Godinez Bravo Diego
Estadística Multivariada
Tarea 4

Ejercicio 1. Sea X vector aleatorio de dimensión $p = 3$, con vector de medias $\mu = (0, 1, 1)'$ y matriz de varianzas y covarianzas

$$\Sigma = \begin{pmatrix} 3 & -4 & 2 \\ -4 & 12 & -4 \\ 2 & -4 & 3 \end{pmatrix}$$

Se sabe que X sigue un modelo factorial de un único factor, con matriz de varianzas específicas $\Psi = \text{diag}(1, 4, 1)$.

a) Escriba el modelo factorial y calcule la matriz de cargas.

b) Calcule las comunalidades y los porcentajes de variación de cada variable explicados por el factor del modelo obtenido en (a).

c) Discútase si la solución a los apartados (a) y (b) es única.

Solución

Algebraicamente el modelo de análisis de factores se establece de la siguiente manera

$$X_1 - \mu_1 = l_{11}F_1 + l_{12}F_2 + ... + l_{1m}F_m + \epsilon_1$$

$$X_2 - \mu_2 = l_{21}F_1 + l_{22}F_2 + ... + l_{2m}F_m + \epsilon_2$$

$$\vdots$$

$$X_p - \mu_p = l_{p1}F_1 + l_{p2}F_2 + ... + l_{pm}F_m + \epsilon_p$$

Considerando que la matriz de covarianzas se puede expresar de la siguiente manera

$$\Sigma = LL' + \Psi$$

Entonces

$$\begin{array}{ccc} 3 = l_{11}^2 + \psi_1 & -4 = l_{11}l_{21} & 2 = l_{11}l_{31} \\ & 12 = l_{21}^2 + \psi_2 & -4 = l_{21}l_{31} \\ & & 3 = l_{31}^2 + \psi_3 \end{array}$$

Utilizando las ecuaciones

$$2 = l_{11}l_{31} \qquad -4 = l_{21}l_{31}$$

Encontramos que

$$l_{31} = \frac{2}{l_{11}}$$

Por lo tanto

$$-4 = l_{21}(\frac{2}{l_{11}})$$

$$l_{21} = -2l_{11}$$

Sustituyendo l_{21}

$$-4 = l_{11}(-2l_{11})$$

$$l_{11} = \sqrt{2}$$

Sustituyendo l_{11}

$$l_{21} = -2l_{11}$$

$$l_{21} = -2\sqrt{2}$$

Por último

$$l_{31} = \frac{2}{l_{11}}$$

$$l_{31} = \frac{2}{\sqrt{2}}$$

Los respectivos valores encontrados de l_{ij} son

$$l_{11} = \sqrt{2} \qquad l_{21} = -2\sqrt{2} \qquad l_{31} = \frac{2}{\sqrt{2}}$$

$$L = \begin{pmatrix} \sqrt{2} \\ -2\sqrt{2} \\ \frac{2}{\sqrt{2}} \end{pmatrix}$$

De manera que el modelo factorial queda expresado de la siguiente manera

$$X_1 = \sqrt{2}F_1 + \epsilon_1$$

$$X_2 = -2\sqrt{2}F_2 + \epsilon_2$$

$$X_3 = \frac{2}{\sqrt{2}}F_3 + \epsilon_3$$

Solución

La contribución de los factores comunes, expresada por la suma de cuadrados de las cargas en esa variable se denomida *comunalidad*.

