Taller #3 Diseño de Experimentos 3007340 - Semestre 02 2021: Fecha de entrega Enero 24 de 2022

Nota 1: Además de lo que pida cada problema, especifique claramente lo siguiente:

1. Característica o variable respuesta estudiada y unidad de medida
2. Estructura de tratamientos (factores de diseño controlables, tipos de efectos y niveles)
3. Unidad experimental y estructura de diseño (completamente aleatorizada o diseño en bloques y quién es el factor de bloqueo)
4. La presentación del taller deberá realizarse usando la plantilla para informes de talleres, siguiendo estrictamente las directrices que allí se dan. **Máximo número de páginas para el informe 30. Pero esta vez no adjuntará la programación R.**

Nota 2: El taller consta de cinco partes, cada una sobre un problema experimental diferente: A cada grupo se le asigna un problema. La parte I es asignada igual a todos los grupos; en el resto de partes cada grupo tendrá una asignación de acuerdo a la Tabla 1 al final de este documento. ***Para la parte IV deben estudiar la Sección 9.9 del Capítulo 9 de Notas de Clase.***

Parte I. Todos los grupos:

1. Este punto es para desarrollar sin la ayuda de R: A continuación, se muestran los resultados de un diseño factorial 2^2 con $n = 2$ réplicas en los puntos factoriales

		Réplicas	
A	**B**	**1**	**2**
-1	-1	55.8	54.4
+1	-1	60.3	60.9
-1	+1	63.9	64.4
+1	+1	67.9	68.5

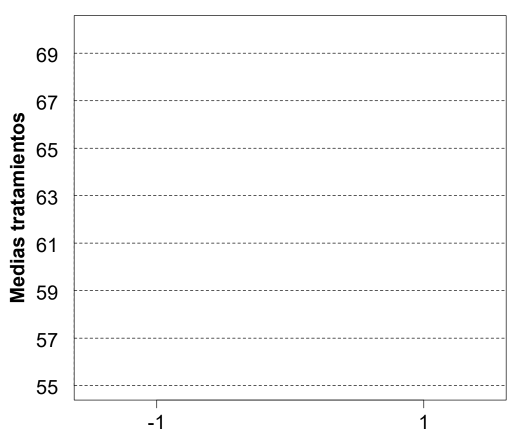
a) Señale los tipos de efectos que se pueden estudiar a través de este diseño.

b) Obtenga para cada efecto el valor de los contrastes de totales, calcule los respectivos efectos y construya la ec. ajustada del modelo de regresión usando variables codificadas ∓ 1.

Yates	total	A	B	AB
(1)				
a				
b				
ab				
	Contraste			
Efectos estimados				
Promedio global estimado:				

Ecuación estimada: $\widehat{Y} =$

c) Construya en el siguiente plano la gráfica de interacción de A vs. B e interprete. Si este análisis indica que la interacción no es significativa replantee modelo de regresión y dé la ec. del modelo predictivo final (imprima esta página, trace manualmente los perfiles con colores diferentes, tome foto y adjunte como imagen en el informe).



d) El diseño factorial es expandido usando $n_c = 4$ réplicas en el centro, quedando el diseño finalmente con los siguientes datos (los factoriales previamente observados y los centrales agregados):

Tratamientos			Réplicas				
i	**A**	**B**	**1**	**2**	**3**	**4**	**Varianzas (S_i^2)**
1=(1)	-1	-1	55.8	54.4	--	--	
2=a	+1	-1	60.3	60.9	--	--	

3=b	-1	+1	63.9	64.4	--	--	
4=ab	+1	+1	67.9	68.5	--	--	
5=centro	0	0	61.5	62.0	61.9	62.4	
$\sum_{i=1}^{5}(n_i-1)S_i^2 =$							

i. Calcule las varianzas muestrales por tratamiento (diligencie la última columna de la tabla anterior) y halle $\sum_{i=1}^{5}(n_i-1)S_i^2$ ¿a cuál suma de cuadrados corresponde esta suma?

ii. A continuación, se presentan resultados R para el análisis del modelo de primer orden

Estimaciones en modelo predictivo

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	61.9917	0.1401	442.3492	0.0000
A	2.3875	0.1716	13.9101	0.0000
B	4.1625	0.1716	24.2516	0.0000
A:B	-0.3625	0.1716	-2.1120	0.0677
R^2_{adj}=0.9861				

ANOVA modelo predictivo

Source	Df	Sum Sq	Mean Sq	F value	Pr($>$F)
A	1	45.6012	45.6012	193.4904	0.0000
B	1	138.6112	138.6112	588.1406	0.0000
A:B	1	1.0512	1.0512	4.4606	0.0677
Residuals	8	1.8854	0.2357		

Escriba la ecuación ajustada del modelo predictivo, en términos de las variables codificadas X_j y pruebe si el modelo predictivo es adecuado; para ello, haga la descomposición apropiada del SSE del modelo en: suma de cuadrados de carencia de ajuste (la única causa posible de carencia de ajuste en este problema es que existan efectos cuadráticos no explicados ya que se han explicado efectos lineales y la interacción de los dos factores) y suma de cuadrados de error puro y realice las pruebas pertinentes (Use un nivel de significancia de 0.05) y evalúe e interprete su R^2_{adj}

Fuente	SS	DF	CM	F0	Fcrítico
Error modelo predictivo					
Error puro					
Carencia de ajuste.					

