

DISEÑO DE BLOQUES COMPLETOS RANDOMIZADO O DISEÑO DE BLOQUE COMPLETO AL AZAR

Un diseño de Bloques Completos Randomizado es aquel cumple con las siguientes condiciones:

- 1) Las unidades experimentales se distribuyen en grupos o bloques, de manera tal que las unidades experimentales dentro de cada bloque sean relativamente homogéneas y que el número de unidades experimentales dentro de un bloque sea igual al número de tratamientos por investigar; y
- 2) Los tratamientos se asignan al azar a las unidades experimentales dentro de cada bloque.

Ejemplo

En los experimentos agrícolas, los bloques puede estar constituido por grupos de parcelas relativamente homogéneas que puede ser agrupados de acuerdo a gradiente de fertilidad, otro porque se encuentra en una pendiente.

Ventajas

Este diseño tiene muchas ventajas, tales como

- 1.- En general es posible agrupar las unidades experimentales de modo que se logre mayor precisión con respecto a un Diseño completamente al azar
- 2.- La única restricción sobre el número de tratamiento por bloque y tratamiento es la disponibilidad de unidades experimentales
- 3.- Si se pierde información de todo un bloque o por contratiempo los datos de un bloque completo es inutilizable estos datos puede omitirse, porque el resto mantiene la misma estructura de un diseño de bloques completos al azar.
- 4.- Si se pierde información de algunas de las unidades estas puede estimarse.

Modelo Aditivo Lineal

El modelo aditivo Lineal del Diseño de Bloques Completo al Azar con una observación por unidad experimental, La observación Y_{ij} puede representarse por el modelo siguiente:

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}; \quad i = 1, 2, \dots, t \text{ y } j = 1, 2, \dots, b$$

donde:

Y_{ij} : es la respuesta obtenida de la unidad experimental del j -ésimo bloque sujeta al

- tratamiento i .
- μ : El efecto de la media común.
- τ_i : El verdadero efecto del i -ésimo tratamiento.
- β_j : El verdadero efecto del j -ésimo bloque.
- ε_{ij} : Es una variable aleatoria no observable llamado error

Para el proceso de inferencia se asume que ε_{ij} es una variable aleatoria independiente que se distribuye normalmente con media cero y variancia común σ^2 .

Modelo I (efectos fijos)

Sea asume que los niveles de los factores son fijados por el investigador y estos efectos son desviaciones con respecto a la media. Entonces se cumple:

$$\sum_{i=1}^t \tau_i = 0, \quad \sum_{j=1}^b \beta_j = 0$$

Modelo II (efectos aleatorios)

Los niveles de los factores son elegidos aleatoriamente de poblaciones grandes. Entonces los τ_i son variables aleatorias independientes distribuidas normalmente con media cero y variancia σ_τ^2 , los β_j son variables aleatorias independientes distribuidas normalmente con media cero y variancia σ_β^2 ,

Modelo III (Modelo mixto)

Los niveles de los tratamientos son fijados por el investigador y los niveles de los bloques son elegidos al azar en este caso se cumple que

$$\sum_{i=1}^t \tau_i = 0;$$

y los β_j son variables aleatorias independientes distribuidas normalmente con media cero y variancia σ_β^2 ,

Cuadro de Datos

tratamientos	Bloques				Total
	1	2	...	b	
1	Y_{11}	Y_{12}	...	Y_{1b}	$Y_{1\cdot}$
2	Y_{21}	Y_{22}	...	Y_{2b}	$Y_{2\cdot}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
t	Y_{t1}	Y_{t2}	...	Y_{tb}	$Y_{t\cdot}$
Total	$Y_{\cdot 1}$	$Y_{\cdot 2}$...	$Y_{\cdot b}$	$Y_{\cdot\cdot}$

