Como $\widehat{\boldsymbol{\beta}} \sim N_p(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$, entonces $\widehat{\beta_i} = \mathbf{e}_i'\widehat{\boldsymbol{\beta}} \sim N(\beta_i, \sigma^2\mathbf{e}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{e}_i)$. Si denotamos $\Sigma_{\widehat{\boldsymbol{\beta}}} = \sigma^2\mathbf{D}$

$$\boxed{\widehat{\beta_i} \sim N(\beta_i, \sigma^2 d_{ii})}$$

siendo d_{ii} el i-ésimo elemento diagonal de $\mathbf{D}$.

Si para un i fijo queremos testear

$$\boxed{H_o : \beta_i = 0 \quad \text{vs.} \quad H_1 : \beta_i \neq 0}$$

tenemos que bajo H_o

$$\frac{\widehat{\beta_i}}{s\sqrt{d_{ii}}} \sim t_{n-p}$$

Por lo tanto, rechazaremos H_o con nivel α si

$$\left|\frac{\widehat{\beta_i}}{s\sqrt{d_{ii}}}\right| > t_{n-p,\frac{\alpha}{2}}$$

En el caso de regresión simple tendríamos

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad 1 \leq i \leq n, \quad \epsilon_i \sim N(0, \sigma^2)$$

Entonces:

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} n & \sum\limits_{i=1}^{n} x_i \\ \sum\limits_{i=1}^{n} x_i & \sum\limits_{i=1}^{n} x_i^2 \end{pmatrix}$$

y la inversa resulta

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{n \sum_{i=1}^{n} x_i^2 - n^2 \overline{x}^2} \begin{pmatrix} \sum\limits_{i=1}^{n} x_i^2 & -\sum\limits_{i=1}^{n} x_i \\ -\sum\limits_{i=1}^{n} x_i & n \end{pmatrix}$$

$$\widehat{\beta_0} = -\overline{x}\,\widehat{\beta_1} + \overline{y}$$

y

$$\widehat{\beta_1} = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n} (x_i - \overline{x})^2}$$

Luego, si queremos testear

$$H_o : \beta_1 = 0 \quad \text{vs.} \quad H_1 : \beta_1 \neq 0$$

el estadístico será

$$T = \left| \frac{\widehat{\beta_1}}{s\sqrt{d_{11}}} \right| = \left| \frac{\widehat{\beta_1}}{s/\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2}} \right|$$

y rechazaremos H_o si

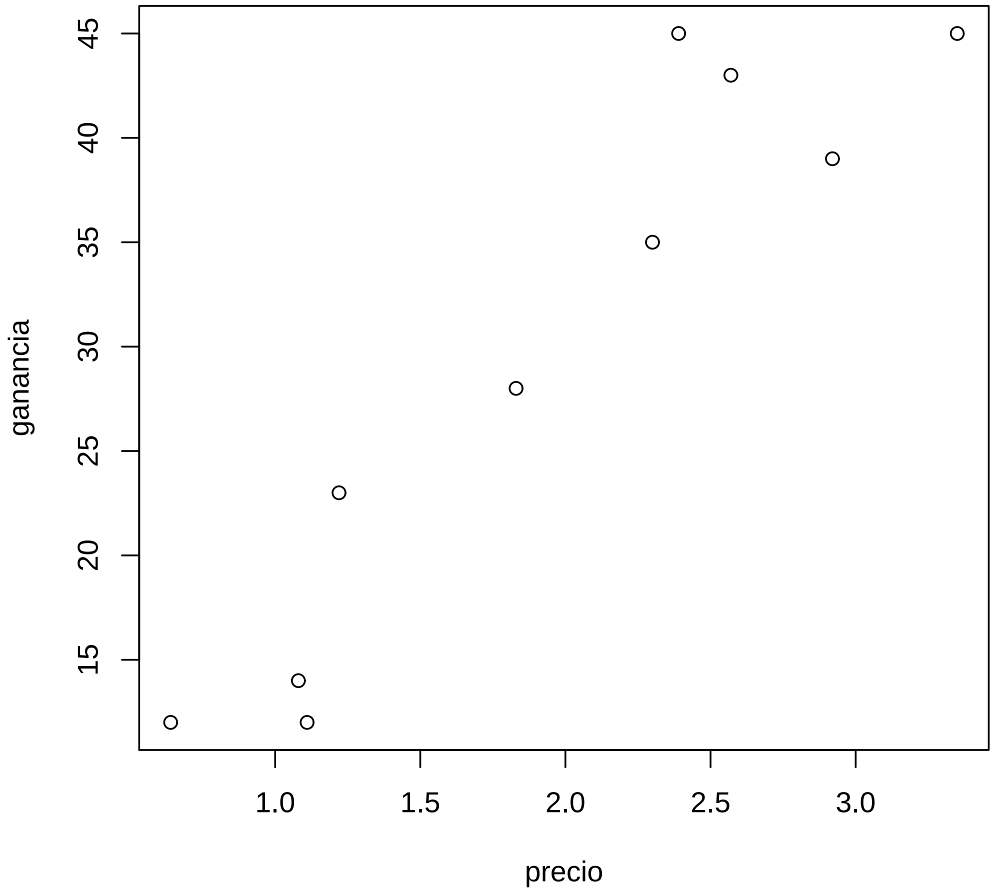
$$\boxed{|T| > t_{n-2,\frac{\alpha}{2}}}$$

