



UNIVERSITAT DE  
BARCELONA

# Análisis discriminante I

Prof. Miquel Salicrú

Prof. Sergi Civit

## Ideas básicas

**Contexto.** Disponemos de una población  $\Pi$  en la que tenemos definida una partición,

$$\Pi = H_1 \oplus H_2 \oplus \dots \oplus H_q$$

**Problema.** Clasificar un individuo  $\omega_0 \in \Pi$  a una de las clases de la partición, basando la decisión en la observación del vector aleatorio  $(X_1, X_2, \dots, X_p)$ .

**Regla de decisión.** Asignar  $\omega_0$  a la clase  $H_i$  si  $X(\omega_0) \in R_i \subset \mathbb{R}^p$

$$\mathbb{R}^p = R_1 \oplus R_2 \oplus \dots \oplus R_q$$

de forma que se optimice la función objetivo

## Aplicaciones 1/5

### Marketing relacional

- Para orientar las campañas de marketing, identificar y caracterizar los patrones de comportamiento en adquisición de productos y servicios (clasificación)
- Clasificar/asignar individuos nuevos (o con cambio de comportamiento) a una de las categorías (patrón) establecidas (discriminación)
- Identificar y seleccionar (si procede) el conjunto mínimo de variables explicativas (reducir dimensión)
- Valorar la eficiencia de las campañas y realizar el seguimiento individual de clientes (control de clientes)

## Aplicaciones 2/5

### Biomedicina

- Identificar los referentes de las categorías (infectado/no infectado, Nivel T1, Nivel T2,...) por métodos de diagnóstico contrastado y protocolizado (clasificación clínica)
- Obtener la expresión del marcador o marcadores que caracterizan el estadio clínico (PSA: cáncer de próstata; CEA: cáncer de colon,..) y clasificar/asignar los pacientes a una de las categorías establecidas (diagnosticar).
- Realizar el seguimiento de pacientes (control con marcadores)
- Identificar y seleccionar el mínimo conjunto de variables que maximizan la capacidad predictiva.

## Aplicaciones 3/5

### Transporte de pasajeros en zona urbana

- Para estimar los recursos necesarios por línea (autobuses y conductores), identificar los patrones de carga y obtener el coeficiente de rotación (clasificar)
- Clasificar/asignar una línea nueva (afectada por cambios en la red) a una de las categorías patrón establecidas (discriminación)
- Realizar el seguimiento de la ocupación por líneas (satisfacción de cliente y coste)

## Aplicaciones 4/5

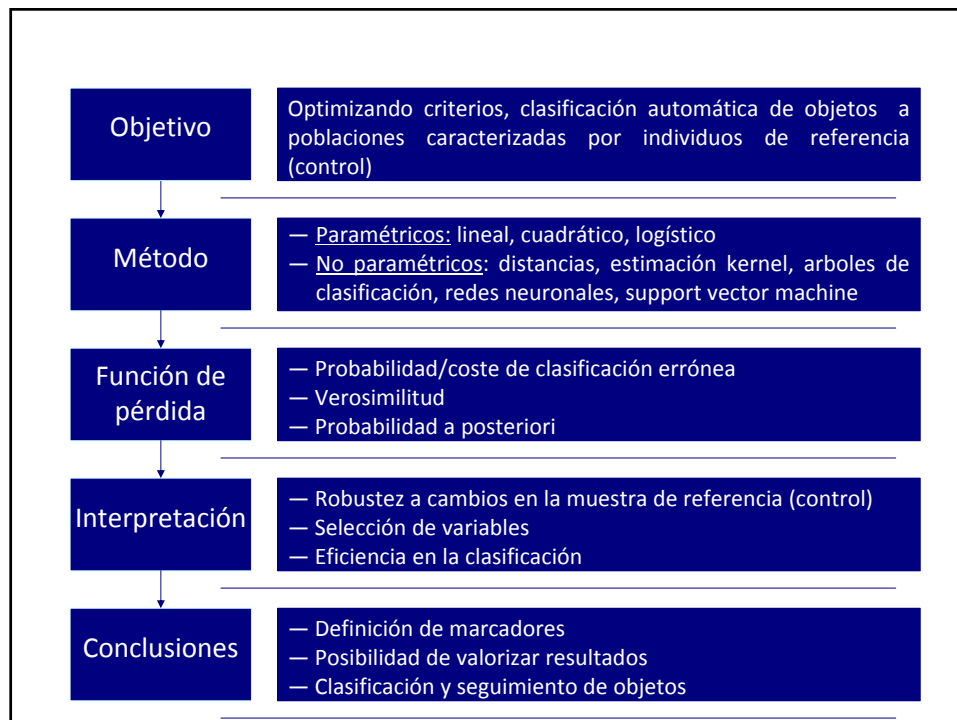
### Gestión de zonas con actividad contaminante potencial

