de las puntuaciones factoriales seleccionado. Cada combinación de estos tres aspectos del análisis dará lugar a un conjunto de puntuaciones factoriales distintas para un sujeto dado. La elección de las puntuaciones factoriales más adecuadas dependerán de los propósitos del investigador. Así, por ejemplo, si se desea realizar un análisis factorial de segundo orden sobre las puntuaciones factoriales, no deberá seleccionarse el método de Anderson-Rubin ya que este método impone la ortogonalidad de las puntuaciones factoriales y el análisis factorial ulterior no tendría sentido. En dicho caso sería recomendable utilizar el método de Regresión con un rotación Obllicua y asegurar así la posibilidad de explorar las relaciones existentes entre los factores. Para solicitar las estimaciones de las puntuaciones factoriales:

- ▶ En el cuadro de diálogo *Análisis factorial* (ver figura 20.1), seleccionar las seis variables con las que venimos trabajando en los últimos ejemplos: *educ*, *catlab*, *salario*, *salini*, *exprev* y *edad*, y trasladarlas a la lista **Variables**.
- ▶ Pulsar en el botón **Puntuaciones...** para acceder al subcuadro de diálogo *Análisis factorial: Puntuaciones factoriales* (ver la figura 20.15). Marcar la opción **Guardar como variables** (al marcar esta opción, aparece seleccionada, por defecto, la opción **Regresión** del apartado **Método**) y la opción **Mostrar la matriz de coeficientes de las puntuaciones factoriales**. Pulsar el botón **Continuar**.
- ▶ Pulsar el botón **Extracción...** para acceder al subcuadro de diálogo *Análisis factorial: Extracción* (ver figura 20.4). Seleccionar la opción *Componentes principales* en el menú emergente del cuadro **Método** e introducir un 2 en el cuadro de texto **Número de factores**. Pulsar el botón **Continuar**.
- ▶ Pulsar el botón **Rotación...** para acceder al subcuadro de diálogo *Análisis factorial: Rotación* (ver figura 20.8) y seleccionar la opción **Ninguno** del apartado **Método**. Pulsar el botón **Continuar**.

Aceptando estas selecciones, el *Visor* ofrece, además de las tablas ya vistas en los ejemplos anteriores (comunalidades, porcentajes de varianza explicada, matriz de componentes, etc.), dos nuevas tablas con información referida a las puntuaciones factoriales solicitadas.

La tabla 20.28 muestra la *matriz de coeficientes para el cálculo de las puntuaciones factoriales*, la cual contiene las ponderaciones que recibe cada variable en el cálculo de las puntuaciones factoriales. Puesto que hemos utilizado el método de extracción de *componentes principales*, las dimensiones obtenidas reciben el nombre de *componentes* (en lugar de *factores*).

**Tabla 20.28.** Matriz de coeficientes para el cálculo de las puntuaciones factoriales.

	Componente	
	1	2
Nivel educativo	.255	-.093
Categoría laboral	.266	.141
Salario actual	.298	.048
Salario inicial	.288	.127
Experiencia previa (meses)	-.057	.500
Edad (años)	-.074	.492

Método de extracción: Análisis de componentes principales.  
Puntuaciones de componentes.

Combinando cada variable con sus correspondientes coeficientes pueden construirse las dos ecuaciones lineales en las que se basa el cálculo de las puntuaciones factoriales:

$$Y_1 = 0,255 \times educ + 0,266 \times catlab + 0,298 \times salario + 0,288 \times salini - 0,057 \times expprev - 0,074 \times edad$$

$$Y_2 = -0,093 \times educ + 0,141 \times catlab + 0,048 \times salario + 0,127 \times salini + 0,500 \times expprev + 0,492 \times edad$$

Las dos puntuaciones factoriales de un sujeto se obtienen sustituyendo cada variable por sus respectivos valores.

La tabla 20.29, obtenida con el procedimiento **Informes > Resúmenes de casos** del menú **Analizar** muestra las puntuaciones factoriales de los 10 primeros sujetos.

**Tabla 20.29.** Listado de las puntuaciones factoriales de los 10 primeros sujetos.

	REGR factor score 1 for analysis 1	REGR factor score 2 for analysis 1
1	1.38309	.88997
2	.28680	-.46629
3	-1.00661	2.35341
4	-1.09351	.85072
5	.28864	.23986
6	-.14950	-.40335
7	.07017	.01337
8	-.64438	-1.03708
9	-.35485	.32729
10	-.72959	1.03906
Total N	10	10

Las puntuaciones factoriales se encuentran en formato *diferencial*, por lo que una puntuación de cero se corresponde con una puntuación factorial igual a la media, las puntuaciones positivas son puntuaciones mayores que la media y las puntuaciones negativas son puntuaciones menores que la media. Si se desea eliminar los signos negativos siempre es posible realizar una transformación de las puntuaciones para cambiar la escala de las nuevas variables.