2. Considere un diseño factorial 2^{5-1}. Si se elige como generador a I=ABCDE,

a) obtenga la matriz de diseño (los tratamientos que constituyen este diseño) y la estructura de Alias.

b) ¿cuáles efectos serán estimables?, es decir ¿cómo se atribuirá la significancia de los contrastes de los efectos aliados? ¿Cuál es la ecuación del modelo de regresión con variables codificadas asociado a esta fracción?

c) ¿Cuál es la resolución del diseño a correr? ¿por qué?

3. Suponga que se desea correr un diseño 2^{5-2} (es decir, fracción 1/4) para ello, se eligen como generadores a ABD y BCE.

a) Escriba la relación definidora correspondiente y diga cuántas fracciones diferentes pueden ser obtenidas y el número de tratamientos distintos que corresponde a cada fracción

b) Usando el método de dos pasos, obtenga una fracción, use la matriz de diseño correspondiente con la notación –1 y +1 e indique con la notación de Yates cuáles tratamientos corresponden a tal fracción.

Notación de Yates	A	B	C	D	E

c) Obtenga la estructura de Alias y con base en ésta defina cuál es la resolución del diseño y dé la ecuación del modelo de regresión con variables codificadas asociado a esta fracción.

Parte II. Experimentos 2^k no replicados y saturados

Problema 1: De acuerdo a la asignación por grupo de la variable respuesta (ver tabla de asignaciones en última página), resolver en R el problema: En el proceso de refinación de pulpa de madera interesa estudiar cómo se afecta la calidad de la fibra al introducirle cargas de material inorgánico con dos diferentes métodos o procesos (mecánico e *in situ*), y ver cómo interactúa al considerar otros factores. Para ello se decide correr un diseño 2^4 no replicado, con los siguientes factores y sus niveles:

Factor	Niveles	
	bajo	**Alto**
A: Proceso	Mecánico	*In situ*
B: Velocidad de agitación (rpm)	2000	3000
C: Tiempo (minutos)	30	60
D: Consistencia de la pulpa (%)	0.5	2.0

Se midieron cuatro variables de respuesta Y_1: Cenizas (%), Y_2: tensión (m), Y_3: Blancura (%), Y_4: Opacidad (%). Los resultados para los 16 tratamientos se muestran a continuación en orden de corrida:

		Variables respuesta medidas				**Complete signos variables codificadas en tratamientos observados en cada corrida**			
Corrida	**Tratamiento**	Y_1	Y_2	Y_3	Y_4	X_1	X_2	X_3	X_4
1	**c**	0.48	579	86.17	77.85				
2	**bcd**	1.46	692	86.50	76.28				
3	**bc**	0.94	581	86.75	76.16				
4	**(1)**	0.49	671	87.22	76.68				
5	**ad**	5.50	653	88.55	79.30				
6	**b**	1.49	867	87.62	77.15				
7	**a**	7.49	496	88.77	79.75				
8	**ab**	11.59	467	88.05	80.35				
9	**ac**	13.23	437	87.30	78.72				
10	**bd**	2.21	631	86.45	78.03				
11	**cd**	5.06	565	85.75	79.42				
12	**abcd**	7.78	549	88.45	80.81				
13	**abd**	11.75	460	88.77	81.43				
14	**abc**	12.57	462	89.42	81.55				
15	**d**	0.72	620	87.67	77.87				
16	**acd**	9.61	568	88.45	78.78				
Variables codificadas X_1 para A, X_2 para B, X_3 para C, X_4 para D									

a) Para la variable respuesta asignada a su grupo, ***analice solo los gráficos de interacciones dobles*** y comente brevemente la forma en que tales factores aparentemente actúan sobre la respuesta de interés.

b) Haga la estimación preliminar de los efectos en el modelo lineal saturado (escriba la ec. de este modelo) e investigue qué efectos influyen significativamente, apóyese en el pareto de efectos y en el gráfico de Daniel (en Daniel use para respuestas Y_1, Y_4 un $\alpha = 0.2$; y para respuestas Y_2, Y_3 use $\alpha = 0.3$)

c) Con base en b) construya la tabla ANOVA aproximada (la del modelo predictivo) eliminando los efectos que aparentan no ser significativos. Lleve al error las interacciones y los efectos que claramente se muestren de poca importancia en los gráficos, pero recuerde además que si retiene una interacción doble, también deben permanecer en el modelo los efectos principales de los dos factores que participan en la interacción. Escriba la ec. teórica y ajustada del del modelo predictivo al que se llega.

d) ¿Cuál tratamiento maximiza la variable respuesta asignada a su grupo? Considere los factores presentes en el modelo predictivo final para dar la respuesta. ***Use los gráficos de superficie de respuesta que sean apropiados para este análisis*** teniendo en cuenta las interacciones que permanezcan en el modelo predictivo. Dé el valor estimado en ese tratamiento usando la ec. estimada del modelo predictivo final.

e) Verifique los supuestos de independencia (realice test de incorrelación con Ljung-Box para m=6 o bien tests ACF con m=6), normalidad (en caso de que no detecte correlaciones en los errores de ajuste) y varianza constante para el error de ajuste del modelo predictivo, usando residuos estudentizados internamente (excepto en tests de incorrelación donde debe usar residuales comunes). Enuncie claramente las hipótesis, estadísticos de prueba y su distribución, criterio de rechazo y conclusión. Recuerde que para las pruebas de incorrelación, los residuos del ajuste deben estar en el orden de corrida y no en el orden de los tratamientos en la matriz de diseño y también, se analizan los gráficos de residuos comunes vs. orden de corrida, residuos estudentizados internamente vs. respuesta estimada y residuos estudentizados internamente vs. variables codificadas (incluso contra aquellas Xj que se hayan eliminado por completo del modelo, para verificar que no haya carencia de ajuste debido a la ausencia de esas variables).