- Identificar y caracterizar los patrones de contaminación por zona (clasificar una muestra de entrenamiento/control)
- Establecer protocolos generales de gestión para las categorías patrón y definir el plan de seguimiento individual
- Clasificar las zonas a una de las categorías establecidas (discriminación)
- Identificar y seleccionar el mínimo conjunto de variables que maximizan la capacidad predictiva (reducir la dimensión)

## Aplicaciones 5/5

### Y en muchos otros ámbitos

- Identificar y caracterizar el **perfil de la dieta** de depredadores para determinar la peligrosidad potencial sobre la especie diana (clasificar y discriminar)
- Identificar y caracterizar **perfiles de riesgo en la contratación de servicios** (financieros, seguros,...) y asignar/clasificar a grupos preestablecidos (clasificación libre o dirigida, y discriminación automática)
- Identificar y caracterizar **perfiles en satisfacción de cliente** (clasificar y discriminar para orientar y valorizar )
- .....



## CLASIFICACIÓN EN DOS POBLACIONES

### Ideas básicas

**Marco.** Disponemos de una población  $\Pi$  en la que tenemos definida una partición en dos categorías,

$$\Pi = H_1 \oplus H_2$$

en las que  $\Pi_1$  y  $\Pi_2$  son las probabilidades a priori de las categorías y  $f_1(x)$ ,  $f_2(x)$  son las funciones de densidad de probabilidad correspondientes a  $(X_1, \dots, X_p)$

**Regla de decisión.** Asignar  $\omega_0$  a la clase  $H_i$  si  $X(\omega_0) \in R_i \subset R^p$

$$R^p = R_1 \oplus R_2$$

optimizando la función objetivo

**Probabilidades de error:**  $P(2/1)$  y  $P(1/2)$ , siendo  $P(j/i)$  la probabilidad de clasificar un objeto a la clase  $j$ , cuando es de la clase  $i$

$$P(j/i) = \int_{R_j} f_i(x) dx$$

## Criterios de optimización 1/2

**Minimizar la probabilidad de clasificación errónea**

$$pce = \pi_1 P(2/1) + \pi_2 P(1/2)$$

**Minimizar el coste total de clasificación**

$$pce = \pi_1 \cdot P(2/1) \cdot C(2/1) + \pi_2 \cdot P(1/2) \cdot C(1/2)$$

siendo  $C(j/i)$  el coste de clasificar un objeto a la clase  $j$ , cuando es de la clase  $i$

**Minimizar la máxima probabilidad de error** (criterio minimax)

$$\max\{P(2/1), P(1/2)\}$$

## Criterios de optimización 2/2

**Maximizar la función de verosimilitud**

$$\frac{f_1(x)}{f_2(x)}$$

**Maximizar la probabilidad a posteriori** (criterio Bayes)

$$\frac{q_1(x_0)}{q_2(x_0)} : \frac{P[\omega_0 \in H_1 / x = x_0]}{P[\omega_0 \in H_2 / x = x_0]} \Leftrightarrow \frac{f_1(x_0)\pi_1}{f_2(x_0)\pi_2}$$

## Probabilidad de clasificación errónea

De aplicar la formula de la probabilidad total,

$$\begin{aligned} pce &= P[X(\omega) \in R_1 / \omega \in H_2] \cdot P[H_2] + P[X(\omega) \in R_2 / \omega \in H_1] \cdot P[H_1] \\ &= \left[ \int_{R_1} f_2(x) dx \right] \cdot \pi_2 + \left[ \int_{R_2} f_1(x) dx \right] \cdot \pi_1 = \pi_1 + \int_{R_1} [\pi_2 \cdot f_2(x) - \pi_1 \cdot f_1(x)] dx \end{aligned}$$

pce se minimiza cuando  $\pi_2 \cdot f_2(x) - \pi_1 \cdot f_1(x) < 0$  para todo  $x \in R_1$