Las puntuaciones factoriales se almacenan de manera automática en el *Editor de datos* y reciben de forma automática un nombre que identifica, por este orden, el método de estimación de las puntuaciones (en el ejemplo, *REGR*), el número del factor al que corresponden las puntuaciones (*factor score 1*) y el número del análisis (*analysis 1*). Este nombre es único y distintivo, de manera que si se solicitan nuevas estimaciones de las puntuaciones, las nuevas puntuaciones se almacenarán al final del archivo de datos con nuevos nombres.

Para interpretar mejor las puntuaciones factoriales de los sujetos conviene solicitar algunos estadísticos descriptivos de las nuevas variables. La tabla 20.30 muestra algunos descriptivos obtenidos con el procedimiento **Estadísticos descriptivos > Descriptivos** del menú **Analizar**.

La media de las nuevas variables vale 0 y su desviación típica 1, lo que significa que podemos interpretar las puntuaciones de los sujetos como si fueran puntuaciones típicas. Inspeccionando la tabla 20.29 podemos apreciar que el primer sujeto recibe una puntuación alta y por encima de la media en el primer factor y también una puntuación moderadamente alta en el segundo. El tercer sujeto puntúa bajo en el primer factor, (se encuentra a una desviación típica por debajo de la media) y muy alto en el segundo (se encuentra a más de dos desviaciones típicas por encima de la media). El séptimo sujeto se encuentra situado en torno a la media en ambos factores.

**Tabla 20.30.** Estadísticos descriptivos de las puntuaciones factoriales.

	N	Mínimo	Máximo	Media	Desv. típ.
REGR factor score 1 for analysis 1	473	-1.5290	4.9499	.0000	1.0000
REGR factor score 2 for analysis 1	473	-1.2124	3.0834	.0000	1.0000
N válido (según lista)	473				

Además de las puntuaciones factoriales, el procedimiento ofrece también la *matriz de varianzas-covarianzas* de las puntuaciones factoriales (tabla 20.31). Lógicamente, esta matriz contiene, en la diagonal principal (es decir, en las casillas 1:1 y 2:2), la varianza de las puntuaciones factoriales de cada componente o factor (que ya sabemos que vale uno) y, fuera de la diagonal principal (es decir, en las casillas 1:2 y 2:1), las covarianzas existentes entre cada par de componentes o factores (covarianza que en nuestro ejemplo vale cero, indicando esto que las puntuaciones factoriales de ambos factores son completamente independientes entre sí: su correlación es nula).

**Tabla 20.31.** Matriz de varianzas-covarianzas de las puntuaciones factoriales.

Componente	1	2
1	1.000	.000
2	.000	1.000

Método de extracción: Componentes principales.  
Puntuaciones de componentes

Esta circunstancia (la independencia completa entre las puntuaciones factoriales) es bastante excepcional cuando las estimaciones se efectúan mediante el método de *regresión*. De hecho, se trata de un caso muy particular. Pero obsérvese que la tabla 20.31 no informa sobre el método de estimación utilizado (a pesar de que en el cuadro de diálogo *Análisis factorial: Puntuaciones* hemos seleccionado el método de estimación *regresión*), sino que se limita a señalar que las puntuaciones factoriales analizadas son las *puntuaciones en las componentes principales*. Cuando la extracción de los factores se realiza con el método *componentes principales*, las puntuaciones factoriales no se obtienen mediante estimación, sino que son directamente calculadas a partir de la solución factorial. Y puesto que la extracción con el método *componentes principales* siempre ofrece una solución ortogonal, las puntuaciones factoriales basadas en esa solución también serán ortogonales.

Sin embargo, cuando se utiliza un método de extracción distinto del de *componentes principales* no es posible obtener directamente las puntuaciones factoriales a partir de la matriz de estructura, sino que deben ser estimadas mediante uno cualquiera de los métodos de estimación disponibles.

La tabla 20.32 ofrece la *matriz de coeficientes* obtenida con el método de extracción *componentes principales*, el método de rotación *Oblimin directo* y el método *regresión* para la estimación de las puntuaciones factoriales.

Puede observarse en la tabla que las nuevas ponderaciones son ligeramente distintas de las obtenidas con la solución no rotada. Pero no tiene sentido estudiar esta matriz de coeficientes pues sólo sirve como instrumento de cálculo en el caso de que se desee estimar las puntuaciones factoriales con alguna herramienta distinta del SPSS. Si se desea estudiar el cambio en la estructura factorial es más adecuado referirse a las matrices de configuración y de estructura.