Donde : $Y_{i\cdot} = \sum_{j=1}^b Y_{ij}$, para $i = 1, 2, \dots, t$; $Y_{\cdot j} = \sum_{i=1}^t Y_{ij}$, para $j = 1, 2, \dots, b$;

$$Y_{\cdot\cdot} = \sum_{i=1}^t \sum_{j=1}^b Y_{ij}$$

Estimación de Parámetro para el Modelo I

Los estimadores de los parámetros pueden ser encontrados aplicando el método de los mínimos cuadrados. Con este método se obtiene:

$$\hat{\mu} = \bar{Y}_{\cdot\cdot} = \frac{1}{tb} \sum_{i=1}^t \sum_{j=1}^b Y_{ij}; \quad \hat{\tau}_i = \bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}, \text{ para } i = 1, 2, \dots, t;$$

$$\hat{\beta}_j = \bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot}, \text{ para } j = 1, 2, \dots, b$$

Residual o residuo

$$e_{ij} = Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot}$$

ANÁLISIS DE VARIANCIA

La variación total puede ser descompuesta de la siguiente forma:

$$\sum_{i=1}^t \sum_{j=1}^b (Y_{ij} - \bar{Y}_{\cdot\cdot})^2 = \sum_{i=1}^t \sum_{j=1}^b (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 + \sum_{i=1}^t \sum_{j=1}^b (\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})^2 + \sum_{i=1}^t \sum_{j=1}^b (Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2$$

donde:

$$SCTotal = \sum_{i=1}^t \sum_{j=1}^b (Y_{ij} - \bar{Y}_{\cdot\cdot})^2 = \sum_{i=1}^t \sum_{j=1}^b Y_{ij}^2 - \frac{Y_{\cdot\cdot}^2}{bt} \text{ es la medida de la variación total.}$$

$SCTrat = \sum_{i=1}^t \sum_{j=1}^b (\bar{Y}_{i\cdot} - \bar{Y}_{..})^2 = \sum_{i=1}^t \frac{Y_{i\cdot}^2}{b} - \frac{Y_{..}^2}{bt}$ es una medida de la variación entre tratamientos.

$SCBloq = \sum_{i=1}^t \sum_{j=1}^b (\bar{Y}_{\cdot j} - \bar{Y}_{..})^2 = \sum_{j=1}^b \frac{Y_{\cdot j}^2}{t} - \frac{Y_{..}^2}{tb}$ es una medida de la variación existente entre bloques

$SCE = \sum_{i=1}^t \sum_{j=1}^b (Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{..})^2 = SCTotal - SCTrat - SCBloq$, es la variación debido a otros factores no considerados en el modelo.

Cuadrados Medios

Los cuadrados Medios se definen como el cociente entre la suma de los cuadrados y sus respectivos grados de libertad:

$$CMBloq = \frac{SCBloq}{b-1}, \quad CMTrat = \frac{SCTrat}{t-1}, \quad CME = \frac{SCE}{(b-1)(t-1)}$$

Luego, se tiene el siguiente cuadro de ANVA

Fuente de Variación	SC	GL	CM	Cuadrados Medios Esperados	
				Modelo I	Modelo II
Bloques	SCBloq	b-1	CMBloq	$\sigma^2 + \frac{t}{b-1} \sum_{j=1}^b \beta_j^2$	$\sigma^2 + t\sigma_\beta^2$
Tratamientos	SCTrat	t-1	CMTrat	$\sigma^2 + \frac{b}{t-1} \sum_{i=1}^t \tau_i^2$	$\sigma^2 + b\sigma_\tau^2$
Error	SCE	(b-1)(t-1)	CME	σ^2	σ^2
Total	SCTotal	bt-1			

Prueba de Hipótesis (Modelo I)

$H_p : \tau_1 = \tau_2 = \dots = \tau_t = 0$ El cual es equivalente $H_p : \mu_1 = \mu_2 = \dots = \mu_t$
 $H_a : \tau_i \neq 0$, para al menos un i $H_a : \text{al menos dos } \mu_i \text{ son diferentes}$
Nivel de Significación α

$$F_c = \frac{CMTrat}{CME} \sim F_{(t-1, (b-1)(t-1))} / \text{si la Hp es cierta}$$