$$h_i^2 = l_{i1}^2 + l_{i2}^2 + ... + l_{im}^2$$

Por lo tanto las comunalidades son

$$\begin{pmatrix} 2 \\ 8 \\ 2 \end{pmatrix}$$

Porcentajes de Variación

La contribución del factor específico, llamada varianza específica esta determinada por

$$Var(X_i) = h_i^2 + \Psi_i$$

Por lo tanto, para X_1

$$Var(X_1) = (\sqrt{2})^2 + 1$$

$$Var(X_1) = 3$$

De manera que el porcentaje de variación para X_1 es

$$\frac{h_1^2}{V(X_1)} = \frac{2}{3} = 0{,}66 = 66\,\%$$

Por otro lado para X_2

$$Var(X_2) = (-2\sqrt{2})^2 + 4$$

$$Var(X_2) = 12$$

De manera que el porcentaje de variación para X_2 es

$$\frac{h_2^2}{V(X_2)} = \frac{8}{12} = 0{,}66 = 66\,\%$$

Por último sea X_3

$$Var(X_3) = (\frac{2}{\sqrt{2}})^2 + 1$$

$$Var(X_3) = 3$$

De manera que el porcentaje de variación para X_3 es

$$\frac{h_3^2}{V(X_3)} = \frac{2}{3} = 0{,}66 = 66\,\%$$

Solución

La respuesta obtenida en el inciso a) **no es única**, el modelo también es consistente con la matriz de cargas

$$L = \begin{pmatrix} -\sqrt{2} \\ 2\sqrt{2} \\ -\frac{2}{\sqrt{2}} \end{pmatrix}$$

Sin embargo, para el caso de las comunalidades obtenidas estas son únicas, ya que al ser calculadas como $h_i^2 = l_{i1}^2 + l_{i2}^2 + ... + l_{im}^2$, el signo no influye en el resultado obtenido.

Ejercicio 2. La matriz

$$R = \begin{pmatrix} 1 & 0{,}69 & 0{,}28 & 0{,}35 \\ & 1 & 0{,}255 & 0{,}195 \\ & & 1 & 0{,}61 \\ & & & 1 \end{pmatrix}$$

Exhibe las correlaciones muestrales entre 4 variables que caracterizan el estado financiero de una empresa.

a) Calcule los valores y vectores propios de R.

b) Plantee el modelo factorial ortogonal con m factores para el vector X que generó estos datos.

c) Mediante el método de componentes principales, en los modelos factoriales con $m = 2$ y $m = 3$ factores, calcule las matrices de cargas, las comunalidades y el porcentaje que supone la comunalidad respecto a la varianza de cada variable.

d) Decida razonadamente entre el modelo con 2 o 3 factores.

e) Para el modelo seleccionado en el apartado d), calcule las correlaciones entre Z_2 y todos los factores. Estime la varianza específica para Z_2.

Solución

Haciendo uso del lenguaje de programación R se obtuvo el cálculo de los vectores y valores propios.

Valores propios

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (2{,}1934 \; 1{,}1139 \; 0{,}4245 \; 0{,}2680)$$

Vectores propios

$$(e_1, e_2, e_3, e_4) = \begin{pmatrix} -0{,}5400 & 0{,}4242 & 0{,}3410 & 0{,}6419 \\ -0{,}4938 & 0{,}5396 & -0{,}3205 & -0{,}6017 \\ -0{,}4796 & -0{,}5123 & -0{,}6539 & 0{,}2825 \\ -0{,}4842 & -0{,}5160 & 0{,}5943 & -0{,}3820 \end{pmatrix}$$

Solución

La matriz de cargas estimada $\hat{L}$ está dada por

$$\hat{L} = [\sqrt{\hat{\lambda}_1}\hat{e}_1 | \sqrt{\hat{\lambda}_2}\hat{e}_2 | ... | \sqrt{\hat{\lambda}_m}\hat{e}_m]$$

De manera similar se utilizó el lenguaje de programación R para facilitar el cálculo de $\hat{L}$. La matriz de cargas $\hat{L}$ se encuentra definida de la siguiente manera

$$\hat{L} = \begin{pmatrix} -0{,}7997 & 0{,}4477 & 0{,}2222 & 0{,}3323 \\ -0{,}7313 & 0{,}5695 & -0{,}2088 & -0{,}3115 \\ -0{,}7104 & -0{,}5407 & -0{,}4260 & 0{,}1462 \\ -0{,}7171 & -0{,}5446 & 0{,}3872 & -0{,}1978 \end{pmatrix}$$

