Análisis Multivariado

ID 033521 - Clase 4901

Lina Maria Acosta Avena

Ciencia de Datos
Departamento de Matemáticas
Pontificia Universidad Javeriana

Semana 13: 14/10/24 – 20/10/24

Introducción

Suponga que se quiere **identificar** el perfil de portadores y no portadores de una determinada enfermedad, con el **objetivo de clasificar a los nuevos pacientes** como portadores probables o improbables de la patología en cuestión.

Introducción

Suponga que se quiere **identificar** el perfil de portadores y no portadores de una determinada enfermedad, con el **objetivo de clasificar a los nuevos pacientes** como portadores probables o improbables de la patología en cuestión.

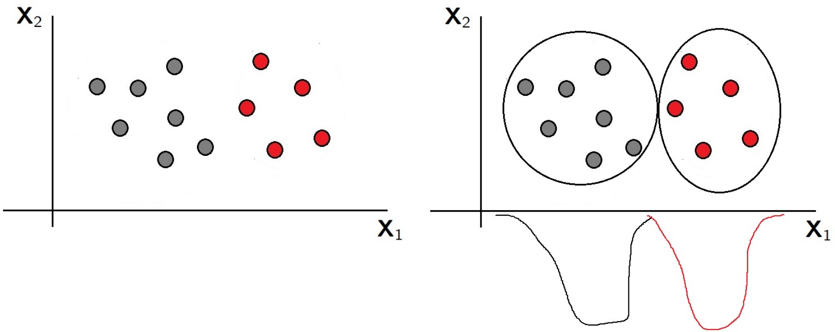
Observe que **se tienen dos grupos previamente definidos, sobre los cuales se clasificará a los nuevos pacientes**.

Introducción

Suponga que se quiere **identificar** el perfil de portadores y no portadores de una determinada enfermedad, con el **objetivo de clasificar a los nuevos pacientes** como portadores probables o improbables de la patología en cuestión.

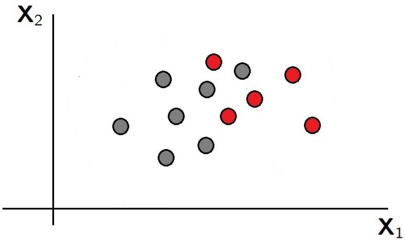
Observe que **se tienen dos grupos previamente definidos, sobre los cuales se clasificará a los nuevos pacientes**. En algunas situaciones **éstos pueden estar bien definidos**, sin embargo, en la práctica ésto no sucede, y en consecuencia habrá mayores probabilidades de clasificar erroneamente a los nuevos pacientes.

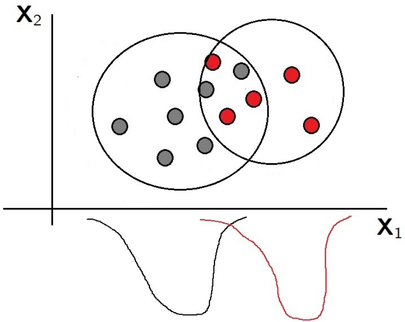
Introducción



Observe que los **grupos** están **bien definidos**, **separados/clasificados** (**no hay intersección** entre los dos grupos - son excluyentes), por lo cuál la **probabilidad de clasificar erroneamente a nuevas unidades investigación/muestrales es casi nula**.

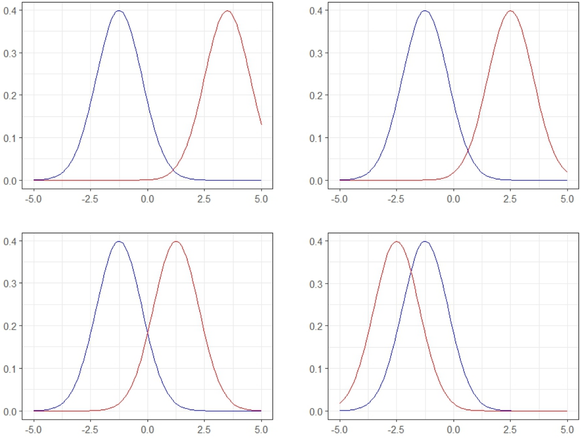
Introducción





En este caso, se observa una **"pequeña" intersección** entre los dos grupos, la cual **aumentaría** (en comparación al caso anterior) la **probabilidad de clasificar erróneamente a las nuevas unidades de investigación/muestrales**.

Introducción



A mayor interesección, mayor es la probabilidad de clasificación errónea de nuevas unidades de investigación/muestreo.

Introducción

Otras situaciones:

- Los profesionales en **botánica** suelen estar interesados en **clasificar** a una **"nueva" planta** en alguna de las **especies conocidas**.
- La **medicina forense** es un área donde el interés suele ser **clasificar individuos** (unidades de investigación o de muestreo) en dos grupos predefinidos: masculino y femenino, con base en algunas medidas determinadas en sus huesos ($X's$).

Introducción

Otras situaciones:

- Los profesionales en **botánica** suelen estar interesados en **clasificar** a una **“nueva” planta** en alguna de las **especies conocidas**.
- La **medicina forense** es un área donde el interés suele ser **clasificar individuos** (unidades de investigación o de muestreo) en dos grupos predefinidos: masculino y femenino, con base en algunas medidas determinadas en sus huesos ($X's$).

El **Análisis Discriminante (AD)** ***(Discriminant Analysis – DA)*** es una técnica que establece **funciones o procedimientos** que **discriminan los dos o más grupos predefinidos, que proporcione pocas clasificaciones incorrectas**.

Introducción

Otras situaciones:

- Los profesionales en **botánica** suelen estar interesados en **clasificar** a una **"nueva" planta** en alguna de las **especies conocidas**.
- La **medicina forense** es un área donde el interés suele ser **clasificar individuos** (unidades de investigación o de muestreo) en dos grupos predefinidos: masculino y femenino, con base en algunas medidas determinadas en sus huesos ($X's$).

El **Análisis Discriminante (AD)** ***(Discriminant Analysis – DA)*** es una técnica que establece **funciones o procedimientos** que **discriminan los dos o más grupos predefinidos, que proporcione pocas clasificaciones incorrectas**. Generalmente, estas funciones son combinaciones lineales de las variables ($X's$).

Introducción

Notas:

- ✓ *Sir Ronald Fisher (1936) propuso el Análisis Discriminante (lineal) para dos grupos/poblaciones*.
- ✓ *C. R. Rao (1948) propuso una generalización (varios grupos) del Análisis Discriminante*.
- ✓ *El Analisis Discriminante es una técnica de aprendizaje supervisado*.

