Lista de Ejercicios No.1 - Diseño de Experimentos

Experimentos de comparaciones simples: Muestras independientes y pareadas

Prof. Nelfi González Alvarez - Escuela de Estadística UNAL-Medellín

1. Aspectos teóricos

Experimentos para comparar dos tratamientos. Muestras independientes y muestras pareadas, realizar Lectura: Secciones 2.4 a 2.6 de Montgomery, D. C. (2013). *Design and analysis of experiments. 8th Edition.* John Wiley & Sons.

1. ¿En qué consisten las muestras pareadas? Dé un ejemplo de datos o muestras pareadas.
2. ¿Cómo se prueba la igualdad de medias de dos tratamientos o poblaciones a través de muestras pareadas y de muestras independientes? Asuma la normalidad.

2. Ejercicios propuestos

Para los siguientes problemas tenga presente postular los supuestos distribucionales necesarios para poder hacer inferencia. Determine en cada caso si se trata de experimentos con muestras independientes o pareadas y resuelva como convenga.

1. Dos máquinas, cada una operada por una persona, son utilizadas para cortar tiras de hule, cuya longitud ideal debe ser de 200 mm. De las inspecciones de una semana (25 piezas) se observa que la longitud media de las 25 piezas para una máquina es de 200.1 mm y para la otra es de 201.2 mm. ¿Es significativa la diferencia entre los dos casos? Argumente.
2. Se desea comprar una gran cantidad de bombillas y se tiene que elegir entre las marcas A y B. Para ello, se compraron 100 focos de cada marca y se encontró que las bombillas probadas de la marca A tuvieron un tiempo de vida medio de 1120 horas con una desviación estándar de 75 horas, mientra que las de la marca B tuvieron un tiempo de vida medio de 1064 horas con una desviación estándar de 82 horas.
   a) ¿Es significativa la diferencia entre los tiempos medios de vida? use una significancia α =0.05.
   b) ¿Con qué tamaño de muestra se aceptaría que los tiempo medios de vida son iguales a un nivel de significancia α =0.05?
3. En un laboratorio bajo condiciones controladas, se evalúo para 10 hombres y 10 mujeres, la temperatura que cada persona encontró más confortable. Los resultados en F fueron lo dsiguientes,

Mujer	75	77	78	79	77	73	78	79	78	80
Hombre	74	72	77	76	76	73	75	73	74	75

   a) ¿Cuáles son los tratamientos que se comparan en este estudio?
   b) ¿Las muestras son dependientes o independientes? Explique.
   c) ¿Es la temperatura media más confortable igual para hombres que para mujeres?
   d) Dibuje los boxplots comparativos e interprete.
4. Se prueban diez piezas en cada nivel de temperatura y se mide el encogimiento sufrido en unidades de porcentaje multiplicado por 10. Los resultados son

Temperatura baja	Temperatura alta
17.2	21.4
17.5	20.9
18.6	19.8
15.9	20.4
16.4	20.6
17.3	21.0
16.8	20.8
18.4	19.9
16.7	21.1
17.6	20.3

   a) ¿La temperatura tiene algún efecto en el encogimiento? Plantee las hipótesis estadísticas correspondientes a este interrogante.
   b) Dé un intervalo de confianza para la diferencia de medias.

c) Compare las varianzas en cada temperatura.

d) Dibuje los boxplots comparativos e interprete.

5. Una empresa transportadora de carga desea escoger la mejor ruta para llevar la mercancía de un depósito a otro. La mayor preocupación es el tiempo de viaje. En el estudio se seleccionaron cinco conductores al alzar de un grupo de 10 y se asignaron a la ruta A; los cinco restantes se asignaron a la ruta B. Los datos obtenidos fueron los siguientes.

Ruta	Tiempo de viaje				
A	18	24	30	21	32
B	22	29	34	25	35

a) ¿Existen diferencias significativas entre las rutas? Plantee y pruebe las hipótesis estadísticas correspondientes.

b) En caso de rechazar H_0 en a), dibuje los boxplots comparativos para decidir cuál ruta es mejor.

c) Sugiera otro modo de obtener los datos (otro diseño alternativo) para este estudio que pudiera lograr una comparación más efectiva.

