

# IN627 Investigación de Mercados

Profesores:  
Rodrigo Niño  
Emilio Polit

1

## 7. Segmentación

2

### Segmentación

*"Los segmentos de mercado consisten en grupos de personas que son similares en términos de cómo ellas responden a un conjunto de variables o en alguna forma que sea útil para los propósitos de la planeación de marketing"*

3  
© 2003 DII

### Motivación

- La segmentación de mercados es uno de los principales objetivos de marketing asociados a la Investigación de Mercados.
- La segmentación corresponde al proceso de dividir un mercado en grupos identificables, similares y significativos, con el propósito de ajustar el marketing-mix a la medida de las necesidades de uno o más segmentos específicos.



4  
© 2003 DII

### Motivación

- ¿qué procedimiento debo seguir para hacer una clasificación de objetos?.
- ¿para qué me puede servir hacer una clasificación de objetos?
- ¿qué técnicas existen para clasificar objetos?
- ¿cuál es la diferencia entre agrupamiento jerárquico y agrupamiento no jerárquico?
- ¿cuántos segmentos debo crear?

5  
© 2003 DII

### Motivación

- La idea central de una buena segmentación es encontrar grupos heterogéneos entre sí, en que los elementos que componen cada grupo sean relativamente homogéneos entre sí.
- En general buscaremos grupos significativos, alcanzables, homogéneos y diferenciados de los demás.



6  
© 2003 DII

# Análisis de conglomerados

(Cluster Analysis)

## ...y su aplicación a la segmentación de mercados...

7

## Conceptos Previos

8

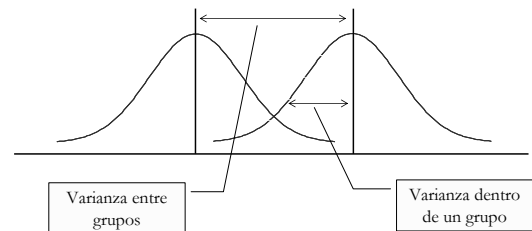
### Análisis de conglomerados

- **Definición:** conjunto de técnicas que se utilizan para clasificar los objetos o casos en grupos relativamente homogéneos llamados conglomerados (clusters).
- Existen diferentes los métodos por los cuales cada objeto es asignado a un segmento.
- Se busca que los objetos dentro cada grupo sean similares entre sí y diferentes a los objetos de los otros grupos.
- Se utiliza un principio de maximización de la varianza entre clusters mientras que minimiza la varianza dentro de un cluster.



9

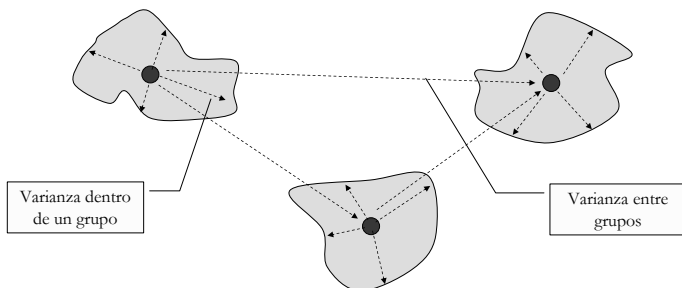
### Gráficamente



$$\max_{c \in C} \left\{ \frac{\text{varianza entre grupos}}{\text{varianza en los grupos}} \right\} C = \text{Conjunto de clusters posibles}$$

© 2003 DII 10

### Gráficamente



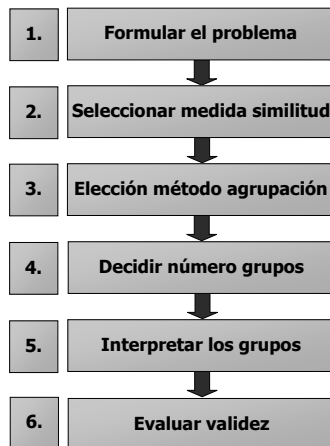
$$\max_{c \in C} \left\{ \frac{\text{varianza entre grupos}}{\text{varianza en los grupos}} \right\} C = \text{Conjunto de clusters posibles}$$

© 2003 DII 11

## Procedimiento para la realización de un análisis de conglomerados

12

## Esquema general



## 1. Formular el problema

- Corresponde a la selección de las variables en las que se basa la agrupación.
- La inclusión de variables irrelevantes puede distorsionar las soluciones.
- El conjunto de variables debe describir la similitud de los objetos de acuerdo a los propósitos de la clasificación.
- La selección de variables se hace en base a los resultados de una etapa exploratoria previa de la investigación..

Ésta es probablemente la parte mas importante en el análisis de conglomerados y requiere el desarrollo de un buen criterio e intuición.



