UNIV. NAC. AGRARIA LA MOLINA SEMESTRE 2018-II

DPTO. ACAD. DE ESTADÍSTICA E INFORMÁTICA Prof. Maehara

**Examen Final de Análisis y Diseños de Experimentos**

El tequila es una bebida que está sujeta a una norma oficial mexicana, y conforme a ésta se debe cumplir con ciertas especificaciones físico-químicas. En un laboratorio de investigación, mediante un diseño factorial 25 no replicado, se estudió la influencia de diversos factores sobre la producción de alcoholes superiores en la etapa de fermentación. Los factores estudiados y los niveles fueron: tipo de cepa, *A*(1, 2), temperatura, *B*(30, 35°C), fuente de nitrógeno, *C*(NH4)2SO4 y urea-, relación carbono/nitrógeno, *D*(62/1, 188/1) y porcentaje de inóculo, *E*(5 y 10%). En la siguiente tabla se muestran los resultados obtenidos en cuanto a alcohol isoamílico (mg/L), que es par te de los alcoholes superiores

(1)=21.4 *d=*42.5 *e=*32.9 *de=*54.0

*a*=16.8 *ad=*21.0 *ae=*17.5 *ade=*21.8

*b=*29.3 *bd=*79.1 *be=*30.0 *bde=*79.9

*ab=*12.7 *abd=*20.0 *abe*=24.1 *abde=*31.5

*c=*27.5 *cd=*48.6 *ce=*26.7 *cde=*47.9

*ac=*22.9 *acd=*27.1 *ace=*11.4 *acde=*15.6

*bc=*35.4 *bcd=*85.2 *bce=*23.9 *bcde=*73.8

*abc=*18.8 *abcd=*26.1 *abce=*18.0 *abcde=*25.4

y<-c(21.4, 16.8, 29.3, 12.7, 27.5, 22.9, 35.4, 18.8, 42.5, 21.0, 79.1, 20.0, 48.6, 27.1, 85.2, 26.1, 32.9, 17.5, 30.0, 24.1, 26.7, 11.4, 23.9, 18.0, 54.0, 21.8, 79.9, 31.5, 47.9, 15.6, 73.8, 25.4)

1. Estime los efectos y grafique en gráfico de probabilidad normal. De acuerdo a los resultados obtenidos, presente el cuadro de ANVA más conveniente, analice y dé sus conclusiones en término de enunciado. (5 puntos)

> y<-c(21.4, 16.8, 29.3, 12.7, 27.5, 22.9, 35.4, 18.8, 42.5, 21.0, 79.1, 20.0, 48.6, 27.1, 85.2, 26.1, 32.9, 17.5, 30.0, 24.1, 26.7, 11.4, 23.9, 18.0, 54.0, 21.8, 79.9, 31.5, 47.9, 15.6, 73.8, 25.4)

> A<-rep(c(-1,1),16)

> B<-rep(c(rep(-1,2),rep(1,2)),8)

> C<-rep(c(rep(-1,4),rep(1,4)),4)

> D<-rep(c(rep(-1,8),rep(1,8)),2)

> E<-c(rep(-1,16),rep(1,16))

> mod<-lm(y~A\*B\*C\*D\*E)

> Estimados<-2\*coefficients(mod)[-1]

> data.frame(Estimados)

