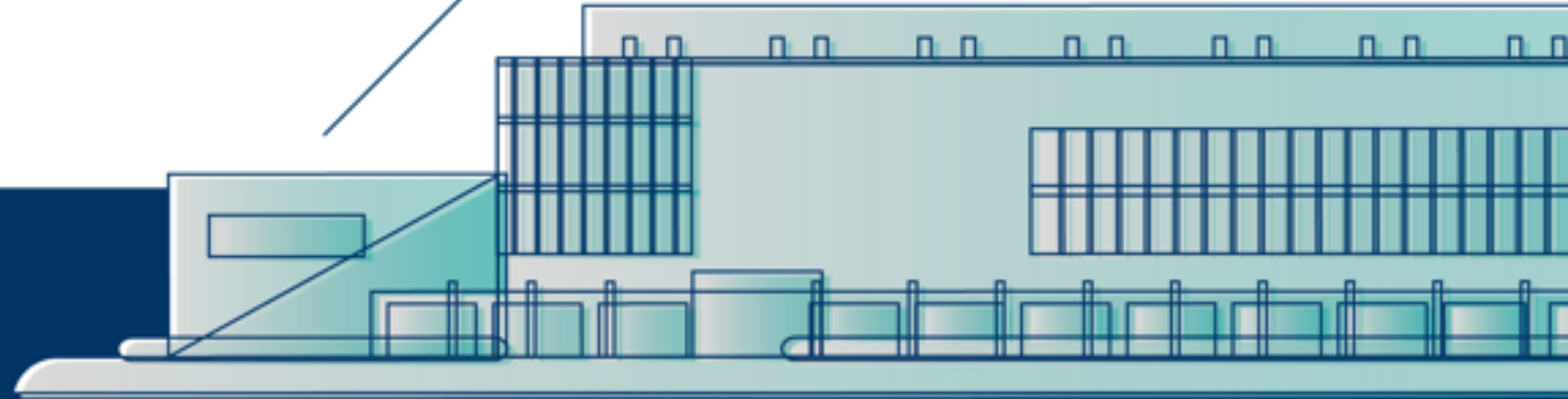




UNIVERSIDAD  
DE MÁLAGA

| [uma.es](http://uma.es)



## Tema 4: La clasificación de consumidores a través del análisis discriminante

# ¿Por qué necesitamos el análisis discriminante?

---

En marketing, a veces no sólo queremos explicar, sino que también necesitamos clasificar.

---

Por ejemplo: ¿Clientes con riesgo o sin riesgo?, ¿Compradores o no compradores? ¿Fidelizados o no fidelizados?

---

Para poder hacer esto, normalmente se necesita un método que use variables métricas para predecir la pertenencia a grupos conocidos.

# El problema de partida

---

En muchas situaciones de investigación comercial, los grupos ya están definidos previamente (por ejemplo: clientes con y sin impago, compradores y no compradores, clientes fieles y no fieles).

---

Para cada individuo se dispone de variables métricas (ingreso, gasto, edad, uso de servicios, etc.).

---

El problema consiste en determinar: qué variables diferencian realmente a los grupos y, cómo clasificar correctamente nuevos individuos en uno u otro grupo a partir de esa información.

# ¿Qué hace exactamente el Análisis Discriminante?



Construye combinaciones lineales de variables métricas.



Maximiza la separación entre grupos conocidos.



Permite entender qué variables discriminan y clasificar nuevos individuos.

# ¿Cuándo tiene sentido usar el Análisis Discriminante?

Cluster: grupos desconocidos.

ANOVA: comparar medias entre grupos conocidos.

Discriminante: clasificar y predecir pertenencia a grupos conocidos.

# Tipos de análisis discriminante

AD simple: variable dependiente binaria.

AD múltiple: variable dependiente con más de dos grupos.