**Tabla 20.32.** Matriz de coeficientes para el cálculo de las puntuaciones factoriales.

	Componente	
	1	2
Nivel educativo	.240	-.148
Categoría laboral	.283	.077
Salario actual	.302	-.021
Salario inicial	.302	.058
Experiencia previa (meses)	.012	.500
Edad (años)	-.006	.496

Método de extracción: Análisis de componentes principales.

Método de rotación: Oblimin con normalización de Kaiser.

Puntuaciones de las componentes.

Donde se pueden apreciar con claridad las diferencias entre las soluciones rotada y no rotada es en la *matriz de varianzas-covarianzas* (tabla 20.33). Ahora, en la solución rotada, las varianzas de las puntuaciones siguen valiendo aproximadamente uno, pero la covarianza entre las puntuaciones ya no es nula, sino que existe una ligera relación negativa entre las puntuaciones de ambas componentes.

**Tabla 20.33.** Matriz de varianzas-covarianzas de las puntuaciones factoriales.

Componente	1	2
1	1.009	-.186
2	-.186	1.009

Método de extracción: Análisis de componentes principales.

Método de rotación: Oblimin con normalización de Kaiser.

Puntuaciones de las componentes.

Por último, veamos qué efecto tiene sobre las puntuaciones factoriales aplicar el método de extracción *ejes principales*, manteniendo el método de rotación *Oblimin directo* y el método de estimación *regresión*. La tabla 20.34 muestra los coeficientes obtenidos. Ahora sí se menciona en una nota a pie de tabla el método de estimación utilizado para obtener las puntuaciones factoriales (*Método de puntuaciones factoriales: Regresión*), pues al utilizar un método de extracción distinto de *componentes principales*, las puntuaciones factoriales no pueden calcularse directamente sino que necesitan ser estimadas. Tal vez sorprenda observar que las variables del primer factor no reciben ponderaciones similares a las obtenidas con el método *componentes principales* (ver tabla 20.32). Esto es debido a que el método *regresión* es muy similar al análisis de regresión múltiple y, cuando existe colinealidad entre las variables, no es necesario incorporar la misma información de manera repetitiva.

**Tabla 20.34.** Matriz de coeficientes para el cálculo de las puntuaciones factoriales.

	Factor	
	1	2
Nivel educativo	.060	-.056
Categoría laboral	.091	.035
Salario actual	.531	-.067
Salario inicial	.356	.060
Experiencia previa (meses)	-.001	.527
Edad (años)	-.019	.440

Método de extracción: Factorización del eje principal.

Método de rotación: Oblimin con normalización de Kaiser.

Método de puntuaciones factoriales: Regresión.

En la matriz de varianzas-covarianzas de las puntuaciones factoriales (tabla 20.35) podemos ver que las puntuaciones factoriales están más relacionadas que antes y que las varianzas de las puntuaciones factoriales de ambos factores son distintas de uno y distintas entre sí. Normalmente, cuando se utiliza el método de estimación *regresión*, las puntuaciones factoriales correspondientes a factores con autovalores más grandes suelen tener una varianza mayor. Además, si se utiliza un método de extracción distinto del de *componentes principales*, las puntuaciones factoriales tienden a mostrar cierta correlación incluso aunque no se haya efectuado una rotación oblicua.

**Tabla 20.35.** Matriz de varianzas-covarianzas de las puntuaciones factoriales.

Factor	1	2
1	.988	-.296
2	-.296	.926

Método de extracción: Factorización del eje principal.

Método de rotación: Oblimin con normalización de Kaiser.

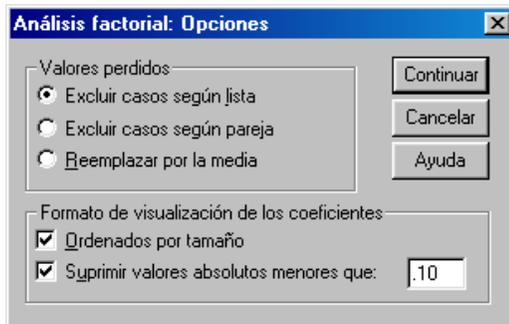
Método de puntuaciones factoriales: Regresión.

## Opciones

El cuadro de diálogo *Análisis factorial: Opciones* permite controlar algunos aspectos relacionados con el tratamiento que deben recibir los valores perdidos y el formato de las tablas de resultados que genera el *Visor de resultados*. Para controlar estos aspectos:

- ▶ Pulsar en el botón Opciones... del cuadro de diálogo *Análisis factorial* (ver figura 20.1) para acceder al subcuadro de diálogo *Análisis factorial: Opciones* que muestra la figura 20.15.