6. Se tienen dos proveedores de una pieza metálica cuyo diámetro ideal o valor objetivo es igual a 20.25cm. Se toman dos muestras de 14 piezas a cada proveedor y los datos obtenidos fueron los siguientes:

Proveedor	Diámetro de las piezas						
1	21.38	20.13	19.12	19.85	20.54	8.00	22.24
	21.94	19.07	18.60	21.89	22.60	18.10	19.25
2	21.51	22.22	21.49	21.91	21.52	22.06	21.51
	21.29	22.71	22.65	21.53	22.22	21.92	20.82

a) Describa un procedimiento de aleatorización para la obtención de estos datos.

b) Pruebe la hipótesis de igualdad de los diámetros de las piezas de los proveesores en cuanto a sus medias.

c) Pruebe la hipótesis de igualdad de varianzas.

d) Si las epecificaciones para el diámetro son 20.25cm$\pm$2.25cm ¿cuál proveedor produce menos piezas defectuosas?

e) ¿Con cuál proveedor se quedaría Ud.? ¿Por qué?

7. En una prueba de dureza, una bola de acero se presiona contra el material al que se mide la dureza. El diámetro de la depresión en el material es la medida de su dureza. Se dispone de dos tipos de bolas de acero y se quiere estudiar su desempeño. Para ello, se prueban ambas bolas con los mismos 10 especímenes elegidos aleatoriamente y los resultados son,

Bola X	75	46	57	43	58	32	61	56	34	65
Bola Y	52	41	43	47	32	49	52	44	57	60

a) Analice paso a paso cómo se hizo el experimento y explique por qué es importante hacerlo asi.

b) Pruebe la hipótesis de que ambas bolas dan la misma medida de dureza.

c) Pruebe la igualdad de las bolas sin considerar que están pareadas. Compare con lso resultados obtenidos en b)

d) ¿En qué situación se esperaría que los análisis en b) y c) den los mismos resultados?

8. Se conduce un experimento para determinar si el uso de un aditivo químico y un fertilizante estándar aceleran el crecimiento de las plantas. En cada una de 10 localidades se estudiaron dos plantas sembradas en condiciones similares. A una planta de cada lote se le aplicó el fertilizante puro y a la otra el fertilizante más el aditivo. Después de cuatro semanas el crecimiento en cms. fue el siguiente,

	Localidad									
Tratamiento	1	2	3	4	5	6	7	8	9	10
Sin aditivo	20	31	16	22	19	32	25	18	20	19
Con aditivo	23	34	15	21	22	31	29	20	24	23

a) ¿Los datos obtenidos apoyan la afirmación de que el aditivo químico acelera el crecimiento de las plantas? Plantee las hipótesis apropiadas y pruébelas usando $\alpha = 0.05$ de significancia.

b) Obtenga un intervalo de confianza del 95 % para la diferencia media.

c) Explique con detalle cómo se pueden asignar aleatoriamente los tratamientos a las plantas en cada localidad utilizando una moneda.

d) Suponga que en cada localidad una planta queda hacia el este y la otra hacia el oeste, realice un asignación aleatoria de los tratamientos a las plantas lanzando una moneda 10 veces.

9. Se realizó un experimento para ver si dos técnicos tienen alguna tendencia a obtener diferentes resultados cuando determinan la pureza de cierto producto. Cada muestra fue dividida en dos porciones y cada técnico determinó la pureza de una de las porciones. Los resultados se muestran a continuación.

	Pureza de las muestras							
Técnico	1	2	3	4	5	6	7	8
1	74.0	73.1	73.5	73.9	71.2	72.5	73.0	74.3
2	73.0	71.3	73.2	71.1	70.3	71.5	73.4	72.4

a) ¿Estos datos deben analizarse en forma pareada? Explique por qué.

b) Formule la hipótesis correcta al problema.

c) Prueba la hipótesis y obtenga conclusiones.

d) Si los técnicos son diferentes, ¿hay alguna evidencia sobre cuál de ellos hace mal el trabajo?

e) ¿Qué recomendaría para lograr mayor uniformidad en las mediciones de los dos técnicos?