Así,

$$R_1 = \{x \in R^p / \frac{f_1(x)}{f_2(x)} > \frac{\pi_2}{\pi_1}\} \quad \text{y} \quad R_2 = \{x \in R^p / \frac{f_1(x)}{f_2(x)} < \frac{\pi_2}{\pi_1}\}$$

## Coste total de clasificación errónea

De

$$\begin{aligned} c_T &= \pi_2 \cdot C(1/2) \cdot \int_{R_1} f_2(x) dx + \pi_1 \cdot C(2/1) \int_{R_2} f_1(x) dx \\ &= \int_{R_1} [\pi_2 \cdot C(1/2) \cdot f_2(x) - \pi_1 \cdot C(2/1) \cdot f_1(x)] dx \end{aligned}$$

$c_T$  se minimiza cuando  $\pi_2 \cdot C(1/2) \cdot f_2(x) - \pi_1 \cdot C(2/1) \cdot f_1(x) < 0$ ,  $\forall x \in R_1$

Así,

$$R_1 = \{x \in R^p / \frac{f_1(x)}{f_2(x)} > \frac{\pi_2 \cdot C(1/2)}{\pi_1 \cdot C(2/1)}\} \quad \text{y} \quad R_2 = \{x \in R^p / \frac{f_1(x)}{f_2(x)} < \frac{\pi_2 \cdot C(1/2)}{\pi_1 \cdot C(2/1)}\}$$

## Probabilidad a posteriori

Criterio Bayes

$$\frac{q_1(x_0)}{q_2(x_0)} = \frac{P[\omega \in H_1 / x = x_0]}{P[\omega \in H_2 / x = x_0]} = \frac{f_1(x_0)\pi_1 / (f_1(x_0)\pi_1 + f_2(x_0)\pi_2)}{f_2(x_0)\pi_2 / (f_1(x_0)\pi_1 + f_2(x_0)\pi_2)} = \frac{f_1(x_0)\pi_1}{\underline{\underline{f_2(x_0)\pi_2}}}$$

ya que

$$\begin{aligned} q_1(x_0) &= \frac{P(x = x_0 / \omega \in H_1) \cdot P(\omega \in H_1)}{P(x = x_0 / \omega \in H_1) \cdot P(\omega \in H_1) + P(x = x_0 / \omega \in H_2) \cdot P(\omega \in H_2)} \\ &= \frac{f_1(x_0)\pi_1}{(f_1(x_0)\pi_1 + f_2(x_0)\pi_2)} \end{aligned}$$

$$q_2(x_0) = 1 - q_1(x_0)$$

## Denominador común

Las regiones de clasificación

$$R_1 = \{x \in R^p / \frac{f_1(x)}{f_2(x)} > c\} \quad \text{y} \quad R_2 = \{x \in R^p / \frac{f_1(x)}{f_2(x)} < c\}$$

También,

- maximizar la verosimilitud:  $\pi_1 = \pi_2 = 1/2$  y  $c=1$
- minimizar la máxima probabilidad de error



## Discriminador lineal 1/9

Población  $\Pi$  en la que tenemos definida una partición en dos categorías,

$$f_1 \approx N_p(\mu_1, \Sigma) \quad \text{y} \quad f_2 \approx N_p(\mu_2, \Sigma)$$

$$\frac{f_1(x)}{f_2(x)} = \frac{(2\pi)^{-p/2} |\Sigma|^{-1/2} e^{-\frac{1}{2}(x-\mu_1)'\Sigma^{-1}(x-\mu_1)}}{(2\pi)^{-p/2} |\Sigma|^{-1/2} e^{-\frac{1}{2}(x-\mu_2)'\Sigma^{-1}(x-\mu_2)}} > \frac{\pi_2}{\pi_1}$$

$$\Leftrightarrow \ln \frac{f_1(x)}{f_2(x)} = \underbrace{-\frac{1}{2}(x-\mu_1)'\Sigma^{-1}(x-\mu_1) + \frac{1}{2}(x-\mu_2)'\Sigma^{-1}(x-\mu_2)}_{F(x)} > \ln \frac{\pi_2}{\pi_1}$$

