

# Análisis de datos multivariantes. Aplicado al rendimiento académico universitario

**María Claudia Morelli, Laura Cánovas y Nora Martinengo**

Dpto. de Biomatemática y Físico-Química  
Facultad de Ciencias Agrarias - UNCuyo

Se repite por doquier que la Estadística es una ciencia útil para el tratamiento de un conjunto de datos de los cuales se pretende obtener alguna conclusión. Tal utilidad no es suficientemente aprovechada ya que no se ha hecho de la Estadística una disciplina científica asequible y de fácil dominio.

Desde la Cátedra de Cálculo Estadístico y Biometría se decidió publicar el libro "Análisis Multivariante aplicado al Rendimiento Académico Universitario", con el propósito de presentar las técnicas multivariadas empleadas en el marco de un trabajo de investigación.



Figura 1. Portada de libro

## ¿Qué es el Análisis Multivariante?

En un sentido amplio, se refiere a todos los métodos estadísticos que analizan el comportamiento e interrelación de múltiples variables. El objetivo del análisis Multivariante es resumir grandes cantidades de datos por medio de pocos parámetros.

Las Técnicas Multivariantes son métodos estadísticos para analizar simultáneamente conjuntos de datos. De hecho, las técnicas que se pueden aplicar varían de acuerdo a la información recogida por las variables, y se clasifican en: **(i) Métodos dependientes:** suponen que las variables analizadas están divididas en dos grupos: las variables dependientes y las variables independientes. El objetivo de los métodos de dependencia consiste en determinar si el conjunto de variables independientes afecta al conjunto de

variables dependientes y de qué forma.

**(ii) Métodos independientes:** No distinguen entre variables dependientes e independientes y su objetivo consiste en identificar qué variables están relacionadas, cómo lo están y por qué.

**(iii) Métodos estructurales:** Suponen que las variables están divididas en dos grupos: el de las variables dependientes y el de las independientes. El objetivo de estos métodos es analizar, no sólo cómo las variables independientes afectan a las variables dependientes, sino también cómo están relacionadas las variables de los dos grupos entre sí.

Para cumplir con nuestros objetivos se decidió utilizar **Componentes principales** y **Análisis de Correspondencia** dentro de los métodos independientes y **Análisis de la Varianza** dentro de los **dependientes**.

Según el autor César Pérez López, se utiliza componentes principales para reducir el elevado número de variables cuantitativas y obtener una cantidad menor de dimensiones (combinación lineal de las variables originales). La técnica análisis de correspondencia fue ideada, en 1973, por el estadístico francés Benzecri para reducir la dimensión de una base de datos formada por variables cualitativas. El **Análisis de la Varianza** se utilizó para calcular las medias y analizar la variabilidad de la variable dependiente (respuesta) para los grupos que forman las variables independientes: si se aprecian diferencias estadísticas significativas se puede concluir que las variables independientes afectan a la variable respuesta.

El libro da cuenta del trabajo realizado en el marco del proyecto de investigación 2013-2015, "*Factores contextuales y personales asociados al rendimiento académico de alumnos de la carrera de Ingeniería Agronómica Universidad Nacional de Cuyo*" donde se analizan ocho variables cuantitativas y gran cantidad de variables institucionales y sociales.

Después de una experiencia rica en contenido estadístico y aplicación se pudo arribar a los siguientes resultados:

1. Para el caso de métodos independientes, **Componentes Principales** se destaca como una de las técnicas más eficientes a la hora de reducir la dimensión ya que sirve para obtener pocas variables ficticias, combinación lineal de muchas variables cuantitativas observadas, con la mínima pérdida de información.
2. Comúnmente las variables cualitativas se presentan en tablas de contingencia que, a lo sumo, son de tres variables (para facilitar la confección y comprensión de tablas). Si se trabaja con múltiples (más de tres) variables, el **Análisis de Correspondencia** es muy útil para reducir la dimensión original de las variables para, así, confeccionar la tabla y facilitar el análisis. Se redujo considerablemente la dimensión tanto de las variables sociales como de las institucionales y se mostraron resultados utilizando herramientas descriptivas (tablas y gráficos).
3. En nuestro caso, para las variables institucionales y sociales se obtuvieron dos componentes y entre las dos sólo explicaron el 46% y 37% del rendimiento, respectivamente. Esto se debe a la baja correlación entre las variables lo que provoca un déficit en la variabilidad de los datos.
4. El **análisis de la varianza** es una técnica que compara más de dos medias de la variable dependiente cuantitativa habiendo fijado las variables independientes que son cualitativas. Esto requiere que la variable dependiente se distribuya normalmente y que haya más de dos repeticiones en los grupos producto de las combinaciones de los niveles de un factor con todos los niveles de los demás factores. En este estudio, debido al tamaño pequeño de las muestras (no se llegaron a procesar todas las encuestas), no se cumplió ese supuesto de normalidad ni el principio de repetición.
5. No se logra determinar, mediante la técnica ADEVA, la influencia de alguna o varias variables institucionales y/o sociales en el rendimiento académico.

**Descriptor:** Estadística como ciencia – Métodos estadísticos – Análisis multivariante

**Palabras claves:** Rendimiento académico

### **Bibliografía**

- Cánovas, L.; Montenegro, A.; Martinengo, N.; Sánchez, M.L.; Maure, E.; Díaz Micari, M. (2013). Aspectos cognitivo-motivacionales que impactan en el desempeño académico y su relación con régimen de enseñanza-aprendizaje, de alumnos de la Facultad de Ciencias Agrarias, UNCuyo. Informe final proyecto subsidiado por SeCyTP. Trabajo en elaboración para su publicación.
- Hair, J; Anderson, R; Tatham, R; Black, W. (1999). Análisis Multivariante. Quinta edición. Prentice Hall; 799p. España.
- Morelli, M.C, Cánovas, L.L, Martinengo, N.B. (2017). Análisis Multivariante aplicado al rendimiento académico universitario. Editorial Académica Española. Alemania. 80pp. Isbn 978-620-2-25630-8.
- Peña, D. (2002). Análisis de datos multivariantes. Editorial Mc Graw Hill, 539p. Madrid.
- Pérez López, C. (2005). Métodos estadísticos avanzados con SPSS. Editorial Thomson, 775p. España.
- SPSS para Windows. Versión 15.01.. (2006). Vista Hotfix Applied. SPSS Inc, 456p.