10. En una empresa lechera se tienen varios silos para almacenar leche (cisternas de 60000 litros). Un aspecto crítico para que se conserve la leche es la temperatura de almacenamiento. Tradicionalmente se han usado termómetros de mercurio (Mer) para verificar que la temperatura sea la adecuada; pero se han comprado termómetros electrónicos (Rtd) para facilitar el proceso de medición. Sin embargo, se duda de las mediciones de estos nuevos dispositivos. Para aclarar dudas y diagnosticar la situación, durante cinco días se toman mediciones con ambos tipos de termómetros en varios silos (a la misma hora). Los datos para los cinco silos se muestran a continuación.

	Medición 1		Medición 2		Medición 3		Medición 4		Medición 5	
Silo	Mer	Rtd	Mer	Rtd	Mer	Rtd	Mer	Rtd	Mer	Rtd
A	4.0	2.6	4.0	2.8	5.0	5.0	0.5	0.0	3.0	2.4
B	5.0	6.4	6.0	6.4	2.0	2.3	4.0	4.2	4.0	4.0
C	4.5	3.3	4.0	1.4	3.5	1.8	2.0	-1.9	3.0	-7.6
D	2.5	3.1	4.0	5.0	6.5	6.6	4.5	2.7	4.0	6.3
E	4.0	0.0	4.0	0.4	3.5	0.6	2.0	-4.0	4.0	-6.3

a) Observe los datos y establezca una conjetura acerca de la confiabilidad de las mediciones con Rtd (del termómetro de mercurio no hay duda).

b) Ignorando el silo, formule la hipótesis de mayor interés en este problema.

c) Pruebe la hipótesis que formuló y obtenga conclusiones.

11. En la fabricación de semiconductores, a menudo se utiliza una sustancia química para quitar el silicio de la parte trasera de las obleas antes de la metalización. En este proceso es importante la rapidez con la que actúa la sustancia. Se han comparado dos soluciones químicas, utilizando para ello dos muestras aleatorias de 10 obleas para cada solución. La rapidez de acción observada es la siguiente (en mils/min):

solución 1		solución 2	
9.9	10.6	10.2	10.0
9.4	10.3	10.6	10.2
9.3	10.0	10.7	10.7
9.6	10.3	10.4	10.4
10.2	10.1	10.5	10.3

a) ¿Los datos apoyan la afirmación de que la rapidez promedio de acción es la misma para ambas soluciones? Para obtener sus conclusiones utilice α=0.05. Calcule el valor P de la prueba.

b) Construya los boxplots comparativos de las dos muestras ¿Estas gráficas apoyan la hipótesis de que las varianzas son iguales? ¿Qué se concluye en relación a las medias? Haga una interpretación práctica de estas gráficas

c) Dé un intervalo de confianza del 95 % para la diferencia de medias e interprete.

12. Se investiga la temperatura de deflexión bajo carga para dos tipos diferentes de tubería de plástico. Para ello se toman dos muestras aleatorias, cada una de 15 especímenes y se registra las temperaturas de deflexión en F. Los resultados son los siguientes,

Tipo 1			Tipo 2		
206	193	192	177	176	198
188	207	210	197	185	188
205	185	194	206	200	189
187	189	178	201	197	203
194	213	205	180	192	192

a) ¿Los datos apoyan la afirmación de que la temperatura de deflexión bajo carga para la tubería tipo 2 es mayor que para la tubería de tipo 1? Utilice α=0.05.

b) Construya los boxplots comparativos de las dos muestras ¿Estas gráficas apoyan la hipótesis de que las varianzas son iguales? Haga una interpretación práctica para estas gráficas.

c) Suponga que la temperatura de deflexión media para la tubería de tipo 2 es mayor que la de la tubería de tipo 1 tanto como $5F$, entonces es importante detectar esta diferencia con una probabilidad de al menos 0.90. Para este problema, ¿Resulta adecuada la selección hecha de $n_1 = n_2 = n = 15$?