Utilizando la matriz de cargas $\hat{L}$, calculamos Ψ de la siguiente manera

$$\Psi = diag(R - \hat{L}\hat{L}')$$

$$\Psi = \begin{pmatrix} 0{,}0000e^{00} & 0 & 0 & 0 \\ 0 & 1{,}4432e^{-15} & 0 & 0 \\ 0 & 0 & 2{,}6645e^{-15} & 0 \\ 0 & 0 & 0 & 9{,}9929e^{-16} \end{pmatrix}$$

Por lo tanto, el modelo de factores ortogonales planteado es el siguiente

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} = \begin{pmatrix} -0{,}79F_1 & 0{,}44F_2 & 0{,}22F_3 & 0{,}33F_4 \\ -0{,}73F_1 & 0{,}56F_2 & -0{,}20F_3 & -0{,}31F_4 \\ -0{,}71F_1 & -0{,}54F_2 & -0{,}42F_3 & 0{,}14F_4 \\ -0{,}71F_1 & -0{,}54F_2 & 0{,}38F_3 & -0{,}19F_4 \end{pmatrix}$$

con

$$\hat{L} = \begin{pmatrix} -0{,}7997 & 0{,}4477 & 0{,}2222 & 0{,}3323 \\ -0{,}7313 & 0{,}5695 & -0{,}2088 & -0{,}3115 \\ -0{,}7104 & -0{,}5407 & -0{,}4260 & 0{,}1462 \\ -0{,}7171 & -0{,}5446 & 0{,}3872 & -0{,}1978 \end{pmatrix}$$

y

$$\Psi = \begin{pmatrix} 0{,}0000e^{00} & 0 & 0 & 0 \\ 0 & 1{,}4432e^{-15} & 0 & 0 \\ 0 & 0 & 2{,}6645e^{-15} & 0 \\ 0 & 0 & 0 & 9{,}9929e^{-16} \end{pmatrix}$$

En general, el número de factores se determina de manera que se explique aproximadamente el 70 % de la varianza total. En este caso podemos observar en el scree plot (**Figura 1**), que al considerar dos factores se explica aproximadamente 80 % de la varianza total.

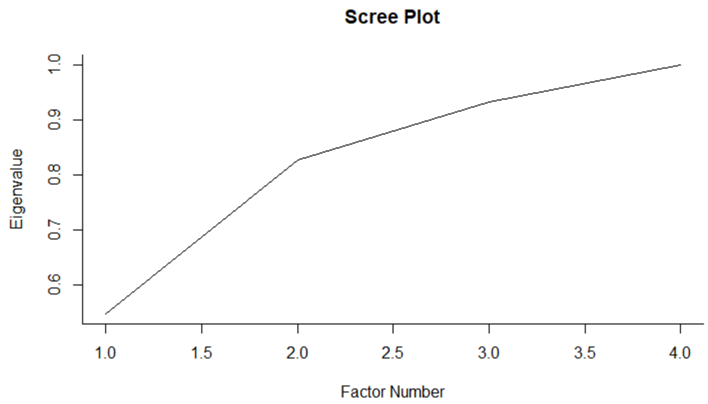

Figura 1. Varianza total explicada por cada uno de los m factores.

Solución

Sea $m = 2$, la matriz de cargas $\hat{L}$ se encuentra definida de la siguiente manera

$$\hat{L} = [\sqrt{\hat{\lambda}_1}\hat{e}_1 | \sqrt{\hat{\lambda}_2}\hat{e}_2] = \begin{bmatrix} -0{,}7997 & 0{,}4477 \\ -0{,}7313 & 0{,}5695 \\ -0{,}7104 & -0{,}5407 \\ -0{,}7171 & -0{,}5446 \end{bmatrix}$$

La contribución de los factores comunes, expresada por la suma de cuadrados de las cargas en esa variable se denomida *comunalidad.*

$$h_i^2 = l_{i1}^2 + l_{i2}^2 + ... + l_{im}^2$$

Haciendo uso del lenguaje de programación R se realizó el cálculo de las comunalidades.

