Diseño de Experimentos - 3007340
DOE - Diseño factoriales 2^k - Parte II: Diseños no replicados

Nelfi González Alvarez
Profesora Asociada Escuela de Estadística
e-mail: ngonzale@unal.edu.co

Facultad de Ciencias, Universidad Nacional de Colombia Sede Medellín

Escuela de Estadística
Semestre 02 de 2021

Contenido I

Contenido

Introducción

- Con k grande puede resultar altamente costoso replicar todos los 2^k tratamientos posibles.
- *Por otra parte, muchos efectos de interacción de alto orden en la práctica no son significativos, por tanto replicar tratamientos conduce a un exceso de información.*
- De acuerdo a Gutiérrez y de la Vara Salazar (2012), en la mayoría de experimentos, con 32 corridas se pueden estudiar los efectos de interés, e incluso, 16 corridas pueden ser suficientes en una primera etapa de experimentación.
- *Con solo una réplica por tratamiento, interacciones de orden alto pueden usarse para construir un SSE aproximado así como sus grados de libertad.*

Tabla 1: Réplicas recomendadas en la familia de diseños 2^k, según Gutiérrez y de la Vara Salazar (2012).

Diseño	Réplicas recomendadas	Número de corridas
2^2	3 ó 4	12, 16
2^3	2	16,
2^4	1 ó 2	16, 32
2^5	Fracción 2^{5-1}	16, 32
2^6	Fracción 2^{6-2} ó 2^{6-1}	16, 32
2^7	Fracción 2^{7-3} ó 2^{7-2}	16, 32

Contenido

Estimación saturada del modelo

- Cada tratamiento posible del factorial 2^k es observado una vez, por tanto el total de observaciones disponibles es $N = 2^k$.
- *El total de efectos posibles: pricipales, más todas las interacciones dobles, más todas las interacciones triples, etc., es igual a $2^k - 1$, de modo que en el MRLM equivalente con variables codificadas, se tienen $p = 2^k$ parámetros, contando al intercepto y a todos los efectos de tatamientos posibles.*
- El modelo anterior puede ser ajustado con un tamaño de muestra igual al número de parámetros a estimar, y el ajuste es saturado, es decir con 0 grados de libertad para el error, y por tanto, no es posible calcular un SSE y menos tener una estimación de la varianza σ^2.
- *Los parámetros también pueden estimarse mediante los contrastes de totales de tratamientos como en el caso del 2^k replicado, teniendo en cuenta que cada total de tratamiento es igual al valor de la única observación del tratamiento ($n = 1$), y por tanto, las estimaciones de los respectivos efectos y las sumas de cuadrados asociadas, son de la forma*

$$\widehat{\alpha}_{ABC\cdots K} = \frac{(\text{Contraste totales})_{ABC\cdots K}}{2^{k-1}}. \tag{1}$$

$$SS(ABC\cdots K) = \frac{[(\text{Contraste totales})_{ABC\cdots K}]^2}{2^k} = 2^{k-2}\left(\widehat{\alpha}_{ABC\cdots K}\right)^2. \tag{2}$$

Contenido

Aproximación del SSE en un 2^k no replicado

1. *Suponer de antemano que las interacciones de tres o más factores no son significativas y enviarlas directamente al error.* Sin embargo verificar esto mediante técnicas gráficas
   - diagramas de Pareto de efectos
   - el gráfico de Daniel
2. *Con los efectos despreciados aproximar el SSE y sus grados de libertad:* Sumar las sumas de cuadrados de estos efectos y sus grados de libertad. Esta aproximación tiene su riesgo dado que no se basa en replicaciones.

Nota 3.1

De acuerdo a Gutiérrez y de la Vara Salazar (2012),

- *Para saber si el MSE resultante es apropiado, compararlo con la varianza que típicamente se ha observado en la respuesta en experimentos previos o en el proceso.*
- *Se deben mandar al error al menos 8 efectos pequeños para tener mayores probabilidades de que esté bien aproximado el SSE.*

Escogencia de efectos a mandar al SSE

Tenga en cuenta que

- *Si manda al error efectos que realmente son significativos, inflará el MSE y restará potencia en la detección de efectos*
- *Si se subestima σ^2, en los tests ANOVA podría detectar efectos falsamente como significativos*
- *Un MSE muy alejado del σ^2 histórico es síntoma de su mala estimación*
- *No se debe aplicar técnicas de aproximación del MSE en experimentos con pocos efectos: los 2^2, 2^3, estos deben correrse con réplicas.*