## Variables de segmentación

- SEGMENTACION PSICOGRAFICA
- SEGMENTACION POR BENEFICIOS
- SEGMENTACION DE USO
- SEGMENTACION DEMOGRAFICA
- SEGMENTACION GEOGRAFICA



- Latente (oculta)
- Compleja

- Manifiesta (visible)
- Simple

## Variables de Segmentación Geográfica

- Es la segmentación más obvia.
- Consiste en segmentar a los consumidores por donde ellos se encuentran (país, ciudad, comuna).
- Todas las empresas que se localizan geográficamente para atender a sus clientes hacen una segmentación geográfica (supermercados, bancos, restaurantes, mall, arriendo de videos, etc)

## Variables de Segmentación Demográfica

- Variables Demográficas:** Edad, sexo, educación, ingreso, ocupación, estado civil, tenencia de bienes, etc.
- Las variables demográficas son las más comúnmente usadas.
- Mientras más desarrollados (competitivos) son los mercados la segmentación por variables demográficas (por ejemplo por nivel socioeconómico es menos efectiva.

## Variables de Segmentación por Uso del Producto

### Variables relacionadas con el hábito de uso/compra de los consumidores con los productos.

- Cantidad de uso (¿cuántas veces al mes compra el producto?)
- Estacionalidad de uso (¿cuándo consume el producto?)
- Patrones de compra (¿con qué otros productos compra el producto?)
- Lealtad de marca (leales y desleales)
- Sensibilidad al precio (sensibles y no sensibles)
- Forma de uso (intensivos y no intensivos)

## Variables de Segmentación por Beneficio Buscado en el Producto

- Un mismo producto puede ser usado por diferentes consumidores por distintas razones.
- Cada uno de ellos busca un BENEFICIO distinto en el producto.
- Requiere la siguiente información:
  - ✓ Atributos o Ventajas Asociados a la Categoría Producto
  - ✓ Evaluación de la Importancia Relativa de ellos
  - ✓ Reagrupamiento de los compradores con evaluaciones similares.
  - ✓ Identificación del tamaño y perfil de los compradores de cada segmento.

© 2003 DII 19

## Variables de Segmentación Psicográfica

- Llamadas también de “**ESTILOS DE VIDA**”
- La segmentación por variables psicográficas parte de la premisa que individuos muy diferentes demográficamente pueden tener comportamientos de compra similares.
- Cada vez más usadas, especialmente en mercados complejos.
- Incorpora variables de: opinión, interés, actividad, estilo de vida.

Actividades	Intereses	Opiniones	Estilos de vida
Trabajo	Familia	Sobre sí mismo	Tradicionalista
Hobbies	Hogar	Asuntos sociales	Ambientalista
Vida social	Trabajo	Política	Vanguardista
Vacaciones	Comunidad	Negocios	Deportista
Placeres	Diversión	Economía	Vida social activa
Clubes	Educación	Habitat	Cultural
Comunidades	Alimentación	Productos	Light
Compras	Medios de comunicación	Futuro	Tolerante
Deporte	Realización	Cultura	Innovador

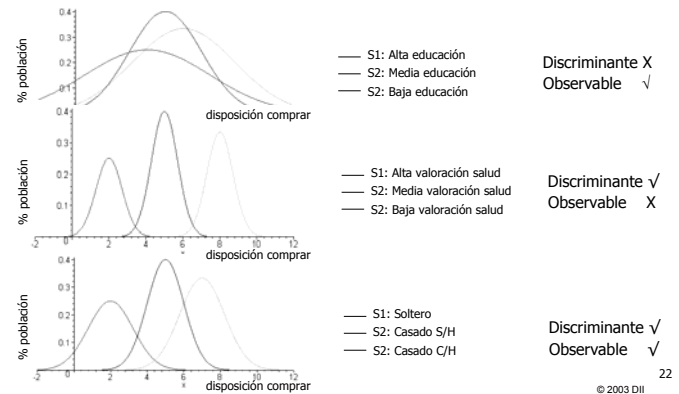
© 2003 DII 20

## Variables de segmentación

- Las variables usadas en la segmentación deben ser **relevantes** para el objetivo buscado.
- Las variables de segmentación deben ser capaces de **establecer diferencias** entre los distintos individuos.
- Deben ser **observables** o fáciles de medir.

© 2003 DII 21

## Ejemplo: venta planes Isapre



© 2003 DII 22

## 2. Seleccionar medida similitud

### Medidas de asociación continua:

- Distancia Euclídeana:  $d_{ij} = \sqrt{\sum_k (x_i^k - x_j^k)^2}$
- Distancia Manhattan:  $d_{ij} = \sum_k |x_i^k - x_j^k|$
- Distancia de Chebychev:  $d_{ij} = \max_k \{|x_i^k - x_j^k|\}$

- Si las unidades de medición de las variables seleccionadas son diferentes, conviene estandarizar las variables.
- Detección de *outliers*.
- Distintas medidas generan distintos resultados. Se recomienda usar varias medidas y comparar resultados.