$$\begin{pmatrix} 0{,}8401593 \\ 0{,}8592981 \\ 0{,}7970610 \\ 0{,}8108875 \end{pmatrix}$$

Sea $m = 3$, la matriz de cargas $\hat{L}$ se encuentra definida de la siguiente manera

$$\hat{L} = [\sqrt{\hat{\lambda}_1}\hat{e}_1|\sqrt{\hat{\lambda}_2}\hat{e}_2|\sqrt{\hat{\lambda}_3}\hat{e}_3] = \begin{bmatrix} -0{,}7997 & 0{,}4477 & 0{,}2222 \\ -0{,}7313 & 0{,}5695 & -0{,}2088 \\ -0{,}7104 & -0{,}5407 & -0{,}4260 \\ -0{,}7171 & -0{,}5446 & 0{,}3872 \end{bmatrix}$$

Haciendo uso del lenguaje de programación R se realizó el cálculo de las comunalidades.

$$\begin{pmatrix} 0{,}8895417 \\ 0{,}9029268 \\ 0{,}9786046 \\ 0{,}9608664 \end{pmatrix}$$

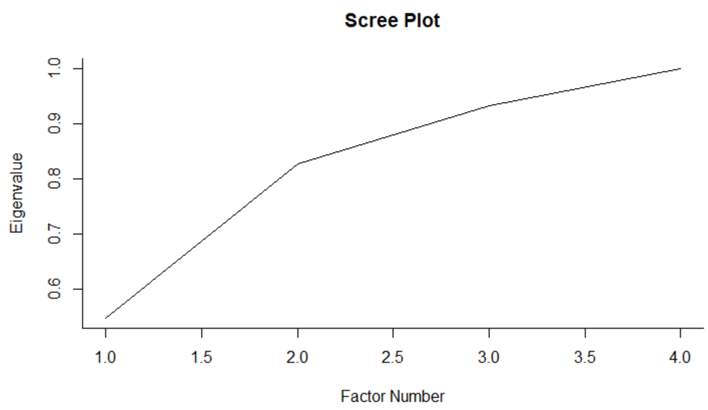

Figura 2. Varianza explicada por cada uno de los m factores. Primeros dos factores explican alrededor del 80 % de la varianza total.

Solución

Recordando los valores propios obtenidos con anterioridad

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (2{,}1934\ 1{,}1139\ 0{,}4245\ 0{,}2680)$$

y de acuerdo al **criterio de Kaiser**, el cual consiste en elegir el número de factores como el número de valores propios de **R** mayores a 1,0, podemos inclinarnos hacia el modelo cuando se consideran $m = 2$ factores.

Retomando lo observado en el scree plot (**Figura 1**), observamos que los primeros dos factores explican alrededor del 80 % de la varianza total, lo que respalda la idea de elegir el modelo con $m = 2$.

Solución

Modelo a considerar $m = 2$.

La correlación entre Z_2 y sus factores esta dada por la matriz de cargas $\hat{L}$

	F1	F2
Z_2	-0.7313	0.5695

Varianza específica para Z_2

$$1 - h_2^2 = 1 - 0{,}8592981 = 0{,}1407$$

Ejercicio 3. Considere la siguiente matriz de varianzas y covarianzas muestral:

$$S = \begin{pmatrix} 0{,}35 & 0{,}15 & -0{,}19 \\ 0{,}15 & 0{,}13 & -0{,}03 \\ -0{,}19 & -0{,}03 & 0{,}16 \end{pmatrix}$$

Asumiendo un modelo de 2 factores, estime los parámetros del modelo mediante el método del factor principal y compare los resultados con los obtenidos después de aplicar el mismo método sobre la matriz R (ejercicio visto en clase). Considere un ϵ de 0.05.