13. Se realizó un experimento para comparar los tiempos medios necesarios para que dos empleados A y B, completen el trámite de las cuentas corrientes personales para nuevos clientes. Se asignaron al azar diez clientes a cada empleado y

se registraron los tiempos de servicio para cada cliente. Los resultados en minutos fueron los siguientes:

Empleado A	Empleado B
27.64	26.75
17.01	29.16
28.97	25.24
19.01	22.97
23.24	30.72
27.07	37.61
27.62	34.62
16.38	31.93
23.99	27.40
25.70	28.36

a) Construya los boxplot comparativos. Obtenga conclusiones tentativas a partir de estas gráficas.

b) ¿Existe evidencia suficiente para indicar una diferencia entre los tiempos medios requeridos para completar los trámites necesarios de una cuenta corriente para un cliente nuevo? Use α =0.1.

c) Suponga que se quisiera estimar la diferencia en los tiempos medios de servicio con una exactitud dentro de 1 minuto, con una probabilidad aproximada del 95 %, suponiendo tamaños de muestra iguales para cada empleado, calcule n.

14. Quince hombre adultos, cuyas edades fluctúan entre 35 y 50 años, participan en un estudio para evaluar el efecto de la dieta y el ejercicio sobre los niveles de colesterol en la sangre. El colesterol total fue medido al inicio en cada sujeto y tres meses después de participar en un programa de ejercicio aeróbico y de haber cambiado a una dieta baja en grasas. Los datos aparecen en la siguiente tabla,

Sujeto	Antes	Después
1	265	229
2	240	231
3	258	227
4	295	240
5	251	238
6	245	241
7	287	234
8	314	256
9	260	247
10	279	239
11	283	246
12	240	218
13	238	219
14	225	226
15	247	233

a) Realice un análisis descriptivo apropiado de estos datos y dé una conclusión preliminar sobre el efecto de la dieta combinada con el ejercicio aeróbico.

b) ¿Cuáles son los tratamientos en este experimento? ¿Estos datos deben analizarse de forma pareada? Explique por qué.

c) ¿Apoyan los datos la hipótesis de que la dieta baja en grasa combinada con el ejercicio aeróbico son de gran valor en la disminución de los niveles de colesterol? Use α =0.05.

15. Diez individuos participan en un programa de modificación de dieta diseñado para estimular la pérdida de peso. En la tabla siguiente se indica el peso de cada participante antes y después de haber participados en el programa.

sujeto	Antes	Después
1	195	187
2	213	195
3	247	221
4	201	190
5	187	175
6	210	197
7	215	199
8	246	221
9	294	278
10	310	285

a) ¿Existe evidencia que apoye la afirmación de que este programa de modificación de dieta es eficaz para reducir el peso? Use α =0.05.

b) ¿Existe evidencia que apoye la afirmación de que este programa de modificación de dieta en particular reduce el peso promedio al menos 10 libras? Use α =0.05.

c) Suponga que si el programa de modificación de dieta da como resultado una pérdida de al menos 10 libras, entonces es importante detectar esto con una probabilidad de 0.90, ¿El uso de 10 sujetos será un tamaño de muestra adecuado? Si no es así ¿cuántos sujetos deben emplearse?

3. Código R para algunos análisis en los problemas con datos de la Sección 2

Código R 3.1. *Problema 3*

```
datos3=data.frame(sexo=factor(rep(c("Mujer","Hombre"),each=10)),TC=scan())
75 77 78 79 77 73 78 79 78 80
74 72 77 76 76 73 75 73 74 75

attach(datos3)

boxplot(TC~sexo,boxwex=0.5) #boxplots de las observaciones según factor sexo
var.test(TC~sexo,alternative="two.sided") #test de homogeneidad de varianza

#test t de igualdad de medias, muestras normales independientes, varianzas poblacionales iguales
t.test(TC~sexo,var.equal =TRUE,alternative="two.sided",paired=FALSE)

detach(datos3)
```