© 2003 DII 23

### Medidas de asociación discreta:

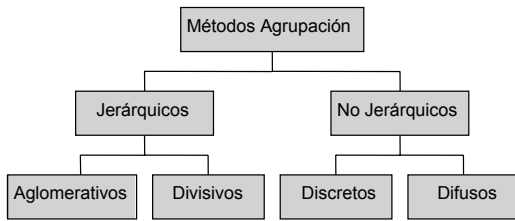
- Consideremos dos objetos  $i$  y  $j$  caracterizados por vectores  $x_i$  y  $x_j$  de  $A$  atributos con valores en  $\{0,1\}$  (cumple o no cumple una condición).
- Sean:
  - $a$ : # atributos en que  $i$  y  $j$  cumplen
  - $b$ : # atributos en que  $i$  si cumple y  $j$  no cumple.
  - $c$ : # atributos en que  $i$  no cumple y  $j$  si cumple.
  - $d$ : # atributos en que  $i$  y  $j$  no cumplen.
$$a + b + c + d = A$$

$$\text{Jaccard} \quad d(i, j) = \frac{a}{a + b + c}$$

$$\text{Sokal-Sneath} \quad d(i, j) = \frac{a + d}{b + c}$$

© 2003 DII 24

### 3. Seleccionar procedimiento agrupación



© 2003 DII 25

### Métodos de asignación

- Corresponden a la lógica en que los objetos se asignan a cada grupo.
- **Métodos jerárquicos:** Los objetos se agrupan (dividen) por partes hasta clasificar todos los objetos.
  - De una iteración a otra, se modifica el valor de pertenencia a grupos de un único objeto.
  - No requiere *a priori* fijar un número de clusters.
- **Métodos no jerárquicos:** Se tiene un número de grupos predefinidos y cada objeto se ubica en un grupo hasta alcanzar estabilidad.
  - De una iteración a otra, se puede modificar el valor de pertenencia a grupos de todos los objetos.
  - Requiere *a priori* fijar un número de clusters.

© 2003 DII 26

### Métodos Jerárquicos

- En cada iteración se determina un criterio de aglomeración/división formando dendogramas/árboles.
- Las estructura de árbol o dendograma nos indica explícitamente la naturaleza de las relaciones entre objetos.
- Distinguimos dos tipo de métodos:
  - **Aglomerativos:** Inicialmente cada objeto es un grupo y en cada iteración se juntan los grupos mas similares.
  - **Divisivos:** Inicialmente se tiene un único grupo y en cada división se separa un grupo.

© 2003 DII 27

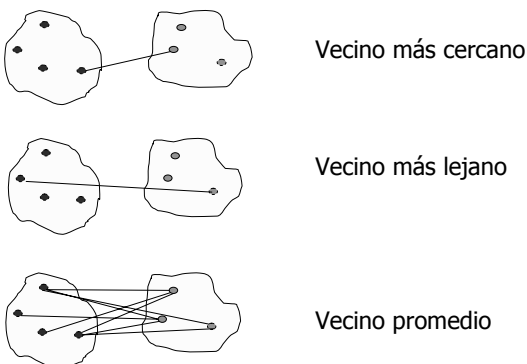
### Métodos Jerárquicos

#### Métodos Aglomerativos

- Esquema general algoritmo:
  1. Cada objeto corresponde a un grupo.
  2. En cada iteración se juntan los dos grupos mas cercanos bajo algún criterio de cercanía entre grupos.
  3. Los dos grupos recién unidos forman un único grupo.
  4. Iterar hasta formar un único grupo.
- El número de grupos puede definirse a posteriori.
- El método jerárquico aglomerativo más utilizado es **Ward**.

© 2003 DII 28

### Criterios de cercanía



© 2003 DII 29

### Métodos Jerárquicos

#### Métodos Aglomerativos

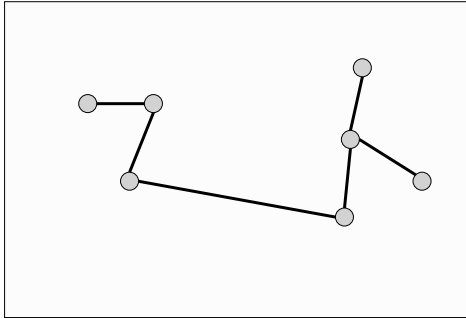
##### Método de Ward

- Este procedimiento trata de identificar grupos de casos, tratando de minimizar la varianza dentro de los grupos.
- Se minimiza la distancia euclideana cuadrada a las medias del conglomerado



© 2003 DII 30

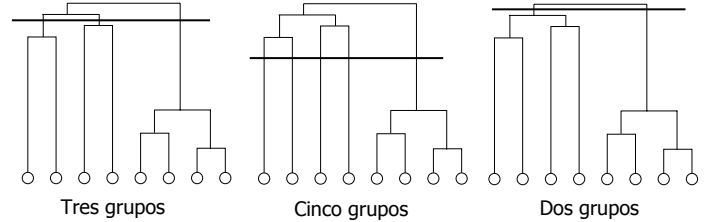
## Ejemplo gráfico



7 clusters  
6 clusters  
5 clusters  
4 clusters  
3 clusters  
2 clusters  
1 cluster

## Dendogramas

- Un dendograma es un árbol en el que el largo de las ramas está asociado inversamente a la fortaleza de la relación.