Nota: Como los bloques son fijados y no cumple con el principio de aleatorización no se puede realizar pruebas de hipótesis sobre los efectos de bloques. En lugar de esto se puede encontrar eficiencia relativa respecto a un diseño completamente al azar, el cual se define:

$$ER = \frac{SCB_{loq} + b(t-1)CME}{\frac{tb-1}{CME}}$$

Si $ER > 1$ entonces el Diseño de Bloques Completos al Azar es más eficiente que un Diseño Completamente al azar.

Ejemplo: Se llevó a cabo un experimento para señalar los méritos de 5 gasolinas. Debido a que es inevitable la variación en eficiencia de vehículo a vehículo, la prueba se realizó un experimento con 5 automóviles, que de aquí en adelante llamaremos bloques. Se dispone de las siguientes descripciones de las 5 gasolinas:

- A: Control
- B: Control + aditivo X elaborado por la compañía I
- C: Control + aditivo Y elaborado por la compañía I
- D: Control + aditivo U elaborado por la compañía II
- E: Control + aditivo V elaborado por la compañía II

Los tipos de gasolinas fueron probadas en cada carro en orden aleatorio. Los datos, en Km/litros, se dan continuación:

Bloques (vehículo)						
Tratamiento Gasolina	1	2	3	4	5	Total
A	8	7	6	6	7	34
B	10	9	8	7	9	43
C	8	8	9	9	10	44
D	9	8	8	8	7	40
E	10	9	8	7	9	43
Total	45	41	39	37	42	204

$$\sum_{i=1}^5 \sum_{j=1}^5 Y_{ij}^2 = 1696, \quad \sum_{i=1}^5 Y_{i.}^2 = 8390, \quad \sum_{j=1}^5 Y_{.j}^2 = 8360$$

$$SCB_{loq} = \sum_{j=1}^b \frac{Y_{.j}^2}{t} - \frac{Y_{..}^2}{tb} = \frac{8360}{5} - \frac{(204)^2}{25} = 7.36$$

$$SCTrat = \sum_{i=1}^t \frac{Y_{i.}^2}{b} - \frac{Y_{..}^2}{bt} = \frac{8390}{5} - \frac{(204)^2}{25} = 13.36$$

$$SCTotal = \sum_{i=1}^t \sum_{j=1}^b Y_{ij}^2 - \frac{Y_{..}^2}{bt} = 1696 - \frac{(204)^2}{25} = 31.36$$

$$SCE = SCTotal - SCTrat - SCBloq = 31.36 - 7.36 - 13.36 = 10.64$$

Fuente de Variación	SC	GL	CM	Fc
Carros	7.36	4	1.84	
Gasolinas	13.36	4	3.34	5.0226
Error	10.64	16	0.665	
Total	31.36	24		

$$H_p : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

$$H_a : \text{al menos dos } \mu_i \text{ son diferentes}$$

$$\alpha = 0.05$$

$$F_c = \frac{CMTrat}{CME} = \frac{3.34}{0.665} = 5.0226$$

$$F_{(0.05,4,4)} = 3.01, \text{ como } F_c > F_{(0.05,4,4)}, \text{ se rechaza la } H_p.$$

```

> gasolina<-read.table("gasolina.txt",header=T)
> rendimiento<-gasolina[,1]
> vehiculo<-factor(gasolina[,2])
> tipos<-gasolina[,3]
> modeg<-lm(rendimiento~vehiculo+tipos)
> anva<-anova(modeg)
> anva
Analysis of Variance Table

Response: rendimiento
          Df Sum Sq Mean Sq F value    Pr(>F)
vehiculo   4   7.360    1.840    2.7669 0.063664 .
tipos      4  13.360    3.340    5.0226 0.008138 **
Residuals 16  10.640     0.665
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> cm<-anva$Mean
> sc<-anva$Sum
> nt<-tapply(rendimiento,vehiculo,length)
> t<-nt[1]
> nb<-tapply(rendimiento,tipos,length)
> b<-nb[1]
> ER<-((sc[1]+b*(t-1)*cm[3])/(t*b-1))/cm[3]
> ER