Tipos de efectos según Gutiérrez y de la Vara Salazar (2012):

- *Los que claramente son significativos y deben retenerse*
- *los que claramente no afectan y deben mandarse al error*
- *efectos intermedios sobre los cuales no es claro si son significativos o no (los que resultan problemáticos para tomar la decisión de mandarlos o no al error)*

Gráfico de probabilidad normal de efectos (Gráfico de Daniel)

Permite selección de efectos a retener bajo los siguientes principios:

- Efectos estimados son sumas de variables a aleatorias normales
- Para efectos no significativos estos estimadores se distribuyen normales con media cero y con misma varianza.
- En un gráfico de probabilidad normal los efectos estimados asociados a efectos poblacionales no significativos se alinearán sobre la recta de probabilidad.
- Por el contrario, los efectos estimados asociados a efectos poblacionales no nulos, se mostrarán como outliers respecto a la recta de probabilidad normal y se asume que esto es evidencia de la significancia de tales efectos.
- *Como algunos efectos pueden ser positivos y otros negativos, para facilitar la evaluación, se prefiere el gráfico de probabilidad media normal (half-normal plot)* en el que se grafican los valores absolutos de los efectos estimados, esta idea se apoya en la propiedad de simetría de la normal y considera de misma importancia efectos estimados de misma magnitud pero de signos contrarios. *Asociado a este gráfico se tienen los métodos de prueba de Lenth, LGB, entre otros, para señalar con un nivel de significancia dado cuáles efectos son efectos activos.*

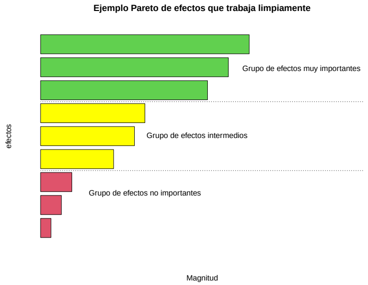
Pareto de efectos (o de los betas del MRLM)

Gráfico de barras donde se representan en orden de mayor a menor, los valores absolutos de

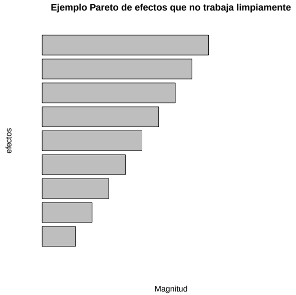
- Los efectos estimados no estandarizados, o bien
- Las estimaciones de MCO de los coeficientes de regresión,

De acuerdo a Gutiérrez y de la Vara Salazar (2012), este gráfico es útil para identificar efectos significativos, cuando este gráfico ***trabaja limpiamente***, es decir, cuando permite delimitar grupo de efectos desde los más importantes a los menos importantes.

En un pareto que trabaja limpiamente es posible identificar "puntos de corte" para los efectos a retener.



En un pareto que no trabaja limpiamente no es posible identificar "puntos de corte" para los efectos a retener.



Otros posibles criterios

De acuerdo a Gutiérrez y de la Vara Salazar (2012), cuando no se logra determinar claramente con los gráficos de pareto y de probabilidad media normal, cuáles efectos mandar al error:

- *Comparar el efecto observado con el error estándar basado en una varianza histórica:* Teniendo en cuenta que con $n = 1$, $\mathrm{Var}[\widehat{\alpha}] = \sigma^2/2^{k-2}$, si $\widehat{\sigma}^2$ es una estimación de la varianza basada en la historia previa del proceso, retener efectos tales que,

$$|\widehat{\alpha}| > 2\frac{\widehat{\sigma}}{\sqrt{2^{k-2}}} \tag{3}$$

- *Efectos que en ANOVAs preliminares tuvieron un valor P cerca de 0.2 o menores no se excluyen del análisis necesariamente:* Esta decisión es más confiable cuando dichos ANOVAS alcanzaron al menos 8 grados de libertad para el error.
- *Los grados de libertad del error:* Deben ser al menos 8 para tener un ANOVA más confiable.
- *El R^2adj de los ANOVAS preliminares:* Si se elimina un efecto y este estadístico decrece, posiblemente tal efecto no deba excluirse. Se requiere que el decrecimiento mencionado sea de cuando menos 3 % para que valga la pena incluir otra vez al efecto.