Solución

El análisis de factores por componentes principales de la matriz de covarianzas S es especificado en términos de sus pares de valores y vectores propios

$$(\hat{\lambda}_1, \hat{e}_1), (\hat{\lambda}_2, \hat{e}_2), ...(\hat{\lambda}_p, \hat{e}_p)$$

con

$$\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq ... \geq \hat{\lambda}_p$$

donde la matriz de cargas estimada $\tilde{L}$ se encuentra definida de la siguiente manera

$$\tilde{L} = [\sqrt{\hat{\lambda}_1}\hat{e}_1 | \sqrt{\hat{\lambda}_2}\hat{e}_2 | ... | \sqrt{\hat{\lambda}_m}\hat{e}_m]$$

Los vectores y valores propios de S representan estimaciones de los verdaderos valores y vectores propios.

Haciendo uso del lenguaje de programación R se obtuvo el cálculo de los vectores y valores propios.

Valores propios

$$(\lambda_1, \lambda_2, \lambda_3) = (0{,}5208\ 0{,}1126\ 0{,}0065)$$

Vectores propios

$$(e_1, e_2, e_3) = \begin{pmatrix} 0{,}8169 & 0{,}0426 & 0{,}5750 \\ 0{,}3488 & 0{,}7575 & -0{,}5517 \\ -0{,}4591 & 0{,}6513 & 0{,}6040 \end{pmatrix}$$

Por lo tanto, la matriz de cargas estimada $\tilde{L}$ se define como

$$\tilde{L} = \begin{bmatrix} 0{,}5896 & 0{,}0143 & 0{,}0464 \\ 0{,}2517 & 0{,}2542 & -0{,}0445 \\ -0{,}3313 & 0{,}2186 & 0{,}0487 \end{bmatrix}$$

Dado que las varianzas específicas estimadas se encuentran dadas por los elementos de la diagonal de la matriz $S - \tilde{L}\tilde{L}'$, entonces la matriz Ψ está definida de la siguiente manera

$$\tilde{\Psi} = \begin{bmatrix} \tilde{\Psi}_1 & 0 & \dots & 0 \\ 0 & \tilde{\Psi}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{\Psi}_p \end{bmatrix}$$

donde

$$\tilde{\Psi}_i = s_{ii} - \sum_{j=1}^{m} \tilde{l}_{ij}^2$$

Y las comunalidades son estimadas por

$$\tilde{h}_i^2 = \tilde{l}_{i2}^2 + \tilde{l}_{i1}^2 + \dots \tilde{l}_{im}^2$$

En este caso se obtuvieron las comunalidades y la matriz $\tilde{\Psi}$ con ayuda del lenguaje de programación R.

Matriz estimada $\tilde{\Psi}$.

$$\tilde{\Psi} = \begin{bmatrix} 1{,}6653e^{-16} & 0 & 0 \\ 0 & 1{,}9428e^{-16} & 0 \\ 0 & 0 & 0{,}0000e^{00} \end{bmatrix}$$

Comunalidades $\tilde{h}_i^2$.

$$\begin{pmatrix} 0{,}35 \\ 0{,}13 \\ 0{,}16 \end{pmatrix}$$

Ejercicio 4. Si suponemos que el modelo factorial ortogonal $X - \mu = LF + \epsilon$, con $Var(\epsilon) = \Psi$ es válido que F y ϵ siguen distribuciones normales, entonces X también sigue una distribución normal y es posible estimar la matriz de cargas por el método de máxima verosimilitud. Considere la matriz de cargas

$$L = \begin{pmatrix} 0{,}9 & 0{,}05 \\ 0{,}8 & 0{,}3 \\ 0{,}2 & 0{,}95 \\ 0{,}3 & 0{,}9 \\ 0{,}7 & 0{,}15 \end{pmatrix}$$

y la matriz de varianzas específicas $\Psi = diag(0{,}2, 0{,}3, 0{,}1, 0{,}2, 0{,}3)$. Tomando $\mu = 0$ genere una muestra de tamaño $n = 1000$ de X y obtenga la estimación de máxima verosimilitud de la matriz de cargas para $m = 2$ factores. Calcule la correspondiente estimación de la matriz de varianzas específicas.