```

1
1.294486
Como $ER > 1$ el uso de bloques ha sido efectivo para reducir el error experimental

Comparaciones Múltiples (Modelo I)

Si se define que $\mu_i = \mu + \tau_i$ entonces un estimador de μ_i está dado por

$$\hat{\mu}_i = \bar{Y}_{i\cdot} = \frac{1}{b} \sum_{j=1}^b Y_{ij},$$

la variancia de $\bar{Y}_{i\cdot}$, para $i = 1, 2, \dots, t$ está dado por:

$$\text{var}[\bar{Y}_{i\cdot}] = \frac{\sigma^2}{b}, \text{ y su estimado está dado por: } S_{\bar{Y}_{i\cdot}}^2 = \frac{CME}{b}$$

la variancia de $\bar{Y}_{i\cdot} - \bar{Y}_{l\cdot}$, para $i \neq l$ y $i, l = 1, 2, \dots, t$, está dado por:

$$\text{var}[\bar{Y}_{i\cdot} - \bar{Y}_{l\cdot}] = \frac{2\sigma^2}{b} \text{ y su estimado está dado por } S_{\bar{Y}_{i\cdot} - \bar{Y}_{l\cdot}}^2 = \frac{2CME}{b}$$

Prueba de t

Hipótesis

Caso A Bilateral	Caso B Unilateral a la Derecha	Caso C Unilateral a La Izquierda
$H_p : \mu_i - \mu_l = k$	$H_p : \mu_i - \mu_l \leq k$	$H_p : \mu_i - \mu_l \geq k$
$H_a : \mu_i - \mu_l \neq k$	$H_a : \mu_i - \mu_l > k$	$H_a : \mu_i - \mu_l < k$

Para $i \neq l$; $i, l = 1, 2, \dots, t$

Nivel de significación α

Estadística de prueba:

$$t_c = \frac{\bar{Y}_{i\cdot} - \bar{Y}_{l\cdot} - k}{S_{\bar{Y}_{i\cdot} - \bar{Y}_{l\cdot}}} \sim t_{(gle)} / H_p \text{ es verdadera}$$

Decisión	Caso A	Caso B	Caso C
Se Acepta H_p	$t_{\left(\frac{\alpha}{2}, gle\right)} \leq t_c \leq t_{\left(1-\frac{\alpha}{2}, gle\right)}$	$t_c \leq t_{(1-\alpha; gle)}$	$t_c \geq t_{(\alpha; gle)}$
Se Rechaza H_p	$t_c < t_{\left(\frac{\alpha}{2}, gle\right)} \text{ ó } t_c > t_{\left(1-\frac{\alpha}{2}, gle\right)}$	$t_c > t_{(1-\alpha; gle)}$	$t_c < t_{(\alpha; gle)}$

Diferencia Mínima de Significación (DMS), también se le conoce con el nombre de diferencia límite de significación

$$H_p : \mu_i = \mu_l$$

$$H_a : \mu_i \neq \mu_l$$

Para $i \neq l, i, l = 1, 2, \dots, t$

Nivel de significación α

Entonces si definimos

$$DMS(i, l) = t_{\left(1 - \frac{\alpha}{2}, GLE\right)} S_{\bar{Y}_{i\bullet} - \bar{Y}_{l\bullet}}$$

Luego, un criterio para examinar si existe diferencia significativa entre medias de tratamiento se puede usar este criterio de la diferencia mínima significativa ($DMS(i, l)$).