**Figura 20.15.** Subcuadro de diálogo Análisis factorial: Opciones.



**Valores perdidos.** Este recuadro permite controlar el tratamiento que se desea dar a valores perdidos.

- **Excluir casos según lista.** Es la opción por defecto. Se excluyen del análisis los sujetos que tengan valores perdidos en cualquiera de las variables trasladadas a la lista **Variab**les del cuadro de diálogo *Análisis factorial* (ver figura 20.1). Es el tratamiento más consistente de todos: sólo se incluyen en el análisis los casos *completos* (es decir, los casos con puntuación válida en todas las variables seleccionadas). Sin embargo, conviene tener en cuenta que esta forma de tratar los valores perdidos puede suponer la pérdida de un gran número de casos y la consiguiente reducción del tamaño efectivo de la muestra.
- **Excluir casos según pareja.** Los sujetos con valor perdido en una variable se excluyen del análisis sólo para el cálculo de los estadísticos en los que esté implicada esa variable. Este método permite aprovechar más cantidad de información que el anterior, pero, puesto que no todas las correlaciones se calculan sobre el mismo número de sujetos, podrían obtenerse matrices de correlaciones inconsistentes imposibles de analizar posteriormente.
- **Reemplazar por la media.** Los valores perdidos en una variable se sustituyen por la *media* de esa variable. Si en una variable existen muy pocos casos con valor perdido, reemplazar el valor perdido por la media no constituye un problema importante. Pero en la medida en que el número de valores perdidos aumenta, la sustitución por la media tiene el efecto de *centrar* las variables disminuyendo su variabilidad.

**Formato de visualización de los coeficientes.** Las opciones de este apartado permiten cambiar algunos aspectos relacionados con el formato de presentación de las tablas.

- Ordenados por tamaño.** Esta opción sirve para ordenar las variables de las tablas de resultados en función de la magnitud (en valor absoluto) de los coeficientes de esas tablas (saturaciones, correlaciones, etc). La ordenación se realiza de forma ascendente: primero las variables con coeficientes más altos. Si no se marca esta opción, las tablas muestran las variables en el mismo orden en el que han sido trasladadas a la lista de **Variables** del cuadro de diálogo *Análisis factorial* (ver figura 20.1).
- Suprimir valores absolutos menores que...** Esta opción permite suprimir de las tablas de resultados los coeficientes cuyo un valor absoluto sea menor que el valor establecido en el cuadro de texto. El valor por defecto es 0,10, pero este valor puede cambiarse introduciendo un valor distinto. Esta opción es de gran ayuda: al desaparecer de la tabla los coeficientes excesivamente pequeños (en valor absoluto), se facilita notablemente la interpretación de los resultados.

## Ejemplo (Análisis factorial > Opciones)

Este ejemplo muestra cómo utilizar algunas de las opciones disponibles en el procedimiento. En concreto, muestra cómo ordenar las variables en la matriz de estructura y cómo suprimir las saturaciones pequeñas.

Utilizaremos las mismas seis variables que hemos venido utilizando en los ejemplos anteriores, *ejes principales* como método de extracción (forzando una solución de 2 factores) y *varimax* como método de rotación. Para acceder a las opciones del procedimiento:

- ▶ En el cuadro de diálogo *Análisis factorial: Opciones* (ver figura 20.15), seleccionar las opciones **Ordenar por tamaño** y **Suprimir valores absolutos menores que 0,10** del apartado Formato de visualización de los coeficientes.

Aceptando estas selecciones, el *Visor de resultados* ofrece además de las comunalidades, porcentajes de varianza explicada, etc., la matriz de la estructura factorial rotada que muestra la tabla 20.35.

**Tabla 20.36.** Matriz de factores rotados (estructura factorial de la solución rotada).

	Factor <sup>a</sup>	
	1	2
Salario actual	.954	
Salario inicial	.935	
Categoría laboral	.810	
Nivel educativo	.675	-.268
Experiencia previa (meses)		.906
Edad (años)		.885

Método de extracción: Factorización del eje principal.

Método de rotación: Varimax con normalización de Kaiser.

a. La rotación ha convergido en 3 iteraciones.

Las opciones que hemos seleccionado sólo tienen efecto sobre las matrices de *configuración* y de *estructura*. Lo primero que podemos observar en la tabla 20.36 es que las saturaciones menores que 0,01 (en valor absoluto) se han suprimido. Dado que las saturaciones muy pequeñas suelen carecer de valor interpretativo, suprimirlas de la tabla permite que nuestra atención pueda centrarse con más facilidad en las saturaciones más relevantes. Esta opción cumple únicamente el propósito de ayudar a interpretar la solución factorial; por tanto, no debe utilizarse si se tiene intención de publicar los resultados.

Además de la supresión de las saturaciones, la tabla 20.36 muestra las variables ordenadas por el tamaño de sus saturaciones. Primero se encuentran las variables que más saturan en el primer factor (empezando por las saturaciones más altas); después, las que más saturan en el segundo factor. De nuevo se trata de una opción cuya única función es la de facilitar la interpretación de la matriz de saturaciones.

< Fin del capítulo 20 >