Estimados

A -2.546250e+01

B 9.850000e+00

C -1.250000e-02

D 2.063750e+01

E 2.865051e-15

A:B -7.037500e+00

A:C -2.262129e-15

B:C 1.250000e-02

A:D -1.485000e+01

B:D 7.962500e+00

C:D -2.242932e-15

A:E -1.250000e-02

B:E -5.045394e-15

C:E -6.112500e+00

D:E 3.750000e-02

A:B:C 2.565384e-15

A:B:D -6.400000e+00

A:C:D -1.250000e-02

B:C:D 6.386167e-15

A:B:E 5.362500e+00

A:C:E 5.214056e-15

B:C:E 1.250000e-02

A:D:E 5.431784e-15

B:D:E 1.250000e-02

C:D:E 3.110746e-16

A:B:C:D 1.250000e-02

A:B:C:E -1.574263e-15

A:B:D:E -2.020022e-15

A:C:D:E -1.250000e-02

B:C:D:E -4.946282e-15

A:B:C:D:E 1.250000e-02

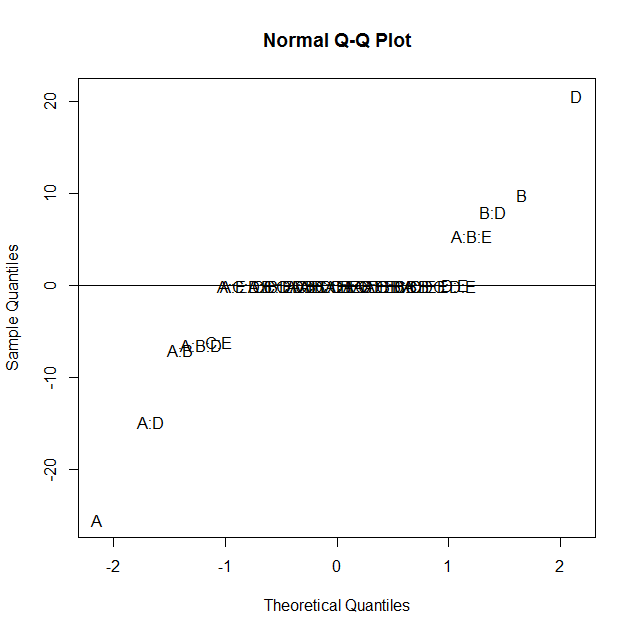
> Nombre<-names(Estimados)

> qq<-qqnorm(Estimados,type="n")

> text(qq$x, qq$y, labels = Nombre)

> Estimados1<-Estimados[-c(1:6, 9, 10, 11, 14, 17, 20)]

> qqline(Estimados1)



> mod1<-lm(y~A+B+D+A\*B+A\*D+B\*D+C\*E+A\*B\*D+A\*B\*E)

> anva1<-aov(mod1)

> summary(anva1)

Df Sum Sq Mean Sq F value Pr(>F)

A 1 5187 5187 4.668e+06 <2e-16 \*\*\*

B 1 776 776 6.986e+05 <2e-16 \*\*\*

D 1 3407 3407 3.067e+06 <2e-16 \*\*\*

C 1 0 0 1.125e+00 0.303

E 1 0 0 0.000e+00 1.000

A:B 1 396 396 3.566e+05 <2e-16 \*\*\*

A:D 1 1764 1764 1.588e+06 <2e-16 \*\*\*

B:D 1 507 507 4.565e+05 <2e-16 \*\*\*

C:E 1 299 299 2.690e+05 <2e-16 \*\*\*

A:E 1 0 0 1.125e+00 0.303

B:E 1 0 0 0.000e+00 1.000

A:B:D 1 328 328 2.949e+05 <2e-16 \*\*\*

A:B:E 1 230 230 2.070e+05 <2e-16 \*\*\*

Residuals 18 0 0

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

****

Donde 

 representa al factor A,  representa al factor C y  representa al factor D

,

 versus , para 

 versus , para  , 

 versus 

 versus 

Los efectos que tienen una influencia altamente significativa sobre la producción de alcohol isoamílico son: A = tipo de cepa, B = temperatura, C = fuente de nitrógeno y urea, D= relación carbono/nitrógeno, E = porcentaje de inóculo; y entre las interacciones

* A = tipo de cepa con B = temperatura
* A = tipo de cepa con D= relación carbono/nitrógeno.
* A = tipo de cepa con E = porcentaje de inóculo.
* B = temperatura con D= relación carbono/nitrógeno
* B = temperatura con E = porcentaje de inóculo
* C = fuente de nitrógeno y urea con E = porcentaje de inóculo.
* D= relación carbono/nitrógeno con E = porcentaje de inóculo.