Esto es, se rechaza H_0 si

$$|\bar{Y}_{i\bullet} - \bar{Y}_{l\bullet}| > DMS(i, l)$$

Para $i \neq l, i, l = 1, 2, \dots, t$

Ejemplo: Con los datos del ejemplo de gasolina, suponga que fue planeado realizar la comparación entre la gasolina D y E. Realice la prueba de t a un nivel de significación $\alpha = 0.05$, para realizar esta comparación

Las medias de los rendimientos está dado por:

$$\bar{Y}_{A\bullet} = 6.8, \quad \bar{Y}_{B\bullet} = 8.6, \quad \bar{Y}_{C\bullet} = 8.8, \quad \bar{Y}_{D\bullet} = 8.0, \quad \bar{Y}_{E\bullet} = 8.6$$

$$H_p : \mu_D = \mu_E \text{ o } H_p : \mu_D - \mu_E = 0$$

$$H_a : \mu_D \neq \mu_E \text{ o } H_a : \mu_D - \mu_E \neq 0$$

$$\alpha = 0.05 \quad T_{(0.975, 16)} = 2.22, \quad S_{\bar{Y}_{D\bullet} - \bar{Y}_{E\bullet}}^2 = \frac{2CME}{b} = \frac{2(0.665)}{5} = 0.266$$

$$t_c = \frac{\bar{Y}_{D\bullet} - \bar{Y}_{E\bullet} - k}{S_{\bar{Y}_{D\bullet} - \bar{Y}_{E\bullet}}} = \frac{8 - 8.6 - 0}{\sqrt{0.266}} = -1.16335. \text{ Se acepta } H_p$$

Con lenguaje R

```
> model<-lm(rendimiento~vehiculo+tipos)
> model
```

Call:

```
lm(formula = rendimiento ~ vehiculo + tipos)
```

Coefficients:

(Intercept)	vehiculo2	vehiculo3	vehiculo4	vehiculo5	tiposb
7.64	-0.80	-1.20	-1.60	-0.60	1.80
tiposc	tiposd	tipose			
2.00	1.20	1.80			

El lenguaje R da unos estimados de efectos para los dos factores. Para el caso de gasolina viene hacer la diferencia de la medias de tratamientos de B, C, D y E con respecto a la media de tratamiento de A, respectivamente.

```
> efect<-model$coefficients
> dmedia<-efect-efect[9]
> dmedia<-dmedia[8]
> dmedia
tiposd
-0.6
> tc<-dmedia/sqrt(cm[3]*(2/5))
> tc
tiposd
-1.16335
> > pvalue<-2*pt(tc,df.residual(model))
> pvalue
tiposd
0.2617441
```

Se acepta H_0

Prueba de Tukey-Cramer (Tukey HSD)

Planteamiento de hipótesis

$$H_0 : \mu_i = \mu_l$$

$$H_a : \mu_i \neq \mu_l \quad \text{Para } i \neq l, i, l = 1, 2, \dots, t$$

Nivel de significación α

Cálculo del Valor Crítico:

$$w = q_\alpha(t, GLE) \frac{1}{\sqrt{2}} S_{\bar{Y}_{i\cdot} - \bar{Y}_{l\cdot}}$$

donde:

$q_\alpha(t, GLE)$ = amplitud estudiantizada para la prueba de Tukey

t = número de tratamiento a comparar

GLE = Grados de libertad del error

Se rechaza H_0 aun nivel de significación α , si

$$|\bar{Y}_{i\cdot} - \bar{Y}_{l\cdot}| > w$$

Ejemplo: Con los datos del ejemplo de gasolina, realice la prueba de Tukey a un nivel de significación $\alpha = 0.05$, para realizar esta comparación

$$H_0 : \mu_i = \mu_{i'}$$

$$H_a : \mu_i \neq \mu_{i'} \text{ para } i, i' = A, B, C, D, E, i \neq i'$$

$\alpha = 0.05$,

```
> library(multcomp)
> nd<-tapply(rendimiento,tipos,length)
> mt<-contrMat(nd,type="Tukey")
> simint(rendimiento~tipos,type="Tukey",cmatrix=mt)
```