Objetivos del Análisis Discriminante

Como vimos, en el Análisis Discriminante los objetivos son:

1. **Clasificar, separar o discriminar grupos**.
   Aquí se trata de encontrar las diferencias entre dos o más grupos a través de una función discriminante, que generalmente es una combinación lineal de las variables, así que, el problema de la discriminación es comprobar si esas variables permiten diferenciar los grupos predefinidos.

2. **Predecir o asignar unidades muestrales en alguno de los grupos predefenidos**.
   Aquí la asignación de las unidades muestrales se hace de acuerdo con una regla de clasificación o de localización, la cuál suele ser la misma de discriminación.

Reglas para Discriminación de $g = 2$ Grupos

¿Cómo construimos una regla de clasificación que minimice la chance de clasificar incorrectamente a las unidades muestrales?

Reglas para Discriminación de $g = 2$ Grupos

Si la distribución de probabilidad de las $X_i's$ de cada grupo es conocida, podemos utilizar **el principio de verosimilitud**. Particularmente, suponga el **caso univariado** y que se tienen **dos grupos**:

$$\text{Grupo 1}: X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$$
$$\text{Grupo 2}: X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$$

Así, para cada x posible de una unidad de investigación podemos calcular la razón de densidades de x en los grupos 1 y 2:

$$\lambda(x) = \frac{f_{X_1}(x)}{f_{X_2}(x)} = \frac{(2\pi\sigma_1^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right\}}{(2\pi\sigma_2^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_2^2}(x-\mu_2)^2\right\}}$$

Reglas para Discriminación de $g = 2$ Grupos

$$\lambda(x) = \left[\frac{\sigma_2^2}{\sigma_1^2}\right]^{1/2} \exp\left\{-\frac{1}{2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \left(\frac{x-\mu_2}{\sigma_2}\right)^2\right]\right\} \qquad (1)$$

Así que

i. Si $\lambda(x) > 1$:

$$\frac{f_{X_1}(x)}{f_{X_2}(x)} > 1 \Longleftrightarrow f_{X_1}(x) > f_{X_2}(x)$$

entonces es razonable clasificar a la unidad de investigación (x) en el Grupo 1.

Reglas para Discriminación de $g = 2$ Grupos

ii. Si $\lambda(x) < 1$:

$$\frac{f_{X_1}(x)}{f_{X_2}(x)} < 1 \Longleftrightarrow f_{X_1}(x) < f_{X_2}(x)$$

entonces es razonable clasificar a la unidad de investigación (x) en el Grupo 2.

iii. Si $\lambda(x) = 1$:

$$\frac{f_{X_1}(x)}{f_{X_2}(x)} = 1 \Longleftrightarrow f_{X_1}(x) = f_{X_2}(x)$$

entonces **la probabilidad de x estar en cualquiera de los grupos es la misma**.

Reglas para Discriminación de $g = 2$ Grupos

Por lo tanto, $\lambda(x)$ dado en la **ecuación** (1) es nuestra **función discriminante**.

Note que apartir de (1) tenemos:

$$-2\ln[\lambda(x)] = -\ln\left[\frac{\sigma_2^2}{\sigma_1^2}\right] + \left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \left(\frac{x-\mu_2}{\sigma_2}\right)^2\right] \quad (2)$$

en ese caso

- Si $\lambda(x) > 1 \Longrightarrow -2\ln[\lambda(x)] < 0$, por lo cual x está más cerca de μ_1 y sería clasificado en el Grupo 1.

Reglas para Discriminación de Grupos

- Si $\lambda(x) < 1 \Longrightarrow -2\ln[\lambda(x)] > 0$, por lo cual x está más cerca de μ_2 y sería clasificado en el Grupo 2.
- Si $\lambda(x) = 1 \Longrightarrow -2\ln[\lambda(x)] = 0$, así que x estaría igualmente cerca de μ_1 que de μ_2.

Así que,

$$-2\ln[\lambda(x)]$$

dado en la **ecuación** (2) también es llamada **función discriminante**.

Reglas para Discriminación de Grupos

Resumiendo

$\lambda(x)$	$-2\ln[\lambda(x)]$	**Decisión**
>1	<0	$\mathbf{x}$ está más cerca de μ_1 (Grupo 1)
<1	>0	$\mathbf{x}$ está más cerca de μ_2 (Grupo 2)
$=1$	$=0$	$\mathbf{x}$ está igualmente cerca de μ_1, μ_2

Reglas para Discriminación de Grupos

Resumiendo

$\lambda(x)$	$-2\ln[\lambda(x)]$	**Decisión**
> 1	< 0	$\mathbf{x}$ está más cerca de μ_1 (Grupo 1)
< 1	> 0	$\mathbf{x}$ está más cerca de μ_2 (Grupo 2)
$= 1$	$= 0$	$\mathbf{x}$ está igualmente cerca de μ_1, μ_2

Podemos extender al caso multivariado!

Reglas de Discriminación para $g = 2$ Grupos

Suponga que se tienen dos poblaciones multivariadas:

$$\text{Grupo 1}: \ \mathbf{X}_1 \sim \mathrm{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$$
$$\text{Grupo 2}: \ \mathbf{X}_2 \sim \mathrm{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$$

luego,

$$\lambda(\mathbf{x}) = \frac{(2\pi)^{p/2}|\boldsymbol{\Sigma}_1|^{-1/2}\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_1)^\top\boldsymbol{\Sigma}_1^{-1}(\mathbf{x}-\boldsymbol{\mu}_1)\right\}}{(2\pi)^{p/2}|\boldsymbol{\Sigma}_2|^{-1/2}\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_2)^\top\boldsymbol{\Sigma}_2^{-1}(\mathbf{x}-\boldsymbol{\mu}_2)\right\}}$$

y

$$\begin{aligned}-2\ln\left[\lambda(\mathbf{x})\right] = {} & (\mathbf{x}-\boldsymbol{\mu}_1)^\top\boldsymbol{\Sigma}_1^{-1}(\mathbf{x}-\boldsymbol{\mu}_1) - (\mathbf{x}-\boldsymbol{\mu}_2)^\top\boldsymbol{\Sigma}_2^{-1}(\mathbf{x}-\boldsymbol{\mu}_2) \\ & + \ln\left(|\boldsymbol{\Sigma}_1|\right) - \ln\left(|\boldsymbol{\Sigma}_2|\right)\end{aligned} \quad (3)$$

Reglas de Discriminación para $g = 2$ Grupos

Así que

- Si $-2\ln[\lambda(\mathbf{x})] < 0$, $\mathbf{x}$ es clasificado en el Grupo 1.
- Si $-2\ln[\lambda(\mathbf{x})] > 0$, $\mathbf{x}$ es clasificado en el Grupo 2.
- Si $-2\ln[\lambda(\mathbf{x})] = 0$, $\mathbf{x}$ puede ser clasificado en el Grupo 1 o 2.