Y la triple interacciones entre

* A = tipo de cepa con B = temperatura y con D= relación carbono/nitrógeno
* A = tipo de cepa con B = temperatura y con E = porcentaje de inóculo

No se encontró diferencias significativas

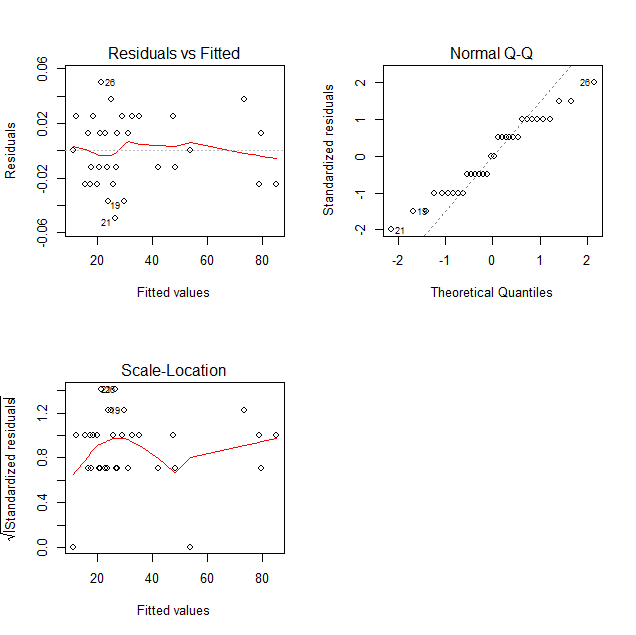
entre A = tipo de cepa con E = porcentaje de inóculo

entre B = temperatura con E = porcentaje de inóculo

> mod3<-lm(y~A+B+C+D+E+A\*E+B\*E+B\*C+C\*D+C\*E+C\*D+C\*E+C\*D\*E)

> par(mfrow=c(2,2))

> plot(mod1)



> ri1<-rstandard(mod1)

> shapiro.test(ri1)

Shapiro-Wilk normality test

data: ri1

W = 0.95359, p-value = 0.1821

> library(car)

> ncvTest(mod1)

Non-constant Variance Score Test

Variance formula: ~ fitted.values

Chisquare = 0.007709173 Df = 1 p = 0.9300341

>

1. Suponga que en este experimento, en lugar de que se haya realizado una réplica completa, debido a limitaciones experimentales de en un solo día solo se puede realizar ocho corrida, razón por la cual se tuvo que confundir las siguientes interacciones ACD y BCE para ,formar los bloques que se deben realizar.

b.1) Muestre el plan experimental distribuyendo las combinaciones de tratamiento en sus respectivos bloques. Deduzca la interacción generalizada. (5 puntos)

> A<-rep(c(-1,1),16)

> B<-rep(c(rep(-1,2),rep(1,2)),8)

> C<-rep(c(rep(-1,4),rep(1,4)),4)

> D<-rep(c(rep(-1,8),rep(1,8)),2)

> E<-c(rep(-1,16),rep(1,16))

> ACD<-A\*C\*D

> BCE<-B\*C\*E

> data.frame(A,B,C,D,E,ACD,BCE)

