

# Analisis de Discriminantes Lineales

Carlos Alessandro Becerril Estrada

## ##INTRODUCCION

El análisis discriminante lineal es una técnica estadística utilizada en el campo del aprendizaje automático y la estadística multivariante para clasificar o predecir la pertenencia de una observación a diferentes grupos o categorías. Su objetivo principal es encontrar una combinación lineal de variables predictoras que maximice la separación entre las clases o grupos conocidos.

En el ADL, se supone que las variables predictoras se distribuyen normalmente dentro de cada grupo y que las covarianzas son iguales entre los grupos. A partir de estos supuestos, se busca encontrar una función lineal que maximice la separación entre los grupos y minimice la variación dentro de los grupos.

El análisis discriminante lineal utiliza una combinación lineal de variables predictoras para proyectar los datos en un espacio de menor dimensión, de manera que se maximice la separación entre los grupos. Esta combinación lineal, también conocida como función discriminante, se construye de tal manera que minimiza la clasificación errónea de las observaciones en los grupos conocidos [Santiago de la Fuente Fernández, 2011]

El ADL se utiliza en una amplia gama de aplicaciones, como la clasificación de enfermedades en medicina, la segmentación de clientes en marketing, el reconocimiento de patrones en la visión por computadora y muchas otras áreas donde se requiere la clasificación precisa de observaciones en grupos predefinidos [Villardón, 2006]

En resumen, el análisis discriminante lineal es una técnica estadística que busca encontrar una función lineal que maximice la separación entre grupos conocidos. Es una herramienta poderosa para la clasificación y predicción en diversas áreas, aprovechando los supuestos de distribución normal y covarianzas iguales entre los grupos.

## ##OBJETIVOS

**Maximizar la separación entre grupos:** El ADL busca encontrar una función discriminante que maximice la separación entre los grupos conocidos, permitiendo una mejor clasificación de las observaciones en las categorías correspondientes.

**Reducir la dimensionalidad:** El ADL proyecta los datos en un espacio de menor dimensión, lo que ayuda a simplificar y representar los datos de manera más eficiente, manteniendo la mayor parte de la información discriminante.

Clasificación y predicción precisas: El ADL tiene como objetivo clasificar correctamente las nuevas observaciones en los grupos conocidos, utilizando la función discriminante encontrada. Se busca minimizar la clasificación errónea y mejorar la precisión en la asignación de observaciones a las categorías correspondientes.

## ##MODELO ESTADISTICO

El ADL presenta como modelo la siguiente  $\delta_x(x) = x_i (\mu_k / (\sigma_x^2)) - (\mu_k^2 / (\sigma_x^2)) + \log(\pi_k)$   $\delta_x(x)$ : Representa la función discriminante lineal para clasificar una observación  $x$  en una categoría específica.  $x_i$ : Es la observación de la variable predictora  $i$ -ésima de la observación  $x$ .  $\mu_k$ : Es el vector de media para la categoría o grupo  $k$ .  $\sigma_x^2$ : Es la matriz de covarianza estimada común para todas las categorías.  $\mu_k^2$ : Es el vector de media para la categoría o grupo  $k$  elevado al cuadrado (elemento a elemento).  $\pi_k$ : Es la probabilidad a priori de la categoría o grupo  $k$ .

## ##EJEMPLOS DE APLICACION

Estos ejemplos ilustran cómo la función ADL puede aplicarse en el ámbito zootecnista, ya sea para clasificar razas de ganado o detectar enfermedades en animales. Es importante tener en cuenta que la función proporcionada es una implementación básica y puede requerir ajustes y personalizaciones adicionales para adaptarse a las características específicas de los datos y los objetivos del análisis en cada caso.

### ###EJEMPLO 1

Clasificación de razas de ganado: Supongamos que tienes datos de diferentes características físicas y de rendimiento de animales de diferentes razas de ganado. El objetivo es clasificar correctamente las observaciones en las razas correspondientes. Puedes utilizar la función ADL para calcular los parámetros estadísticos de las variables predictoras en cada raza y, luego, clasificar nuevas observaciones en función de estas características utilizando la etiqueta generada.

### ###EJEMPLO 2

Detección de enfermedades en animales: En el campo de la salud animal, la función ADL puede ser útil para clasificar animales como “enfermos” o “sanos” en función de variables relacionadas con su estado de salud, como temperatura, recuento de células sanguíneas o niveles de ciertos biomarcadores. Al calcular los parámetros estadísticos de estas variables en los grupos de animales enfermos y sanos, puedes utilizar la función ADL para clasificar nuevas observaciones y detectar posibles enfermedades en animales.

## ##CONCLUSIONES

Estos ejemplos ilustran cómo la función ADL puede aplicarse en el ámbito zootecnista, ya sea para clasificar razas de ganado o detectar enfermedades en animales. Es importante tener en cuenta que la función proporcionada es una implementación básica y puede requerir ajustes y personalizaciones adicionales para adaptarse a las

características específicas de los datos y los objetivos del análisis en cada caso. Es importante destacar que la implementación de la función ADL presentada es básica y puede requerir adaptaciones adicionales según las características específicas de los datos y los objetivos del análisis. Además, el éxito del ADL depende de los supuestos subyacentes, como la normalidad de las variables predictoras dentro de cada grupo y la igualdad de covarianzas entre grupos. Es importante destacar que la implementación de la función ADL presentada es básica y puede requerir adaptaciones adicionales según las características específicas de los datos y los objetivos del análisis. Además, el éxito del ADL depende de los supuestos subyacentes, como la normalidad de las variables predictoras dentro de cada grupo y la igualdad de covarianzas entre grupos.

##LITERATURA CITADA Santiago de la Fuente Fernández. (2011). Análisis Discriminante. Villardón, J. L. V. (2006). Análisis discriminante: introducción. Universidad de Salamanca: Departamento de estadística.