## Métodos Jerárquicos

### Métodos Divisivos

- Esquema general algoritmo:
  - Todos los objetos corresponde a un grupo.
  - Cada grupo se separa bajo algún criterio de maximización de varianza entre grupos.
  - Dividir cada uno de los grupos hasta que:
    - Todos los grupos sean tan homogéneos que no vale la pena seguir dividiendo.
    - Los grupos son tan pequeños que no vale la pena seguir dividiendo.
- El número de grupos puede definirse a priori.
- Junto de decidir que subgrupo separar se debe especificar con qué criterio se asignará cada objeto a cada subconjunto generado.

## Métodos Jerárquicos

### Métodos Divisivos

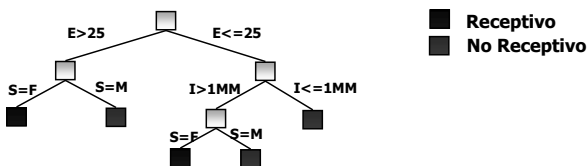
- El proceso genera junto con la estructura de árbol, **reglas explícitas** de clasificación para los objetos.
- El método jerárquico divisivo más utilizado es **Cart**.

#### Método de Cart

- Este procedimiento trata de clasificar las variables independientes de modo de discriminar mejor la variable dependiente.
- Cada nodo del árbol se divide para maximizar la *pureza* de los nodos hijos.

## Árboles de clasificación

- Idea:** Dada una clasificación de objetos se busca determinar una estructura jerárquica de reglas que permitan discriminar de la mejor manera posible entre las categorías.
- Ejemplo:** En las bases de datos de la empresa tengo una gran información de clientes (edad, ingreso, sexo) y se desea saber el perfil de las personas que serán receptivas a la realización de una campaña de promoción.



## Métodos No Jerárquicos

- Algoritmos iterativos que en cada iteración ubican a los objetos en el grupo más cercano a él.
- El número de grupos está predefinido.
- Al asignar todos los objetos en cada iteración, los algoritmos son mucho más eficientes para problemas complejos.
- Distinguimos dos tipos de métodos:
  - Discretos:** Cada objeto sólo puede pertenecer a un único grupo.
  - Difusos:** Cada objeto tiene un grado de pertenencia a cada uno de los grupos.

## Métodos No Jerárquicos

### Métodos Discretos

- El método no jerárquico discreto más utilizado es **K-Means**.
- K-means es una de las técnicas de clustering más utilizada en la segmentación aplicada a la Investigación de Mercados .
- El número de grupos debe definirse a priori.

© 2003 DII 37

## Métodos No Jerárquicos

### Métodos Discretos

#### Método de K-Means

- Esquema general algoritmo:
  - Se tiene un conjunto de  $N$  objetos y  $K$  grupos.
  - Antes de la primera iteración se eligen arbitrariamente los centros de cada grupo.
  - En cada iteración se asigna cada objeto a su grupo más cercano y luego se recalculan los centros de cada grupo con los nuevos elementos asignados.
  - Iterar hasta que los cambios en los centros de cada grupo no sean significativos.
- El método entrega los elementos que pertenecen a cada grupo y los centros de cada grupo.

© 2003 DII 38

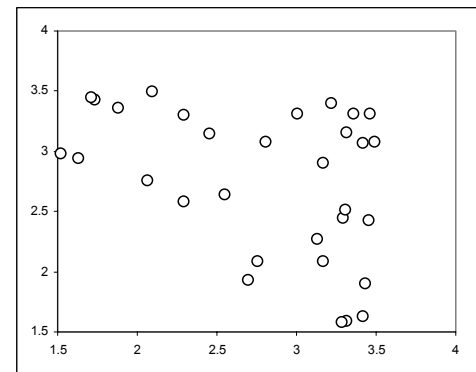
### Formulación matemática K-means

- Entradas
  - $X$  conjunto de  $N$  objetos
  - $K$  número de grupos
- Salidas
  - $S_1, \dots, S_K$   $K$  conjuntos
  - $Z_1, \dots, Z_K$  Los centros de cada grupo
- Inicialización.
  - $t=0$
  - Elegir arbitrariamente  $Z_j(t)$ .
- Asignación y actualización de centros.
  - Asignar  $X_i$  al grupo mas cercano para todo  $i=1 \dots N$ .
  - Recalcular  $Z_j$   $j=1 \dots K$
  - $t=t+1$
- Criterio de parada.
  - Si  $Z_j(t) - Z_j(t+1) < \epsilon$  para todo  $j$ , parar.