# Las 3 fases del análisis discriminante

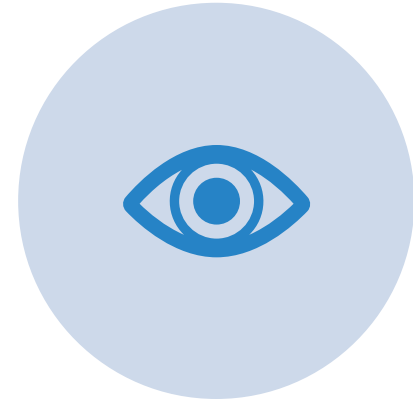
---



ESPECIFICACIÓN Y  
VALIDACIÓN DEL MÉTODO



SIGNIFICACIÓN DE LAS  
FUNCIONES



CLASIFICACIÓN E  
INTERPRETACIÓN

# FASE 1: REQUISITOS PREVIOS



Muestra representativa.



Inclusión de variables con argumentación teórica.



Tipo de medida: la variable dependiente debe ser categórica y las variables independientes métricas.

# FASE 1: SUPUESTOS DE PARTIDA

01

Normalidad multivariante (variables independientes).

02

Igualdad de matrices de covarianzas intra-grupos (M de Box).

03

Ausencia de multicolinealidad (Matriz de estructura de correlaciones vs. Matriz de coeficientes estandarizados de la función discriminante).

# Fase 2: ¿La función discrimina?

Lambda de Wilks:

Cerca de 0: buena separación.  
Cerca de 1: mala separación.



Contraste  $\chi^2$  asociado.

Hipótesis nula: igualdad de las puntuaciones medias entre grupos de la función discriminante.

# Fase 3: Interpretación y clasificación

---

**Probabilidades a priori:** decidir sobre la pertinencia de mantener las mismas probabilidades de clasificación para los grupos.

---

**ANOVAS univariados:** determinar la significación individual de cada variable independiente por separado.

---

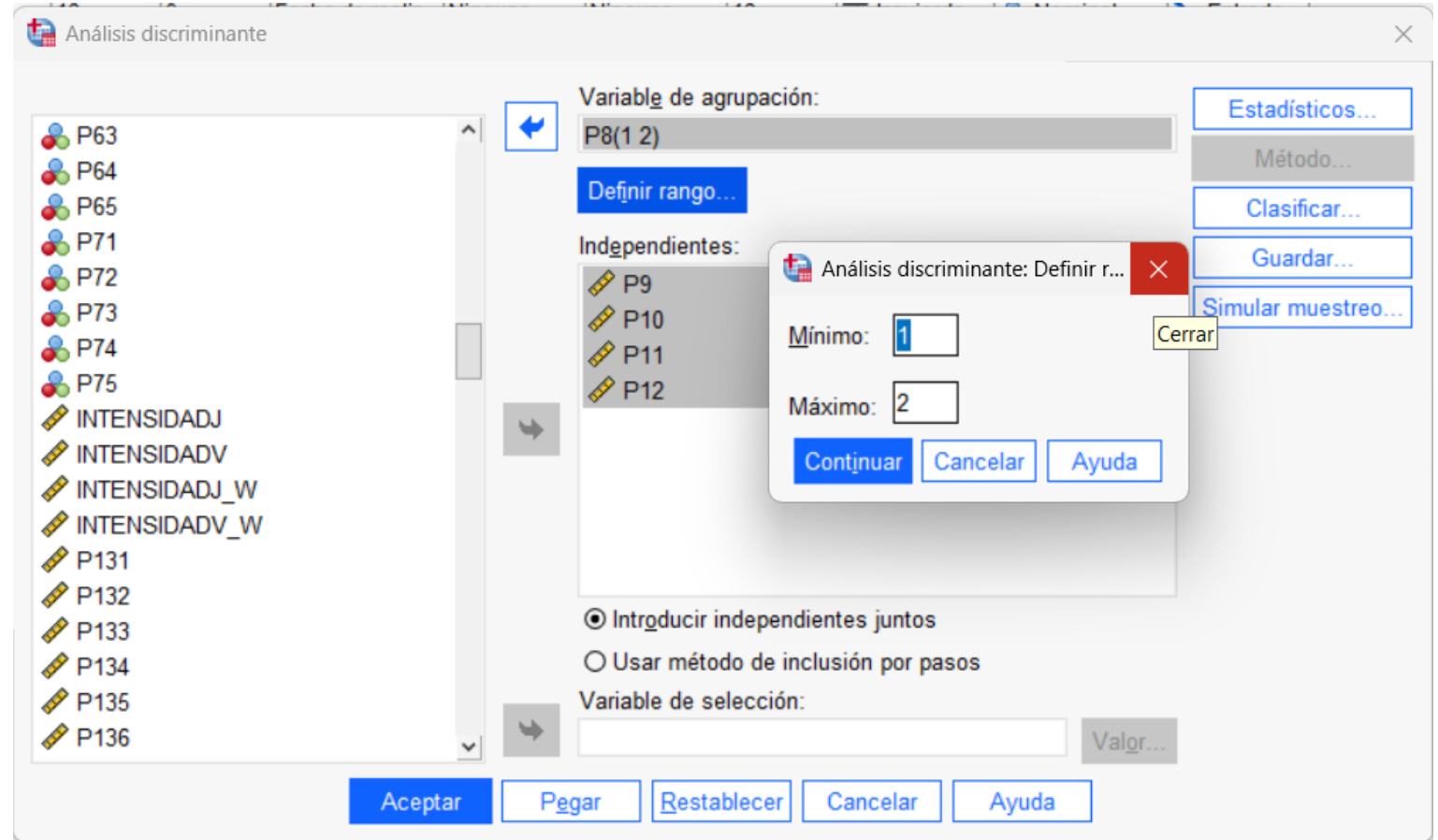
**Funciones lineales de Fischer:** interpretar la clasificación final.