Veamos un ejemplo: Precio del papel.

Y: ganancia en 1972

x: precio de papel en 1973

```
> Ejemplo Precio del Papel
   precio  ganancia
        x      y
 1   1.83   28
 2   3.35   45
 3   0.64   12
 4   2.30   35
 5   2.39   45
 6   1.08   14
 7   2.92   39
 8   1.11   12
 9   2.57   43
10   1.22   23
```


ganancia
15
20
25
30
35
40
45
1.0
1.5
2.0
2.5
3.0
precio

```
> sal.lm
Coefficients:
 (Intercept)         x
    2.027775      14.20517

Degrees of freedom: 10 total; 8 residual
Residual standard error: 5.025083

> summary(sal.lm)

Call: lm(formula = y ~ x, x = T)
Residuals:
    Min        1Q     Median      3Q       Max
 -5.796   -4.222    0.1386    2.952     9.022
```

```
Coefficients:
                 Value    Std. Error    t value    Pr(>|t|)
(Intercept)   2.0278     3.9383          0.5149     0.6206
          x  14.2052     1.8565          7.6516     0.0001

Residual standard error: 5.025 on 8 degrees of freedom
Multiple R-Squared: 0.8798
F-statistic: 58.55 on 1 and 8 degrees of freedom, the p-value
is 0.00006008

Correlation of Coefficients:
  (Intercept)
            x  -0.915
 X'X=
                   (Intercept)       x
(Intercept)         10.00          19.4100
          x         19.41          45.0013
```

```
(X'X)^(-1)=
              (Intercept)         x
(Intercept)    0.6142273  -0.264929
          x   -0.2649290   0.136491

> matriz de covarianza de coeficientes

             (Intercept)         x
(Intercept)    15.510133 -6.689844
          x    -6.689844  3.446597
```

También podríamos interesarnos realizar in I. de C. para la esperanza de una nueva observación independiente de las demás que cumpla el modelo

$$y_i = \beta_1 x_1 + \beta_1 x_2 + \ldots + \beta_p x_p + \epsilon_i$$

en $\mathbf{x}_o = (\mathbf{x}_{o1}, \mathbf{x}_{o2}, \ldots, \mathbf{x}_{op})'$ donde $\epsilon_i \sim N(0, \sigma^2)$ independientes.

Como $E(y_o) = \mathbf{x}_o' \boldsymbol{\beta}$, podemos estimarlo por $\widehat{E(y_o)} = \mathbf{x}_o' \widehat{\boldsymbol{\beta}} = \widehat{y}_o$

Por lo tanto, de acuerdo con lo que hemos visto

$$\widehat{y}_o = \mathbf{x}_o' \widehat{\boldsymbol{\beta}} \sim N(\mathbf{x}_o' \boldsymbol{\beta}, \sigma^2 \mathbf{x}_o' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_o)$$

y es independiente de

$$\frac{(n-p)s^2}{\sigma^2} \sim \chi^2_{n-p}$$

por lo tanto

$$T = \frac{\widehat{y}_o - \mathbf{x}_o' \boldsymbol{\beta}}{s\sqrt{\mathbf{x}_o' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_o}} \sim t_{n-p}$$

En consecuencia,

$$\boxed{\widehat{y}_o \pm t_{n-p,\frac{\alpha}{2}}\; s\sqrt{\mathbf{x}_o'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_o}}$$

es un intervalo de nivel exacto $1-\alpha$.