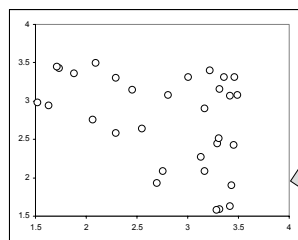
© 2003 DII 39

### Ejemplo Gráfico

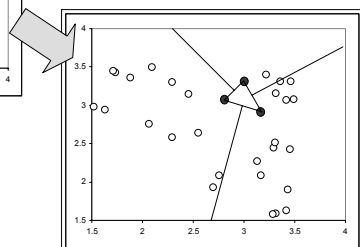


© 2003 DII 40

### Iteración 1

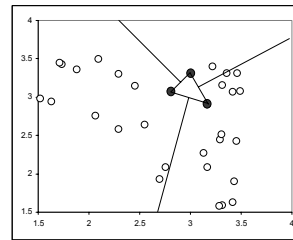


- Elegir los centros de los tres clusters aleatoriamente
- Localizar cada punto en su centro de cluster más cercano

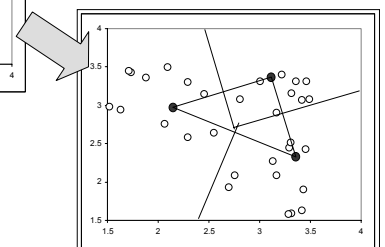


© 2003 DII 41

### Iteración 2

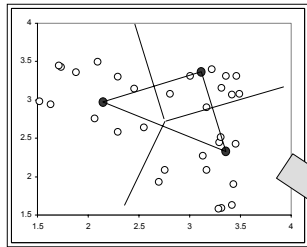


- Calcular nuevamente los centros de los clusters desde los centroides escogidos en la iteración 1.
- Localizar cada punto en el centroide que está más cerca a él.

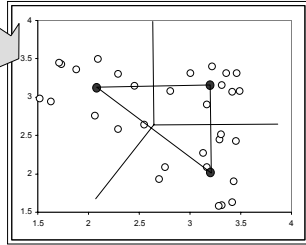


© 2003 DII 42

## Iteración 3



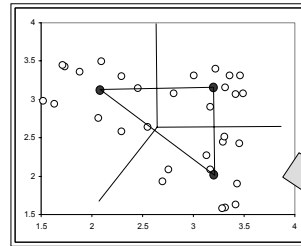
- Recalcular los centros como los centroides encontrados en la iteración 2.
- Localizar cada punto en el centroide que está más cerca a él.



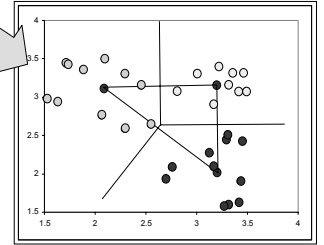
© 2003 DII

43

## Iteración 4 y etapa final



- Recalcular los centros como los centroides de los clusters desde la iteración 3.
- Nada cambió!!
- Ok, está listo.



© 2003 DII

44

## Métodos No Jerárquicos

### Métodos Difusos

- La lógica difusa se basa en la idea que las cosas no tienen un valor de verdad absoluto (verdadero o falso) sino que tienen grados de verdad.
- Los elementos tienen un grado de pertenencia a los distintos conjuntos.
  - Formalmente, sea  $X$  un conjunto tradicional. Un conjunto difuso  $Y$  se define como:
 
$$Y = \{(x, u_Y(x)); x \in X\}$$
  - Donde  $u_Y(x)$  es el grado de pertenencia de el elemento  $x$  al conjunto difuso  $Y$ .
- El método no jerárquico difuso más utilizado es **Fuzzy C-Means**.

© 2003 DII

45

## Métodos No Jerárquicos

### Métodos Difusos

#### Método de Fuzzy C-Means

- Es muy similar al algoritmo de K medias, pero en cada iteración asigna un grado de pertenencia de acuerdo a las distancias a los centroides.
- Los centroides se calculan ponderando el valor de los atributos por los grados de pertenencia de los objetos.
- El número de grupos debe definirse a priori.
- Se debe definir el grado o parámetro de difusividad  $m$  que va desde 1 (k-means) a  $\infty$  (completamente difusos).

© 2003 DII

46

## Métodos No Jerárquicos

### Métodos Difusos

#### Método de Fuzzy C-Means

- Esquema general algoritmo:
  1. Antes de la primera iteración se elige arbitrariamente el grado de pertenencia de cada objeto a cada grupo,  $u_{ij} \in [0,1]$  ( $u_{ij}$  = grado de pertenencia de objeto  $i$  a grupo  $j$ ), donde  $\sum_j u_{ij} = 1$
  2. En cada iteración se determinan los centros de los grupos y se actualizan los grados de pertenencia, utilizando una función con un grado de difusividad asociado.
  3. Iterar hasta que los cambios en los centros de cada grupo no sean significativos.
- El método entrega el grado de pertenencia de cada elemento a cada grupo y los centros de cada grupo.

