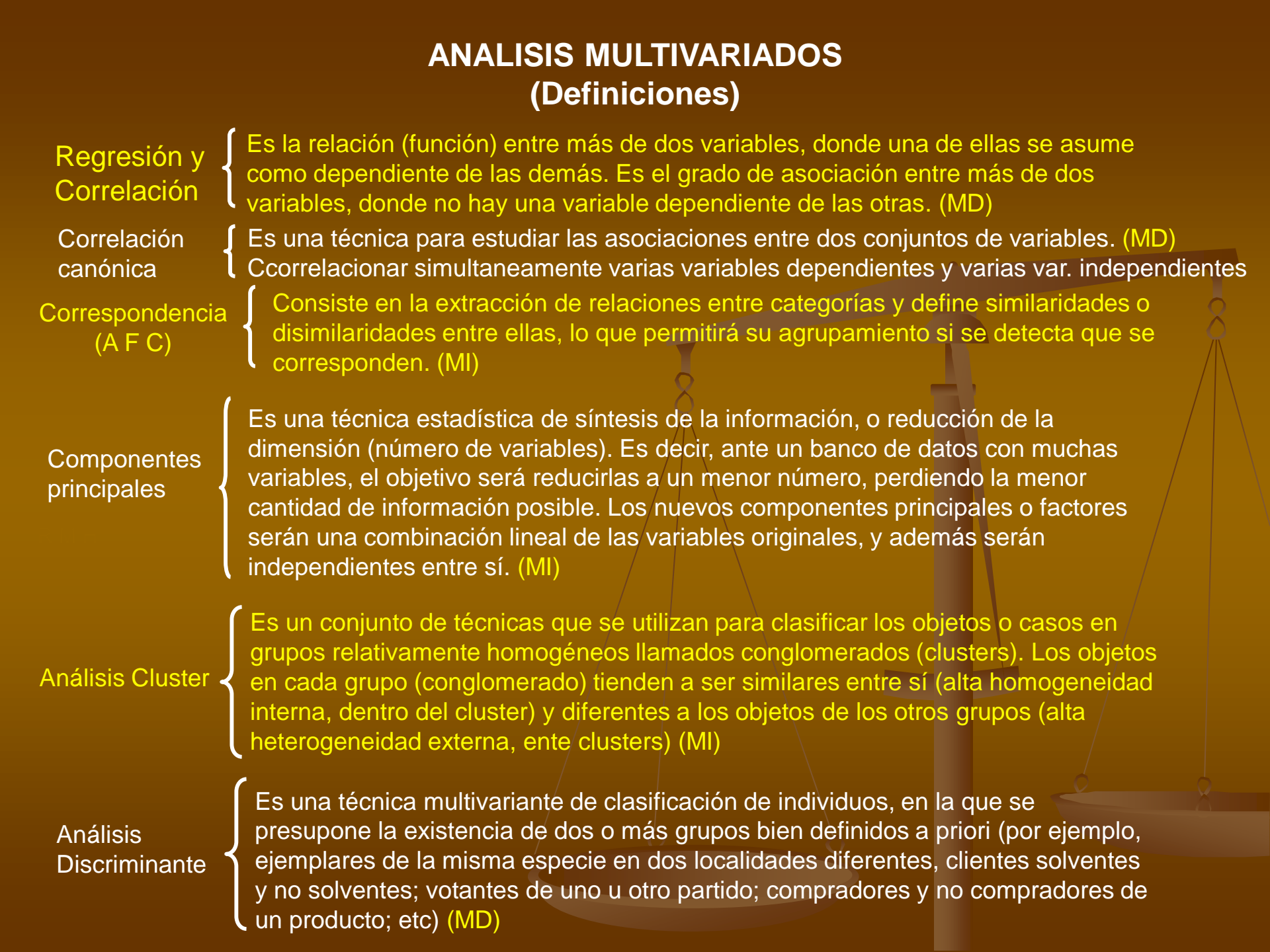


# ANÁLISIS MULTIVARIADOS

## (Definiciones)

- 
- Regresión y Correlación** { Es la relación (función) entre más de dos variables, donde una de ellas se asume como dependiente de las demás. Es el grado de asociación entre más de dos variables, donde no hay una variable dependiente de las otras. (MD)
- Correlación canónica** { Es una técnica para estudiar las asociaciones entre dos conjuntos de variables. (MD)  
Ccorrelacionar simultaneamente varias variables dependientes y varias var. independientes
- Correspondencia (A F C)** { Consiste en la extracción de relaciones entre categorías y define similitudes o disimilitudes entre ellas, lo que permitirá su agrupamiento si se detecta que se corresponden. (MI)
- Componentes principales** { Es una técnica estadística de síntesis de la información, o reducción de la dimensión (número de variables). Es decir, ante un banco de datos con muchas variables, el objetivo será reducirlas a un menor número, perdiendo la menor cantidad de información posible. Los nuevos componentes principales o factores serán una combinación lineal de las variables originales, y además serán independientes entre sí. (MI)
- Análisis Cluster** { Es un conjunto de técnicas que se utilizan para clasificar los objetos o casos en grupos relativamente homogéneos llamados conglomerados (clusters). Los objetos en cada grupo (conglomerado) tienden a ser similares entre sí (alta homogeneidad interna, dentro del cluster) y diferentes a los objetos de los otros grupos (alta heterogeneidad externa, ente clusters) (MI)
- Análisis Discriminante** { Es una técnica multivariante de clasificación de individuos, en la que se presupone la existencia de dos o más grupos bien definidos a priori (por ejemplo, ejemplares de la misma especie en dos localidades diferentes, clientes solventes y no solventes; votantes de uno u otro partido; compradores y no compradores de un producto; etc) (MD)

# ANÁLISIS MULTIVARIADOS (Condiciones)

Regresión

En la relación de las variables independientes (todas con distribución normal), con la variable dependiente (también normal), se establece un error.

Correspondencias  
(A F C)

En este análisis se ordenan los datos en una tabla disyuntiva completa (Z) que consta de un conjunto de individuos  $I=1,2,\dots,n$  (filas), un conjunto de variables o caracteres cualitativos  $J_1, \dots, J_K, \dots, J_Q$  (columnas) y un conjunto de modalidades excluyentes  $1, \dots, m_k$  para cada carácter cualitativo.