Colapsación o proyección del diseño

- En la depuración del modelo preliminar, es posible que un factor resulte sin efectos significativos: No participa en interacciones significativas y tampoco tiene efectos principales significativos.
- Puede decidirse respecto a este factor
  - *Mandar al error todos sus efectos y sus grados de libertad, o bien*
  - *Eliminar por completo del análisis a ese factor y partir de un modelo preliminar sin éste factor, es decir, cambiar el experimento 2^k no replicado por un experimento 2^{k-1} (no confundir con el diseño fracción 1/2, 2^{k-1}) con dos réplicas en cada punto del diseño resultante.*
  - *En general si se eliminan h factores del diseño factorial, los datos se convierten en los de un diseño factorial 2^{k-h} completo con 2^h réplicas en cada punto del diseño.*

Nota 3.2

Hay que ser cautos con esta recomendación pues a pesar que un factor no sea significativo, hizo parte de los tratamientos aplicados y en cierto modo su presencia durante la experimentación restringió la asignación de las unidades experimentales y de las corridas.

Contenido

El mejor ANOVA y modelo predictivo

En el proceso de la construcción de un SSE aproximado, se pretende llegar a un modelo donde queden solo efectos significativos, mandando al error los que no lo son. Como en el análisis preliminar pueden existir efectos sobre los cuales no es clara la decisión de retenerlos o no, el mejor ANOVA debe construirse de forma secuencial,

1. *Primero mandar al error los efectos claramente no significativos*
2. Una vez retirados, revalorar la significancia de los efectos sobre los cuales se tienen dudas.
3. *En esta segunda evaluación, retirar los términos que no resulten significativos*
4. Siguiendo estos pasos, se llega al "mejor ANOVA".
5. *sobre el modelo del "mejor ANOVA" debe evaluarse el R^2 y el R^2_{adj}. Para fines predictivos se recomienda un R^2_{adj} de al menos 0.7*

Si no se logra un modelo predictivo con un $R^2_{adj} \geq 0.7$, determinar posibles razones,

- *No se consideraron los factores con suficiente influencia para explicar la variabilidad de la respuesta*
- Se tomaron muy cercanos los niveles alto y bajo de factores cuantitativos
- *No se controlaron apropiadamente otros factores influenciales sobre la respuesta*
- Errores experimentales y de medición altos.

Contenido

Ejemplo 2^4 no replicado, Sección 9.8, pág. 344

Se prueba la resistencia al fuego de cierto tipo de telas después de aplicarles tratamientos contra el fuego. Se consideran 4 factores:

- A: tipo de tela
- B: tipo de tratamiento contra el fuego.
- C: Condición de lavado (nivel bajo es sin lavar; el nivel alto es después de una lavada)
- D: Método de prueba
- Y: Número de pulgadas de tela quemada en una muestra de prueba de tamaño estándar

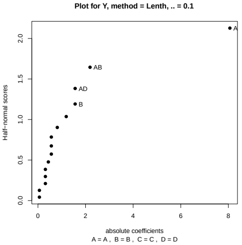
Todos los factores se corren en dos niveles. Los datos son (no están en orden de observación)

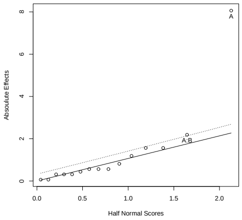
Código Yates	Y
(1)	42
a	31
b	45
ab	29
c	39
ac	28
bc	46
abc	32
d	40
ad	30
bd	50
abd	25
cd	40
acd	25
bcd	50
abcd	23
Promedio	35.9375

ir a archivo MONTGOM12-19tablas.pdf.

```
> modelo1=lm(Y~A*B*C*D) #variables codificadas como -1, 1
> summary(modelo1)
Call: lm.default(formula = Y ~ A * B * C * D)
Residuals: ALL 16 residuals are 0: no residual degrees of freedom!
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  35.9375        NaN     NaN      NaN
A            -8.0625        NaN     NaN      NaN
B             1.5625        NaN     NaN      NaN
C            -0.5625        NaN     NaN      NaN
D            -0.5625        NaN     NaN      NaN
A:B          -2.1875        NaN     NaN      NaN
A:C          -0.3125        NaN     NaN      NaN
B:C           0.8125        NaN     NaN      NaN
A:D          -1.5625        NaN     NaN      NaN
B:D           0.0625        NaN     NaN      NaN
C:D          -0.3125        NaN     NaN      NaN
A:B:C         0.3125        NaN     NaN      NaN
A:B:D        -1.1875        NaN     NaN      NaN
A:C:D        -0.5625        NaN     NaN      NaN
B:C:D        -0.4375        NaN     NaN      NaN
A:B:C:D       0.0625        NaN     NaN      NaN
Residual standard error: NaN on 0 degrees of freedom
Multiple R-squared:      1,     Adjusted R-squared:     NaN
F-statistic:   NaN on 15 and 0 DF,  p-value: NA
```