A B C D E ACD BCE

1 -1 -1 -1 -1 -1 -1 -1

2 1 -1 -1 -1 -1 1 -1

3 -1 1 -1 -1 -1 1 1

4 1 1 -1 -1 -1 -1 1

5 -1 -1 1 -1 -1 1 1

6 1 -1 1 -1 -1 -1 1

7 -1 1 1 -1 -1 -1 -1

8 1 1 1 -1 -1 1 -1

9 -1 -1 -1 1 -1 -1 -1

10 1 -1 -1 1 -1 1 -1

11 -1 1 -1 1 -1 1 1

12 1 1 -1 1 -1 -1 1

13 -1 -1 1 1 -1 1 1

14 1 -1 1 1 -1 -1 1

15 -1 1 1 1 -1 -1 -1

16 1 1 1 1 -1 1 -1

17 -1 -1 -1 -1 1 -1 1

18 1 -1 -1 -1 1 1 1

19 -1 1 -1 -1 1 1 -1

20 1 1 -1 -1 1 -1 -1

21 -1 -1 1 -1 1 1 -1

22 1 -1 1 -1 1 -1 -1

23 -1 1 1 -1 1 -1 1

24 1 1 1 -1 1 1 1

25 -1 -1 -1 1 1 -1 1

26 1 -1 -1 1 1 1 1

27 -1 1 -1 1 1 1 -1

28 1 1 -1 1 1 -1 -1

29 -1 -1 1 1 1 1 -1

30 1 -1 1 1 1 -1 -1

31 -1 1 1 1 1 -1 1

32 1 1 1 1 1 1 1

> tequila1<-read.table("tequila1.txt",T)

> trat<-c("(1)","a","b","ab","c","ac","bc","abc","d","ad","bd","abd","cd","acd","bcd","abcd","e","ae","be","abe","ce","ace","bce","abce","de","ade","bde","abde","cde","acde","bcde","abcde")

> tequila<-data.frame(tequila1,trat)

> tequilab1<-tequila[tequila$Bloque=="1",]

> tequilab2<-tequila[tequila$Bloque=="2",]

> tequilab3<-tequila[tequila$Bloque=="3",]

> tequilab4<-tequila[tequila$Bloque=="4",]

> tequilab1

A B C D E ACD BCE Bloque trat

1 -1 -1 -1 -1 -1 -1 -1 1 (1)

8 1 1 1 -1 -1 -1 -1 1 abc

10 1 -1 -1 1 -1 -1 -1 1 ad

15 -1 1 1 1 -1 -1 -1 1 bcd

19 -1 1 -1 -1 1 -1 -1 1 be

22 1 -1 1 -1 1 -1 -1 1 ace

28 1 1 -1 1 1 -1 -1 1 abde

29 -1 -1 1 1 1 -1 -1 1 cde

> tequilab2

A B C D E ACD BCE Bloque trat

2 1 -1 -1 -1 -1 1 -1 2 a

7 -1 1 1 -1 -1 1 -1 2 bc

9 -1 -1 -1 1 -1 1 -1 2 d

16 1 1 1 1 -1 1 -1 2 abcd

20 1 1 -1 -1 1 1 -1 2 abe

21 -1 -1 1 -1 1 1 -1 2 ce

27 -1 1 -1 1 1 1 -1 2 bde

30 1 -1 1 1 1 1 -1 2 acde

> tequilab3

A B C D E ACD BCE Bloque trat

3 -1 1 -1 -1 -1 -1 1 3 b

6 1 -1 1 -1 -1 -1 1 3 ac

12 1 1 -1 1 -1 -1 1 3 abd

13 -1 -1 1 1 -1 -1 1 3 cd

17 -1 -1 -1 -1 1 -1 1 3 e

24 1 1 1 -1 1 -1 1 3 abce

26 1 -1 -1 1 1 -1 1 3 ade

31 -1 1 1 1 1 -1 1 3 bcde

> tequilab4

A B C D E ACD BCE Bloque trat

4 1 1 -1 -1 -1 1 1 4 ab

5 -1 -1 1 -1 -1 1 1 4 c

11 -1 1 -1 1 -1 1 1 4 bd

14 1 -1 1 1 -1 1 1 4 acd

18 1 -1 -1 -1 1 1 1 4 ae

23 -1 1 1 -1 1 1 1 4 bce

25 -1 -1 -1 1 1 1 1 4 de

32 1 1 1 1 1 1 1 4 abcde

|  |  |  |  |
| --- | --- | --- | --- |
| **blo1** | **blo2** | **blo3** | **blo4** |
| (1) | a | b | ab |
| abc | bc | ac | c |
| ad | d | abd | bd |
| bcd | abcd | cd | acd |
| be | abe | e | ae |
| ace | ce | abce | bce |
| abde | bde | ade | de |
| cde | acde | bcde | abcde |

b.2) Obtenga el cuadro de ANVA más conveniente para analizar estos resultados. Realice las pruebas de hipótesis sobre los efectos que son significativos. Concluya en término de enunciados. (5 puntos)