Componentes  
principales  
(ACP)

Los únicos requerimientos previos para la aplicación del ACP son:  
a) Continuidad en las variables.  
b) El número  $n$  de individuos o elementos observados debe ser mayor que el número  $p$  de variables originales.  
Por otra parte, el ACP tiene la ventaja de no exigir supuestos tales como la normalidad u homoscedasticidad.

Análisis Cluster  
(AC)

El AC es un objetivo metodológico para cuantificar las características de un conjunto de observaciones. Por ello, tiene fuertes propiedades matemáticas, pero no fundamentos estadísticos. Los requisitos de normalidad, linealidad y homocedasticidad (tan relevantes en otras técnicas), tienen poca consistencia en el AC.

Análisis  
Discriminante

Es una técnica estadística de la rama del análisis multivariante, en la cual la variable dependiente es indicadora y no numérica como en el análisis de regresión. El modelo se construye basado en un set de observaciones para las cuales se conocen las clases. Este set de observaciones es algunas veces conocido como el training set.

# ANÁLISIS MULTIVARIABLES (a considerar)

Regresión  
Correlación

Es importante analizar la magnitud, signo y significancia de los coeficientes. Su valor indica el grado de asociación entre las variables.

Correspondencias  
(A F C)

Es centrado y el centro de gravedad de las modalidades de una variable coincide con el del conjunto J, y con el origen, las modalidades de cada variable están centradas en torno al origen, no pudiendo tener todas el mismo signo.

Componentes  
Principales

Un aspecto clave en ACP es la interpretación de los factores, ya que ésta no viene dada a priori, sino que será deducida tras observar la relación de los factores con las variables iniciales (habrá, pues, que estudiar tanto el signo como la magnitud de las correlaciones).

Análisis Cluster

No jerárquicos

Asignan los casos a grupos diferenciados que el propio análisis configura, sin que unos dependan de otros.

Jerárquicos

Aquéllos que configuran grupos con estructura arborescente, de forma que clusters de niveles más bajos van siendo englobados en otros de niveles superiores.

Análisis  
Discriminante

Explicar

La pertenencia de cada caso del archivo patrón a uno u otro grupo, en función de las variables de su perfil, para comprobar su pertenencia.

Predecir

A qué grupo más probable habrá de pertenecer un nuevo individuo del que únicamente se conoce su perfil de variables.