Simultaneous confidence intervals: user-defined contrasts

Call:

```
simint.formula(formula = rendimiento ~ tipos, type = "Tukey",
  cmatrix = mt)
```

95 % confidence intervals

	Estimate	2.5 %	97.5 %
b-a	1.8	0.005	3.595
c-a	2.0	0.205	3.795
d-a	1.2	-0.595	2.995
e-a	1.8	0.005	3.595
c-b	0.2	-1.595	1.995
d-b	-0.6	-2.395	1.195
e-b	0.0	-1.795	1.795
d-c	-0.8	-2.595	0.995
e-c	-0.2	-1.995	1.595
e-d	0.6	-1.195	2.395

Prueba de Dunnett (comparaciones de todas las medias de tratamientos con un control o testigo)

$$H_p : \mu_i = \mu_1$$

$$H_a : \mu_i \neq \mu_1, \text{ para } i = 2, \dots, t$$

Donde: μ_1 es la media del tratamiento testigo o de control

Nivel de significación α

Valor Crítico:

$$d' = t_{Dunnett}(\alpha, p, GLE) S_{\bar{Y}_{i\cdot} - \bar{Y}_{1\cdot}}, \text{ para } i = 2, \dots, t$$

donde :

$$t_{Dunnett}(\alpha, t, GLE) = t \text{ de Dunnett con un nivel de significación } \alpha.$$

p = número de tratamiento a comparar con el control

GLE = Grados de libertad del error

Se rechaza H_0 aun nivel de significación α , si

$$|\bar{Y}_{i\cdot} - \bar{Y}_{1\cdot}| > d', \text{ para } i = 2, \dots, t$$

Ejemplo: En el ejemplo de la gasolina suponga que A es el tratamiento Control. Realice la prueba de Dunnett a un nivel $\alpha = 0.05$

$$H_p : \mu_i = \mu_A$$

$$H_a : \mu_i \neq \mu_A, \text{ para } i = B, C, D, E$$

$$\bar{Y}_{A\bullet} = 6.8, \quad \bar{Y}_{B\bullet} = 8.6, \quad \bar{Y}_{C\bullet} = 8.8, \quad \bar{Y}_{D\bullet} = 8.0, \quad \bar{Y}_{E\bullet} = 8.6;$$

$$S_{\bar{Y}_{i\bullet} - \bar{Y}_{A\bullet}}^2 = \frac{2CME}{b} = \frac{2(0.665)}{5} = 0.266$$

$$d' = t_{Dunnett}(0.5, 4, 16) S_{\bar{Y}_{i\bullet} - \bar{Y}_{A\bullet}} = (2.34)(\sqrt{0.266}) = 1.206859$$

Comparación	$ \bar{Y}_{i\bullet} - \bar{Y}_{A\bullet} $	$d' = t_{Dunnett}(0.5, 4, 16) S_{\bar{Y}_{i\bullet} - \bar{Y}_{A\bullet}}$	
B-A	1.8	1.206859	significativo
C-A	2.0	1.206859	significativo
D-A	1.2	1.206859	No significativo
E-A	1.8	1.206859	significativo

```
> mt<-contrMat(nd,type="Dunnett")
> simint(rendimiento~tipos,type="Dunnett",cmatrix=mt)
```

Simultaneous confidence intervals: user-defined contrasts

Call:

```
simint.formula(formula = rendimiento ~ tipos, type = "Dunnett",
  cmatrix = mt)
```

95 % confidence intervals

	Estimate	2.5 %	97.5 %
b-a	1.8	0.209	3.391
c-a	2.0	0.409	3.591
d-a	1.2	-0.391	2.791
e-a	1.8	0.209	3.391

Prueba de t con contraste:

Suponga que se desea probar la Hipótesis

$$H_p : \sum_{i=1}^t C_i \mu_i = 0$$

$$H_a : \sum_{i=1}^t C_i \mu_i \neq 0$$

a un nivel de significación α

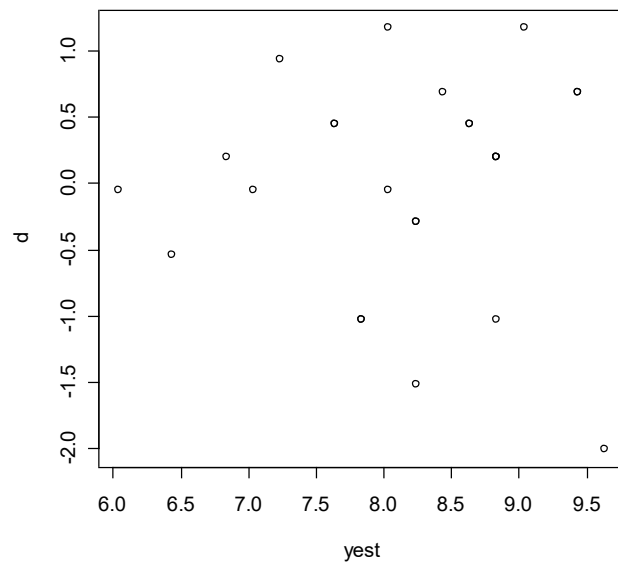
Estadística de Prueba

$$t = \frac{Q}{\sqrt{bCME \sum_{i=1}^t C_i^2}} \sim t_{(GLE)} / H_0 \text{ es verdadera, siendo } Q = \sum_{i=1}^t C_i Y_{i.} = \sum_{i=1}^t b C_i \bar{Y}_{i.}$$

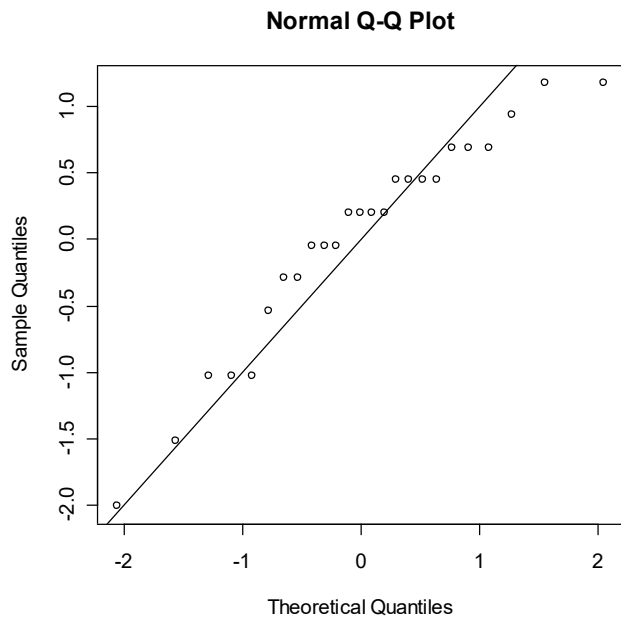
Luego se acepta H_0 si $t_{\left(\frac{\alpha}{2}, GLE\right)} \leq t_c \leq t_{\left(1-\frac{\alpha}{2}, GLE\right)}$, caso contrario se rechaza.

Análisis de residuales

```
> e<-residuals(modeg)
> yest<-predict(modeg)
> d<-e/sqrt(deviance(modeg)/df.residual(modeg))
> plot(yest,d)
```



```
> qqnorm(d)
> qqline(d)
```



```
> shapiro.test(d)
```

Shapiro-Wilk normality test

data: d

W = 0.9362, p-value = 0.1207

De acuerdo al gráfico de residuales es probable que no se cumpla con el supuesto de homogeneidad de variancia una alternativa es realizar transformaciones.