Lectura de los datos y librerías R necesarias

```
library(daewr)
library(rsm)
library(pid)
library(FrF2)

rm(list=ls(all=TRUE))
#Datos ingresados en orden de corrida
problema1=data.frame(scan(what=list(Corrida=0,X1=0,X2=0,X3=0,X4=0,Y1=0,Y2=0,Y3=0,Y4=0)))
1 -1 -1 1 -1 0.48 579 86.17 77.85
2 -1 1 1 1 1.46 692 86.50 76.28
3 -1 1 1 -1 0.94 581 86.75 76.16
4 -1 -1 -1 -1 0.49 671 87.22 76.68
5 1 -1 -1 1 5.50 653 88.55 79.30
6 -1 1 -1 -1 1.49 867 87.62 77.15
7 1 -1 -1 -1 7.49 496 88.77 79.75
8 1 1 -1 -1 11.59 467 88.05 80.35
9 1 -1 1 -1 13.23 437 87.30 78.72
10 -1 1 -1 1 2.21 631 86.45 78.03
11 -1 -1 1 1 5.06 565 85.75 79.42
12 1 1 1 1 7.78 549 88.45 80.81
13 1 1 -1 1 11.75 460 88.77 81.43
14 1 1 1 -1 12.57 462 89.42 81.55
15 -1 -1 -1 1 0.72 620 87.67 77.87
16 1 -1 1 1 9.61 568 88.45 78.78

#escriba aquí resto de programa R usado para respuesta asignada

detach(problema1)
```

Parte III. Experimento 2^k no replicado en puntos factoriales, pero con réplicas en punto al centro

El problema a resolver será de acuerdo a la asignación a los grupos (ver tabla de asignación en la última página):

Problema 2: En una planta de potencia se corrió un experimento factorial 2^4 con tres replicaciones al centro ($n_c = 3$) con el objetivo de hacer más eficaz la operación de una máquina de absorción. La eficacia de la máquina se mide en toneladas de refrigeración entre el flujo de vapor (*tr/fv*). Los factores a controlados fueron:

A: Flujo de vapor. Alto (+1): 4.50, Medio (0): 3.25, Bajo (-1): 2.00;
B: Temperatura de agua helada. Alto (+1): 6.50, Medio (0): 5.25, Bajo (-1): 4.00;
C: Temperatura de agua de enfriamiento. Alto (+1): 27, Medio (0): 25, Bajo (-1): 23
D: Presión diferencial. Alto (+1): 1.7, Medio (0): 1.4, Bajo (-1): 1.1.

Coordenadas del tratamiento en el punto central son (A=3.25, B=5.25, C=25, D=1.4). Tenga en cuenta que el hecho de que se incluyan replicaciones al centro del diseño no cambia el modelo factorial 2^4, es decir, un factorial con 4 factores cada uno con dos niveles al que se adiciona un punto central.
Los datos recolectados se presentan a continuación, en el orden de corrida:

Corrida	¿Tratamiento?	A	B	C	D	Eficacia (tr/fv)
1		4.50	6.50	23	1.1	99
2		3.25	5.25	25	1.4	105
3		2.00	4.00	23	1.7	99
4		2.00	4.00	27	1.7	79
5		4.50	6.50	27	1.7	86
6		2.00	6.50	27	1.1	85
7		4.50	4.00	23	1.1	90
8		2.00	4.00	23	1.1	95
9		4.50	4.00	27	1.7	79
10		4.50	4.00	27	1.1	82
11		4.50	6.50	27	1.1	83
12		4.50	6.50	23	1.7	97
13		3.25	5.25	25	1.4	101
14		3.25	5.25	25	1.4	98
15		2.00	6.50	23	1.7	108
16		2.00	6.50	23	1.1	111
17		2.00	4.00	27	1.1	89
18		4.50	4.00	23	1.7	91
19		2.00	6.50	27	1.7	88

Observaciones resaltadas son los resultados de las réplicas en puntos en el centro

1. Sólo con las observaciones de los16 puntos factoriales (es decir, sin considerar los puntos centrales) realice lo siguiente:

a) Utilice la notación de Yates y anote en la segunda columna de la tabla el código correspondiente a cada una de las corridas, y asegúrese que se corrieron todos los tratamientos del 2^4 correspondientes al diseño factorial completo. ***Construya los gráficos de interacciones dobles para la exploración de los datos*** y comente brevemente la forma en que los factores parecen actuar sobre la respuesta de interés***. Analice también la triple interacción BCD.***

b) Determine el mejor ANOVA para estos datos eliminando los efectos que aparentan no ser significativos, pero parta inicialmente del modelo saturado con todos los posibles efectos y luego seleccione los que aparentan ser significativos usando el gráfico Pareto de efectos y de Daniel (con $\alpha = 0.1$). ¿Cuál es la calidad del ajuste del modelo predictivo? (use el R^2_{adj} para dar respuesta a esta pregunta).