Gráfico probabilidad media normal ejemplo 2^4 no replicado, Sección 9.8, pág. 344






```
> halfnormal(modelo1,code=T,alpha=0.1,linelwd=2,linecol=2,
+            pch.set = c(19, 16, 8))
Significant effects (alpha=0.1, Lenth method):
  A A:B A:D   B
  A  AB  AD   B
```

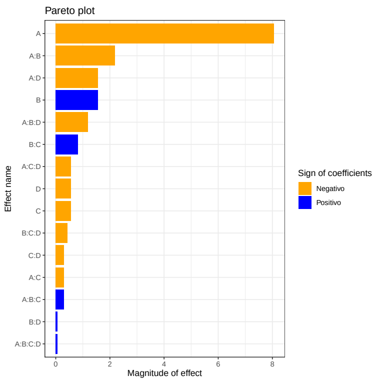
El modelo prelimininar que incluye todos los efectos es:

$$Y = \beta_0 + \sum_{j=1}^{4} \beta_j X_j + \sum_{j=1}^{4} \sum_{k>j}^{4} \beta_{jk} X_j X_k + \sum_{j=1}^{4} \sum_{k>j}^{4} \sum_{l>j,k}^{4} \beta_{jkl} X_j X_k X_l$$

$$+ \beta_{1234} X_1 X_2 X_3 X_4 + E, \text{ con } E_i \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)$$

- $X_1 = -1, +1$, para A; $X_2 = -1, +1$, para B
- $X_3 = -1, +1$, para C; $X_4 = -1, +1$, para D

```
> LGB(coef(modelo1)[-1],rpt=T,alpha=0.1)
Effect Report
Label      Half Effect    Sig(.05)
A          -8.0625        yes
B           1.5625        no
C          -0.5625        no
D          -0.5625        no
A:B        -2.1875        yes
A:C        -0.3125        no
B:C         0.8125        no
A:D        -1.5625        no
B:D         0.0625        no
C:D        -0.3125        no
A:B:C       0.3125        no
A:B:D      -1.1875        no
A:C:D      -0.5625        no
B:C:D      -0.4375        no
A:B:C:D     0.0625        no
Lawson, Grimshaw & Burt Rn Statistic =  1.849003
95th percentile of Rn =  1.122
```



Nota: Las funciones R: `LGB()` de la librerías `daewr`, `halfnormal()` de la librería `DoE.base` y la función `paretoPlot()` de la librería `pid`, grafican no los efectos, sino los betas del MRLM, pero recuerde que estos son la mitad de los efectos

1. De acuerdo al gráfico de probabilidad media normal con el método de Lenth usando un nivel de significancia del 10 % y el pareto, los efectos claramente importantes son: A, AB, AD y B
2. Pero por principio de jeraquía: si permanece un determinado efecto de interacción en un modelo, entonces todos los efectos individuales y de interacción de menor orden que se pueden formar con los factores que participan en esa interacción, deberán permanecer también en el modelo.
3. Por tanto, se mantendrán en el modelo los efectos A, B, D, AB y AD

Nuevo modelo:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_4 X_4 + \beta_{12} X_1 X_2 + \beta_{14} X_1 X_4 + E,\ E_i \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)$$

```
> modelo4=lm(Y~A*B+A*D)
> summary(modelo4)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  35.9375     0.5653  63.576 2.26e-14 ***
A            -8.0625     0.5653 -14.263 5.67e-08 ***
B             1.5625     0.5653   2.764  0.01999 *
D            -0.5625     0.5653  -0.995  0.34316
A:B          -2.1875     0.5653  -3.870  0.00311 **
A:D          -1.5625     0.5653  -2.764  0.01999 *
---
Residual standard error: 2.261 on 10 degrees of freedom
Multiple R-squared:  0.9591,    Adjusted R-squared:  0.9387
F-statistic: 46.94 on 5 and 10 DF,  p-value: 1.27e-06

> anova(modelo4)
Analysis of Variance Table
Response: Y
          Df  Sum Sq Mean Sq  F value    Pr(>F)
A          1 1040.06 1040.06 203.4352 5.668e-08 ***
B          1   39.06   39.06   7.6406   0.01999 *
D          1    5.06    5.06   0.9902   0.34316
A:B        1   76.56   76.56  14.9756   0.00311 **
A:D        1   39.06   39.06   7.6406   0.01999 *
Residuals 10   51.12    5.11
```

