Capítulo 4

Análisis Factorial

4.1. Aspectos prácticos

1. Mostrar que la matriz de covarianzas:

$$\rho = \begin{bmatrix} 1.00 & 0.63 & 0.45 \\ 0.63 & 1.00 & 0.35 \\ 0.45 & 0.35 & 1.00 \end{bmatrix}$$

   de $p = 3$ variables aleatorias estandarizadas Z_1, Z_2 y Z_3 pueden ser generadas por el modelo factorial:

$$\begin{aligned} Z_1 &= 0.9F_1 + \epsilon_1 \\ Z_2 &= 0.7F_1 + \epsilon_2 \\ Z_3 &= 0.5F_1 + \epsilon_3 \end{aligned}$$

   donde $Var(F_1) = 1$, $Cov(\epsilon, F_1) = 0$ y $\Psi = Cov(\epsilon) = diag(0.19, 0.51, 0.75)$

2. Utilizar la información del ejercicio anterior, para:

   a) Calcular las comunalidades $h_i^2, i = 1, 2, 3$ e interpretar estas cantidades.

   b) Calcular $Corr(Z_i, F_1)$ para $i = 1, 2, 3$. ¿Cuál de las variables tiene la mayor ponderación en el factor? ¿Cuál es la razón?.

3. Los eigenvalores y los eigenvectores de la matriz del ejercicio 1 son:

$$\begin{aligned} \lambda_1 &= 1.96 & a_1' &= (0.625, 0.593, 0.507) \\ \lambda_2 &= 0.68 & a_2' &= (-0.219, -0.491, 0.843) \\ \lambda_3 &= 0.36 & a_3' &= (0.749, -0.638, -0.177) \end{aligned}$$

   a) Asumir un modelo factorial con $m = 1$, calcular la matriz de cargas factoriales L y la matriz de varianzas específicas Ψ utilizando la solución por el método de componentes principales. Además determinar la proporción de varianza explicada por el factor?
   b) Asumir un modelo factorial con $m = 2$, calcular la matriz de cargas factoriales L y la matriz de varianzas específicas Ψ utilizando la solución por el método de componentes principales. Además determinar la proporción de varianza explicada por los dos factores?
   c) Comparar las matrices residuales de los incisos a) y b). ¿Cuál de los modelos es más conveniente?.

4. En relación a los datos del ejercicio 3:

   a) Asumir un modelo factorial con $m = 1$, calcular la matriz de cargas factoriales L y la matriz de varianzas específicas Ψ utilizando la solución por el método del Factor Principal. Además determinar la proporción de varianza explicada por el factor?
   b) Asumir un modelo factorial con $m = 2$, calcular la matriz de cargas factoriales L y la matriz de varianzas específicas Ψ utilizando la solución por el método del Factor Principal. Además determinar la proporción de varianza explicada por los dos factores?
   c) Comparar las matrices residuales de los incisos a) y b). ¿Cuál de los modelos es más conveniente?.

5. *(La parametrización del modelo factorial no necesariamente es única)*. Sea el modelo factorial con $p = 2$ y $m = 1$. Mostrar que:

$$\sigma_{11} = l_{11}^2 + \psi_1, \qquad \sigma_{12} = \sigma_{21} = l_{11}l_{21}$$
$$\sigma_{22} = l_{21}^2 + \psi_2$$

   y para σ_{11}, σ_{22} y σ_{12} dados, existe una infinidad de opciones para L y Ψ.

6. *(Única pero impropia solución: caso Heywood)* Considera el modelo factorial con $m = 1$ para una población con matriz de covarianzas:

$$\Sigma = \begin{bmatrix} 1.0 & 0.4 & 0.9 \\ 0.4 & 1.0 & 0.7 \\ 0.9 & 0.7 & 1.0 \end{bmatrix}$$

Mostrar que existe una única opción para L y Ψ, pero que el valor para $\psi_3 < 0$ no es admisible.

4.2. Aspectos teóricos

7. Sea x un vector aleatorio con media μ_x y matriz de covarianzas Σ. Si x es transformado a $y = Ax + b$ donde A es no-singular. Mostrar que: $z - \mu_z = L_z F + \epsilon_z$.

8. Indeterminación con factores correlado e incorrelados.

   a) Sea F de $m \times 1$ el vector de factores con $E[F] = 0$ y $Cov[F] = I$. Si H es cualquier matriz no singular, entonces los factores del modelo factorial que incluye a H están correlados.

   b) Sea F de $m \times 1$ el vector de factores con $E[F] = 0$ y $Cov[F] = AA'$. Si A es una matriz no singular, entonces los factores del modelo factorial que incluyen a la matriz A están incorrelados.

9. Demostrar utilizando la relación fundamental que:

   a) Si la varianza específica es igual a la varianza de una variable, la fila de la matriz de carga correspondiente a dicha variable debe tener todos los elementos nulos.

   b) Si la covarianza entre dos variables es cero, las filas correspondientes a estas variables en la matriz de carga son ortogonales.

   c) Si las variables están estandarizadas la correlación entre las variables i y j es el producto escalar de las filas de la matriz de carga correspondientes a estas variables.

   d) Si las variables están estandarizadas, la correlación entre la variable i y el factor j es el término (ij) de la matriz de carga.

10. Demostrar utilizando la relación fundamental que si las varianzas de los componenetes específicos son idénticos, las columnas de la matriz de carga son vectores propios de la matriz de covarianzas, y obtener los valores propios.

11. Indicar cuál será el número máximo de factores incorrelados que podemos estimar con diez variables. ¿Y si los factores están correlados?(Ayuda: Revisar la sección 12.2.5. del Libro de Peña).

12. Demostrar que si cada columna de la matriz de carga tiene un único elemento no nulo, el modelo factorial no está identificado.

13. Demostrar que si todas las variables tienen perturbaciones con la misma varianza, $\Psi = \psi_0 \, I_p$, y suponiendo $LL' = diagonal = D$, las columnas de L son directamente vectores propios de la matriz Σ, con valores propios $d_i + \psi_0$, donde d_i es el término i-ésimo de la diagonal de D.

14. Sean los fatores F_a y F_b, además de las matrices de rotación:

$$T_1 = \begin{bmatrix} cos(\theta) & -sen(\theta) \\ sen(\theta) & cos(\theta) \end{bmatrix} \qquad y \qquad T_2 = \begin{bmatrix} cos(\theta) & sen(\theta) \\ sen(-\theta) & cos(\theta) \end{bmatrix}$$

   a) Si se desean rotar los factores F_a y F_b, determine el ángulo de rotación utilizando el método cuartimax y la matriz de rotación T_1.
   b) Si se desean rotar los factores F_a y F_b, determine el ángulo de rotación utilizando el método cuartimax y la matriz de rotación T_2.
   c) Considere F_a, F_b y F_c, que fueron rotados en pares utilizando el método cuartimax, tal que el ángulo de rotación entre F_a y F_b es θ_{ab} y el ángulo de rotación entre F_b y F_c es θ_{bc}. Determine la matriz T de Rotación.

4.3. Aplicación Práctica

15. Lea con detenimiento ***las Perspectivas y una Estrategia para el Análisis Factorial*** desarrollado en las páginas $557 - 558$ del libro de Jhonson y Wichern. Posteriormente plasme esa teoría de tal manera que sea comprensible y fácil de aplicar a cualquier conjunto de datos.