2. Considerando todas las 19 observaciones, es decir tanto las observaciones de los puntos factoriales como del punto central:

a) Ajuste el modelo predictivo hallado en el numeral 1b) y continúe examinando significancia de efectos; depure el modelo dejando sólo los que resulten significativos. En el modelo final, realice la prueba para evaluar la carencia de ajuste y la significancia de efectos cuadráticos. Tenga en cuenta que dado que hay $n_c = 3$ replicaciones en el centro, es posible construir en la tabla ANOVA una descomposición para el SSE del modelo en la suma de cuadrados de efectos cuadráticos puros (ver ec. 14 en documento de clase "Experimentos factoriales 2^k") y la suma de cuadrados de carencia de ajuste por efectos eliminados, y realizar los correspondientes tests ¿Hay presencia de efectos de curvatura? ¿hay carencia de ajuste por efectos eliminados?

b) Determine el mejor tratamiento: Sólo considerando los factores que permanecen con algún efecto significativo en el modelo predictivo final, ***analice los gráficos de superficie de respuesta para determinar el mejor tratamiento*** (¿qué interesa? ¿maximizar o minimizar la respuesta media?) y haga la predicción de la eficacia esperada en el tratamiento óptimo con el modelo predictivo.

c) En el modelo final, verifique los supuestos de independencia (realice test de incorrelación con Ljung-Box para m=6 o bien tests ACF con m=6), normalidad (en caso de que no detecte correlaciones en los errores de ajuste) y varianza constante, para el error de ajuste del modelo predictivo, usando residuos estudentizados internamente (excepto en tests de incorrelación donde debe usar residuales comunes). Enuncie claramente las hipótesis, estadísticos de prueba y su distribución, criterio de rechazo y conclusión. Recuerde que para las pruebas de incorrelación, los residuos del ajuste deben estar en el orden de corrida y no en el orden de los tratamientos en la matriz de diseño, también se deben analizar los gráficos de residuos vs. orden de corrida, residuos estudentizados internamente vs. respuesta estimada y residuos estudentizados internamente vs. variables codificadas (incluso contra aquellas X_j que se hayan eliminado por completo del modelo, para verificar que no haya carencia de ajuste debido a la ausencia de esas variables).

Lectura de los datos y librerías R necesarias y análisis sólo con puntos factoriales, modelo saturado

```
rm(list=ls(all=TRUE))
library(daewr)
library(rsm)
library(pid)
library(FrF2)
#Lectura todos los datos
problema2=data.frame(scan(what=list(Corrida=0,A=0,B=0,C=0,D=0,Y=0,PuntoCentral="")))
1 4.50 6.50 23 1.1 99 no
2 3.25 5.25 25 1.4 105 si
3 2.00 4.00 23 1.7 99 no
4 2.00 4.00 27 1.7 79 no
5 4.50 6.50 27 1.7 86 no
6 2.00 6.50 27 1.1 85 no
7 4.50 4.00 23 1.1 90 no
8 2.00 4.00 23 1.1 95 no
9 4.50 4.00 27 1.7 79 no
10 4.50 4.00 27 1.1 82 no
11 4.50 6.50 27 1.1 83 no
12 4.50 6.50 23 1.7 97 no
13 3.25 5.25 25 1.4 101 si
14 3.25 5.25 25 1.4 98 si
15 2.00 6.50 23 1.7 108 no
16 2.00 6.50 23 1.1 111 no
17 2.00 4.00 27 1.1 89 no
18 4.50 4.00 23 1.7 91 no
19 2.00 6.50 27 1.7 88 no

#Crear variables Xj que corresponden a los factores con niveles codificados: nivel alto=+1, bajo=-1 y medio=0
#Función usuario para crear las Xj en unidades codificadas de -1, 0, y 1
codificación=function(factor){
fcod=(factor-mean(factor))/((max(factor)-min(factor))/2)
fcod
}

X1<-codificación(factor=problema2$A)
X2<-codificación(factor=problema2$B)
X3<-codificación(factor=problema2$C)
X4<-codificación(factor=problema2$D)

problema2=data.frame(problema2,X1,X2,X3,X4)

################################################################################
#Punto 1. Análisis sólo con puntos factoriales (excluyendo los puntos centrales)
################################################################################
problema2f=problema2[problema2$PuntoCentral=="no",] #data.frame sin los puntos centrales

#Gráficos de interacción sólo con puntos factoriales
with(problema2f,{
interaction.plot(X1,X2,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=F,xlab="A")
legend("topright",legend=c("B=-1","B=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

win.graph()
interaction.plot(X1,X3,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=F,xlab="A")
legend("topright",legend=c("C=-1","C=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

win.graph()
interaction.plot(X1,X4,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=F,xlab="A")
legend("topright",legend=c("D=-1","D=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

win.graph()
interaction.plot(X2,X3,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=F,xlab="B")
legend("topleft",legend=c("C=-1","C=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

win.graph()
interaction.plot(X2,X4,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=F,xlab="B")
legend("topleft",legend=c("D=-1","D=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

win.graph()
interaction.plot(X3,X4,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=F,xlab="C")
legend("topright",legend=c("D=-1","D=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

#ANÁLISIS INTERACCIÓN BCD
win.graph()
interaction.plot(X2[X4=="-1"],X3[X4=="-1"],Y[X4=="-1"],type="b",pch=c(1,2),col=c("black","red"),
                 lwd=4,cex=2,cex.lab=1.5,legend=F,xlab="B",main="B*C en D=-1",ylab="Y")
legend("topleft",legend=c("C=-1","C=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

win.graph()
interaction.plot(X2[X4=="1"],X3[X4=="1"],Y[X4=="1"],type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,
                 cex.lab=1.5,legend=F,xlab="B",main="B*C en D=1",ylab="Y")
legend("topleft",legend=c("C=-1","C=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)
})

#Modelo saturado sin usar puntos centrales
modelo1f=lm(Y~X1*X2*X3*X4,data=problema2f)
summary(modelo1f)
```

```
win.graph()
paretoPlot(modelo1f,negative=c("Negativo","orange"),positive=c("Positivo","blue"))

win.graph()
halfnormal(modelo1f,code=T,alpha=0.1,linelwd=2,linecol=2,err.points=T, pch.set = c(19, 16, 8))

#Depurar modelo con base en análisis de pareto de efectos y gráfico de Daniel
#Con el análisis anterior se define modelo eliminando los efectos no significativos para comprobar
#significancia de los que quedan.
#Recuerde que si se retiene una triple interacción, entonces deben retenerse todos los efectos de interacciones
#dobles y efectos principales de los factores involucrados en la interacción triple

#escriba código R para estimación modelo final depurado con puntos factoriales

########################################
#punto 2 (Análisis con todos los datos)
########################################
#Después de identificar los efectos que deben quedar según análisis con puntos factoriales, ajuste ese
#modelo usando ahora a las variables Y, X1, X2, X3, X4 según los efectos a considerar y usando todos los datos
#guardados en data.frame "problema2"

attach(problema2)

#escriba aquí programa R


detach(problema2)
```