Solución

Asumiendo que los factores comunes F y los factores específicos ϵ son normalmente distribuidos, se pueden obtener los estimadores de máxima verosimilitud de las cargas de los factores y de las varianzas específicas.

Cuando F_j y ϵ_j siguen una distribución normal multivariada las observaciones $X_j = \mu + LF_j + \epsilon_j$ siguen una distribución normal p-dimensional con vector de medias μ, y matriz de covarianzas $\Sigma = LL' + \Psi$.

Por lo tanto, la matriz S esta definida de la siguiente manera

$$S = \begin{pmatrix} 1{,}0125 & 0{,}735 & 0{,}2275 & 0{,}315 & 0{,}6375 \\ 0{,}735 & 1{,}030 & 0{,}4450 & 0{,}510 & 0{,}6050 \\ 0{,}2275 & 0{,}4450 & 1{,}0425 & 0{,}915 & 0{,}2825 \\ 0{,}315 & 0{,}510 & 0{,}915 & 1{,}100 & 0{,}3450 \\ 0{,}6375 & 0{,}605 & 0{,}2825 & 0{,}3450 & 0{,}8125 \end{pmatrix}$$

Haciendo uso del lenguaje de programación R se simuló una muestra de una distribución normal multivariada con $n = 1000$ a partir de la cual se realizó la estimación de máxima verosimilitud utilizando la función *factanal* para $m = 2$.

Sea

$$\Sigma = LL' + \Psi$$

S se encuentra definido como

$$S = \begin{pmatrix} 0{,}9999 & 0{,}7300 & 0{,}2513 & 0{,}3512 & 0{,}7096 \\ 0{,}7300 & 0{,}9998 & 0{,}4185 & 0{,}4774 & 0{,}6761 \\ 0{,}2513 & 0{,}4185 & 1{,}0001 & 0{,}8607 & 0{,}3164 \\ 0{,}3512 & 0{,}4774 & 0{,}8607 & 0{,}9995 & 0{,}3901 \\ 0{,}7096 & 0{,}6761 & 0{,}3164 & 0{,}3901 & 0{,}9999 \end{pmatrix}$$

con

$$\hat{L} = \begin{pmatrix} 0{,}880 & 0{,}114 \\ 0{,}791 & 0{,}298 \\ 0{,}158 & 0{,}985 \\ 0{,}292 & 0{,}827 \\ 0{,}781 & 0{,}196 \end{pmatrix}$$

y

$$\hat{\Psi} = \begin{pmatrix} 0{,}21225 & 0 & 0 & 0 & 0 \\ 0 & 0{,}2853 & 0 & 0 & 0 \\ 0 & 0 & 0{,}0050 & 0 & 0 \\ 0 & 0 & 0 & 0{,}2303 & 0 \\ 0 & 0 & 0 & 0 & 0{,}3515 \end{pmatrix}$$

Ejercicio 5. Una empresa está evaluando la calidad de su personal de ventas para lo cual seleccionó una muestra aleatoria de 50 vendedores y evaluó en cada uno de ellos 3 medidas de rendimiento: crecimiento de ventas, rentabilidad de ventas y ventas de nuevas cuentas. Estas medidas se han convertido a una escala, en la que 100 indica desempeño 'promedio'. Además, a los 50 individuos se les aplicaron 4 pruebas, que pretendían medir la creatividad, el razonamiento mecánico, el razonamiento abstracto y la capacidad matemática, respectivamente. Las $n = 50$ observaciones sobre las $p = 7$ variables se muestran en el archivo '*datosvendedores*'.

(a) Asumiendo un modelo ortogonal de factores para las variables estandarizadas, obtenga la solución por máxima verosimilitud de L y Ψ para $m = 2$ y $m = 3$ factores, considerando una rotación varimax, e interprete las soluciones con $m = 2$ y $m = 3$ factores.