> Bloque<-as.factor(tequila1$Bloque)

> mod<-lm(y~Bloque+A\*B\*C\*D\*E)

> betasest<-coefficients(mod)

> data.frame(betasest)

betasest

(Intercept) 3.338750e+01

Bloque2 1.250000e-02

Bloque3 2.500000e-02

Bloque4 1.250000e-02

A -1.273125e+01

B 4.925000e+00

C -6.250000e-03

D 1.031875e+01

E 3.544729e-15

A:B -3.518750e+00

A:C -3.051561e-16

B:C 6.250000e-03

A:D -7.425000e+00

B:D 3.981250e+00

C:D -3.084128e-15

A:E NA

B:E -1.080359e-15

C:E -3.056250e+00

D:E 1.875000e-02

A:B:C NA

A:B:D -3.200000e+00

A:C:D -6.250000e-03

B:C:D 1.121512e-15

A:B:E 2.681250e+00

A:C:E 2.060685e-15

B:C:E NA

A:D:E 3.912965e-17

B:D:E 6.250000e-03

C:D:E 2.212113e-15

A:B:C:D 6.250000e-03

A:B:C:E -4.477462e-15

A:B:D:E -3.773128e-15

A:C:D:E -6.250000e-03

B:C:D:E -2.368264e-15

A:B:C:D:E 6.250000e-03

> Efectos<-2\*(betasest[-c(1:4,16,20,26)])

> data.frame(Efectos)

Efectos

A -2.546250e+01

B 9.850000e+00

C -1.250000e-02

D 2.063750e+01

E 7.089458e-15

A:B -7.037500e+00

A:C -6.103121e-16

B:C 1.250000e-02

A:D -1.485000e+01

B:D 7.962500e+00

C:D -6.168255e-15

B:E -2.160717e-15

C:E -6.112500e+00

D:E 3.750000e-02

A:B:D -6.400000e+00

A:C:D -1.250000e-02

B:C:D 2.243024e-15

A:B:E 5.362500e+00

A:C:E 4.121370e-15

A:D:E 7.825930e-17

B:D:E 1.250000e-02

C:D:E 4.424226e-15

A:B:C:D 1.250000e-02

A:B:C:E -8.954924e-15

A:B:D:E -7.546257e-15

A:C:D:E -1.250000e-02

B:C:D:E -4.736527e-15

A:B:C:D:E 1.250000e-02

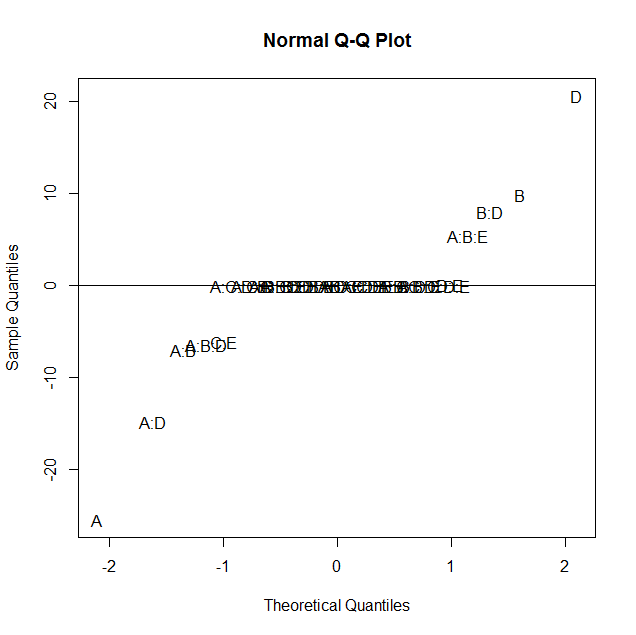
> qq<-qqnorm(Efectos,type="n")

> nombre<-names(Efectos)

> text(qq$x, qq$y, labels = nombre)

> efectos1<-efectos[-c(1,2,3,5,6,9,10,13,16,20)]

> qqline(efectos1)