Problema 3. Se llevó a cabo un experimento en un proceso químico que produce un polímero. Cuatro factores fueron estudiados:
A: temperatura, nivel bajo: 100°C, nivel alto: 120°C
B: Concentración del catalizador, nivel bajo: 4%, nivel alto 8%
C: tiempo, nivel bajo 20min, nivel alto 30min
D: presión, nivel bajo 60psi, nivel alto: 75psi
Se corrieron además $n_c = 4$ réplicas en el punto central del diseño
Como respuesta se observó el peso molecular. La matriz de diseño y los datos son mostrados a continuación (*Los resultados observados están listados en orden de corrida*)

orden de corrida	¿Tratamiento?	Y	A	B	C	D
1		2515	0	0	0	0
2		2300	-1	1	-1	1
3		2610	-1	-1	1	-1
4		2625	-1	-1	1	1
5		2500	0	0	0	0
6		2400	-1	-1	-1	1
7		2390	1	-1	-1	1
8		2510	1	1	-1	-1
9		2410	1	-1	-1	-1
10		2520	1	1	-1	1
11		2625	1	-1	1	-1
12		2475	0	0	0	0
13		2315	-1	1	-1	-1
14		2400	-1	1	1	-1
15		2500	-1	1	1	1
16		2400	0	0	0	0
17		2750	1	1	1	-1
18		2400	-1	-1	-1	-1
19		2630	1	-1	1	1
20		2710	1	1	1	1

Observaciones resaltadas son los resultados de las réplicas en puntos en el centro

1. Sólo con los 16 datos de los puntos factoriales (no considere aún los puntos centrales):
a) Utilice la notación de Yates y anote en la segunda columna de la tabla el código correspondiente a cada una de las corridas, y asegúrese que se corrieron todos los tratamientos correspondientes al diseño 2^4 completo. ***Construya solo los gráficos de interacciones dobles para la exploración de los datos*** y comente brevemente la forma en que los factores parecen actuar sobre la respuesta de interés.
b) Determine el mejor ANOVA para estos datos, pero considere inicialmente todos los posibles efectos y luego seleccione aquellos que aparentan ser significativos, usando el gráfico Pareto de efectos y de Daniel (con $\alpha = 0.1$). ¿Cuál es la calidad del ajuste del modelo predictivo? (use el R^2_{adj} para dar respuesta a esta pregunta).

2.Considerando las 20 observaciones, es decir, tanto los puntos factoriales y los centrales:
a) Ajuste de nuevo el modelo predictivo hallado en el numeral 1b). Realice la prueba para evaluar la carencia de ajuste y la significancia de efectos cuadráticos. Dado que hay $\boldsymbol{n_c = 4}$ réplicas al centro, es posible construir en la tabla ANOVA una descomposición para el SSE del modelo en la suma de cuadrados de efectos cuadráticos puros (ver ec. 14 en documento de clase "Experimentos factoriales 2^k") y la suma de cuadrados de carencia de ajuste por efectos eliminados, y realizar los correspondientes tests ¿Hay presencia de efectos de curvatura? ¿hay carencia de ajuste por efectos eliminados?