(b) A partir de las estimaciones de los parámetros obtenga las comunalidades, las varianzas específicas y $\hat{L}\hat{L}' + \hat{\Psi}$ para las soluciones en $m = 2$ y $m = 3$ factores. Compare los resultados. ¿Qué elección de m prefiere en este punto? ¿Por qué?

(c) Realice una prueba de $H_0 : \Sigma = LL' + \Psi$ versus $H_1 : \Sigma \neq LL' + \Psi$ para $m = 2$ y $m = 3$. A partir de estos resultados y de la parte (b), ¿qué elección de m parece ser la adecuada?

(d) De acuerdo al número de factores elegido en (c), calcule las puntuaciones de los factores (factor scores) para los vendedores mediante: i) mínimos cuadrados ponderados y ii) mediante el enfoque de regresión. ¿Existe algún patrón de agrupamiento de los vendedores de acuerdo a sus puntuaciones factoriales?, si es así, ¿cómo se caracterizan los vendedores de cada grupo, de acuerdo a la interpretación de los factores?

Solución

Sea $m = 2$.

Haciendo uso del lenguaje de programación R se realizó la estimación de máxima verosimilitud utilizando la función *factanal* para $m = 2$.

	Factor1	*Factor2*
Crecimiento.de.ventas.x1.	0.852	0.452
Rentabilidad.de.ventas.x2.	0.868	0.419
Ventas.de.nueva.cuenta.x3.	0.717	0.602
Creatividad.x4.	0.148	0.987
Razonamiento.mecánico..x5.	0.501	0.525
Razonamiento.abstracto.x6.	0.619	0
Capacidad.matemática.x7.	0.946	0.277

Matriz $\hat{\Psi} = \text{diag}(0.0691, 0.0703, 0.1233, 0.0050, 0.4735, 0.6136, 0.0288)$

Sea $m = 3$.

Haciendo uso del lenguaje de programación R se realizó la estimación de máxima verosimilitud utilizando la función *factanal* para $m = 3$.

	Factor1	*Factor2*	*Factor3*
Crecimiento.de.ventas.x1.	0.793	0.374	0.438
Rentabilidad.de.ventas.x2.	0.911	0.317	0.185
Ventas.de.nueva.cuenta.x3.	0.651	0.544	0.438
Creatividad.x4.	0.255	0.964	0
Razonamiento.mecánico..x5.	0.542	0.465	0.207
Razonamiento.abstracto.x6.	0.299	0	0.950
Capacidad.matemática.x7.	0.917	0.180	0.298

Matriz $\hat{\Psi}$ = diag(0.0385, 0.0344, 0.0881, 0.0050, 0.4466, 0.0050, 0.0375)

Solución

Sea $m = 2$.

Haciendo uso del lenguaje de programación R se realizó el cálculo de las comunalidades y de la matriz $\hat{L}\hat{L}' + \hat{\Psi}$.

Comunalidades obtenidas

$$\begin{pmatrix} 0{,}9302 \\ 0{,}9289 \\ 0{,}8764 \\ 0{,}9960 \\ 0{,}5266 \\ 0{,}3831 \\ 0{,}9716 \end{pmatrix}$$

$$\hat{L}\hat{L}' + \hat{\Psi} = \begin{pmatrix} 0{,}9993996 & 0{,}9289240 & 0{,}8829880 & 0{,}572220 & 0{,}664152 & 0{,}5273880 & 0{,}931196 \\ 0{,}9289240 & 0{,}9993654 & 0{,}8745940 & 0{,}542017 & 0{,}654843 & 0{,}5372920 & 0{,}937191 \\ 0{,}8829880 & 0{,}8745940 & 0{,}9998018 & 0{,}700290 & 0{,}675267 & 0{,}4438230 & 0{,}845036 \\ 0{,}5722200 & 0{,}5420170 & 0{,}7002900 & 1{,}001073 & 0{,}592323 & 0{,}0916120 & 0{,}413407 \\ 0{,}6641520 & 0{,}6548430 & 0{,}6752670 & 0{,}592323 & 1{,}000211 & 0{,}3101190 & 0{,}619371 \\ 0{,}5273880 & 0{,}5372920 & 0{,}4438230 & 0{,}091612 & 0{,}310119 & 0{,}9967996 & 0{,}585574 \\ 0{,}9311960 & 0{,}9371910 & 0{,}8450360 & 0{,}413407 & 0{,}619371 & 0{,}5855740 & 1{,}00046 \end{pmatrix}$$