© 2003 DII

47

## Formulación matemática Fuzzy -C-means

- Supongamos que disponemos de  $n$  objetos que queremos agrupar en  $c$  clases difusas.
- Sean:
  - $u_{ij}$  = grado de pertenencia de objeto  $i$  a clase  $j$ .
  - $x_i$  = vector de atributos del objeto  $i$ .
  - $c_j$  = centro de la clase  $j$ .
  - $d^2(x, c) =$  distancia entre vectores  $x$  y  $c$ .
  - $m =$  parámetro de difusividad ( $1 < m < \infty$ ).
- $m$  permite ajustar que tan difusos serán los conjuntos resultantes.

© 2003 DII

48



## Formulación matemática Fuzzy -C-means

### ■ Inicialización:

- Determinar arbitrariamente una matriz  $U=\{u_{ij}\}$  tal que:
  - $u_{ij} \in [0,1]$
  - $\sum_j u_{ij} = 1$ .

### ■ Paso genérico:

- Determinar los centros de las clases: 
$$c_j = \frac{\sum_{i=1}^n u_{ij} \cdot x_i}{\sum_{i=1}^n u_{ij}}$$
- Actualizar los grados de pertenencia: 
$$u_{ij} = \left\{ \sum_{h=1}^c \left( \frac{d(x_i, c_j)}{d(x_i, c_h)} \right)^{\frac{2}{m-1}} \right\}^{-1}$$

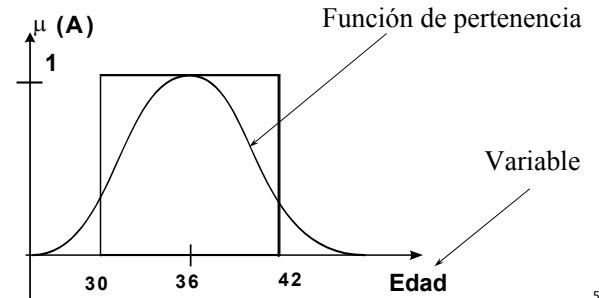
### ■ Criterio de parada:

- Si  $Z_i(t) - Z_i(t+1) < \epsilon$  para todo  $i$ , parar.

© 2003 DII 49

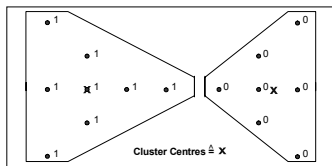
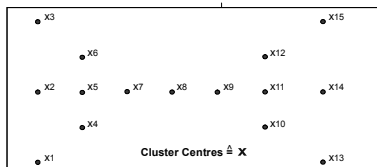
## Ejemplo Gráfico

“Cliente joven”



© 2003 DII 50

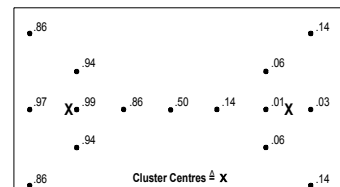
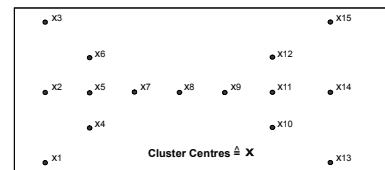
## Ejemplo Gráfico



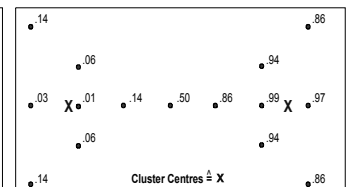
Grupos estrictos

© 2003 DII 51

## Ejemplo Gráfico



Grupo difuso 1



Grupo difuso 2

© 2003 DII 52

## 4. Decidir número de grupos

### ■ No existe regla general. Algunos lineamientos:

- Consideraciones estratégicas o comerciales.
- Consideraciones presupuestarias.
- En conglomerados jerárquicos, observar los cambios en las distancias en las que los grupos se combinan (ej: uso de dendograma).
- En conglomerados jerárquicos puede graficarse la razón (varianza entre grupos)/(número de grupos) y detectar puntos de cambios marcados.
- Los tamaños de los grupos deben ser significativos.

© 2003 DII 53

## 5. Interpretar y elaborar un perfil de los conglomerados

- Se busca una semántica que defina a los objetos del grupo.
- Dos enfoques complementarios:
  - Análisis de los centroides de cada grupo y comparación con el de los otros grupos.
  - Análisis discriminante para determinar las variables que marcan diferencias significativas.



© 2003 DII 54

## 6. Evaluar validez del conglomerado

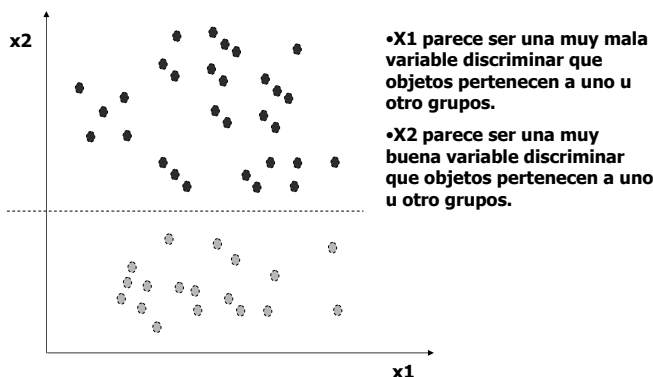
- Algunos chequeos mínimos de validez:
  - Analizar la estabilidad de la clasificación haciendo el análisis para selecciones distintas de variables.
  - Utilizar varios métodos y comparar resultados.
  - Dividir los datos en dos partes: hacer análisis de conglomerados en ambas y comparar resultados.
  - Para clasificación no jerárquica, cambiar valor de centroides iniciales.
  - Para clasificación no jerárquica, cambiar el número de cluster.