b) Sólo considerando los factores que permanecen con algún efecto en el modelo predictivo, ***analice los gráficos de superficie de respuesta para determinar el mejor tratamiento*** considerando que el proceso es mejor cuando se maximiza el peso molecular. Estime la respuesta en el mejor tratamiento de acuerdo al análisis anterior; utilice para ello la ecuación ajustada del modelo predictivo con base en las variables codificadas para los factores.

c) En el modelo final verifique los supuestos de independencia (realice test de incorrelación con Ljung-Box para m=6 o bien tests ACF con m=6), normalidad (en caso de que no detecte correlaciones en los errores de ajuste) y varianza constante para el error de ajuste del modelo predictivo, usando residuos estudentizados internamente (excepto en tests de incorrelación donde debe usar residuales comunes). Enuncie claramente las hipótesis, estadísticos de prueba y su distribución, criterio de rechazo y conclusión. Recuerde que para las pruebas de incorrelación, los residuos del ajuste deben estar en el orden de corrida y no en el orden de los tratamientos en la matriz de diseño, además, también se analizan los gráficos de residuos comunes vs. orden de corrida, residuos estudentizados internamente vs. respuesta estimada y residuos estudentizados internamente vs. variables codificadas (incluso contra aquellas X_j que se hayan eliminado por completo del modelo, para verificar que no haya carencia de ajuste debido a esas variables omitidas).

Lectura de los datos y librerías R necesarias y análisis sólo con puntos factoriales, modelo saturado

```
library(daewr)
library(rsm)
library(pid)
library(FrF2)
rm(list=ls(all=TRUE))
#Datos en orden de corrida. Se ingresan directamente las variables Xj codificadas en lugar de A, B, C, D
problema3=data.frame(scan(what=list(Corrida=0,Y=0,X1=0,X2=0,X3=0,X4=0,PuntoCentral="")))
1 2515 0 0 0 0 si
2 2300 -1 1 -1 1 no
3 2610 -1 -1 1 -1 no
4 2625 -1 -1 1 1 no
5 2500 0 0 0 0 si
6 2400 -1 -1 -1 1 no
7 2390 1 -1 -1 1 no
8 2510 1 1 -1 -1 no
9 2410 1 -1 -1 -1 no
10 2520 1 1 -1 1 no
11 2625 1 -1 1 -1 no
12 2475 0 0 0 0 si
13 2315 -1 1 -1 -1 no
14 2400 -1 1 1 -1 no
15 2500 -1 1 1 1 no
16 2400 0 0 0 0 si
17 2750 1 1 1 -1 no
18 2400 -1 -1 -1 -1 no
19 2630 1 -1 1 1 no
20 2710 1 1 1 1 no

##########################################################################
#Análisis sólo con puntos factoriales (excluyendo los puntos centrales)
##########################################################################
problema3f=problema3[problema3$PuntoCentral=="no",]

#Gráficos de interacción sólo con puntos factoriales
with(problema3f,{
win.graph()
interaction.plot(X1,X2,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=,xlab="A")
legend("topleft",legend=c("B=-1","B=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

win.graph()
interaction.plot(X1,X3,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=,xlab="A")
legend("topleft",legend=c("C=-1","C=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

win.graph()
interaction.plot(X1,X4,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=,xlab="A")
legend("topleft",legend=c("D=-1","D=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

win.graph()
interaction.plot(X2,X3,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=,xlab="B")
legend("bottomleft",legend=c("C=-1","C=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

win.graph()
interaction.plot(X2,X4,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=,xlab="B")
legend("bottomleft",legend=c("D=-1","D=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)

win.graph()
interaction.plot(X3,X4,Y,type="b",pch=c(1,2),col=c("black","red"),lwd=4,cex=2,cex.lab=1.5,legend=,xlab="C")
legend("topleft",legend=c("D=-1","D=+1"),col=1:2,pch=1:2,lwd=2,lty=c(2,1),bty="n",cex=2)
})

#Modelo saturado sin usar puntos centrales
modelo1f=lm(Y~X1*X2*X3*X4,data=problema3f)
summary(modelo1f)
anova(modelo1f)

win.graph()
paretoPlot(modelo1f,negative=c("Negativo","orange"),positive=c("Positivo","blue"))

win.graph()
halfnormal(modelo1f,code=T,alpha=0.1,linelwd=2,linecol=2,err.points=T, pch.set = c(19, 16, 8))
```

```
#Corra el modelo eliminando los efectos no significativos para comprobar
#significancia de los que quedan. El modelo resulta
#recuerde que si se retiene una interacción doble, también debe retenerse
#los efectos principales de los dos factores que intervienen en la interacción

#escriba código R para estimación modelo final depurado con puntos factoriales


########################################
#punto 2 (Análisis con todos los datos)
########################################
#Después de identificar los efectos que deben quedar según análisis con puntos factoriales, ajuste ese
#modelo usando ahora a las variables Y, X1, X2, X3, X4 según los efectos a considerar y usando todos los datos
#guardados en data.frame "problema3"

attach(problema3)

#escriba aquí programa R


detach(problema3)
```

Parte IV. 2^k no replicado, en bloques incompletos (para este punto deben estudiar la sección 9.9 de notas de clase): Cada grupo, de acuerdo a variable asignada (ver tabla de asignación en la última página), deberá realizar la programación R apropiada, incluso para la lectura correcta de los datos.