---

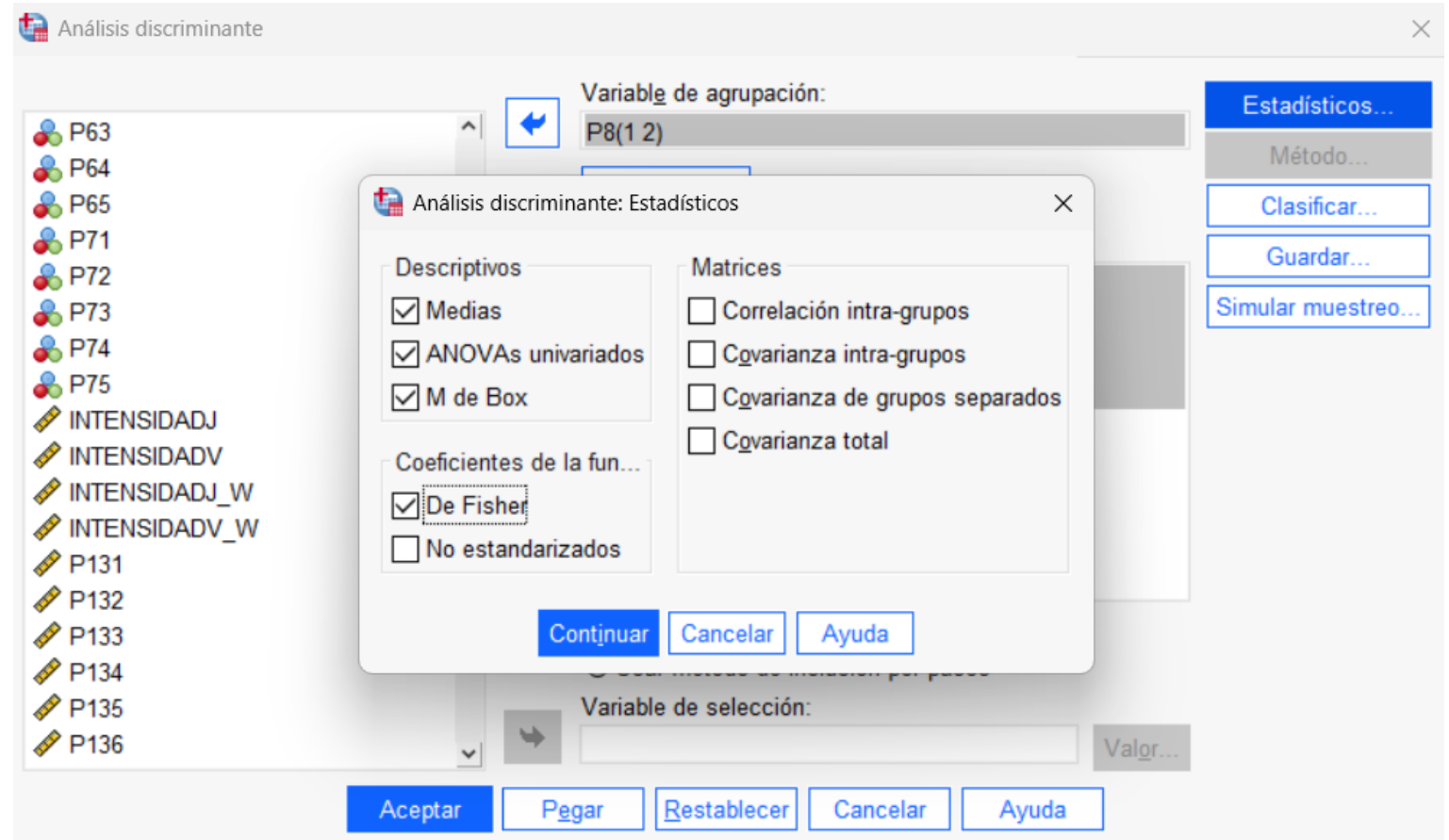
**Matriz de clasificación:** analizar el porcentaje de casos bien clasificados y comparar con la validación cruzada.