> mod4<-lm(y~Bloque+A+B+D+A\*B+A\*D+B\*D+C\*E+A\*B\*D+A\*B\*E)

> anva4<-aov(mod4)

> summary(anva4)

Df Sum Sq Mean Sq F value Pr(>F)

Bloque 3 0 0 7.110e-01 0.559

A 1 5187 5187 4.426e+06 <2e-16 \*\*\*

B 1 776 776 6.623e+05 <2e-16 \*\*\*

D 1 3407 3407 2.908e+06 <2e-16 \*\*\*

C 1 0 0 1.067e+00 0.317

E 1 0 0 0.000e+00 1.000

A:B 1 396 396 3.381e+05 <2e-16 \*\*\*

A:D 1 1764 1764 1.505e+06 <2e-16 \*\*\*

B:D 1 507 507 4.328e+05 <2e-16 \*\*\*

C:E 1 299 299 2.551e+05 <2e-16 \*\*\*

B:E 1 0 0 0.000e+00 1.000

A:B:D 1 328 328 2.796e+05 <2e-16 \*\*\*

A:B:E 1 230 230 1.963e+05 <2e-16 \*\*\*

Residuals 16 0 0

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

****

Donde



Para j=1,2,3,y 4



 representa al factor A,  representa al factor C y  representa al factor D

,

 versus , para 

 versus , para  , 

 versus 

 versus 

Los efectos que tienen una influencia altamente significativa sobre la producción de alcohol isoamílico son: A = tipo de cepa, B = temperatura, C = fuente de nitrógeno y urea, D= relación carbono/nitrógeno, E = porcentaje de inóculo; y entre las interacciones

* A = tipo de cepa con B = temperatura
* A = tipo de cepa con D= relación carbono/nitrógeno.
* A = tipo de cepa con E = porcentaje de inóculo.
* B = temperatura con D= relación carbono/nitrógeno
* B = temperatura con E = porcentaje de inóculo
* C = fuente de nitrógeno y urea con E = porcentaje de inóculo.
* D= relación carbono/nitrógeno con E = porcentaje de inóculo.

Y la triple interacciones entre

* A = tipo de cepa con B = temperatura y con D= relación carbono/nitrógeno
* A = tipo de cepa con B = temperatura y con E = porcentaje de inóculo

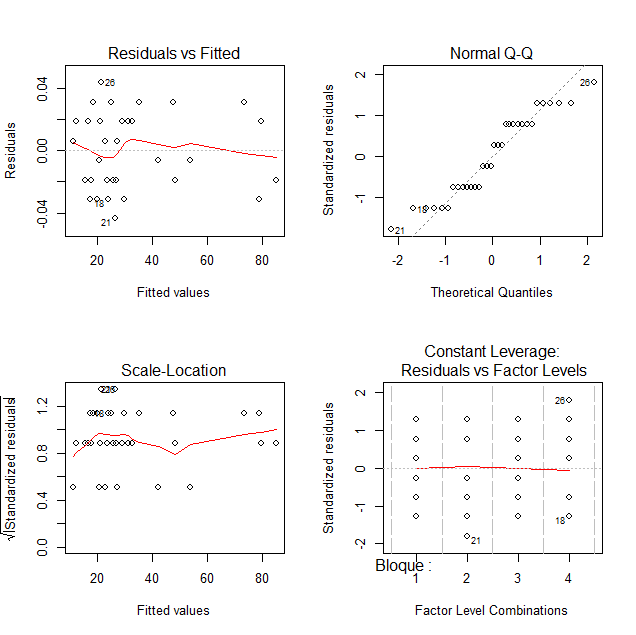
No se encontró diferencias significativas

entre A = tipo de cepa con E = porcentaje de inóculo

entre B = temperatura con E = porcentaje de inóculo

> par(mfrow=c(2,2))

> plot(mod4)



> par(mfrow=c(2,2))

> plot(mod4)

> ri4<-rstandard(mod4)

> shapiro.test(ri4)

Shapiro-Wilk normality test

data: ri4

W = 0.92225, p-value = 0.02394

> library(car)

