

ESTADÍSTICA MULTIVARIANTE

UGR, GRADO EN MATEMÁTICAS
Curso Académico 2023-2024

José Miguel Angulo Ibáñez (*jmangulo@ugr.es*)

Departamento de Estadística e Investigación Operativa
Universidad de Granada

- ▶ **TEMA 1. Distribución Normal Multivariante**
Motivación sobre la DNM

- La DNM es una de las distribuciones más conocidas en la Estadística, jugando un papel predominante en muchas áreas de aplicación.
- En el Análisis Multivariante, por ejemplo, la mayor parte de los procedimientos de inferencia para el análisis de datos vector-
valuados se han desarrollado bajo la suposición de normalidad.

- La DNM constituye la extensión natural, al contexto multivariante, de la distribución normal univariante. En particular, las distribuciones marginales univariantes de la DNM son normales.

Más aún, las distribuciones marginales de cualquier dimensión de la DNM son también DNM.

- Muchos procedimientos basados en la suposición de normalidad (multivariante, en general), particularmente procedimientos basados en los momentos de primer y segundo orden, cumplen ciertas propiedades de optimalidad.

En determinados casos, los procedimientos presentan cierta robustez frente a desviaciones de la normalidad en la distribución teórica subyacente o en los propios datos. Una justificación viene dada, por ejemplo, por el TCL acerca de la DNM asintótica del vector de medias muestrales, de modo que, para grandes muestras, la DNM puede usarse como aproximación.

En otros casos, la DNM sirve como modelo de referencia para evaluar posibles desviaciones de la normalidad.

- Técnicamente, la DNM ha sido estudiada en profundidad y ofrece ciertas ventajas en su tratamiento analítico, también desde el punto de vista estadístico-probabilístico:
 - La DNM viene determinada de forma única por los momentos de primer y segundo orden (μ , Σ).
 - Si dos subvectores de un vector aleatorio con DNM tienen correlaciones cruzadas nulas, entonces dichos subvectores son mutuamente independientes.
 - La familia de DNMs es cerrada bajo transformaciones lineales (...).
 - La familia de DNMs es cerrada bajo combinaciones lineales de vectores independientes (...).
 - Las distribuciones condicionadas (internamente) en una DNM son también DNMs. Sus momentos de primer y segundo orden se obtienen de forma sencilla a partir de los correspondientes a la DNM de referencia.
 - ...

- La DN bivalente comenzó a estudiarse a mediados del s.XIX. Tuvo un gran impacto a partir de una publicación de Francis Galton (1888), sobre análisis de correlación en genética.
- Karl Pearson (1896), a partir del trabajo de Galton, estableció su formulación matemática de forma rigurosa.
- Francis Edgeworth (1892) introdujo, en el estudio de la regresión múltiple y el análisis de correlación parcial y múltiple, elementos fundamentales para el desarrollo de la DNM.
- Durante el s.XX se sucedieron importantes avances en relación con la fundamentación teórica de la DNM, la inferencia estadística con referencia a la DNM (incluyendo el desarrollo y estudio de otras distribuciones relacionadas, e. g. T^2 de Hotelling, Wishart), así como su contextualización en familias de distribuciones más generales (e. g. distribuciones simétricas de contornos elípticos).