Asimismo, podríamos estar interesados en la predicción de y_o, una nueva observación que cumpla el modelo, y en un intervalo para ella, que llamaremos intervalo de predicción.

Observemos que el predictor de y_o es $\widehat{y}_o = \mathbf{x}_o'\widehat{\boldsymbol{\beta}}$. En efecto,
$E(\widehat{y}_o - y_o) = 0$. ¿Qué distribución tiene $\widehat{y}_o - y_o$?

Tenemos que

$$\begin{aligned}
\widehat{y}_o &\sim N(\mathbf{x}_o'\boldsymbol{\beta}, \sigma^2\mathbf{x}_o'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_o) \\
y_o &\sim N(\mathbf{x}_o'\boldsymbol{\beta}, \sigma^2)
\end{aligned}$$

y dado que y_o es independiente de las restantes y_i's con las que estimamos, entonces por la independencia entre estas dos normales queda que

$$\widehat{y}_o - y_o \sim N(0, \sigma^2(1 + \mathbf{x}_o'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_o))$$

Por lo tanto, el intervalo de predicción de nivel $1 - \alpha$ estará dado por

$$\boxed{\widehat{y}_o \pm t_{n-p,\frac{\alpha}{2}}\ s\sqrt{1 + \mathbf{x}_o'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_o}}$$

Ejemplo. Los siguientes son datos que corresponden a 10 porcentajes y_i de una sustancia que fueron medidos en experiencias de laboratorio y que se desean relacionar con la temperatura x_i a la que fueron realizados dichas experiencias.

i	x	y
1	100	45
2	110	52
3	120	54
4	130	63
5	140	62
6	150	68
7	160	75
8	170	76
9	180	92
10	190	88

La tabla con los estadísticos calculados es:

Coeficiente	Estimación	Error Standard	Valor de t
β_0	-4.47273	5.63433	-0.79
β_1	0.49636	0.03812	13.02
s	3.46213	g.l.=8	