© 2003 DII 55

## Análisis Discriminante

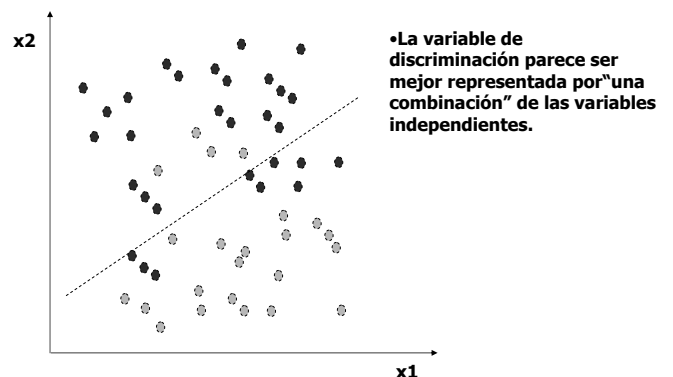
56

### Motivación



© 2003 DII 57

### Motivación



© 2003 DII 58

### Definición

- Técnica de análisis de datos en que la variable dependiente es categórica y las variables independientes son de intervalo.
  - Ej 1: Una empresa de confecciones quiere determinar que elementos permiten distinguir a los clientes que prefieren determinadas marcas de ropa
    - ⇒ Variable dependiente: Marca de ropa preferida por el cliente.
    - ⇒ Variables independientes: Ingreso, edad, puntaje en encuesta sobre importancia de la calidad, puntaje en encuesta sobre estar a la moda.
  - Ej 2: Un profesor de un curso desea saber que factores inciden en la aprobación de un alumno:
    - ⇒ Variable dependiente: 1 si alumno aprueba, 0 si alumno reprueba.
    - ⇒ Variables independientes: Notas semestres anteriores, comuna en que vive, etc.

© 2003 DII 59

### Usos de análisis discriminante en Investigación de Mercados

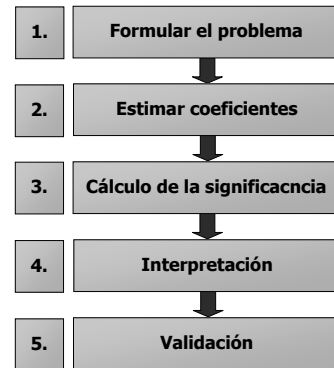
- Algunos ejemplos:
  - ¿cómo se diferencian los consumidores de que muestran lealtad respecto a quienes no la muestran?
  - ¿qué características psicográficas ayudan a diferenciar entre los compradores que consideran el precio en la elección y aquellos que no lo consideran.
  - ¿Los diversos segmentos del mercado difieren en sus hábitos de uso de medios de comunicación?

© 2003 DII 60

# Procedimiento para la realización de un análisis discriminante

61

## Esquema general



## 1. Formular el problema

- Identificación de la variable dependiente y de las variables independientes.
- Si la variable dependiente es de intervalo o relación debe categorizarse.
- En esta etapa se sugiere dividir las observaciones en dos grupos:
  - **Datos de análisis:** sobre ellos se calibran los parámetros y se encuentra la función discriminante.
  - **Datos de validación:** se utilizan para la validación y la estimación de la capacidad predictiva de la solución.



© 2003 DII 63

## 2. Estimar coeficientes

- Los procedimientos de calibración dependen del método de análisis discriminante que se esté utilizando.
  - **Método directo:** El cálculo de la función discriminante se realiza de modo que todos los indicadores se incluyen simultáneamente.
  - **Método discreto:** Los indicadores entran secuencialmente al análisis en base a su capacidad de discriminación.
  - Existen otros enfoques como el de "support vector machines".

© 2003 DII 64

## 3. Determinación del grado de significancia.

- No tendría sentido interpretar el análisis si las funciones discriminantes no fueran significativas.
- La idea básica de los test de hipótesis asociados a este problema es "tratar" de rechazar la hipótesis nula que no existe diferencia en las funciones discriminantes de los distintos grupos.
- Algunos procedimientos básicos:
  - Comparación entre los valores de la función discriminante para los centroides de los grupos.
  - Test de hipótesis estadístico.

© 2003 DII 65

## 4. Interpretación

- Procedimiento similar a la caracterización de los grupos del análisis de conglomerados:
  - Analizar las características de los centroides.
  - Comparar los valores de la función discriminante para cada uno de los grupos.
  - Analizar los coeficientes de la función discriminante.