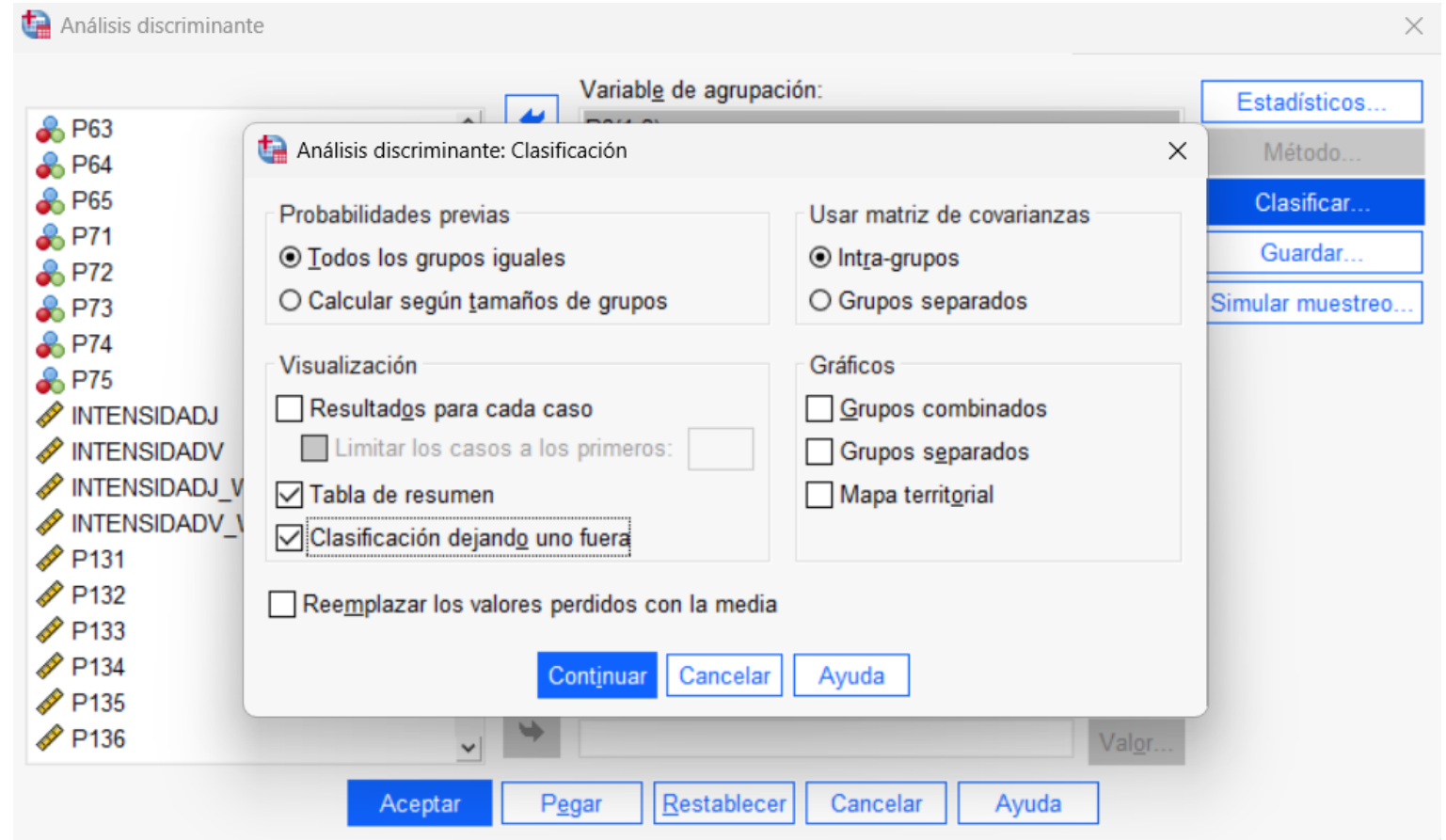
# Aplicación práctica con SPSS



# Aplicación práctica con SPSS



# Aplicación práctica con SPSS



# Ejercicio 1: Clientes compradores ecommerce (clientes\_ecommerce.sav)

- Una empresa de comercio electrónico quiere identificar qué características diferencian a los clientes que han realizado al menos una compra en el último año de aquellos que no han comprado. Para una muestra de clientes, se dispone de la siguiente información: gasto medio mensual en la web (€), número medio de visitas mensuales, tiempo medio de permanencia en la web (segundos), antigüedad como cliente (años). Además, se conoce si cada cliente ha comprado o no ha comprado durante el último año. Con esta información:
  - Justifique el uso de un método estadístico adecuado para clasificar a los clientes en compradores y no compradores.
  - Aplique el método seleccionado y valore la calidad del modelo.
  - Concluya de manera clara qué perfil caracteriza a los clientes compradores.

## Ejercicio 2: Riesgo de impago de clientes de entidad financiera (riesgo\_impago.sav)

- Una entidad financiera desea identificar las características que diferencian a los clientes que han incurrido en impago de aquellos que no han tenido incidencias en el pago de sus préstamos. Para cada cliente se dispone de: ingresos mensuales (€), ratio deuda/ingresos (%), antigüedad laboral (años) y saldo medio en tarjeta de crédito (€). Además, se conoce si el cliente ha tenido o no impago en el pasado. En base a esta información, se pide:
  - Justificar el uso de un método adecuado para clasificar a los clientes según su riesgo de impago.
  - Analizar e interpretar el modelo.
  - Concluir sobre qué variables estás más asociadas al riesgo de impago.

# Ejercicio 3: Clasificación de clientes según riesgo de siniestro (riesgo\_siniestro.sav)

- Una aseguradora quiere clasificar a sus clientes entre aquellos con alta probabilidad de siniestro y aquellos con baja probabilidad de siniestro, con el objetivo de ajustar sus políticas de riesgo. Para cada cliente, se dispone de información sobre: edad del asegurado, número de siniestros previos, coste medio de los siniestros (€), antigüedad en la compañía (años), kilómetros recorridos al año. Además, se sabe que en la cartera de la aseguradora el número de clientes sin siniestro es muy superior al de clientes con siniestro.
  - Con esta información, justifique el uso de un análisis multivariante para clasificar a los clientes según el riesgo de siniestro.
  - Evalúe la solución propuesta.
  - Concluya si el modelo es útil para apoyar decisiones de gestión del riesgo.



UNIVERSIDAD  
DE MÁLAGA

| [uma.es](http://uma.es)