Intervalos de Estimación y de Predicción

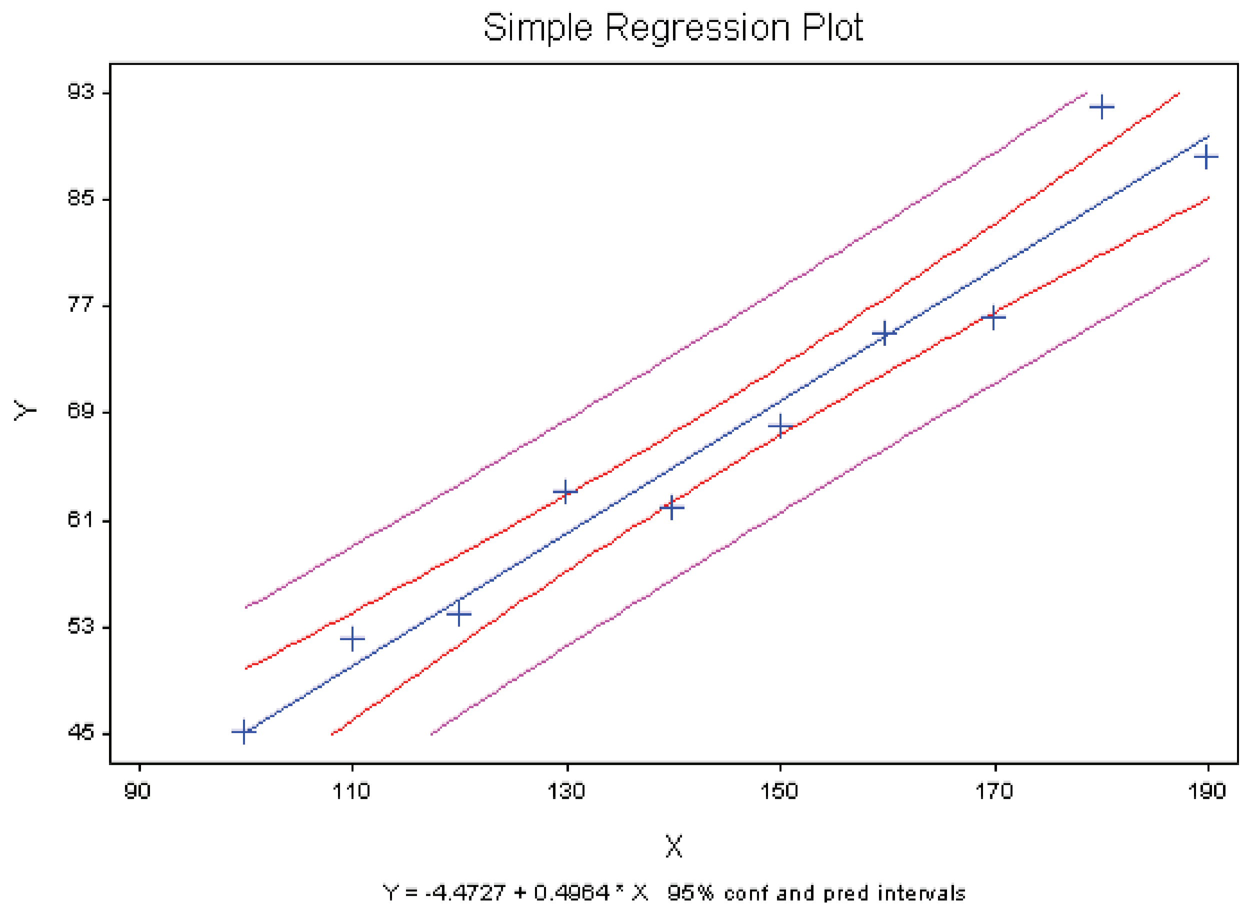


Tabla de Resultados

```
UNWEIGHTED LEAST SQUARES LINEAR REGRESSION OF Y

PREDICTOR
VARIABLES     COEFFICIENT    STD ERROR     STUDENT'S T      P
---------     -----------    ---------     -----------    ------
CONSTANT        -4.47273      5.63433         -0.79       0.4502
X                0.49636      0.03812         13.02       0.0000

R-SQUARED              0.9549      RESID. MEAN SQUARE (MSE)     11.9864
ADJUSTED R-SQUARED     0.9493      STANDARD DEVIATION           3.46213

SOURCE         DF       SS            MS           F          P
----------    ---    ----------    ----------    -----    ------
REGRESSION      1      2032.61       2032.61     169.58    0.0000
RESIDUAL        8      95.8909       11.9864
TOTAL           9      2128.50

CASES INCLUDED 10    MISSING CASES 0
```

- El valor estimado de $\widehat{\beta}_1 \simeq 0{,}5$, $\Rightarrow$ esperamos que el porcentaje aumente 0.5 por cada incremento de un grado en la temperatura.

- $s_{\hat{\beta}_1} = 0{,}03812$

- Si testeamos $H_0 : \beta_1 = 0$ $t = \frac{0{,}49636}{0{,}038112} = 13{,}02$ y $t_{8,0{,}025} = 2{,}306004$

$\Rightarrow$ los datos nos dan evidencia suficiente al nivel 5 % como para concluir que la pendiente es no nula.

Observemos que en el gráfico la recta ajustada *está encerrada* entre 2 curvas interiores y 2 exteriores. Las exteriores corresponden al intervalo de predicción de nivel 0.95 y las interiores a los intervalos de confianza de nivel 0.95 para la media.

Notemos que **el nivel de confianza 0.95 se aplica a cada punto y no es global**

Supongamos que queremos plantear un test de nivel α para

$$\boxed{H_o : \mathbf{C}\boldsymbol{\beta} = \boldsymbol{\delta} \text{ vs. } H_1 : \mathbf{C}\boldsymbol{\beta} \neq \boldsymbol{\delta}}$$

siendo $\text{rg}(\mathbf{C}) = q$, $\mathbf{C} \in \Re^{q \times p}$.