Código R 3.2. *Problema 4*

```
datos4=data.frame(Temperatura=factor(rep(c("baja","alta"),times=10)),encogimiento=scan())
17.2 21.4
17.5 20.9
18.6 19.8
15.9 20.4
16.4 20.6
17.3 21.0
16.8 20.8
18.4 19.9
16.7 21.1
17.6 20.3

attach(datos4)

boxplot(encogimiento~Temperatura,boxwex=0.5)

var.test(encogimiento~Temperatura,alternative="two.sided")
t.test(encogimiento~Temperatura,var.equal =TRUE,alternative="two.sided",paired=FALSE)

detach(datos4)
```

Código R 3.3. *Problema 5*

```
datos5=data.frame(Ruta=factor(rep(c("A","B"),each=5)),TV=scan())
18 24 30 21 32
22 29 34 25 35

attach(datos5)

boxplot(TV~Ruta,boxwex=0.5)
medias5=sapply(split(TV,Ruta),mean);medias5 #calcunado media muestral según factor ruta
var5=sapply(split(TV,Ruta),var);var5 #calculando varianza muestral según ruta

var.test(TV~Ruta,alternative="two.sided")
t.test(TV~Ruta,var.equal =TRUE,alternative="two.sided",paired=FALSE)

detach(datos5)
```

Código R 3.4. *Problema 6*

```
datos6=data.frame(Proveedor=factor(rep(1:2,each=14)),Diámetro=scan())
21.38 20.13 19.12 19.85 20.54 8.00 22.24
21.94 19.07 18.60 21.89 22.60 18.10 19.25
21.51 22.22 21.49 21.91 21.52 22.06 21.51
21.29 22.71 22.65 21.53 22.22 21.92 20.82

attach(datos6)

boxplot(Diámetro~Proveedor,boxwex=0.5)
abline(h=c(20.25-2.25,20.25,20.25+2.25),lty=2) #líneas horizontales pasando por las especificaciones de diámetros
var6=sapply(split(Diámetro,Proveedor),var);var6

var.test(Diámetro~Proveedor,alternative="two.sided") #¿varianzas iguales?
t.test(Diámetro~Proveedor,var.equal=FALSE,alternative="two.sided",paired=FALSE)

detach(datos6)
```

Código R 3.5. *Problema 7*

```
datos7=data.frame(Bola=factor(rep(c("X","Y"),each=10)),Diámetro=scan())
75 46 57 43 58 32 61 56 34 65
52 41 43 47 32 49 52 44 57 60

attach(datos7)

boxplot(Diámetro~Bola,boxwex=0.5) #¿será válido este análisis?
difer=datos7[,2][Bola=="X"]-datos7[,2][Bola=="Y"] #diferencias por pares
boxplot(difer,boxwex=0.5,xlab="Diferencias por pares") #distribución de la diferencia de pares de observaciones

#¿será válida la siguiente comparación de varianzas y medias según muestras normales independientes?
var.test(Diámetro~Bola,alternative="two.sided")
t.test(Diámetro~Bola,var.equal =TRUE,alternative="two.sided",paired=FALSE)

#Comparación de medias con muestras pareadas
t.test(Diámetro~Bola,var.equal =TRUE,alternative="two.sided",paired=TRUE)

detach(datos7)
```

Código R 3.6. *Problema 8*

```
datos8=data.frame(Aditivo=factor(rep(c("Sin Aditivo","Con Aditivo"),each=10)),Crecimiento=scan())
20 31 16 22 19 32 25 18 20 19
23 34 15 21 22 31 29 20 24 23

attach(datos8)

boxplot(Crecimiento~Aditivo,boxwex=0.5) #¿será válido este análisis?
difer=datos8[,2][Aditivo=="Con Aditivo"]-datos8[,2][Aditivo=="Sin Aditivo"] #diferencia por pares de observaciones
boxplot(difer,boxwex=0.5,xlab="Diferencias por pares") #distribución de las diferencias pareadas

#Test t para muestras pareadas
t.test(Crecimiento~Aditivo,var.equal =TRUE,alternative="two.sided",paired=TRUE)

detach(datos8)
```