Problema 4: Una aleación de titanio-níquel es usada en componentes de motores de turbinas de aviones. El agrietamiento es un problema potencialmente serio en la parte final, pues puede conducir a falla catastrófica. Una prueba se realiza para determinar el efecto de cuatro factores sobre las grietas. Estos factores son

A: Temperatura de colada
B: Contenido de titanio
C: Método de tratamiento térmico
D: Cantidad de refinador de grano usado.

Suponga que sólo 8 de las 16 corridas podían hacerse en un solo día de modo que cada día debía considerarse como un bloque.

1. Usando como efecto a confundir con los bloques a ABCD, halle la conformación de los bloques, es decir, indique cuáles tratamientos se deben correr en cada día, use la notación de yates para identificar los tratamientos

2. Los datos recolectados se muestran a continuación (Ud. trabajará con Y1 ó Y2 según la asignación a su grupo):

Yates	Y1	Y2	A	B	C	D
(1)	7.037	6.376	-1	-1	-1	-1
a	14.707	15.219	1	-1	-1	-1
b	11.635	12.089	-1	1	-1	-1
ab	17.273	17.815	1	1	-1	-1
c	10.403	10.151	-1	-1	1	-1
ac	4.368	4.098	1	-1	1	-1
bc	9.360	9.253	-1	1	1	-1
abc	13.440	12.923	1	1	1	-1
d	8.561	8.951	-1	-1	-1	1
ad	16.867	17.052	1	-1	-1	1
bd	13.876	13.658	-1	1	-1	1
abd	19.824	19.639	1	1	-1	1
cd	11.846	12.337	-1	-1	1	1
acd	6.125	5.904	1	-1	1	1
bcd	11.190	10.935	-1	1	1	1
abcd	15.635	15.053	1	1	1	1

Para el conjunto de datos de la respuesta asignada a su grupo:

a) ***Construya solo los gráficos de interacciones dobles para la exploración de los datos*** y comente brevemente la forma en que tales factores parecen actuar sobre la respuesta de interés. ***Analice además la triple interacción ABC.***

b) Considerando el diseño en bloques, determine el mejor ANOVA para estos datos, eliminando los efectos que aparentan no ser significativos, pero parta inicialmente del modelo saturado con todos los posibles efectos y bloques y luego, seleccione los efectos que resulten significativos usando el gráfico Pareto de efectos y de Daniel (use $\alpha = 0.2$, además tenga en cuenta que efectos de bloques e intercepto no se consideran en la gráfica Pareto ni en gráfico Daniel). ¿Cuál es la calidad del ajuste del modelo predictivo? (use el R^2_{adj} para dar respuesta a esta pregunta).

c) Identifique los efectos significativos e interprételos; además, determine con la seudo prueba F si fue útil el bloqueo. Encuentre también el mejor tratamiento (tenga en cuenta que la respuesta tiene que ver con una medida de la dimensión de las grietas en la estructura del motor). Estime la respuesta en el mejor tratamiento de acuerdo al análisis anterior. Utilice para ello la ecuación ajustada del modelo predictivo con base en las variables codificadas para los factores.

d) Verifique los supuestos, normalidad y varianza constante para los errores del modelo predictivo final usando residuos estudentizados internamente.

Parte V. Experimento 2^{k-p}

Problema a realizar según asignación en la tabla de la página final. Cada grupo deberá realizar la programación R apropiada incluso para la lectura correcta de los datos. ***CUIDADO: la función FrF2 tal vez no produzca por defecto, la fracción que es usada en el problema 5, así que mejor no la use con tal fin!,*** ud. debe descubrir cuál fue o fueron los generadores para la fracción usada e ingresar los datos con la función data.frame combinada con scan(what=list(…)) definiendo de una vez los valores de las variables codificadas tomando valores -1 ó +1 de acuerdo a los tratamientos observados.

Problema 5: A continuación, se muestran los tratamientos y los resultados obtenidos de un experimento factorial 2^{5-1}:

(1)=700	**de**=2515
ae=1317	**ad**=2507
be=468	**bd**=2247
ab=424	**abde**=2232
ce=580	**cd**=2031
ac=2247	**acde**=2314
bc=446	**bcde**=2262
abce=468	**abcd**=2299

a) ¿Cuál es el generador de esta fracción factorial? (defina las dos fracciones posibles y determine cuál fue la que se usó en este caso) ¿Cuál es la resolución de este diseño (halle la estructura de alias para resolver esta pregunta)? ¿Qué significa?

b) ¿Obtenga un diagrama Pareto de efectos (recuerde que los efectos que se ingresan inicialmente en el modelo saturado se definen de acuerdo a la estructura de alias, y por tanto, según la resolución del diseño aplicado) y un gráfico de probabilidad normal de efectos (use $\alpha = 0.3$)¿Cuáles efectos parecen importantes? Obtenga el mejor ANOVA

c) Si lo que se quiere es maximizar, ¿cuál es el mejor tratamiento?

d) ¿Puede colapsar el diseño factorial fraccionado en un factorial completo? Argumente.

e) Elimine completamente del modelo los efectos que menos impacto tuvieron, colapse el diseño 2^{5-1} en un diseño 2^3, ¿Cuántas réplicas tiene el diseño colapsado? Analice en detalle el diseño colapsado y obtenga conclusiones. ¿Las conclusiones obtenidas en el diseño colapsado son las mismas del diseño fraccionado?

