



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

TAREA 8

EJERCICIOS

1. Considera de nuevo los datos en el fichero `sons.dat`, en el lugar habitual, que se reproducen a continuación. Fueron descritos y su normalidad examinada en una tarea anterior (Tarea 3).

Una cuestión de interés sería la siguiente: ¿se parecen en cuanto a sus medidas las cabezas del primogénito y segundo hijos de una pareja?. Observa que ello *no* implica necesariamente correlación entre las longitudes de cabeza (X_1 e Y_1), ni entre las anchuras (Y_1 e Y_2). Podría ocurrir que se parecieran en proporción, o que a primogénitos de cabeza alargada siguieran hermanos de cabeza ancha, o . . .

Un método bastante general de investigar la cuestión consistiría en examinar si hay alguna función lineal de las medidas del primer hijo que esté relacionada con alguna función lineal de las medidas del segundo hijo.

Haz este análisis y comenta sus resultados.

2. Toma de nuevo los datos en `Bizkaia.dge` y has un análisis de correlación canónica entre (alternativamente):
 - a) Dos elecciones diferentes.
 - b) Una elección y una batería de indicadores de tu elección tomados entre las restantes variables.

Sobre el análisis que hayas efectuado contrasta la hipótesis de que la dimensión del espacio generado por las variables canónicas con correlación canónica no nula es $\min(p, q)$, $\min(p, q) - 1, \dots, 1, 0$.

AYUDAS, SUGERENCIAS Y COMPLEMENTOS

1. Dependiendo de las variables que selecciones en el Ejercicio 2 tendrás más o menos facilidad para interpretar los resultados, tarea que con datos de elevada dimensionalidad nunca suele ser fácil.
2. Sea $s = \min p, q$ en la notación empleada en clase. Bajo la hipótesis H_0 de que las últimas $s - k$ correlaciones canónicas son nulas, el estadístico

$$\sum_{i=k+1}^s \log_e(1 - \lambda_i) \quad (1)$$

tiene, al multiplicarlo por una constante adecuada, distribución lambda de Wilks o (asintóticamente) χ^2 . Por ejemplo, asintóticamente,

$$- \left(N - \frac{1}{2}(p + q + 3) \right) \sum_{i=k+1}^s \log_e(1 - \lambda_i) \sim \chi_{(p-k)(q-k)}^2, \quad (2)$$

Cuadro 1: Medidas de cabeza del primer y segundo hijos de una muestra de 25 familias.

Primer hijo		Segundo hijo		Primer hijo		Segundo hijo	
\bar{X}_1	\bar{X}_2	\bar{Y}_1	\bar{Y}_2	\bar{X}_1	\bar{X}_2	\bar{Y}_1	\bar{Y}_2
191	155	179	145	174	150	185	152
195	149	201	152	190	159	195	157
181	148	185	149	188	151	187	158
183	153	188	149	163	137	161	130
176	144	171	142	195	155	183	158
208	157	192	152	186	153	173	148
189	150	190	149	181	145	182	146
197	159	189	152	175	140	165	137
188	152	197	159	192	154	185	152
192	150	187	151	174	143	178	147
179	158	186	148	176	139	176	143
183	147	174	147	197	167	200	158
–	–	–	–	190	163	187	150

Fuente: Tomados de [6] (Tabla 3.9, p. 91), que a su vez los toma de [3].

donde N es el número de observaciones y los λ_i las correlaciones canónicas muestrales al cuadrado (ver por ej. [5], p. 448).

Tal estadístico permite contrastar el número de parejas de variables canónicas con correlación canónica no nula (algo así como el número de “facetas” o aspectos incorrelados entre sí que relacionan los dos grupos de variables).

3. **Lectura recomendada.** Al margen de tus apuntes de clase puedes mirar las secciones relevantes de cualquiera de los manuales ya citados en tareas anteriores, como [4], [7], [5], [2]. Además de ellos, tienes una pequeña monografía sobre correlación canónica, [8], que puedes encontrar de interés. En [1] puedes ver una de las pocas aplicaciones directas del método de correlación canónica que he encontrado.

Referencias

- [1] D.A. Aaker. *Multivariate Analysis in Marketing: Theory and Application*. Wadsworth, 1971.
- [2] C.M. Cuadras. *Métodos de Análisis Multivariante*. Eunibar, Barcelona, 1981.
- [3] G.P. Frets. Heredity of head form in man. *Genetica*, 3:193–384, 1921.
- [4] R.A. Johnson and D.W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice Hall, 1992.
- [5] W.J. Krzanowski. *Principles of Multivariate Analysis: A User's Perspective*. Oxford, 1988. Signatura: 519.23 KRZ.
- [6] A.C. Rencher. *Methods of Multivariate Analysis*. Wiley, 1995.
- [7] G.A.F. Seber. *Multivariate Observations*. Wiley, New York, 1984.
- [8] B. Thompson. *Canonical Correlation Analysis*. SAGE, 1984.