## Discriminador lineal 2/9

### Criterio de clasificación

Asignar  $\omega$  a la categoría  $H_1$  si  $F(x) > \ln(\pi_2/\pi_1)$

Asignar  $\omega$  a la categoría  $H_2$  si  $F(x) < \ln(\pi_2/\pi_1)$

Siendo (  $F(x)$  reescrita expresión )

$$F(x) = a'(x - \frac{1}{2}(\mu_1 + \mu_2)) \quad \text{y} \quad a = \Sigma^{-1}(\mu_1 - \mu_2)$$

## Discriminador lineal 3/9

### Parámetros desconocidos: estimación

Medias.  $\hat{\mu}_1 = \bar{X}_1$  y  $\hat{\mu}_2 = \bar{X}_2$

Matriz de varianzas-covarianzas común  $\hat{\Sigma} = \frac{(n_1 - 1) \cdot S_1 + (n_2 - 1) \cdot S_2}{n_1 + n_2 - 2}$

En este caso,

$$F(x) = \hat{a}'(x - \frac{1}{2}(\hat{\mu}_1 + \hat{\mu}_2)) \quad \text{y} \quad \hat{a} = \hat{\Sigma}^{-1}(\hat{\mu}_1 - \hat{\mu}_2)$$

## Discriminador lineal 4/9

**Ejemplo 1.** *Mytilicola intestinalis* es un copépodo parásito del mejillón, que en fase larval presenta diferentes estadios de crecimiento. El primer estadio (nauplios) y el segundo estadio (*Metanauplius*) son difíciles de distinguir. Sobre una muestra de  $n_1 = 76$  y  $n_2 = 91$  copépodos que se pudieron identificar al microscopio como del primer y segundo estadio respectivamente, se midieron las variables  $l$  = longitud,  $a$  = anchura.

La descriptiva de las poblaciones es:

$$\bullet \quad \bar{X}_1 = (219.5, 138.1) \quad \bar{X}_2 = (241.6, 147.8)$$

$$\bullet \quad S_1 = \begin{pmatrix} 409.9 & -1.3 \\ -1.3 & 306.2 \end{pmatrix} \quad S_2 = \begin{pmatrix} 210.9 & 57.9 \\ 57.9 & 152.8 \end{pmatrix}$$

## Discriminador lineal 5/9

### Discriminador

$$F(x) = \hat{a}'(x - \frac{1}{2}(\hat{\mu}_1 + \hat{\mu}_2)) = \left( \begin{pmatrix} 301.4 & 31.0 \\ 31.0 & 222.5 \end{pmatrix}^{-1} \begin{pmatrix} -22.1 \\ -9.7 \end{pmatrix} \right)^t \left( \begin{pmatrix} l \\ a \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 461.1 \\ 285.9 \end{pmatrix} \right)$$

$$F(l, a) = -0.06985 \cdot l - 0.03386 \cdot a + 20.9452$$

### Clasificar el individuo con (l, a)=(221, 143)

$$F(221, 143) = -0.06985 \cdot 221 - 0.03386 \cdot 143 + 20.9452 = 0.7331$$

$$\text{Valor de corte proporcional: } \ln((91/(91+76))/(76/(91+76))) = 0.18$$

**Decisión:** clasificar al grupo nauplios

## Discriminador lineal 6/9

### Probabilidades de error (aparente) y coste de clasificación errónea

- Ámbito metodológico: probabilidad de error o coste con distribuciones exactas o asintóticas del discriminador en cada una de las categorías .

$$P(i/j) = P[F(x) < \ln(\frac{\pi_i}{\pi_j}) / \omega \in H_j]$$

- Ámbito aplicado: estimación por métodos computacionales
  - Resustitución: estimación optimista
  - **Leave-one-out cross-validation**: estimación más realista si la muestra de entrenamiento/control es representativa