Código R 3.7. *Problema 9*

```
datos9=data.frame(Técnico=factor(rep(c(1,2),each=8)),Pureza=scan())
74.0 73.1 73.5 73.9 71.2 72.5 73.0 74.3
73.0 71.3 73.2 71.1 70.3 71.5 73.4 72.4

attach(datos9)

boxplot(Pureza~Técnico,boxwex=0.5) #será válido este análisis?
difer=datos9[,2][Técnico==1]-datos9[,2][Técnico==2] #diferencias entre pares de observaciones
boxplot(difer,boxwex=0.5,xlab="Diferencias por pares") #Distribución de las diferencias pareadas
#Test de igualdad de medias de muestras pareadas ¿será apropiado?
t.test(Pureza~Técnico,var.equal =TRUE,alternative="two.sided",paired=TRUE)

#Test t de con H1:mu_d>0 con muestras pareadas
t.test(Pureza~Técnico,var.equal =TRUE,alternative="greater",paired=TRUE)

detach(datos9)
```

Código R 3.8. *Problema 10*

```
datos10=data.frame(Termómetro=factor(rep(c("Mer","Rtd"),times=25)),Temperatura=scan())
4.0 2.6 4.0 2.8 5.0 5.0 0.5 0.0 3.0 2.4
5.0 6.4 6.0 6.4 2.0 2.3 4.0 4.2 4.0 4.0
4.5 3.3 4.0 1.4 3.5 1.8 2.0 -1.9 3.0 -7.6
2.5 3.1 4.0 5.0 6.5 6.6 4.5 2.7 4.0 6.3
4.0 0.0 4.0 0.4 3.5 0.6 2.0 -4.0 4.0 -6.3

attach(datos10)

plot(datos10) #¿será apropiado este análisis?
difer=datos10[,2][Termómetro=="Mer"]-datos10[,2][Termómetro=="Rtd"]
boxplot(difer,boxwex=0.5,xlab="Diferencias por pares")

#Test t para comparar medias con muestras pareadas
t.test(Temperatura~Termómetro,var.equal=TRUE,alternative="two.sided",paired=TRUE)

detach(datos10)
```

Código R 3.9. *Problema 11*

```
datos11=data.frame(solución=factor(rep(c(1,1,2,2),times=5)),rapidez=scan())
9.9 10.6 10.2 10.0
9.4 10.3 10.6 10.2
9.3 10.0 10.7 10.7
9.6 10.3 10.4 10.4
10.2 10.1 10.5 10.3

attach(datos11)

#para análisis con muestras independientes
boxplot(rapidez~solución,boxwex=0.5)
var.test(rapidez~solución,alternative="two.sided")

#cuál de los siguientes dos tests es más apropiado?
#test t de igualdad de medias con muestras independientes de poblaciones
#normales con igual varianza
t.test(rapidez~solución,var.equal =TRUE,alternative="two.sided",paired=FALSE)

#test t de igualdad de medias con muestras independientes de poblaciones
#normales con diferente varianza
t.test(rapidez~solución,var.equal =FALSE,alternative="two.sided",paired=FALSE)

detach(datos11)
```

Código R 3.10. *Problema 12*

```
datos12=data.frame(Tipo.Tub=factor(rep(c(1,1,1,2,2,2),times=5)),Temperatura=scan())
206 193 192 177 176 198
188 207 210 197 185 188
205 185 194 206 200 189
187 189 178 201 197 203
194 213 205 180 192 192

attach(datos12)

boxplot(Temperatura~Tipo.Tub,boxwex=0.5)
medias12=sapply(split(Temperatura,Tipo.Tub),mean);medias12
var12=sapply(split(Temperatura,Tipo.Tub),var);var12

var.test(Temperatura~Tipo.Tub,alternative="two.sided")
t.test(Temperatura~Tipo.Tub,var.equal =TRUE,alternative="less",paired=FALSE)

#Cálculo de potencia  pedido
n=15
S2p=(n-1)*sum(var12)/(2*n-2) #estimación pooled de varianza común

ncp=-5*sqrt(15/(2*S2p)) #parámetro de no centraldad
pt(-qt(0.05,df=28,lower.tail=F),df=28,ncp=ncp)

#o bien
pot.t=power.t.test(n=n,delta=5,sd=sqrt(S2p),type = "two.sample",alternative ="one.sided")
pot.t

detach(datos12)
```