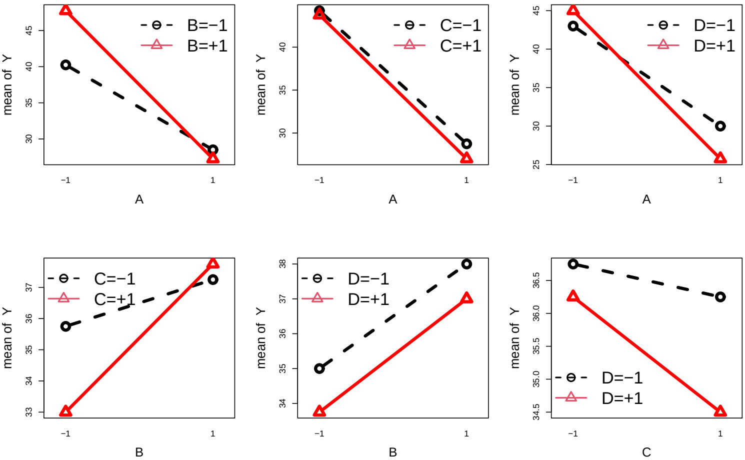


Figura 1: Interacciones dobles. Como todos los factores son cualitativos, no se hace análisis de superficies ni contornos de respuesta. La respuesta óptimo es buscando tratamiento que minimice su valor (menor número de pulgadas quemadas en la tela) es con $A = +1$, $B = -1$, $D = +1$ y C debe fijarse en su nivel más económico pues no resulta significativo en el ANOVA.