## Discriminador lineal 7/9

**Cálculo metodológico:**  $P(i/j) = P[F(x) < \ln(\frac{\pi_i}{\pi_j}) / \omega \in H_j]$

$$F(x) = (\mu_1 - \mu_2)' \Sigma^{-1} x - \frac{1}{2} (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 + \mu_2) \approx N_p(E(F(x)), \text{VAR}(F(X)))$$

$$E(F(x)) = (\mu_1 - \mu_2)' \Sigma^{-1} \mu_i - \frac{1}{2} (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 + \mu_2) = \begin{cases} \frac{1}{2} \Delta^2 & \text{si } \omega \in H_1 \\ -\frac{1}{2} \Delta^2 & \text{si } \omega \in H_2 \end{cases}$$

$$\text{VAR}(F(x)) = (\mu_1 - \mu_2)' \Sigma^{-1} \Sigma \Sigma^{-1} (\mu_1 - \mu_2) = \Delta^2$$

siendo,

$$\Delta^2 = (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2)$$

## Discriminador lineal 8/9

**Probabilidad de clasificación errónea**

$$P(2/1) = P[F(x) < \ln(\frac{\pi_2}{\pi_1}) / \omega \in H_1]$$

$$= P\left[\frac{F(x) - \frac{1}{2} \Delta^2}{\Delta} < \frac{\ln(\pi_2 / \pi_1) - \frac{1}{2} \Delta^2}{\Delta}\right] = \Phi\left(\frac{\ln(\pi_2 / \pi_1) - \frac{1}{2} \Delta^2}{\Delta}\right)$$

y

$$P(1/2) = P\left[\frac{F(x) + \frac{1}{2} \Delta^2}{\Delta} > \frac{\ln(\pi_2 / \pi_1) + \frac{1}{2} \Delta^2}{\Delta}\right] = \Phi\left(\frac{-\ln(\pi_2 / \pi_1) - \frac{1}{2} \Delta^2}{\Delta}\right)$$

## Discriminador lineal 9/9

### **Mytilicola intestinalis** (Ejemplo 1)

- Distancia Mahalanobis  $\Delta^2 = (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2) = 1.8722$
- Probabilidad a priori  $\ln(\pi_2 / \pi_1) = 0.1801$
- Probabilidades de clasificación errónea

$$P(2/1) = \phi\left(\frac{0.1801 - 0.9361}{1.3683}\right) = 0.29$$

$$P(1/2) = \phi\left(\frac{-0.1801 - 0.9361}{1.3683}\right) = 0.21$$

- pce y  $c_T$  aplicando fórmulas

## Aproximación por distancias

### **Método de la distancia**

$$\Pi = H_1 \oplus H_2 \oplus \dots \oplus H_q$$

en la que tenemos definida una distancia  $d_{ij}$

### **Criterio de clasificación**

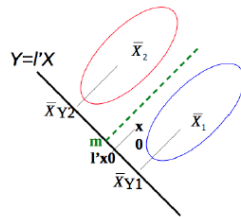
Asignar  $\omega$  a la categoría  $H_1$  si  $d(\omega, H_1) = \min \{d(\omega, H_1), \dots, d(\omega, H_q)\}$

Nota. Con **distancia Mahalanobis** se obtiene el criterio descrito en normalidad multivariante con matrices de varianzas-covarianzas comunes (discriminador lineal)

## Aproximación geométrica

### Eje de máxima variabilidad

$$Y = l'X = (\bar{X}_1 - \bar{X}_2)' S_p^{-1} X \quad \hat{m} = \frac{1}{2} (\bar{X}_1 - \bar{X}_2)' S_p^{-1} (\bar{X}_1 - \bar{X}_2)$$



Let  $x_0$  be an observation, the classification rule consists on,

- Assign  $x_0$  to  $\pi_1$  if  $(\bar{X}_1 - \bar{X}_2)' S_p^{-1} x_0 - \hat{m} \geq 0$
- Assign  $x_0$  to  $\pi_2$  if  $(\bar{X}_1 - \bar{X}_2)' S_p^{-1} x_0 - \hat{m} < 0$