Sea $\Psi = \mathbf{C}'\boldsymbol{\beta}$. Sabemos que $\widehat{\Psi} \sim N_q(\Psi, \sigma^2 \mathbf{A}^* \mathbf{A}^{*'}) = N_q(\Psi, \sigma^2 \mathbf{B})$. Por lo tanto, tenemos que

$$(1) : \quad Q = \frac{1}{q}(\widehat{\Psi} - \boldsymbol{\delta})' \mathbf{B}^{-1} (\widehat{\Psi} - \boldsymbol{\delta})$$

es independiente de

$$(2) : \quad s^2 = \frac{\|\mathbf{Y} - \widehat{\mathbf{Y}}\|^2}{n - r}$$

Veremos que

$$E(Q) = \sigma^2 + \eta^2$$

y que $\eta^2 = 0$ sólo cuando H_o es cierta.

Bajo H_o, (1) y (2) son estimadores insesgados de σ^2, es decir que bajo H_o esperamos que

$$\frac{(1)}{(2)} \simeq 1,$$

pero si H_o no es cierta, esperamos que

$$\frac{(1)}{(2)} > 1.$$

Luego, el cociente $\dfrac{(\widehat{\Psi} - \boldsymbol{\delta})'\mathbf{B}^{-1}(\widehat{\Psi} - \boldsymbol{\delta})}{qs^2}$ nos dará una idea de la *veracidad de* H_O, de manera que rechazaremos H_o si el cociente es grande.

¿ Cuán grande?

Bajo H_o

$$\frac{(\widehat{\Psi} - \boldsymbol{\delta})' \mathbf{B}^{-1} (\widehat{\Psi} - \boldsymbol{\delta})}{\sigma^2} \sim \chi^2_q$$

independiente de

$$\frac{(n-r)s^2}{\sigma^2} \sim \chi^2_{n-r}$$

En consecuencia:

$$F = \frac{(\widehat{\Psi} - \boldsymbol{\delta})' \mathbf{B}^{-1} (\widehat{\Psi} - \boldsymbol{\delta})}{q s^2} \sim \mathcal{F}_{q,n-r}$$

Rechazaremos H_o si

$$F > \mathcal{F}_{q,n-r,\alpha}$$

Veamos dos situaciones frecuentes para el caso de rango completo.

1. Una hipótesis simple.

$\mathbf{C} = \mathbf{c}$ consiste en una sola fila, de manera que $\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}$ es un escalar, con lo cual el estadístico resulta

$$F = \frac{(\mathbf{c}'\widehat{\boldsymbol{\beta}} - \delta)^2}{s^2\, \mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}$$

que bajo H_o tiene distribución $F_{1,n-p}$

En función de la relación entre las distribuciones t y F podríamos utilizar la distribución t de Student y

$$\text{rechazamos } H_o \text{ si } \left| \frac{c'\widehat{\boldsymbol{\beta}} - \delta}{s\, \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \right| > t_{n-p,\alpha/2}$$

2. Tests para k coeficientes iguales a 0.

$H_o : \Psi = \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$, donde $\mathbf{C} = \begin{pmatrix} e_{i_1} \\ \cdot \\ \cdot \\ e_{i_k} \end{pmatrix}$, para $i_1 \leq 1 < \ldots < i_k \leq p$.

El numerador será:

$$(\mathbf{C}\widehat{\boldsymbol{\beta}})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\widehat{\boldsymbol{\beta}})$$

donde $\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'$ es una submatriz de $\mathbf{D} = (\mathbf{X}'\mathbf{X})^{-1}$ que sólo involucra los coeficientes correspondientes a aquellos β_i presentes en la hipótesis a testear.