Problema 6: En una empresa panificadora existen problemas con la simetría y el color del pan integral. Los responsables del proceso sospechan que el problema se origina desde la fase de fermentación. En ésta se combina agua, harina, cierta cantidad de levadura más una serie de ingredientes. Al final de la fermentación se obtiene lo que se llama "esponja líquida" la cual debe cumplir una serie de parámetros de calidad:
ATT: Acidez total titulable mayor a 6.0
PH: Mayor a 4.8.
Sin embargo, no se ha cumplido con dichas exigencias de calidad; se han hecho algunos intentos experimentales con un factor a la vez, pero los resultados han sido malos. En busca de una mejor estrategia experimental se decide utilizar un diseño factorial fraccionado 2^{6-2} para investigar el efecto de seis factores en las variables ATT y pH. Los primeros cinco factores (con sus respectivos niveles entre paréntesis) se refieren a cierta cantidad que se agrega en la fermentación:
A: Levadura (17, 19),
B: Sal (2.5, 3.7),
C: Fosfato (2.0, 3.6),
D: Sulfato (1.5, 2.2),
E: Cloruro (0.89, 1.20).
El sexto factor es **F:** Temperatura inicial del agua (22, 26). Los datos obtenidos se muestran en la siguiente tabla:

Orden Corrida	Matriz de diseño						Variables respuesta	
	A	B	C	D	E	F	ATT	pH
9	-1	-1	-1	-1	-1	-1	6.2	4.86
5	+1	-1	-1	-1	+1	-1	5.6	4.86
6	-1	+1	-1	-1	+1	+1	5.8	4.85
1	+1	+1	-1	-1	-1	+1	5.8	4.99
14	-1	-1	+1	-1	+1	+1	5.7	4.94
10	+1	-1	+1	-1	-1	+1	6.4	4.74
13	-1	+1	+1	-1	-1	-1	6.4	4.83
12	+1	+1	+1	-1	+1	-1	6.6	4.85
11	-1	-1	-1	+1	-1	+1	5.3	4.81
3	+1	-1	-1	+1	+1	+1	6.6	4.81
15	-1	+1	-1	+1	+1	-1	5.2	4.98
16	+1	+1	-1	+1	-1	-1	5.5	4.98
8	-1	-1	+1	+1	+1	-1	6.9	4.84
4	+1	-1	+1	+1	-1	-1	7.1	4.85
2	-1	+1	+1	+1	-1	+1	6.7	4.96
7	+1	+1	+1	+1	+1	+1	6.9	4.84

a) Identifique la fracción usada, lo cual implica que debe determinar cuáles fueron los dos contrastes de definición usados y la fracción tomada (clave: vea en notas de clase ejemplos dados para k=6 y p=2) ¿Cuál es la resolución de este diseño y qué significa ésta? Escriba la estructura alias reducida ¿Cómo se pueden interpretar efectos de interacción que son alias?

b) ¿Cuáles efectos explican el comportamiento de la variable respuesta asignada? Encuentre el mejor ANOVA e interprete resultados. Recuerde que debe definir primero con la estructura de alias cuáles efectos van en el modelo de regresión saturado, correrlo y con la ayuda de gráfico Pareto y de probabilidad media normal (use $\alpha = 0.2$) decidir cuáles mandar al error.

c) Determine las condiciones de operación que conlleven a maximizar la variable respuesta asignada (se trata de señalar el mejor tratamiento dado que este experimento no es de superficie de respuesta). Recuerde también que el tratamiento a recomendar debe cumplir con los parámetros de calidad exigidos.

d) Verifique los supuestos en el modelo final: independencia (realice test de incorrelación con Ljung-Box para m=6 o bien tests ACF con m=6), normalidad (en caso de que no detecte correlaciones en los errores de ajuste) y varianza constante, para los errores de ajuste del modelo predictivo, usando residuos estudentizados internamente (excepto en tests de incorrelación donde debe usar residuales comunes). Enuncie claramente las hipótesis, estadísticos de prueba y su distribución, criterio de rechazo y conclusión. Recuerde que para las pruebas de incorrelación, los residuos del ajuste deben estar en el orden de corrida y no en el orden de los tratamientos en la matriz de diseño, además, también se analizan los gráficos de residuos comunes vs. orden de corrida, residuos estudentizados internamente vs. respuesta estimada y residuos estudentizados internamente vs. variables codificadas (incluso contra aquellas Xj que se hayan eliminado por completo del modelo, para verificar que no haya carencia de ajuste debido a esas variables omitidas).

TABLA 1: ASIGNACIONES A LOS GRUPOS

	Parte II				**Parte III**		**Parte IV**		**Parte V**		
	Problema 1						**Problema 4**		**Problema 5**	**Problema 6**	
grupo	**Y1**	**Y2**	**Y3**	**Y4**	**problema 2**	**problema 3**	**Y1**	**Y2**		**ATT**	**pH**
1	x				x		x		x		
2		x				x		x		x	
3			x		x		x		x		
4				x		x		x			x
5		x			x			x		x	
6	x					x	x		x		
7			x			x		x			x
8				x	x		x		x		
9	x					x	x			x	