Observación: La expresión dada en la ***ecuación (3)*** *es llamada* ***Función Discriminante Cuadrática****.*

Particularmente, si en la ecuación (3)

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$$

Reglas de Discriminación para $g = 2$ Grupos

sigue que

$$-2\ln[\lambda(\mathbf{x})] = (\mathbf{x}-\boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_1) - (\mathbf{x}-\boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_2)$$

o equivalentemente la **Función Discriminante de Fisher**

$$fd(\mathbf{x}) = \underbrace{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}^{-1}}_{\mathbf{b}^\top}\mathbf{x} - \frac{1}{2}\underbrace{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}^{-1}}_{\mathbf{b}^\top}(\boldsymbol{\mu}_1+\boldsymbol{\mu}_2)$$

$$= \mathbf{b}^\top \mathbf{x} - \underbrace{\frac{1}{2}\mathbf{b}^\top(\boldsymbol{\mu}_1+\boldsymbol{\mu}_2)}_{\mathrm{C}}$$

Reglas de Discriminación para $g = 2$ Grupos

donde

$$\mathbf{b}^{\top} = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^{\top}\boldsymbol{\Sigma}^{-1} \qquad ; \qquad \mathsf{C} = \frac{1}{2}\mathbf{b}^{\top}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)$$

es decir, se construye una función lineal que separa los dos grupos (a partir de un punto pertenece a uno y apartir de otro pertenece a otro):

- Si $fd(\mathbf{x}) > 0$

$$\mathbf{b}^{\top}\mathbf{x} > \underbrace{\frac{1}{2}\mathbf{b}^{\top}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)}_{C}$$

una unidad muestral con observaciones $\mathbf{x}$ es clasificado en el Grupo 1.

Reglas de Discriminación para $g = 2$ Grupos

- Si $fd(\mathbf{x}) < 0$

$$\mathbf{b}^\top \mathbf{x} < \underbrace{\frac{1}{2}\mathbf{b}^\top(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)}_{c}$$

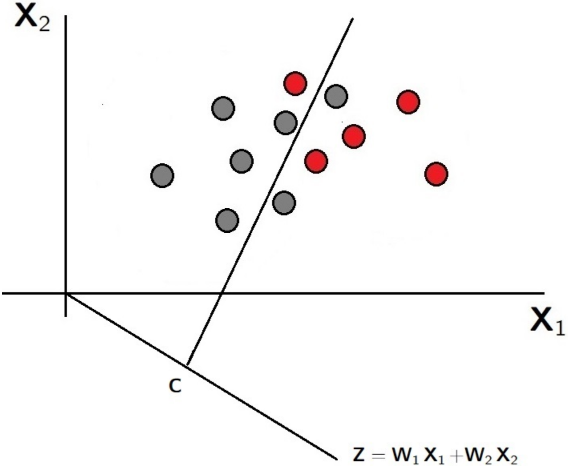
  una unidad muestral con observaciones $\mathbf{x}$ es clasificado en el Grupo 2.

- Si $fd(\mathbf{x}) = 0$

$$\mathbf{b}^\top \mathbf{x} = \underbrace{\frac{1}{2}\mathbf{b}^\top(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)}_{c}$$

  una unidad muestral con observaciones $\mathbf{x}$ puede ser clasificado en cualquiera de los dos grupos.

Reglas de Discriminación para $g = 2$ Grupos



Reglas de Discriminación para $g = 2$ Grupos

Sabemos que

$$\boldsymbol{\mu}_1,\ \boldsymbol{\mu}_2,\ \boldsymbol{\Sigma}_1,\ \boldsymbol{\Sigma}_2$$

son **desconocidos**, así que usamos sus correspondientes **estimaciones**

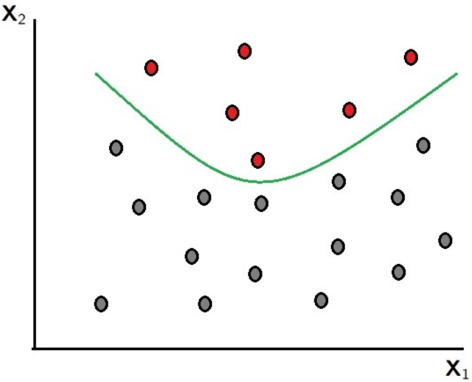
$$\overline{\mathbf{x}}_1,\ \overline{\mathbf{x}}_2,\ \mathbf{S}_1,\ \mathbf{S}_2$$

y por lo tanto, la **función de discriminante cuadrática** queda dada por

$$\begin{aligned} -2\ln\left[\widehat{\lambda}(\mathbf{x})\right] &= (\mathbf{x}-\overline{\mathbf{x}}_1)^\top \mathbf{S}_1^{-1}(\mathbf{x}-\overline{\mathbf{x}}_1) - (\mathbf{x}-\overline{\mathbf{x}}_2)^\top \mathbf{S}_2^{-1}(\mathbf{x}-\overline{\mathbf{x}}_2) \\ &\quad + \ln(\mathbf{S}_1) - \ln(\mathbf{S}_2) \end{aligned} \tag{4}$$

Reglas de Discriminación para $g = 2$ Grupos

Hay situaciones donde es necesaria una función **Discriminante Cuadrática**



Reglas de Discriminación para $g = 2$ Grupos

Example

Considere que el proceso de selección en una escuela está compuesto por **dos fases**: reprobados (aprobaron sólo la fase 1) y aprobados (aprobaron ambas fases). Así que se tienen **dos grupos** de estudiantes:

$$\begin{aligned} \text{Grupo 1} &: \text{ estudiantes que sólo aprobaron la fase 1} \\ \text{Grupo 2} &: \text{ estudiantes que aprobaron } \underline{\text{ambas fases}} \end{aligned}$$

Considere que se tienen las notas de matemáticas (x_1) y lengua extranjera (x_2):

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 16 \\ 15 \end{bmatrix} \qquad \boldsymbol{\mu}_2 = \begin{bmatrix} 18 \\ 16 \end{bmatrix} \qquad \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma} = \begin{bmatrix} 4 & 5 \\ 5 & 9 \end{bmatrix}$$

Reglas de Discriminación para $g = 2$ Grupos

Luego,

$$\mathbf{b}^\top = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}^{-1} = [-2 \quad 1]^\top \begin{bmatrix} 4 & 5 \\ 5 & 9 \end{bmatrix}^{-1} = [-1.182 \quad 0.545]^\top$$

y por lo tanto, la **Función de Fisher** queda dada por:

$$\mathbf{b}^\top \mathbf{x} = [-1.182 \quad 0.545]^\top \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= -1.182x_1 + 0.545x_2$$

Note que cuanto **más grande** sea el valor de x_1 y **más pequeño** sea el valor de x_2, $\mathbf{b}^\top$ **x** será **más negativa**.