Así supongamos que tenemos 5 coeficientes $\beta_1, \ldots, \beta_5$ y queremos testear

$$\begin{aligned} H_o : \beta_1 &= 0 \\ \beta_3 &= 0 \\ \beta_5 &= 0 \end{aligned}$$

luego,

$$\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}' = \begin{pmatrix} d_{11} & d_{13} & d_{15} \\ d_{13} & d_{33} & d_{35} \\ d_{15} & d_{35} & d_{55} \end{pmatrix}$$

y en el numerador tendremos

$$\left(\widehat{\beta_1}, \widehat{\beta_3}, \widehat{\beta_5}\right) \begin{pmatrix} d_{11} & d_{13} & d_{15} \\ d_{13} & d_{33} & d_{35} \\ d_{15} & d_{35} & d_{55} \end{pmatrix}^{-1} \begin{pmatrix} \widehat{\beta_1} \\ \widehat{\beta_3} \\ \widehat{\beta_5} \end{pmatrix}$$

Test de Cociente de Verosimilitud

El test de F también puede motivarse como test de cociente de verosimilitud.

Sea Ω el conjunto de supuestos generales y supongamos que bajo este modelo testeamos la hipótesis H, llamemos $\omega = \Omega \cap H$. Así, por ejemplo, si

$$\Omega : \quad \mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I}) \quad \boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{p-1})'$$

y

$$H : \beta_1 = \ldots = \beta_{p-1} = 0$$

entonces $\omega = \Omega \cap H$ equivale a

$$\mathbf{Y} \sim N_n(\beta_0, \sigma^2 \mathbf{I}) \, .$$

Si $p(\mathbf{y})$ es la función de densidad o de probabilidad de $\mathbf{Y}$ definimos λ el estadístico del cociente de verosimilitud como

$$\lambda = \frac{\text{máx}_\omega \, p(\mathbf{y})}{\text{máx}_\Omega \, p(\mathbf{y})}$$

Notemos que $0 \leq \lambda \leq 1$ ya que $\omega \in \Omega$ y por lo tanto $\text{máx}_\omega \, p(\mathbf{y}) \leq \text{máx}_\Omega \, p(\mathbf{y})$.

H será rechazada cuando $\max_\omega p(\mathbf{y})$ es mucho más chico que $\max_\Omega p(\mathbf{y})$, por lo tanto rechazaremos H si $\lambda < \lambda_\alpha$.

Existen dos formas equivalentes de plantear las hipótesis:

- 1)

$$\begin{aligned} \Omega &: \mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}) \quad \text{rg} = r \\ H &: \Psi_1 = \Psi_2 = \ldots = \Psi_q = 0 \end{aligned}$$

  donde $\{\Psi_i\}$ son l.i. funciones estimables

- 2)

$$\begin{aligned} \Omega &: \mathbf{Y} \sim N_n(\boldsymbol{\eta}, \sigma^2\mathbf{I}) \quad \boldsymbol{\eta} \in \mathcal{V}_r \\ H &: \boldsymbol{\eta} \in \mathcal{V}_{r-q} \end{aligned}$$

  donde $\mathcal{V}_r$ es un subespacio de dimensión r en $\Re^n$ y $\mathcal{V}_{r-q}$ es un subespacio de dimensión $r-q$ en $\mathcal{V}_r$.

$\mathcal{V}_r$ es el espacio generado por las columnas de $\mathbf{X}$ y $\mathcal{V}_{r-q}$ es el espacio al cual está restringido $\boldsymbol{\eta}$ a yacer al imponerle las restricciones $\Psi_1 = \Psi_2 = \ldots = \Psi_q = 0$.

Las dos formas son equivalentes, nosotros probaremos que • 1) $\Longrightarrow$ • 2).

Tenemos que $\mathbf{Y} \sim N_n(\boldsymbol{\eta}, \sigma^2\mathbf{I}) \quad \boldsymbol{\eta} \in \mathcal{V}_r$. Llamemos $\mathbf{C}$ a la matriz tal que $\Psi = \mathbf{C}\boldsymbol{\beta}$. Luego:

$$\begin{aligned} V_\omega &= \{\mathbf{v} = \mathbf{X}\boldsymbol{\beta} \text{ tal que } \mathbf{C}\boldsymbol{\beta} = \mathbf{0}\} = \{\mathbf{v} = \mathbf{X}\boldsymbol{\beta} \text{ tal que } \mathbf{A}^*\mathbf{X}\boldsymbol{\beta} = \mathbf{0}\} \\ &= \{v = \mathbf{X}\boldsymbol{\beta} \text{ tal que } \mathbf{A}^*\mathbf{v} = \mathbf{0}\} \end{aligned}$$

$\mathbf{A}^* = \begin{pmatrix} \mathbf{a}_1^* \\ . \\ . \\ . \\ \mathbf{a}_q^* \end{pmatrix}$. Es decir, $\mathbf{v} \perp \mathbf{a}_i^* \quad 1 \leq i \leq q$.