## Discriminador cuadrático 1/3

Población  $\Pi$  en la que tenemos definida una partición en dos categorías,

$$f_1 \approx N_p(\mu_1, \Sigma_1) \quad \text{y} \quad f_2 \approx N_p(\mu_2, \Sigma_2)$$

$$\frac{f_1(x)}{f_2(x)} = \frac{(2\pi)^{-p/2} |\Sigma_1|^{-1/2} e^{-\frac{1}{2}(x-\mu_1)'\Sigma_1^{-1}(x-\mu_1)}}{(2\pi)^{-p/2} |\Sigma_2|^{-1/2} e^{-\frac{1}{2}(x-\mu_2)'\Sigma_2^{-1}(x-\mu_2)}} > \frac{\pi_2}{\pi_1}$$

$\Leftrightarrow$

$$\ln \frac{f_1(x)}{f_2(x)} = \frac{1}{2} \ln \left( \frac{|\Sigma_2|}{|\Sigma_1|} \right) - \frac{1}{2} (x - \mu_1)' \Sigma_1^{-1} (x - \mu_1) + \frac{1}{2} (x - \mu_2)' \Sigma_2^{-1} (x - \mu_2) > \ln \frac{\pi_2}{\pi_1}$$

$\underbrace{\hspace{10em}}_{Q(x)}$

## Discriminador cuadrático 2/3

### Criterio de clasificación

Asignar  $\omega$  a la categoría  $H_1$  si  $Q(x) > \ln(\Pi_2/\Pi_1)$

Asignar  $\omega$  a la categoría  $H_2$  si  $Q(x) < \ln(\Pi_2/\Pi_1)$

siendo

$$Q(x) = \frac{1}{2} \ln\left(\frac{|\Sigma_2|}{|\Sigma_1|}\right) - \frac{1}{2}(x - \mu_1)' \Sigma_1^{-1}(x - \mu_1) + \frac{1}{2}(x - \mu_2)' \Sigma_2^{-1}(x - \mu_2)$$

## Discriminador cuadrático 3/3

### Parámetros desconocidos: estimación

Medias.  $\hat{\mu}_1 = \bar{X}_1$  y  $\hat{\mu}_2 = \bar{X}_2$

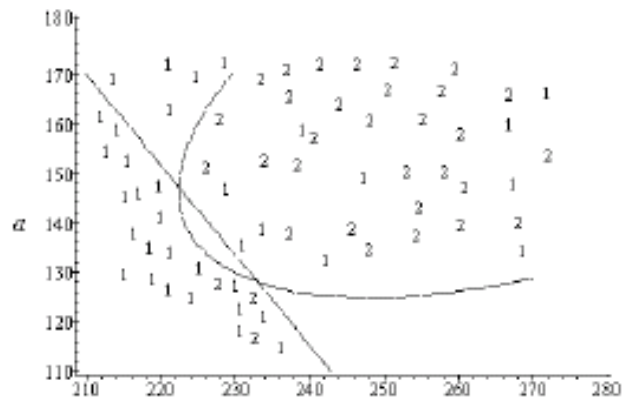
Matriz de varianzas-covarianzas  $\hat{\Sigma}_1 = S_1$  y  $\hat{\Sigma}_2 = S_2$

En este caso,

$$Q(x) = \frac{1}{2} \ln\left(\frac{|\hat{\Sigma}_2|}{|\hat{\Sigma}_1|}\right) - \frac{1}{2}(x - \hat{\mu}_1)' \hat{\Sigma}_1^{-1}(x - \hat{\mu}_1) + \frac{1}{2}(x - \hat{\mu}_2)' \hat{\Sigma}_2^{-1}(x - \hat{\mu}_2)$$

## Interpretación geométrica

### Lineal vs cuadrático



## Discriminador logístico 1/4

Población  $\Pi$  en la que tenemos definida una partición en dos categorías y  $f_1(x)$ ,  $f_2(x)$  son las funciones de densidad de probabilidad correspondientes a  $(X_1, \dots, X_p)$  que satisfacen la condición:

$$\ln \frac{f_1(x)}{f_2(x)} = b_0 + b \cdot x = b_0 + b_1 x_1 + \dots + b_p x_p$$

Aplicabilidad: Variables discretas (multinomial) y combinación de discretas y continuas (normal multivariante o parecidas)

### Criterio de clasificación

Asignar  $\omega$  a la categoría  $H_1$  si  $\ln(f_1(x)/f_2(x)) > \ln(\Pi_2/\Pi_1)$