Reglas de Discriminación para $g = 2$ Grupos

La constante que va delimitar la región de clasificación es

$$C = \frac{1}{2}\mathbf{b}^{\top}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)$$

$$= \frac{1}{2}[-1.182 \quad 0.545]^{\top} \begin{bmatrix} 34 \\ 31 \end{bmatrix}$$

$$= -11.65$$

Reglas de Discriminación para $g = 2$ Grupos

Suponga que un estudiante obtuvo en matemática $x_1 = 18.5$ y en lengua extranjera $x_2 = 17.5$, entonces su **score** en la función discriminante es:

$$-1.182(18.5) + 0.545(17.5) = -12.33$$

Observe que

$$-12.33 < -11.65$$

así que ese estudiante será ubicado en el Grupo 2 (aprobado) .

Reglas de Discriminación para $g = 2$ Grupos

Example

Un grupo de **49 personas de edad avanzada** que participaron en un estudio fueron **clasificadas** mediante una **evaluación psiquiátrica** en una de las siguientes categorías: **senil o no senil**. Se realizó una prueba de inteligencia adulta a cada uno de los participantes del estudio, la cuál reveló grandes **diferencias entre los dos grupos en algunas partes de la prueba**. En consecuencia, los investigadores decidieron considerar algunas partes de la prueba (subpruebas) con el el fin de considerar una regla de decisión. Las mediciones de las subpruebas son:

Reglas de Discriminación para $g = 2$ Grupos

Variable	**Subprueba**	**No Senil** $(n_1 = 37)$	**Senil** $(n_2 = 12)$
X_1	Información	12.57	8.75
X_2	Similaridades	9.57	5.33
X_3	Aritmética	11.49	8.50
X_4	Hab. Artística	7.97	4.75

Asuma que los datos de cada grupo sigue una distribución normal 4-variada con $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$. Se tiene que:

$$\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2 = \begin{bmatrix} 3.82 \\ 4.24 \\ 2.99 \\ 3.22 \end{bmatrix} \qquad \mathbf{S} = \begin{bmatrix} 11.2553 & 9.4042 & 7.1489 & 3.3830 \\ 9.4042 & 13.5318 & 7.3830 & 2.5532 \\ 7.1489 & 7.3830 & 11.5744 & 2.6170 \\ 3.3830 & 2.5532 & 2.6170 & 5.8085 \end{bmatrix}$$

Pontificia Universidad JAVERIANA Bogotá

Luego,

$$
\begin{aligned}
\mathbf{b}^\top &= (\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2)^\top \mathbf{S}^{-1} \\
&= \begin{bmatrix} 3.82 & 4.24 & 2.99 & 3.22 \end{bmatrix} \begin{bmatrix} 11.2553 & 9.4042 & 7.1489 & 3.3830 \\ 9.4042 & 13.5318 & 7.3830 & 2.5532 \\ 7.1489 & 7.3830 & 11.5744 & 2.6170 \\ 3.3830 & 2.5532 & 2.6170 & 5.8085 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 0.03 & 0.2 & 0.01 & 0.44 \end{bmatrix}
\end{aligned}
$$

y por lo tanto, la **función de Fisher** queda dada por:

$$
\begin{aligned}
\mathbf{b}^\top \mathbf{x} &= \begin{bmatrix} 0.03 & 0.2 & 0.01 & 0.44 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} \\
&= 0.03X_1 + 0.2X_2 + 0.01X_3 + 0.44X_4
\end{aligned}
$$

Reglas de Discriminación para $g = 2$ Grupos

La constante que va delimitar la región de clasificación es

$$C = \frac{1}{2}\mathbf{b}^\top(\overline{\mathbf{X}}_1 + \overline{\mathbf{X}}_2)$$

$$= \frac{1}{2}\begin{bmatrix} 0.03 & 0.2 & 0.01 & 0.44 \end{bmatrix}\begin{bmatrix} 10.66 \\ 7.45 \\ 9.99 \\ 6.36 \end{bmatrix}$$

$$= 4.7081$$

Reglas de Discriminación para $g = 2$ Grupos

Por lo tanto,

- Si
$$\mathbf{b}^\top \mathbf{x} > 4.7081$$
el individuo se asigna en el grupo no senil (Grupo 1)
- Si
$$\mathbf{b}^\top \mathbf{x} < 4.7081$$
se asigna al grupo senil (Grupo 2)

Considere que un individuo obtuvo los puntajes $[10 \quad 8 \quad 7 \quad 5]$, así que

$$0.03(10) + 0.2(8) + 0.01(7) + 0.44(5) = 4.17$$

Reglas de Discriminación para $g = 2$ Grupos

Observe que

$$4.17 < 4.7081$$

así que ese individuo debe ser considerado en el grupo senil.

Observaciones: Aunque la función discrinante se asemeja al análisis de regresión lineal, esisten varias diferencias:

- *En el análisis de regresión lineal se asume que la variable dependiente (Y) sigue una distribución normal y las variables independientes ($X's$) son consideradas fijas. En cambio, en el análisis discriminante, las $X's$ se asumen normales e Y asume valores 0 y 1 (dos grupos) es fija.*

Reglas de Discriminación para $g = 2$ Grupos

- *El objetivo en el análisis de regresión es predecir* $\mathrm{E}[Y \mid \boldsymbol{x}]$, *mientras que en el análisis discriminante se pretende encontrar una combinación de las* X'*s que minimicen la probabilidad de clasificar incorrectamente unidades muestrales en sus respectivos grupos.*

Reglas de Discriminación para $g = 2$ Grupos

Para construir la función discriminante usamos el **principio de verosimilitud** y la función quedaba dada por

$$\lambda(x) = \frac{f_{X_1}(x)}{f_{X_2}(x)}$$

En ese caso, se asumió que **los grupos eran igualmente probables**.

Reglas de Discriminación para $g = 2$ Grupos

Para construir la función discriminante usamos el **principio de verosimilitud** y la función quedaba dada por

$$\lambda(x) = \frac{f_{X_1}(x)}{f_{X_2}(x)}$$

En ese caso, se asumió que **los grupos eran igualmente probables**. Sin embargo, en algunas situaciones se pueden considerar **probabilidades apriori para los grupos**. Por ejemplo, apartir de un diagnóstico clínico, para **algún grupo de personas**, se puede considerar que la **gripa tiene mayor probabilidad de ocurrencia que la polio**.