Como $\text{rg} = q$ entonces $\mathbf{a}_1^*, \ldots, \mathbf{a}_q^*$ son l.i. Por lo tanto, $\mathbf{v} \in \mathcal{V}^{\perp}_{<\mathbf{a}_1^* \ldots \mathbf{a}_q^*>}$: complemento ortogonal de $\mathcal{V}_{<\mathbf{a}_1^* \ldots \mathbf{a}_q^*>}$ en $\mathcal{V}_r$.

Además, tenemos que

$$r = \dim(\mathcal{V}_{<\mathbf{a}_1^* \ldots \mathbf{a}_q^*>}) + \dim(\mathcal{V}^{\perp}_{<\mathbf{a}_1^* \ldots \mathbf{a}_q^*>})$$

por lo tanto,

$$\dim(\mathcal{V}^{\perp}_{<\mathbf{a}_1^* \ldots \mathbf{a}_q^*>}) = dim(\mathcal{V}_\omega) = r - q$$

Calculemos λ. Para ello deberemos calcular el máximo de de $p(\mathbf{y})$ en c/u de los subespacios.

Veremos que $\lambda = \left(\frac{\|\mathbf{y}-\widehat{\boldsymbol{\eta}}\|^2}{\|\mathbf{y}-\widehat{\boldsymbol{\eta}}_\omega\|^2}\right)^{n/2}$ y por lo tanto rechazamos H_0 si

$$\lambda = \left(\frac{\|\mathbf{y}-\widehat{\boldsymbol{\eta}}\|^2}{\|\mathbf{y}-\widehat{\boldsymbol{\eta}}_\omega\|^2}\right)^{n/2} < k_\alpha$$

Si aplicamos a este cociente la función $g(t) = \frac{n-r}{q}\left(t^{-2/n} - 1\right)$, resulta

$$\begin{aligned} F &= \frac{n-r}{q} \frac{\|\mathbf{y}-\widehat{\boldsymbol{\eta}}_\omega\|^2 - \|\mathbf{y}-\widehat{\boldsymbol{\eta}}\|^2}{\|\mathbf{y}-\widehat{\boldsymbol{\eta}}\|^2} \\ &= \frac{1}{q} \frac{\|\mathbf{y}-\widehat{\boldsymbol{\eta}}_\omega\|^2 - \|\mathbf{y}-\widehat{\boldsymbol{\eta}}\|^2}{s^2} \end{aligned}$$

Como veremos

$$F = \frac{1}{q} \frac{\|\widehat{\boldsymbol{\eta}}_\omega - \widehat{\boldsymbol{\eta}}\|^2}{s^2}$$

Luego, rechazaremos H si

$$\frac{1}{q} \frac{\|\widehat{\boldsymbol{\eta}}_\omega - \widehat{\boldsymbol{\eta}}\|^2}{s^2} > \lambda_\alpha$$

Una interpretación intuitiva para este test es que $\|\mathbf{y} - \widehat{\boldsymbol{\eta}}_\omega\|^2$ y $\|\mathbf{y} - \widehat{\boldsymbol{\eta}}\|^2$ miden cuan bien ajustan los modelos ω y Ω, respectivamente. Por lo tanto, su cociente compara el ajuste de ω con el de Ω y rechazamos H si este cociente es grande:

$$F > \lambda_\alpha$$

¿ Qué distribución tiene el cociente F?

