Examen Final de Análisis y Diseños de Experimentos

Jaime Gomez Marin

#### 1.- En un artículo de Solid State Technology (“Diseño ortogonal para optimización de procesos y su aplicación en el grabado químico con plasma”) se describe la aplicación de diseños factoriales en el desarrollo de un proceso de grabado químico con nitruros en un dispositivo de grabado químico con plasma para una sola oblea El proceso usa C2F6 como gas de reacción. Cuatro factores son de interés: el entrehierro ánodo-cátodo (A), la presión en la cámara del reactor (B), el flujo del gas C2F6 (C) y la potencia aplicada al cátodo (D). La respuesta de interés es la rapidez de grabado para el nitruro de silicio. Se corre una sola réplica de un diseño 24; los datos se muestran enseguida:

| (-) bajo 1 | (+) alto

———-|——-|—————– A (cm) | 0.80 | 1.20 B (mTorr) | 450 | 550 C (SCCM) | 125 | 200 D (W) | 275 | 325

1. 550 a = 669 b = 604 ab = 650 c = 633 ac = 642 bc = 601 abc = 635 d = 1037 ad = 749 bd = 1052 abd = 868 cd = 1075 acd = 860 bcd = 1063 abcd = 759

y=c(550,669,604,650,633,642,601,635,1037,749,1052,868,1075,860,1063,759)

#### a).Estimar los efectos de los factores y obtenga una gráfica de probabilidad normal de los efectos estimados de estos factores. ¿Que efectos parecen ser grande? ( 6 puntos)

#  
y=c(550,669,604,650,633,642,601,635,1037,749,1052,868,1075,860,1063,759)  
  
#  
A <- rep(c(rep(-1, 1) , rep(1, 1)) , 8)  
B <- rep(c(rep(-1, 2) , rep(1, 2)) , 4)  
C <- rep(c(rep(-1, 4) , rep(1, 4)) , 2)  
D <- rep(c(rep(-1, 8) , rep(1, 8)) , 1)  
  
#  
printmat("y ",y)

## [1] "y : [1] 550 669 604 650 633 642 601 635 1037 749 1052 868 1075 860"  
## [2] "y : [15] 1063 759"

printmat("A ",A)

## [1] "A : [1] -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1 -1 1"

printmat("B ",B)

## [1] "B : [1] -1 -1 1 1 -1 -1 1 1 -1 -1 1 1 -1 -1 1 1"

printmat("C ",C)

## [1] "C : [1] -1 -1 -1 -1 1 1 1 1 -1 -1 -1 -1 1 1 1 1"

printmat("D ",D)

## [1] "D : [1] -1 -1 -1 -1 -1 -1 -1 -1 1 1 1 1 1 1 1 1"

mod <- lm(y~A\*B\*C\*D)  
Estimados <- 2\*coefficients(mod)[-1] # Se retira la columna del intercepto  
# Estimados  
round(Estimados, 0)

## A B C D A:B A:C B:C A:D B:D   
## -98 2 11 310 -4 -21 -40 -150 3   
## C:D A:B:C A:B:D A:C:D B:C:D A:B:C:D   
## 2 -12 8 9 -22 -36

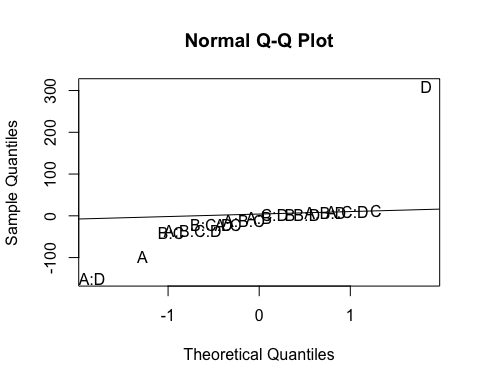
Hay que buscar los estimados que mas aportan al modelo

data.frame(round(Estimados, 0))

## round.Estimados..0.  
## A -98  
## B 2  
## C 11  
## D 310  
## A:B -4  
## A:C -21  
## B:C -40  
## A:D -150  
## B:D 3  
## C:D 2  
## A:B:C -12  
## A:B:D 8  
## A:C:D 9  
## B:C:D -22  
## A:B:C:D -36

Quienes mas aportan son: 1,4,6,7,8,14,15

qq <- qqnorm(Estimados,type="n")  
Efectos <- names(Estimados)  
text(qq$x, qq$y, labels = Efectos)  
Estimados1 <- Estimados[-c(1,4,6,7,8,14,15)] # SON LOS VALORES QUE MAS APORTAN AL MODELO  
# Estimados1 <- Estimados  
qqline(Estimados1)

 De la grafica se puede apreciar que los siguiente efectos son los mas grandes en orden de influencia:

* D : La potencia aplicada al cátodo
* A\*D : El entrehierro ánodo-cátodo y la potencia aplicada al cátodo
* A : El entrehierro ánodo-cátodo
* B\*C : La presión en la cámara del reacto y el flujo del gas C2F6

Hay mas efectos como: A*B*C + B*C*D , pero los anteriores tiene un mayor efecto.