© 2003 DII 66

## 5. Validación

- Recordemos que tenemos dos conjuntos disjuntos de datos:
  - Datos de análisis.
  - Datos de validación.
- Se calibran los coeficientes de acuerdo a los datos de análisis.
- Se testea la capacidad de usar la función discriminante como predictor del grupo en que debiera ser clasificada
- Los resultados suelen representarse en una *matriz de confusión*.

© 2003 DII 67

## Matriz de confusión

		Grupo 1	Grupo 2	Grupo 3	Grupo 4	%Error	Total
Valor	Grupo 1	112	20	18	84	0,52	234,00
	Grupo 2	23	395	3	11	0,09	432,00
	Grupo 3	1	26	497	3	0,05	546,00
	Grupo 4	5	12	7	332	0,07	356,00

De 432 que eran tipo 2,  
91% fueron  
correctamente predichos.

© 2003 DII 68

## Los distintos tipos de análisis asociados a la segmentación de mercados pueden integrarse...

### Ejemplo: La industria del chocolate en Chile

69

## El estudio

- Mercado barras de 200g.
- Etapla cualitativa: 3 focus group a los segmentos a priori detectados como relevantes (mujeres jóvenes, mujeres adultas y hombres adultos).
- Etapla cuantitativa: 224 encuestas en supermercados a compradores de chocolates.



70

## La encuesta

Manifieste su grado de acuerdo con las siguientes proposiciones (-3 muy en desacuerdo, 3 muy de acuerdo).

	-3	-2	-1	0	1	2	3
El chocolate es un alimento							
El consumo de chocolate produce sensación agradable							
El chocolate provoca problemas en la piel.							
El chocolate es un postre y se consume después de las comidas							
El consumo de chocolate produce placer							
El chocolate provoca obesidad							
El chocolate es un buen regalo para ocasiones especiales							
Como chocolate cuando estoy solo(a)							
El chocolate aporta energía							

© 2003 DII 71

## K-medias

- Grupo 1 (24%):
  - El chocolate es un alimento y entrega energía.
  - Comer chocolate es agradable y les da placer.
  - Les gusta comer chocolate cuando están solos.
  - No creen que provoque obesidad ni que cause problemas a la piel.
  - Les gusta regalar chocolates en ocasiones especiales.
- "Amantes del chocolate"

© 2003 DII 72

## K-medias

- Grupo 2 (32%):
  - Consideran que provoca obesidad y que causa problemas a la piel.
  - No gustan comer después de las comidas.
  - No es relevante desde el punto de vista del placer.
  - No les gusta comer chocolate cuando están solos.
- “Consumidor por costumbre”

© 2003 DII 73

## K-medias

- Grupo 3 (4%):
  - Debilmente creen que es un alimento y que entrega energía.
  - Consideran que provoca obesidad y que causa problemas a la piel.
  - Piensan que comer chocolates produce mucho placer.
  - Les gusta comer chocolate cuando están solos.
- “Consumidor culpable”.

© 2003 DII 74

## Análisis discriminante

### Matriz de confusión

		Predicted Group Membership				Total
		1.00	2.00	3.00		
Original	Count	1.00	50	4	0	54
		2.00	7	59	6	72
		3.00	1	9	88	98
	%	1.00	92.6	7.4	.0	100.0
		2.00	9.7	81.9	8.3	100.0
		3.00	1.0	9.2	89.8	100.0
Cross-validated <sup>a</sup>	Count	1.00	50	4	0	54
		2.00	8	58	6	72
		3.00	1	9	88	98
	%	1.00	92.6	7.4	.0	100.0
		2.00	11.1	80.6	8.3	100.0
		3.00	1.0	9.2	89.8	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 87.9% of original grouped cases correctly classified.

c. 87.5% of cross-validated grouped cases correctly classified.

© 2003 DII 75

## Análisis discriminante

- El grupo de consumidores culpables está fuertemente relacionado con el segmento de mujeres jóvenes, y en menor grado con las mujeres adultas.
- Esto concuerda con las hipótesis planteadas en la etapa exploratoria de la investigación, ya que fueron estos segmentos los que plantearon las creencias negativas con respecto al consumo del producto, aunque reconocieron actitudes positivas hacia él.

© 2003 DII 76

## Análisis discriminante

- Los amantes del chocolate principalmente mujeres adultas y en menor grado mujeres jóvenes y hombres adultos.
- Esto también concuerda con los resultados de la etapa exploratoria, ya que fueron los segmentos adultos los que se mostraron menos preocupados o involucrados con las creencias negativas hacia el producto.

© 2003 DII 77

## Análisis discriminante

- Los consumidores por costumbre son principalmente hombres jóvenes y en menor grado hombres adultos.
- Esto resulta coherente con las conclusiones previas, además que concuerda con los porcentajes de consumo y compra del producto por parte de los segmentos, obtenidos anteriormente.

© 2003 DII 78