Tenemos que $\mathcal{V}_{r-q} \in \mathcal{V}_r \in \Re^n$. Tomemos una base ortonormal de $\mathcal{V}_{r-q}$: $\{\boldsymbol{\alpha}_{q+1}, \ldots \boldsymbol{\alpha}_r\}$ y la extendemos a una base ortonormal de $\mathcal{V}_r$: $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_q, \boldsymbol{\alpha}_{q+1}, \ldots, \boldsymbol{\alpha}_r\}$ y finalmente a una de $\Re^n$: $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \ldots, \boldsymbol{\alpha}_n\}$.

$$\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_q, \boldsymbol{\alpha}_{q+1}, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \ldots, \boldsymbol{\alpha}_n$$

Por lo tanto,

$$\mathbf{y} \in \Re^n : \mathbf{y} = \sum_{j=1}^{n} z_j \boldsymbol{\alpha}_j$$

y tenemos que

$$\boldsymbol{\alpha}_i' \mathbf{y} = \sum_{j=1}^{n} z_j \boldsymbol{\alpha}_i' \boldsymbol{\alpha}_j = z_i \boldsymbol{\alpha}_i' \boldsymbol{\alpha}_i = z_i \quad \forall i = 1, \ldots, n$$

Luego, si definimos a $\mathbf{T}$ como la matriz que tiene filas $\boldsymbol{\alpha}_i'$, entonces

$$\mathbf{z} = \mathbf{T}\mathbf{y}$$

Observemos que bajo el modelo Ω

$$E(z_i) = \begin{cases} \boldsymbol{\alpha}_i' \boldsymbol{\eta} = \xi_i & \text{si } 1 \leq i \leq r \\ 0 & \text{si } r+1 \leq i \leq n \end{cases}$$

$$\Sigma_{\mathbf{z}} = \mathbf{T} \Sigma_{\mathbf{y}} \mathbf{T}' = \sigma^2 \mathbf{I}$$

Bajo el modelo ω, tenemos que $\boldsymbol{\eta} = E(\mathbf{Y}) \in \mathcal{V}_{r-q}$, es decir $\boldsymbol{\alpha}_i'\boldsymbol{\eta} = 0$ para $i = 1, \ldots, q$.

$$E(z_i) = \begin{cases} 0 & \text{si } 1 \leq i \leq q \\ \xi_i & \text{si } q+1 \leq i \leq r \\ 0 & \text{si } r+1 \leq i \leq n \end{cases}$$

Entonces podemos escribir:

$$\begin{aligned} \Omega &: \mathbf{z} \sim N_n(\boldsymbol{\xi}, \sigma^2\mathbf{I}) \quad \xi_i = 0 \quad i \geq r+1 \\ \omega &: \mathbf{z} \sim N_n(\boldsymbol{\xi}, \sigma^2\mathbf{I}) \quad \xi_i = 0 \quad 1 \leq i \leq q \text{ y } i \geq r+1 \end{aligned}$$

Utilizando la notación de Scheffé tendremos

$$\begin{aligned} \mathcal{S}_\Omega &= \|\mathbf{y} - \widehat{\boldsymbol{\eta}}\|^2 = \sum_{i=r+1}^{n} z_i^2 \\ \mathcal{S}_\omega &= \|\mathbf{y} - \widehat{\boldsymbol{\eta}}_\omega\|^2 = \sum_{i=1}^{q} z_i^2 + \sum_{i=r+1}^{n} z_i^2 \end{aligned}$$

y además

$$\mathcal{S}_\omega - \mathcal{S}_\Omega = \sum_{i=1}^{q} z_i^2$$

$$\|\mathbf{y}-\widehat{\boldsymbol{\eta}}_\omega\|^2-\|\mathbf{y}-\widehat{\boldsymbol{\eta}}\|^2=\|\widehat{\boldsymbol{\eta}}-\widehat{\boldsymbol{\eta}}_\omega\|^2=\sum_{i=1}^{q} z_i^2$$

Además, bajo H tenemos que $\frac{\mathcal{S}_\omega-\mathcal{S}_\Omega}{\sigma^2}\sim\chi_q$ independiente de s^2 y en consecuencia

$$\frac{1}{q}\frac{\mathcal{S}_\omega-\mathcal{S}_\Omega}{s^2}\sim F_{q,n-r}$$

$$\text{Rechazamos } H \text{ si } \frac{1}{q}\frac{\mathcal{S}_\omega-\mathcal{S}_\Omega}{s^2}>F_{q,n-r,\alpha}$$

Observación: Puede demostrarse que este test es equivalente al tests de F ya visto.

Interpretación Geométrica.