#### b).- Realice el análisis de varianza para confirmar los resultados obtenidos en (a), Obtenga la ecuación de regresión estimada (4 puntos)

# Con la información de los efectos detectados en el apartado anterior procedemos a construimos nuestro modelo  
mod1 <- lm(y ~ A + D + A\*D + B\*C + A\*C + A\*B\*C + B\*C\*D )

# Realizamos nuestro analisis de varianza  
anva1 <- aov(mod1)  
summary(anva1)

## Df Sum Sq Mean Sq F value Pr(>F)   
## A 1 38318 38318 19.510 0.021540 \*   
## D 1 384090 384090 195.559 0.000792 \*\*\*  
## B 1 18 18 0.009 0.929648   
## C 1 495 495 0.252 0.650161   
## A:D 1 89850 89850 45.747 0.006603 \*\*   
## B:C 1 6440 6440 3.279 0.167858   
## A:C 1 1785 1785 0.909 0.410748   
## A:B 1 68 68 0.035 0.864199   
## D:B 1 39 39 0.020 0.896786   
## D:C 1 11 11 0.005 0.946156   
## A:B:C 1 564 564 0.287 0.629211   
## D:B:C 1 1871 1871 0.952 0.401085   
## Residuals 3 5892 1964   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Donde:

= 1, si el nivel del factor de i se encuentra alto

= 0, si el nivel del factor de i se encuentra bajo

i = 1 , representa el factor de A

i = 2 , representa el factor de B

i = 3 , representa el factor de C

i = 4 , representa el factor de D

De los cuatro factores son de interés: el entrehierro ánodo-cátodo (A), la presión en la cámara del reactor (B), el flujo del gas C2F6 (C) y la potencia aplicada al cátodo (D) se ha llegado a las siguientes conclusiones

a).- Los efectos que tiene influenciencia altamente muy significativa sobre la rapidez de grabado para el nitruro de silicio son:

* D = La potencia aplicada al catodo

b).- Los efectos que tiene influenciencia muy significativa sobre la rapidez de grabado para el nitruro de silicio son:

* entre A = El entrehierro ánodo-cátodo con D = La potencia aplicada al catodo

c).- Los efectos que tiene influenciencia significativa sobre la rapidez de grabado para el nitruro de silicio son:

* A = El entrehierro ánodo-cátodo

d).- No se encontró diferencias significativas

* B = la presión en la cámara del reactor
* C = el flujo del gas C2F6
* entre B = la presión en la cámara del reactor con C = el flujo del gas C2F6
* entre A = El entrehierro ánodo-cátodo con C = el flujo del gas C2F6
* entre A = El entrehierro ánodo-cátodo con B = la presión en la cámara del reactor
* entre D = La potencia aplicada al catodo con C = el flujo del gas C2F6
* entre A = El entrehierro ánodo-cátodo , B = la presión en la cámara del reactor y C = el flujo del gas C2F6
* entre D = La potencia aplicada al catodo , B = la presión en la cámara del reactor y C = el flujo del gas C2F6

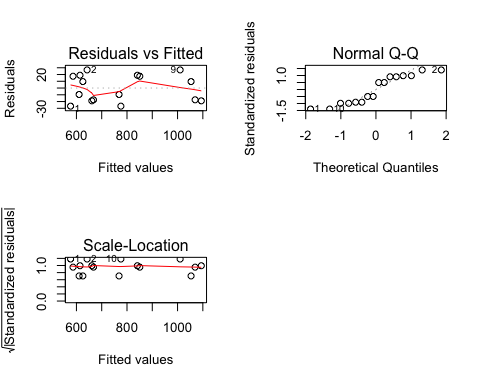
# Obtenemos los parametros  
coefficients(mod1)

## (Intercept) A D B C A:D   
## 777.9375 -48.9375 154.9375 1.0625 5.5625 -74.9375   
## B:C A:C A:B D:B D:C A:B:C   
## -20.0625 -10.5625 -2.0625 1.5625 0.8125 -5.9375   
## D:B:C   
## -10.8125