Asignar  $\omega$  a la categoría  $H_2$  si  $\ln(f_1(x)/f_2(x)) < \ln(\Pi_2/\Pi_1)$



## Discriminador logístico 2/4

### Probabilidades a posteriori

$$q_1(x_0) = P[\omega \in H_1 / x = x_0]$$

$$= \frac{P(x = x_0 / \omega \in H_1) \cdot P(\omega \in H_1)}{P(x = x_0 / \omega \in H_1) \cdot P(\omega \in H_1) + P(x = x_0 / \omega \in H_2) \cdot P(\omega \in H_2)}$$

$$= \frac{f_1(x_0)\pi_1}{(f_1(x_0)\pi_1 + f_2(x_0)\pi_2)} = \frac{\exp\{\ln(\pi_1 / \pi_2) + b_0 + b_1x_1 + \dots + b_px_p\}}{\exp\{\ln(\pi_1 / \pi_2) + b_0 + b_1x_1 + \dots + b_px_p\} + 1}$$

y

$$q_2(x_0) = 1 - q_1(x_0)$$

Asignar  $\omega$  a la población  $H_i$  si  $q_i(x_0) > \frac{1}{2}$  o c de interés (probabilidad a posteriori mayor)

## Rendimiento según c: ejemplo

Rendimiento en predicción - Porcentaje Correcto

C	VERDADERO	FALSO	Total
0,0	100,00	0,00	49,32
0,05	97,26	68,00	82,43
0,1	97,26	73,33	85,14
0,15	97,26	82,67	89,86
0,2	97,26	88,00	92,57
0,25	97,26	92,00	94,59
0,3	95,89	92,00	93,92
0,35	95,89	92,00	93,92
0,4	94,52	94,67	94,59
0,45	94,52	96,00	95,27
0,5	94,52	97,33	95,95
0,55	93,15	98,67	95,95
0,6	93,15	98,67	95,95
0,65	93,15	98,67	95,95
0,7	91,78	98,67	95,27
0,75	90,41	98,67	94,59
0,8	90,41	98,67	94,59
0,85	89,04	98,67	93,92
0,9	80,82	100,00	90,54
0,95	69,86	100,00	85,14
1,0	0,00	100,00	50,68

## Discriminador logístico 3/4

**Parámetros desconocidos: estimación máximo verosímil**

$$L(x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}) = q_1(x_{11}) \cdot \dots \cdot q_1(x_{1n_1}) \cdot q_2(x_{21}) \cdot \dots \cdot q_2(x_{2n_2})$$

↓

$$\frac{\partial \ln L(x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2})}{\partial b_j} = \sum_{j=1}^{n_1} (1 - q_1(x_{1j})) \cdot x_{1j} - \sum_{j=1}^{n_2} q_2(x_{1j}) \cdot x_{2j} = 0$$

$b_0, b_1, \dots, b_p$  son solución del sistema de ecuaciones (resolución por métodos de aproximación computacional)

## Discriminador logístico 4/4

### Aplicación

- Distribuciones normales multivariantes con matriz de varianzas-covarianzas común
- Combinación multivariante de variables dicotómicas independientes
- Distribución multinomial y distribuciones multivariantes discretas que siguen el modelo log-lineal
- Mixturas de distribuciones continuas y discretas de las formas anteriores  $(X^{(1)}, X^{(2)})$  con  $X^{(1)} \sim N_p$  y  $X^{(2)}$  q-vector de variables binarias independientes

## CLASIFICACIÓN EN MÁS DE DOS POBLACIONES

### Generalización de conceptos

$$\Pi = H_1 \oplus H_2 \oplus \dots \oplus H_q$$

#### Criterios de optimización

- Probabilidad de clasificación errónea (generalizada)
- Coste total de clasificación errónea (generalizada)
- Minimizar la máxima probabilidad de error (generalizada)
- Maximizar la probabilidad a posteriori (generalizada)

#### Criterio de clasificación (generalizado)

Asignar  $\omega$  a la categoría  $H_1$  si  $\Pi_1 f_1(x) = \max \{ \Pi_1 f_1(x), \dots, \Pi_q f_q(x) \}$