> ncvTest(mod4)

Non-constant Variance Score Test

Variance formula: ~ fitted.values

Chisquare = 0.000345894 Df = 1 p = 0.9851616

>

1. Suponga que se desea realizar solo una fracción del diseño 25con generadores D=AB y E=AC, genere la estructura de alias, obtenga la fracción a realizar y analice la fracción obtenida. Compare con las conclusiones obtenidas en (a) y (b) (5 puntos)

Estructura de alias

I=ABD=ACE=BCDE

A=BD=CE=ABDE

B=AD=CDE=ABCE

C=AE=BDE=ABCD

D=AB=BCE=ACDE

E=AC=BCD=ABDE

BC=DE=ACD=ABE

BE=CD=ADE=ABC

> A=rep(c(-1,1),4)

> B=rep(c(rep(-1,2),rep(1,2)),2)

> C=c(rep(-1,4),rep(1,4))

> D=A\*B

> E=A\*C

> tequilaf<-data.frame(A,B,C,D,E)

> tequilaf

A B C D E

1 -1 -1 -1 1 1

2 1 -1 -1 -1 -1

3 -1 1 -1 -1 1

4 1 1 -1 1 -1

5 -1 -1 1 1 -1

6 1 -1 1 -1 1

7 -1 1 1 -1 -1

8 1 1 1 1 1

> trat1<-c("de","a","be","abd","cd","ace","bc","abcde")

> tequila3<-data.frame(tequilaf,trat1)

> tequila3

A B C D E trat1

1 -1 -1 -1 1 1 de

2 1 -1 -1 -1 -1 a

3 -1 1 -1 -1 1 be

4 1 1 -1 1 -1 abd

5 -1 -1 1 1 -1 cd

6 1 -1 1 -1 1 ace

7 -1 1 1 -1 -1 bc

8 1 1 1 1 1 abcde

> yf<-c(54,16.8,30,20,48.6,11.4,35.4,25.4)

> mod3<-lm(yf~A+B+C+D+E+BC+CD)

> betaest<-coefficients(mod3)

> efectos<-2\*betaest[-1]

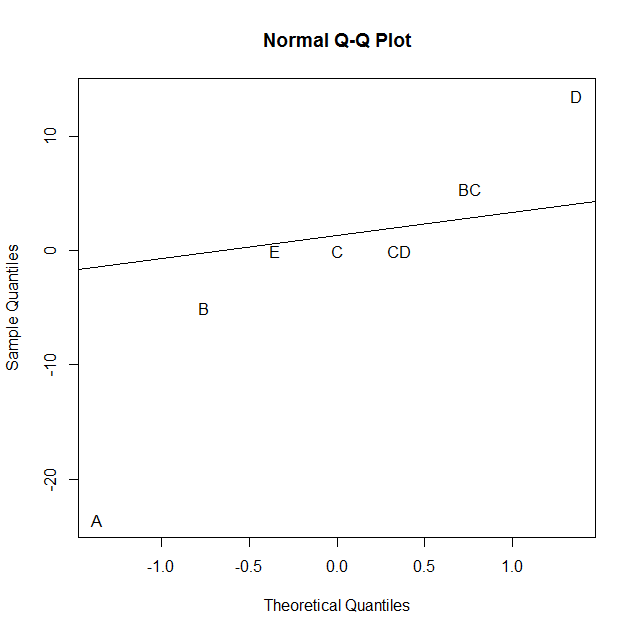
> nombre<-names(efectos)

> qq<-qqnorm(efectos,type="n")

> text(qq$x, qq$y, labels = nombre)

> efectos1<-efectos[-c(1,2,4,5)]

> qqline(efectos1)



A=BD=CE=ABDE

B=AD=CDE=ABCE

C=AE=BDE=ABCD

D=AB=BCE=ACDE

E=AC=BCD=ABDE

BC=DE=ACD=ABE

BE=CD=ADE=ABC

Nombre<-c("A+BD+CE","B+AD","C+AE","D+AB","E+AC","BC+DE","BE+CD")