Del analisis de varianza se detecto que los efectos mas importantes son de A, D y A:D

par(mfrow=c(2,2))  
plot(mod1)

## hat values (leverages) are all = 0.8125  
## and there are no factor predictors; no plot no. 5



ri1<-rstandard(mod1)  
shapiro.test(ri1)

##   
## Shapiro-Wilk normality test  
##   
## data: ri1  
## W = 0.87299, p-value = 0.03022

library(car)

## Loading required package: carData

ncvTest(mod1)

## Non-constant Variance Score Test   
## Variance formula: ~ fitted.values   
## Chisquare = 0.005323703, Df = 1, p = 0.94183

##Analisis Factorial 2x4

#### 2.)- Un teléfono inteligente es un teléfono móvil que ofrece una capacidad de cómputo y conectividad más avanzadas que un “teléfono con funciones” básica contemporáneo. Los datos siguientes son las calificaciones de seis teléfonos inteligentes de cada uno de los cuatro proveedores, tres de los cuales cuestan $ 150 o más, otros tres cuestan menos de $ 150. Las calificaciones tienen un valor máximo de 100 y un mínimo de 0.

Proveedor   
"AT&T((b1)" "Sprint

(b2)" “T-Mobile (b3)” “Verizon (b4)” Costo ￼ $150 76 74 72 75 (a1) 74 69 71 73 69 68 71 73 Costo < $150 69 69 71 72 (a2) 67 64 71 71 64 60 70 70

Datos en archivo teléfono.txt #### a).- Realice el Análisis de Variancia, Pruebe la hipótesis correspondientes y de las conclusiones y recomendaciones correspondiente. (4 puntos)

telefono<-read.table("telefono.txt",T)  
str(telefono)

## 'data.frame': 24 obs. of 3 variables:  
## $ calificaciones: int 76 74 69 69 67 64 74 69 68 69 ...  
## $ costo : Factor w/ 2 levels "a1","a2": 1 1 1 2 2 2 1 1 1 2 ...  
## $ proveedor : Factor w/ 4 levels "b1","b2","b3",..: 1 1 1 1 1 1 2 2 2 2 ...

Calificación<-telefono[,1]  
Costo<-as.factor(telefono[,2])  
Proveedor<-as.factor(telefono[,3])  
mod<-lm(Calificación~Costo\*Proveedor)  
anva<-aov(mod)  
summary(anva)

## Df Sum Sq Mean Sq F value Pr(>F)   
## Costo 1 92.04 92.04 13.893 0.00183 \*\*  
## Proveedor 3 81.12 27.04 4.082 0.02490 \*   
## Costo:Proveedor 3 33.46 11.15 1.683 0.21053   
## Residuals 16 106.00 6.63   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Hipotesis del efecto principal de Costo

: , para i = 1,2,3

: al menos dos , para i = 1,2,3

Se rechaza la , a un nivel de significacion del 1%, se ha encontrado suficientes evidencia estadistica para rechazar la de que las categorias de los Costos tenga un efecto significativo sobre la Calificación del teléfono.Por lo tanto, se puede aceptar de que existe diferencias significativas entre al menos dos categorias de costos de teléfono tienen un efecto significativo sobre su calificación

## Hipotesis del efecto principal de Proveedor

: = 0 , para j = 1,2,3

: al menos dos , para j = 1,2,3

Se rechaza la , a un nivel de significacion del 5%, se ha encontrado suficientes evidencia estadistica para rechazar la de que las proveedores tenga un efecto significativo sobre la calificación del teléfono.Por lo tanto, se puede aceptar de que existe diferencias significativas entre al menos dos proveedores tienen un efecto significativo sobre la calificación del teléfono.

## Hipotesis del efecto de interacción entre el Costo y el Proveedor

: , para i,j = 1,2,3

: al menos dos , para i,j = 1,2,3

Se acepta , a un nivel de significacion del 10%, no se ha encontrado suficiente evidencia estadistica para rechazar la ,es decir que no existe ning'un efecto de la interaci'on entre la categoria del costo y el proveedor influyan sobre la calificación del teléfono.