Reglas de Discriminación para $g = 2$ Grupos

Para construir la función discriminante usamos el **principio de verosimilitud** y la función quedaba dada por

$$\lambda(x) = \frac{f_{X_1}(x)}{f_{X_2}(x)}$$

En ese caso, se asumió que **los grupos eran igualmente probables**. Sin embargo, en algunas situaciones se pueden considerar **probabilidades apriori para los grupos**. Por ejemplo, apartir de un diagnóstico clínico, para **algún grupo de personas**, se puede considerar que la **gripa tiene mayor probabilidad de ocurrencia que la polio**. En estos casos usamos la regla de **discriminación de Bayes**.

Reglas de Discriminación para $g = 2$ Grupos

Suponga que se tienen dos grupos:

$$\text{Grupo 1} : f_{X_1}(x)$$
$$\text{Grupo 2} : f_{X_2}(x)$$

además que

$$p_1 : \text{Prob. de que } x \text{ provenga del Grupo1}$$
$$p_2 : \text{Prob. de que } x \text{ provenga del Grupo 2}$$

Es natural pensar en

$$\begin{cases} \text{Asignar x en el Grupo 1} & , \text{si } p_1 > p_2 \\ \\ \text{Asignar x en el Grupo 2} & , \text{si } p_1 < p_2 \end{cases}$$

Reglas de Discriminación para $g = 2$ Grupos

La regla de Bayes, localiza x en el grupo con más alta probabilidad condicional. Por ejemplo, la probabilidad de asignar x al grupo 1 (G_1):

$$P[G_1|x] = \frac{P[G_1 \cap x]}{P[x]}$$

$$= \frac{P[G_1]\,P[x|G_1]}{P[G_1]\,P[x|G_1] + P[G_2]\,P[x|G_2]}$$

$$= \frac{p_1 f_{X_1}(x)}{p_1 f_{X_1}(x) + p_2 f_{X_2}(x)}$$

Análogamente tenemos que

$$P[G_2|x] = \frac{p_2 f_{X_2}(x)}{p_1 f_{X_1}(x) + p_2 f_{X_2}(x)}$$

Reglas de Discriminación para $g = 2$ Grupos

así que se

$$\begin{cases} \text{Asigna x al Grupo 1} & , \text{si } p_1 f_{X_1}(x) > p_2 f_{X_2}(x) \\ \\ \text{Asignar x al Grupo 2} & , \text{si } p_1 f_{X_1}(x) < p_2 f_{X_2}(x) \end{cases}$$

Reglas de Discriminación para $g = 2$ Grupos

así que se

$$\begin{cases} \text{Asigna x al Grupo 1} & , \text{si } p_1 f_{X_1}(x) > p_2 f_{X_2}(x) \\ \\ \text{Asignar x al Grupo 2} & , \text{si } p_1 f_{X_1}(x) < p_2 f_{X_2}(x) \end{cases}$$

Cuando las **variables son discretas o son una mezcla de discretas y continuas**, se recomienda que la discriminación se haga a través del **modelo logístico**, por lo que la técnica es llamada **Discriminación Logística**.

Reglas de Discriminación para $g = 2$ Grupos

Suponga que se tienen dos grupos:

$$\text{Grupo 1}: \ \mathbf{X}_1 \sim \mathrm{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$$
$$\text{Grupo 2}: \ \mathbf{X}_2 \sim \mathrm{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$$

Cuando $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$, la función lineal del vector observado $\mathbf{x}$ queda dada por:

$$\begin{aligned}
\ln\left[\frac{f_{\mathbf{X}_1}(\mathbf{x})}{f_{\mathbf{X}_2}(\mathbf{x})}\right] &= \underbrace{-\frac{1}{2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)}_{\alpha} + \underbrace{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}^{-1}}_{\boldsymbol{\beta}^\top}\mathbf{x} \\
&= \alpha + \boldsymbol{\beta}^\top \mathbf{x}
\end{aligned}$$

el cuál se conoce como **modelo logístico**.

Reglas de Discriminación para $g = 2$ Grupos

La regla para asignar $\mathbf{x}$ queda dada por:

$$\begin{cases} \text{Asignar } \mathbf{x} \text{ al Grupo 1} & , \text{si } \alpha + \boldsymbol{\beta}^\top \mathbf{x} > \ln\left[\frac{p_1}{p_2}\right] \\ \\ \text{Asignar } \mathbf{x} \text{ al Grupo 2} & , \text{si } \alpha + \boldsymbol{\beta}^\top \mathbf{x} < \ln\left[\frac{p_1}{p_2}\right] \end{cases}$$

donde p_1 y p_2 son las probabilidades apriori de los grupos 1 y 2, respectivamente. Si $p_1 = p_2$,

$$\ln\left[\frac{p_1}{p_2}\right] = 0$$

Reglas de Discriminación para $g = 2$ Grupos

Para $\mathbf{x}$ dado, la probabilidad de que pertenezca al Grupo 1 (G_1) es

$$P\left[G_1 | \mathbf{x}\right] = \frac{p_1 f(\mathbf{x} | G_1)}{p_1 f(\mathbf{x} | G_1) + p_2 f(\mathbf{x} | G_2)}$$

$$= \frac{\exp\left\{\ln\left[\frac{p_1}{p_2}\right] + \alpha + \boldsymbol{\beta}^\top \mathbf{x}\right\}}{1 + \exp\left\{\ln\left[\frac{p_1}{p_2}\right] + \alpha + \boldsymbol{\beta}^\top \mathbf{x}\right\}}$$

$$= \frac{\exp\left\{\alpha_0 + \boldsymbol{\beta}^\top \mathbf{x}\right\}}{1 + \exp\left\{\alpha_0 + \boldsymbol{\beta}^\top \mathbf{x}\right\}}$$

Reglas de Discriminación para $g = 2$ Grupos

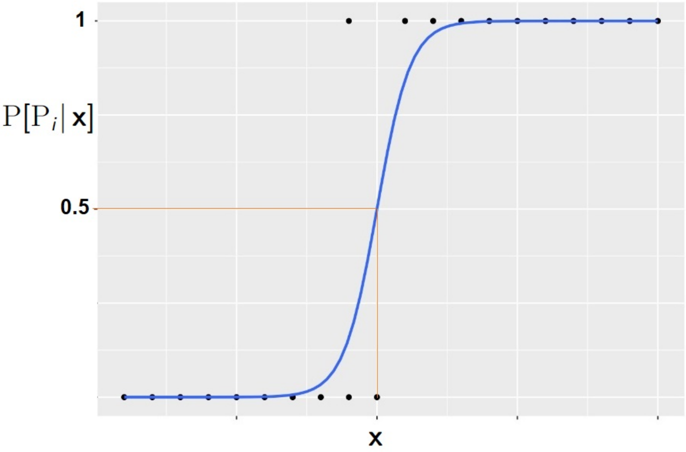
$$P\left[G_1 | \mathbf{x}\right] = \frac{1}{1 + \exp\left\{-(\alpha_0 + \boldsymbol{\beta}^\top \mathbf{x})\right\}} \qquad \alpha_0 = \ln\left[\frac{p_1}{p_2}\right] + \alpha$$

de ahí

$$P\left[G_2 | \mathbf{x}\right] = 1 - P\left[G_1 | \mathbf{x}\right] = \frac{1}{1 + \exp\left\{\alpha_0 + \boldsymbol{\beta}^\top \mathbf{x}\right\}}$$