$$\hat{\Psi} = diag(0{,}0691, 0{,}0703, 0{,}1233, 0{,}0050, 0{,}4735, 0{,}6136, 0{,}0288)$$

Sea $m = 3$.

Haciendo uso del lenguaje de programación R se realizó el cálculo de las comunalidades y de la matriz $\hat{L}\hat{L}' + \hat{\Psi}$.

Comunalidades obtenidas

$$\begin{pmatrix} 0{,}9605 \\ 0{,}9646 \\ 0{,}9115 \\ 0{,}9943 \\ 0{,}5528 \\ 0{,}9919 \\ 0{,}9620 \end{pmatrix}$$

$$\hat{L}\hat{L}' + \hat{\Psi} = \begin{pmatrix} 0{,}9991407 & 0{,}9220110 & 0{,}9115430 & 0{,}562751 & 0{,}6943820 & 0{,}653207 & 0{,}9250250 \\ 0{,}9220110 & 0{,}9991157 & 0{,}8465390 & 0{,}537893 & 0{,}6794620 & 0{,}448139 & 0{,}9475770 \\ 0{,}9115430 & 0{,}8465390 & 0{,}9997028 & 0{,}690421 & 0{,}6964680 & 0{,}610749 & 0{,}8254110 \\ 0{,}5627510 & 0{,}5378930 & 0{,}6904210 & 0{,}999321 & 0{,}5864700 & 0{,}076245 & 0{,}4073550 \\ 0{,}6943820 & 0{,}6794620 & 0{,}6964680 & 0{,}586470 & 0{,}9994585 & 0{,}358708 & 0{,}6424000 \\ 0{,}6532070 & 0{,}4481390 & 0{,}6107490 & 0{,}076245 & 0{,}3587080 & 0{,}996901 & 0{,}5572830 \\ 0{,}9250250 & 0{,}9475770 & 0{,}8254110 & 0{,}407355 & 0{,}6424000 & 0{,}557283 & 0{,}9996028 \end{pmatrix}$$

con

$$\hat{\Psi} = diag(0{,}0385, 0{,}0344, 0{,}0881, 0{,}0050, 0{,}4466, 0{,}0050, 0{,}0375)$$

Solución

En ambos casos ($m = (2, 3)$), se utilizó la función *factanal* para obtener el valor del estadístico y el p valor.

Para $m = 2$ se obtuvo los siguientes valores.

Valor del estadístico de Bartlett

$$117{,}2006$$

p valor

$$1{,}253644e^{-21}$$

Dado que el valor p obtenido es menor que el valor crítico se rechaza la hipótesis nula.

Para $m = 3$ se obtuvo los siguientes valores.

Valor del estadístico de Bartlett

$$62{,}18039$$

p valor

$$2{,}010435e^{-13}$$

Al igual que en el caso anterior, dado que el valor p obtenido es menor que el valor crítico se rechaza la hipótesis nula.

Considerando lo anterior, el modelo más adecuado sería cuando $m = 3$, esto ya que existe una menor diferencia entre el p valor y el valor del estadístico.

Solución

Se utilizó la función *factanal* para obtener las puntuaciones de los factores para el modelo con $m = 3$.

Puntuaciones de los factores mediante mínimos cuadrados.

Factor1	Factor2	Factor3
0.7872	-0.3690	0.4917

Puntuaciones de los factores mediante regresión.

Factor1	Factor2	Factor3
0.7872	-0.3690	0.4917

Ambos enfoques nos arrojan los mismos valores para los *factor scores*.