#### b).- Teniendo en cuenta los resultados obtenido en (a), realice la prueba de Tukey correspondiente

(6 puntos)

De la parte a, se ha visto que la categoria de costo y los proveedores tiene efecto sobre la calificación del teléfono, se procedera a realizar la prueba de Tukey para la Cada uno de ellos

library(multcomp)

tHSD <- TukeyHSD(anva)  
summary(tHSD)

## Length Class Mode   
## Costo 4 -none- numeric  
## Proveedor 24 -none- numeric  
## Costo:Proveedor 112 -none- numeric

tHSD

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = mod)  
##   
## $Costo  
## diff lwr upr p adj  
## a2-a1 -3.916667 -6.144249 -1.689084 0.0018334  
##   
## $Proveedor  
## diff lwr upr p adj  
## b2-b1 -2.500000 -6.7516077 1.751608 0.3643388  
## b3-b1 1.166667 -3.0849410 5.418274 0.8601187  
## b4-b1 2.500000 -1.7516077 6.751608 0.3643388  
## b3-b2 3.666667 -0.5849410 7.918274 0.1039730  
## b4-b2 5.000000 0.7483923 9.251608 0.0185696  
## b4-b3 1.333333 -2.9182743 5.584941 0.8063964  
##   
## $`Costo:Proveedor`  
## diff lwr upr p adj  
## a2:b1-a1:b1 -6.3333333 -13.6093432 0.9426765 0.1129895  
## a1:b2-a1:b1 -2.6666667 -9.9426765 4.6093432 0.8976115  
## a2:b2-a1:b1 -8.6666667 -15.9426765 -1.3906568 0.0140595  
## a1:b3-a1:b1 -1.6666667 -8.9426765 5.6093432 0.9911124  
## a2:b3-a1:b1 -2.3333333 -9.6093432 4.9426765 0.9451981  
## a1:b4-a1:b1 0.6666667 -6.6093432 7.9426765 0.9999756  
## a2:b4-a1:b1 -2.0000000 -9.2760099 5.2760099 0.9752592  
## a1:b2-a2:b1 3.6666667 -3.6093432 10.9426765 0.6614247  
## a2:b2-a2:b1 -2.3333333 -9.6093432 4.9426765 0.9451981  
## a1:b3-a2:b1 4.6666667 -2.6093432 11.9426765 0.3896498  
## a2:b3-a2:b1 4.0000000 -3.2760099 11.2760099 0.5673822  
## a1:b4-a2:b1 7.0000000 -0.2760099 14.2760099 0.0638225  
## a2:b4-a2:b1 4.3333333 -2.9426765 11.6093432 0.4752993  
## a2:b2-a1:b2 -6.0000000 -13.2760099 1.2760099 0.1484555  
## a1:b3-a1:b2 1.0000000 -6.2760099 8.2760099 0.9996326  
## a2:b3-a1:b2 0.3333333 -6.9426765 7.6093432 0.9999998  
## a1:b4-a1:b2 3.3333333 -3.9426765 10.6093432 0.7517157  
## a2:b4-a1:b2 0.6666667 -6.6093432 7.9426765 0.9999756  
## a1:b3-a2:b2 7.0000000 -0.2760099 14.2760099 0.0638225  
## a2:b3-a2:b2 6.3333333 -0.9426765 13.6093432 0.1129895  
## a1:b4-a2:b2 9.3333333 2.0573235 16.6093432 0.0075749  
## a2:b4-a2:b2 6.6666667 -0.6093432 13.9426765 0.0852319  
## a2:b3-a1:b3 -0.6666667 -7.9426765 6.6093432 0.9999756  
## a1:b4-a1:b3 2.3333333 -4.9426765 9.6093432 0.9451981  
## a2:b4-a1:b3 -0.3333333 -7.6093432 6.9426765 0.9999998  
## a1:b4-a2:b3 3.0000000 -4.2760099 10.2760099 0.8321663  
## a2:b4-a2:b3 0.3333333 -6.9426765 7.6093432 0.9999998  
## a2:b4-a1:b4 -2.6666667 -9.9426765 4.6093432 0.8976115

$Costo diff lwr upr p adj a2-a1 -3.916667 -6.144249 -1.689084 0.0018334

A un nivel de significación del 0.1%, se a encontrados suficiente evidencia para afirmar que la media de la calificación de un teléfono que cuesta menos de 150 dolares es diferente a la media de la calificación de un teléfono que cuesta mas de 150 dolares.

$Proveedor diff lwr upr p adj

b4-b2 5.000000 0.7483923 9.251608 0.0185696

A un nivel de significación del 1%, se a encontrados suficiente evidencia para afirmar que la media de la calificaciòn de un teléfono del proveedor Verizon es diferente a la media de la calificación de un teléfono del proveedor Sprint

Para los otros casos no existe diferencias significativa