La **estimación** de α y $\boldsymbol{\beta}$ se hace por **máxima verosimilitud** o **mí-mimos cuadrados ponderado**, sin embargo, éstas son aproximadas (usando **métodos numéricos** como Newton Raphson) desde que los **sistemas de ecuaciones** resultantes de éstos métodos, son **no linea-les**.

Reglas de Discriminación para $g = 2$ Grupos



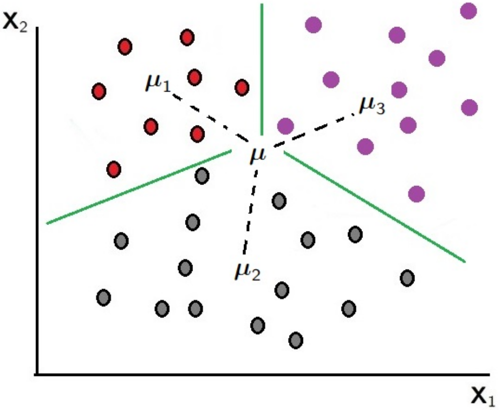
Se asigna $\mathbf{x}$ en la población i, si $P[\mathrm{Pob}_i|\,\mathbf{x}] \geq 0.5 \; i = 1, 2.$

Reglas de Discriminación para $g = 2$ Grupos

Existen otras técnicas de discriminación:

- Discriminación probit.
- Clasificación por medio del vecino más cercano.
- Clasificación mediante redes neuronales (ver Johnson and Wichern (2013) - Applied Multivariate Statistical Analysis, pp. 647) .

Reglas de Discriminación para g Grupos



Reglas de Discriminación para g Grupos

Considere que se observaron $\mathbf{X}_{p\times 1}$ variables en las unidades muestrales de g grupos diferentes, con vector de medias poblacionales $\boldsymbol{\mu_i}$ y matriz de varianzas y covarianzas poblacionales $\mathbf{\Sigma}_i$, $i = 1, 2, \ldots, g$. También considere que no es posible garantizar el supuesto de normalidad en los datos, pero que es razonable asumir que

$$\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \cdots = \mathbf{\Sigma}_g = \mathbf{\Sigma}$$

Sea

- El vector de medias de las medias poblacionales: $\overline{\boldsymbol{\mu}}$.
- La **suma de productos cruzados entre los grupos**:

$$\mathrm{B} = \sum_{i=1}^{g}(\boldsymbol{\mu}_i - \overline{\boldsymbol{\mu}})(\boldsymbol{\mu}_i - \overline{\boldsymbol{\mu}})^\top$$

Reglas de Discriminación para g Grupos

- La combinación lineal: $\mathbf{Y} = \mathbf{a}^\top \mathbf{X}$.
- La variable indicadora

$$\pi_i = \begin{cases} 1 & \text{, la observación pertenece al grupo } i \\ 0 & \text{, en otro caso} \end{cases}$$

i=1,2,. . . , g. Luego

$$\mu_{i,Y} = \mathsf{E}\left[\mathbf{Y} \,|\pi_i\right] = \mathbf{a}^\top \mathsf{E}\left[\mathbf{X} \,|\pi_i\right] = \mathbf{a}^\top \boldsymbol{\mu}_i$$
$$\mathsf{Var}\left[\mathbf{Y} \,|\pi_i\right] = \mathbf{a}^\top \mathsf{Cov}\left[\mathbf{X}\right] \mathbf{a} = \mathbf{a}^\top \boldsymbol{\Sigma}\, \mathbf{a}$$

- La media general de $\mathbf{Y}$:

$$\overline{\mu}_y = \mathsf{E}[Y] = \frac{1}{g}\sum_{i=1}^{g} \mu_{i,Y} = \frac{1}{g}\sum_{i=1}^{g} \mathbf{a}^\top \boldsymbol{\mu}_i = \mathbf{a}^\top \frac{1}{g}\sum_{i=1}^{g} \boldsymbol{\mu}_i$$

Reglas de Discriminación para g Grupos

- La razón[1]:

$$\frac{\sum_{i=1}^{g}(\mu_{i,Y}-\overline{\mu}_y)^2}{\sigma_y^2}=\frac{\sum_{i=1}^{g}(\mathbf{a}^\top\boldsymbol{\mu}_i-\mathbf{a}^\top\overline{\mu}_i)^2}{\mathbf{a}^\top\boldsymbol{\Sigma}\,\mathbf{a}}=\frac{\mathbf{a}^\top\,\mathrm{B}\,\mathbf{a}}{\mathbf{a}^\top\boldsymbol{\Sigma}\,\mathbf{a}}$$

  mide la variabilidad entre grupos sobre la variabilidad intra grupos. El objetivo es buscar **a** que maximice esa razón para así tener la mayor discriminación posible.

[1]Suma de cuadrado de las distancias de las medias de la poblaciones a la media general dividida por la varianza de Y

Reglas de Discriminación para g Grupos

Como

$$\boldsymbol{\mu}_i \qquad \boldsymbol{\Sigma}$$

son desconocidos, usamos las respectivas estimaciones

$$\bar{\mathbf{x}}_i \qquad \mathbf{S}_i \quad i = 1, \ldots, g$$

luego

$$\mathrm{B} = \sum_{i=1}^{g} (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})^\top \qquad \mathrm{W} = \sum_{i=1}^{g} (n_i - 1)\,\mathbf{S}_i$$

Note que

$$\mathrm{W} = [n_1 + n_2 + \cdots + n_g - g]\,\mathbf{S}_{\text{pooled}}$$

Reglas de Discriminación para g Grupos

donde

$$\mathbf{S}_{\text{pooled}} = \frac{(n_1 - 1)\,\mathbf{S}_1 + (n_2 - 1)\,\mathbf{S}_2 + \cdots + (n_g - 1)\,\mathbf{S}_g}{n_1 + n_2 + \cdots + n_g - g}$$