Código R 3.11. *Problema 13*

```
datos13=data.frame(Empleado=factor(rep(c("A","B"),times=10)),Tiempo=scan())
27.64 26.75
17.01 29.16
28.97 25.24
19.01 22.97
23.24 30.72
27.07 37.61
27.62 34.62
16.38 31.93
23.99 27.40
25.70 28.36

attach(datos13)

boxplot(Tiempo~Empleado,boxwex=0.5)
var.test(Tiempo~Empleado,alternative="two.sided")

#cuál de los siguientes dos tests es más apropiado?
#test t de igualdad de medias con muestras independientes de poblaciones
#normales con igual varianza
t.test(Tiempo~Empleado,var.equal =TRUE,alternative="two.sided",paired=FALSE)

#test t de igualdad de medias con muestras independientes de poblaciones
#normales con diferente varianza
t.test(Tiempo~Empleado,var.equal =FALSE,alternative="two.sided",paired=FALSE)

detach(datos13)
```

Código R 3.12. *Problema 14*

```
datos14=data.frame(Trat=factor(rep(c("Antes","Despues"),times=15)),Colesterol=scan())
265 229
240 231
258 227
295 240
251 238
245 241
287 234
314 256
260 247
279 239
283 246
240 218
238 219
225 226
247 233

attach(datos14)

boxplot(Colesterol~Trat,boxwex=0.5) #¿será apropiado este análisis?
difer=datos14[,2][Trat=="Antes"]-datos14[,2][Trat=="Despues"] #diferencia por pares de observaciones

boxplot(difer,boxwex=0.5,xlab="Diferencias por pares") #distribución diferencia de las observaciones pareadas

#¿cuál de las siguientes dos pruebas es más apropiada?
#test t de igualdad de medias con muestras pareadas
t.test(Colesterol~Trat,alternative="two.sided",paired=TRUE)

#test t para con H1:mu_d>0  con muestras pareadas
#mu_d es la media poblacional de las diferencias en observaciones pareadas "antes-después"
t.test(Colesterol~Trat,alternative="greater",paired=TRUE)

detach(datos14)
```

Código R 3.13. *Problema 15*

```
datos15=data.frame(Trat=factor(rep(c("Antes","Despues"),times=10)),Pérdida.Peso=scan())
195 187
213 195
247 221
201 190
187 175
210 197
215 199
246 221
294 278
310 285

attach(datos15)

boxplot(Pérdida.Peso~Trat,boxwex=0.5) #¿será apropiado este análisis?
difer=datos15[,2][Trat=="Antes"]-datos15[,2][Trat=="Despues"]
boxplot(difer,boxwex=0.5,xlab="Diferencias por pares")

#¿cuál de los siguientes dos tests es el apropiado?
#Test t con muestras pareadas con H0:mu_d<=0 vs. H1:mu_d>0
t.test(Pérdida.Peso~Trat,alternative="greater",paired=TRUE)

#test t con muestras pareadas conH0:mu_d>=10 vs.  H1:mu_d<10
t.test(Pérdida.Peso~Trat,alternative="less",paired=TRUE,mu=10)

#Cálculo de potencia test t con n=10 ¿cuál de los siguientes dos cálculos es apropiado para pregunta  c)?
power.t.test(n = 10, delta = 10, sd =sd(difer),sig.level = 0.05,type ="paired",alternative ="two.sided")

power.t.test(n = 10, delta = 10, sd =sd(difer),sig.level = 0.05,type ="paired",alternative ="one.sided")

detach(datos15)
```

Referencias

[1] Gutiérrez Pulido, H. y De la Vara Salazar, R. (2004). *Análisis y Diseño de Experimentos*. McGraw-Hill.

[2] Montgomery, D. C. (1994). *Probabilidad y Estadística Aplicada a Ingeniería*. McGraw-Hill.