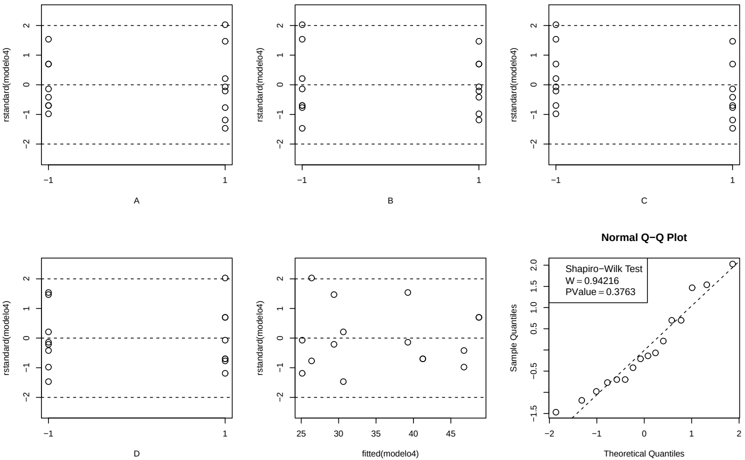


Figura 2: Gráficos de residuos estudentizados del modelo predictivo. No se observa evidencia fuerte en contra de los supuestos sobre errores: Media cero, varianza constante y normalidad

```r
library(rsm)
library(pid)
library(daewr)
library(FrF2)
MONTGOMERY12.19=data.frame(scan(what=list(Y=0,A=0,B=0,C=0,D=0)))
42 -1 -1 -1 -1
31 1 -1 -1 -1
45 -1 1 -1 -1
29 1 1 -1 -1
39 -1 -1 1 -1
28 1 -1 1 -1
46 -1 1 1 -1
32 1 1 1 -1
40 -1 -1 -1 1
30 1 -1 -1 1
50 -1 1 -1 1
25 1 1 -1 1
40 -1 -1 1 1
25 1 -1 1 1
50 -1 1 1 1
23 1 1 1 1

attach(MONTGOMERY12.19)

modelo1=lm(Y~A*B*C*D)
summary(modelo1)
paretoPlot(modelo1,negative=c("Negativo","orange"),
           positive=c("Positivo","blue"))
LGB(coef(modelo1)[-1],rpt=T,alpha=0.1)
halfnormal(modelo1,code=T,alpha=0.1,linelwd=2,linecol=2,
            pch.set = c(19, 16, 8))

modelo4=lm(Y~A*B+A*D)
summary(modelo4)
anova(modelo4)

interaction.plot(A,B,Y,type="b",pch=c(1,2),col=c("black","red"),
                  lwd=4,cex=2,cex.lab=1.5,legend=F)
legend("topright",legend=c("B=-1","B=+1"),col=1:2,pch=1:2,
       lwd=2,lty=c(2,1),bty="n",cex=2)
interaction.plot(A,C,Y,type="b",pch=c(1,2),col=c("black","red"),
                  lwd=4,cex=2,cex.lab=1.5,legend=F)
legend("topright",legend=c("C=-1","C=+1"),col=1:2,pch=1:2,
       lwd=2,lty=c(2,1),bty="n",cex=2)
interaction.plot(A,D,Y,type="b",pch=c(1,2),col=c("black","red"),
                  lwd=4,cex=2,cex.lab=1.5,legend=F)
legend("topright",legend=c("D=-1","D=+1"),col=1:2,pch=1:2,
       lwd=2,lty=c(2,1),bty="n",cex=2)
interaction.plot(B,C,Y,type="b",pch=c(1,2),col=c("black","red"),
                  lwd=4,cex=2,cex.lab=1.5,legend=F)
legend("topleft",legend=c("C=-1","C=+1"),col=1:2,pch=1:2,
       lwd=2,lty=c(2,1),bty="n",cex=2)
interaction.plot(B,D,Y,type="b",pch=c(1,2),col=c("black","red"),
                  lwd=4,cex=2,cex.lab=1.5,legend=F)
legend("topleft",legend=c("D=-1","D=+1"),col=1:2,pch=1:2,
       lwd=2,lty=c(2,1),bty="n",cex=2)
interaction.plot(C,D,Y,type="b",pch=c(1,2),col=c("black","red"),
                  lwd=4,cex=2,cex.lab=1.5,legend=F)
legend("bottomleft",legend=c("D=-1","D=+1"),col=1:2,pch=1:2,
        lwd=2,lty=c(2,1),bty="n",cex=2)

#Validación de supuestos sobre modelo4
shapiro.test(rstandard(modelo4))
stripchart(rstandard(modelo4)~A,vertical=TRUE,ylim=c(-2.5,2.5),
           pch=1,cex=1.5,xlab="A")
abline(h=c(-2,0,2),lty=2)
stripchart(rstandard(modelo4)~B,vertical=TRUE,ylim=c(-2.5,2.5),
           pch=1,cex=1.5,xlab="B")
abline(h=c(-2,0,2),lty=2)
stripchart(rstandard(modelo4)~C,vertical=TRUE,ylim=c(-2.5,2.5),
           pch=1,cex=1.5,xlab="C")
abline(h=c(-2,0,2),lty=2)
stripchart(rstandard(modelo4)~D,vertical=TRUE,ylim=c(-2.5,2.5),
           pch=1,cex=1.5,xlab="D")
abline(h=c(-2,0,2),lty=2)
plot(fitted(modelo4),rstandard(modelo4),ylim=c(-2.5,2.5),cex=1.5)
abline(h=c(-2,0,2),lty=2)
qqnorm(rstandard(modelo4),cex=1.5)
qqline(rstandard(modelo4),lty=2)
legend("topleft",legend=c("Shapiro-WilkTest",
       expression(W==0.94216),expression(PValue==0.3763)),cex=1.1)
detach(MONTGOMERY12.19)
```

Dean, A., Voss, D., and Draguljić, D. (2017). *Design and Analysis of Experiments*, 2nd Edition. Springer.

Gutiérrez Pulido, H. y de la Vara Salazar, R. (2012). *Análisis y Diseño de Experimentos*, 3ª Edición. McGraw-Hill.

Kuehl, R. O. (2001). *Diseño de Experimentos. Principios Estadísticos de Diseño y Análisis de Investigación*, 2ª Edición. Thomson Learning.

Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W. (2005). *Applied Linear Statistical Models*, 5th Edition. McGraw-Hill Irwin.

Montgomery, D. C. (2020). *Design and Analysis of Experiments*, 10th Edition. John Wiley & Sons, Inc.