Entonces, la misma constante $\hat{\mathbf{a}}$ que maximiza

$$\frac{\hat{\mathbf{a}}^\top \mathrm{B}\, \hat{\mathbf{a}}}{\hat{\mathbf{a}}^\top \mathbf{S}_{\text{pooled}}\, \hat{\mathbf{a}}}$$

maximiza

$$\frac{\hat{\mathbf{a}}^\top \mathrm{B}\, \hat{\mathbf{a}}}{\hat{\mathbf{a}}^\top \mathrm{W}\, \hat{\mathbf{a}}}$$

Reglas de Discriminación para g Grupos

Se puede mostrar que el autovector $\hat{\mathbf{e}}_1$ asociado al mayor autovalor $\hat{\lambda}_1$ de $\mathrm{W}^{-1}\,\mathrm{B}$ llevan al máximo de $\frac{\hat{\mathbf{a}}^\top\,\mathrm{B}\,\hat{\mathbf{a}}}{\hat{\mathbf{a}}^\top\,\mathrm{W}\,\hat{\mathbf{a}}}$

Reglas de Discriminación para g Grupos

Se puede mostrar que el autovector $\hat{\mathbf{e}}_1$ asociado al mayor autovalor $\hat{\lambda}_1$ de $\mathrm{W}^{-1}\,\mathrm{B}$ llevan al máximo de $\frac{\hat{\mathbf{a}}^\top\,\mathrm{B}\,\hat{\mathbf{a}}}{\hat{\mathbf{a}}^\top\,\mathrm{W}\,\hat{\mathbf{a}}}$

Discriminantes Lineales de Fisher

Sean $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_s$ los autovalores de $\mathrm{W}^{-1}\,\mathrm{B}$ y $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \ldots, \hat{\mathbf{e}}_s$ los autovectores correspondientes. Entonces, el vector de coeficientes $\hat{\mathbf{a}}$ que maximiza $\frac{\hat{\mathbf{a}}^\top\,\mathrm{B}\,\hat{\mathbf{a}}}{\hat{\mathbf{a}}^\top\,\mathrm{W}\,\hat{\mathbf{a}}}$ es $\hat{\mathbf{a}} = \hat{\mathbf{e}}$. La combinación lineal $\hat{\mathbf{a}}_1^\top\,\mathbf{X} = \hat{\mathbf{e}}_1^\top\,\mathbf{X}$ es llamada **primer discriminante lineal.**

La combinación lineal $\hat{\mathbf{a}}_2^\top\,\mathbf{X} = \hat{\mathbf{e}}_2^\top\,\mathbf{X}$ es llamada **segundo discriminante lineal**, y así en adelante.

Reglas de Discriminación para g Grupos

¿Cómo hacer las clasficaciones?

Reglas de Discriminación para g Grupos

¿Cómo hacer las clasficaciones? Sea

$$\hat{Y}_k = \hat{\mathbf{a}}_k^\top \mathbf{X}$$

el $k-$ésimo discriminante lineal.

Reglas de Discriminación para g Grupos

¿Cómo hacer las clasficaciones? Sea

$$\hat{Y}_k = \hat{\mathbf{a}}_k^\top \mathbf{X}$$

el $k-$ésimo discriminante lineal. Para clasificar una observación $\mathbf{x}$ en el $l-$ésimo grupo, se verifica que

$$(\hat{y}_k - \overline{y}_l)^2 \leq (\hat{y}_k - \overline{y}_i)^2 \quad \forall\, i \neq l$$

o equivalentemente

$$(\hat{\mathbf{a}}_k^\top \mathbf{x} - \hat{\mathbf{a}}_k^\top \overline{\mathbf{x}}_l)^2 \leq (\hat{\mathbf{a}}_k^\top \mathbf{x} - \hat{\mathbf{a}}_i^\top \overline{\mathbf{x}}_i)^2 \quad \forall\, i \neq l$$

En la práctica se usa el **primer discriminante lineal** porque tiene **mayor poder de discriminación** y **menos clasificasiones incorrectas.**

Reglas de Discriminación para g Grupos

Example

Considere que se tienen $p = 2$ variables y $g = 3$ grupos cada una con 3 observaciones. Los datos observados son

Grupos	Variable 1	Variable 2
1	-2	5
1	0	3
1	-1	1
2	0	6
2	2	4
2	1	2
3	1	-2
3	0	0
3	-1	-4

Reglas de Discriminación para g Grupos

Obs: Example 11.10 from Johnson and Wichern (2013), Applied Multivariate Statistical Analysis, pp. 612

Observe que tenemos $g = 3$ grupos de tamaños $n_1 = n_2 = n_3 = 3$, y que

$$\mathbf{x}_1 = \begin{bmatrix} -2 & 5 \\ 0 & 3 \\ -1 & 1 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} 0 & 6 \\ 2 & 4 \\ 1 & 2 \end{bmatrix} \qquad \mathbf{x}_3 = \begin{bmatrix} 1 & -2 \\ 0 & 0 \\ -1 & -4 \end{bmatrix}$$

$$\mathbf{S}_1 = \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \qquad \mathbf{S}_2 = \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \qquad \mathbf{S}_3 = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$$

$$\bar{\mathbf{x}}_1 = \begin{bmatrix} -1 \\ 3 \end{bmatrix} \qquad \bar{\mathbf{x}}_2 = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \qquad \bar{\mathbf{x}}_3 = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$

Reglas de Discriminación para g Grupos

$$\bar{\mathbf{x}} = \begin{bmatrix} 0 & 5/3 \end{bmatrix}^\top$$

$$B = \sum_{i=1}^{3} (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})^\top$$

$$= \begin{bmatrix} -1 \\ -4/3 \end{bmatrix} \begin{bmatrix} -1 & -4/3 \end{bmatrix} + \begin{bmatrix} 1 \\ 7/3 \end{bmatrix} \begin{bmatrix} 1 & 7/3 \end{bmatrix} + \begin{bmatrix} 0 \\ 11/3 \end{bmatrix} \begin{bmatrix} 0 & 11/3 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 \\ 1 & 62/3 \end{bmatrix}$$

