Modelos de Rango Incompleto
Modelos de Clasificación Simple
Diseño Completamente Aleatorizado

Johan Steven Aparicio

Escuela de Estadística
Facultad de Ingeniería - Universidad del Valle
johan.aparicio@correounivalle.edu.co

Mayo, 2022

Tabla de Contenido

Teorema 6

Hay exactamente t funciones estimables linealmente independientes; donde t es el rango de X.

Se debe demostrar que hay exactamente t vectores linealmente independientes para los cuales se cumple que $X'Xr = \lambda$.

Sean q vectores que satisfacen $X'Xr = \lambda$,

$$\lambda_1, \lambda_2, ..., \lambda_q, (q > t)$$

Demostrarlo

Teorema 7

Con cualquier solución de β se obtiene la misma estimación de σ^2.

Dado que $X'X\hat{\beta} = X'Y$ tiene infinitas soluciones, sean $\hat{\beta}$ y $\tilde{\beta}$ dos soluciones cualesquiera, entonces:

$$\begin{aligned} X'X\hat{\beta} &= X'Y \\ -X'X\tilde{\beta} &= X'Y \end{aligned}$$

$$\overline{\hphantom{X'X(\hat{\beta}-\tilde{\beta}) = X'Y}}$$

$$X'X(\hat{\beta} - \tilde{\beta}) = 0$$

$$X'X(\hat{\beta} - \tilde{\beta}) = X'(X\hat{\beta} - X\tilde{\beta})$$

Premultiplicado por $(\hat{\beta} - \tilde{\beta})'$

$$\begin{aligned} (\hat{\beta} - \tilde{\beta})'X'(X\hat{\beta} - X\tilde{\beta}) &= 0 \\ (X\hat{\beta} - X\tilde{\beta})'(X\hat{\beta} - X\tilde{\beta}) &= 0 \end{aligned}$$

La anterior ecuación es una forma cuadrática que tiene como matriz asociada a la matriz identidad I, de tal forma que su producto es igual a cero si y solo si,

$$(X\hat{\beta} - X\tilde{\beta}) = 0$$

De aquí que:

$$X\hat{\beta} = X\tilde{\beta}$$

Lo que implica que con cualquier solución $\hat{\beta}$ o $\tilde{\beta}$ del sistema, se obtendrá la misma estimación de $\hat{\sigma}^2$, donde:

$$\hat{\sigma}^2 = \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{n - t} = \frac{SCE}{\text{g.l. del error}}$$

La hipótesis que interesa en este tipo de experimentos y modelos es la de la igualdad de efectos sobre la variable respuesta de los niveles del factor que se controla.

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_t \text{ vs } H_a : \tau_i \neq \tau_j, i \neq j$$

Una manera de reexpresar esta hipótesis en términos de combinaciones lineales de los parámetros (contrastes) del factor que se estudia es:

$$\begin{aligned} H_0 &: \tau_1 = \tau_2 = 0 \\ &: \tau_1 + \tau_2 - 2\tau_3 = 0 \\ &\vdots \\ &: \tau_1 + \tau_2 + \cdots + \tau_{t-1} + (t-1)\tau_t = 0 \end{aligned}$$

Estas combinaciones lineales son estimables insesgadamente y los vectores asociados a ellas son linealmente independientes, lo que determina un conjunto de contrastes ortogonales.

Con la finalidad de contrastar H_0 se debe encontrar solución al SEN. El SEN en este tipo de modelos y experimentos contiene t + 1 ecuaciones, con t ecuaciones linealmente independientes $(X'X\hat{\beta} = X'Y)$.

Si a este sistema agregamos otra ecuación $(\sum r_i\hat{\tau}_i = 0)$ independiente a las demás, se obtendría:

$$
\begin{array}{cccccccc}
n\hat{\mu} & +r_1\hat{\tau}_1 & +r_2\hat{\tau}_2 & +\cdots & +r_t\hat{\tau}_t & = & Y_{..} \\
r_1\hat{\mu} & +r_1\hat{\tau}_1 & & & & = & Y_{1.} \\
r_2\hat{\mu} & & +r_2\hat{\tau}_2 & & & = & Y_{2.} \\
\vdots & & & & & & \\
r_t\hat{\mu} & & & & +r_t\hat{\tau}_t & = & Y_{t.} \\
 & r_1\hat{\tau}_1 & +r_2\hat{\tau}_2 & +\cdots & +r_t\hat{\tau}_t & = & 0
\end{array}
$$

Este sistema produce una única solución:

$$\begin{aligned}
\hat{\mu} &= \bar{Y}_{..} \\
\hat{\tau}_1 &= \bar{Y}_{1.} - \hat{\mu} = \bar{Y}_{1.} - \bar{Y}_{..} \\
&\vdots \\
\hat{\tau}_t &= \bar{Y}_{t.} - \bar{Y}_{..}
\end{aligned}$$

Con base a esta solución se construye la tabla ANOVA.

La tabla ANOVA se obtiene desarrollando los siguientes cálculos:

$$R(\beta) = \hat{\beta}'X'Y = \begin{bmatrix} \bar{Y}_{..} & \bar{Y}_{1.} - \bar{Y}_{..} & \cdots & \bar{Y}_{t.} - \bar{Y}_{..} \end{bmatrix} \begin{bmatrix} Y_{..} \\ Y_{1.} \\ \vdots \\ Y_{t.} \end{bmatrix}$$

Resolver

$$= R(\mu) + R(\tau)$$

La reducción de τ corregida por μ:

$$R(\tau/\mu) = R(\beta) - R(\mu)$$

Tabla de análisis de varianza para el modelo con un solo factor

Fuente de Variabilidad	Grados de Libertad	Suma de Cuadrados	Cuadrado Medio	F_0
τ	$t-1$	$\sum_{i=1}^{t}\frac{y_{i.}^2}{r_i}-\frac{y_{..}^2}{n}$	$\frac{SC_\tau}{t-1}$	$\frac{CM_\tau}{CM_{Error}}$
Error	$n-t$	$Y'Y-\sum_{i=1}^{t}\frac{y_{i.}^2}{r_i}$	$\frac{SC_{Error}}{n-t}$	
Total	$n-1$	$Y'Y-\frac{y_{..}^2}{n}$		

En una estación experimental de propiedad de la CVC, se realizó un experimento para estudiar los efectos comparativos de tres tipos de fertilizantes en la producción de frijol. En la tabla siguiente se presentan los datos obtenidos sobre producción (en kilogramos por parcela).

Fertilizante	Replicaciones				
	1	2	3	4	5
Gallinaza (G)	79	80	83	82	81
Cal Agricola (CA)	80	78	79	80	82
Calfos (F)	82	84	83	84	82

1. Plantee:
    - Plantee los objetivos del estudio.
    - Plantee la hipótesis del estudio.
    - Determine factores, niveles y tratamientos.
    - Determine cuál es la unidad experimental.
2. Encuentre:
    - X , Y
    - $X'Y$, $X'X$
    - $\hat{\beta}' = \begin{bmatrix} \hat{\mu} & \hat{\vec{\tau}}' \end{bmatrix}$
    - $\hat{\sigma}^2$
    - $\hat{\tau}_1 - \hat{\tau}_2$
3. Construya la tabla de Análisis de Varianza, plantee las hipótesis y analice los resultados obtenidos.