Reglas de Discriminación para g Grupos

$$W = \sum_{i=1}^{3}\sum_{j=1}^{n_i}(\mathbf{x}_{ij} - \overline{\mathbf{x}}_i)(\mathbf{x}_{ij} - \overline{\mathbf{x}}_i)^\top = (n_1 + n_2 + n_3 - 3)\,\mathbf{S}_{pooled}$$

$$= (n_1 + n_2 + n_3 - 3)\frac{(n_1 - 1)\,\mathbf{S}_1 + (n_2 - 1)\,\mathbf{S}_2 + (n_3 - 1)\,\mathbf{S}_3}{n_1 + n_2 + n_3 - 3}$$

$$= (n_1 - 1)\,\mathbf{S}_1 + (n_2 - 1)\,\mathbf{S}_2 + (n_3 - 1)\,\mathbf{S}_3$$

$$= 2\,[\mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3]$$

$$= \begin{bmatrix} 6 & -2 \\ -2 & 24 \end{bmatrix}$$

Reglas de Discriminación para g Grupos

$$\mathrm{W}^{-1}\,\mathrm{B} = \begin{bmatrix} 0.3571 & 0.4667 \\ 0.0714 & 0.9 \end{bmatrix}$$

$$\mathbf{e}_1 = \hat{\mathbf{a}}_1 = \begin{bmatrix} 0.386 \\ 0.495 \end{bmatrix}$$

$$\mathbf{e}_2 = \hat{\mathbf{a}}_2 = \begin{bmatrix} 0.938 \\ -0.112 \end{bmatrix}$$

Observación:Éstos son autovectores estandarizados

Reglas de Discriminación para g Grupos

y por lo tanto, los **discriminantes lineales de Fisher** son

$$\hat{Y}_1 = \hat{\mathbf{a}}_1^\top \mathbf{X} = 0.386X_1 + 0.495X_2$$
$$\hat{Y}_2 = \hat{\mathbf{a}}_2^\top \mathbf{X} = 0.938X_1 - 0.112X_2$$

Considere que se tiene una **nueva observación** $\begin{bmatrix}1 & 3\end{bmatrix}$, luego

$$\hat{Y}_1 = 0.386(1) + 0.495(3) = 1.871$$
$$\hat{Y}_2 = 0.938(1) - 0.112(3) = 0.602$$

Para saber en cual de los grupos clasificamos a la nueva obervación, debemos saber cuál es el grupo más cercano a ella.

Reglas de Discriminación para g Grupos

Debemos determinar

$$\min_{k\in\{1,2,3\}} \sum_{j=1}^{2} (\hat{y}_j - \overline{y_{kj}})^2 = \min_{k\in\{1,2,3\}} \sum_{j=1}^{2} (\hat{y}_j - \hat{\mathbf{a}}_k^\top \overline{\mathbf{x}}_k)^2$$

Para $k = 1$ tenemos

$$\sum_{j=1}^{2} (\hat{y}_j - \overline{y_{1j}})^2 = \sum_{j=1}^{2} (\hat{y}_j - \hat{\mathbf{a}}_j^\top \overline{\mathbf{x}}_1)^2$$

$$= (\hat{y}_1 - \hat{\mathbf{a}}_1^\top \overline{\mathbf{x}}_1)^2 + (\hat{y}_2 - \hat{\mathbf{a}}_2^\top \overline{\mathbf{x}}_1)^2$$

$$= (1.87 - 1.1)^2 + (0.60 + 1.27)^2$$

$$= 4.09$$

Reglas de Discriminación para g Grupos

donde

$$\hat{\mathbf{a}}_1^\top \overline{\mathbf{x}}_1 = \begin{bmatrix} 0.386 & 0.495 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = 1.099 \approx 1.1$$

$$\hat{\mathbf{a}}_2^\top \overline{\mathbf{x}}_1 = \begin{bmatrix} 0.938 & -0.112 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = -1.274 \approx -1.27$$

$$\min_{k \in \{1,2,3\}} \sum_{j=1}^{2} (\hat{y}_j - \overline{y_{kj}})^2 = \begin{cases} 4.09 & \text{si } k = 1 \\ 0.26 & \text{si } k = 2 \\ 8.32 & \text{si } k = 3 \end{cases}$$

Observe que la **menor distancia** es cuando $k = 2$, así que la nueva observación será clasificada en el **Grupo 2**.

Reglas de Discriminación para g Grupos

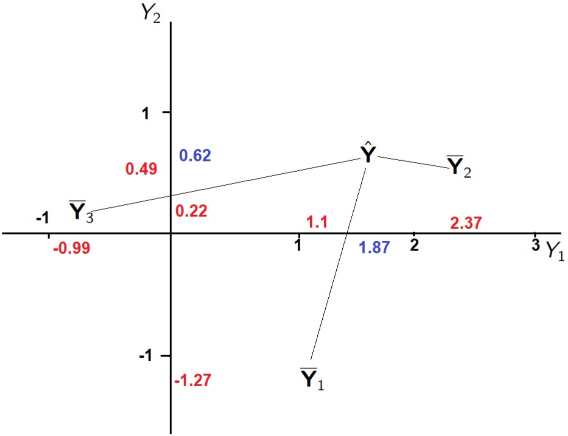
Gráficamente:

$$\overline{\mathbf{Y}}_1 = \begin{bmatrix} 0.386 & 0.495 \\ 0.938 & -0.112 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1.1 \\ -1.27 \end{bmatrix}$$

$$\overline{\mathbf{Y}}_2 = \begin{bmatrix} 0.386 & 0.495 \\ 0.938 & -0.112 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 2.37 \\ 0.49 \end{bmatrix}$$

$$\overline{\mathbf{Y}}_3 = \begin{bmatrix} 0.386 & 0.495 \\ 0.938 & -0.112 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \end{bmatrix} = \begin{bmatrix} -0.99 \\ 0.22 \end{bmatrix}$$

Reglas de Discriminación para g Grupos



Reglas para Discriminación de Grupos

En `R`:

- Función Discriminante **Lineal**:
  función **`lda`** del paquete `MASS`
- Función Discriminante **Cuadrática**
  función **`qda`** del paquete `MASS`

Reglas de Discriminación para g Grupos

```
# ----- Ejemplo 3 ----- #
x<-scan()
1 -2 5
1  0 3
1 -1 1
2  0 6
2  2 4
2  1 2
3  1 -2
3  0 0
3 -1 -4
D<-matrix(x,ncol=3,byrow=TRUE)
Datos<-data.frame(D)
colnames(Datos)<-c("Poblacion", "Variable1","Variable2")
```

Reglas de Discriminación para g Grupos

```
# ----- Sigma 1 = Sigma 2 ----- #
library(biotools)
boxM(Datos[,2:3], grouping = Datos[,1])

# ----- Analisis Discriminante Lineal ----- #
library(MASS)
m <- lda(Poblacion ~ Variable1 + Variable2,Datos)
m

new.obs<- data.frame(Variable1 = 1, Variable2 = 3)
predict(m, newdata = new.obs)
```

Reglas de Discriminación para g Grupos

```
# ----- Datos Iris ----- #
data("iris")
head(iris)
m<-lda(Species~., data=iris)
m
plot(m)
pred<-predict(m)
pred$class    # clasif. por grupo predichas
clasif<-table(pred$class,iris$Species)
```