

Problemas de Inferencia No Paramétrica

12 de mayo de 2024

1. En una explotación agrícola se quiere comparar el efecto de 4 dietas sobre la producción de leche. Para ello se seleccionan al azar 24 vacas y después de 4 semanas se anota la cantidad de leche producida. Se obtienen los siguientes valores para los rangos:

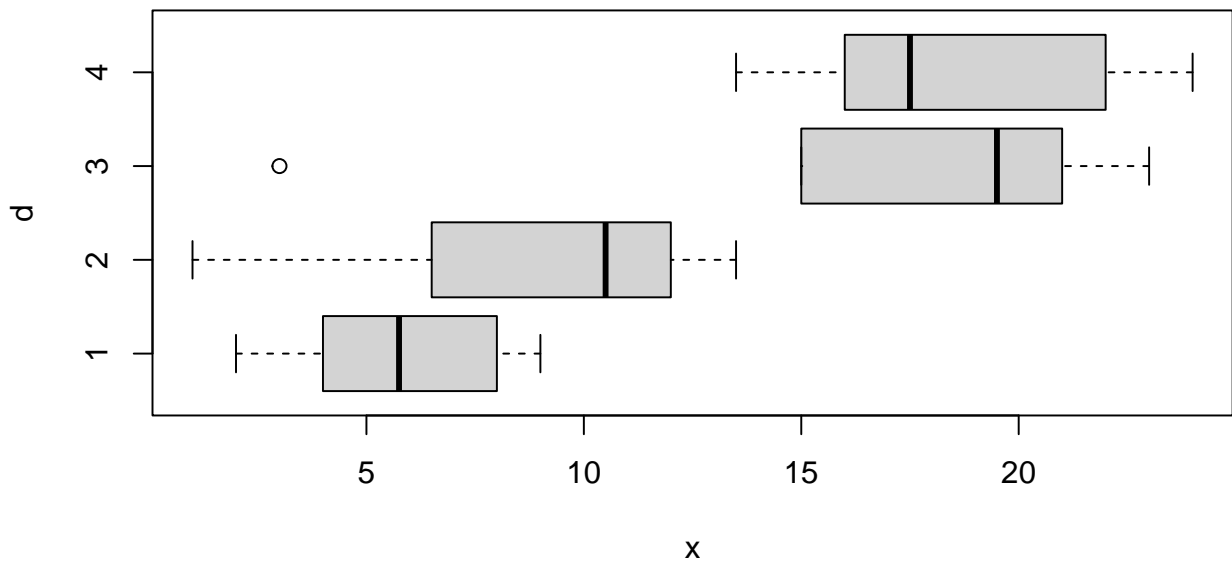
Dieta						
1	2	4	5	8	6,5	9
2	6,5	12	10	11	13,5	1
3	3	23	15	19	20	21
4	13,5	16	18	24	17	22

¿Qué conclusiones se obtienen sobre el efecto de las diferentes dietas?

Ayuda para R:

```
c ( 2, 4, 5, 8, 6.5, 9, 6.5, 12, 10, 11, 13.5, 1,
    3, 23, 15, 19, 20, 21, 13.5, 16, 18, 24, 17, 22)
```

```
x <- c ( 2, 4, 5, 8, 6.5, 9, 6.5, 12, 10, 11, 13.5, 1,
         3, 23, 15, 19, 20, 21, 13.5, 16, 18, 24, 17, 22)
d <- rep (1:4, each = 6)
boxplot (x ~ d, horizontal=TRUE)
```



A simple vista parece haber diferencias notables entre los efectos de las dos primeras dietas y las dos últimas, además de haber un valor atípico en la dieta 3 que, en principio, mantendremos tal cual. Para ver si es adecuado un enfoque paramétrico (análisis de varianza) haremos primeramente pruebas de gaussianidad:

```

by (x, d, shapiro.test)
## d: 1
##
## Shapiro-Wilk normality test
##
## data: dd[x, ]
## W = 0.97806, p-value = 0.9415
##
## -----
## d: 2
##
## Shapiro-Wilk normality test
##
## data: dd[x, ]
## W = 0.89469, p-value = 0.3435
##
## -----
## d: 3
##
## Shapiro-Wilk normality test
##
## data: dd[x, ]
## W = 0.80841, p-value = 0.06986
##
## -----
## d: 4
##
## Shapiro-Wilk normality test
##
## data: dd[x, ]
## W = 0.95336, p-value = 0.7674

```

Al nivel habitual de significación $\alpha = 0,05$ no se puede rechazar gaussianidad en ninguna muestra (y menos aún si se aplicase una corrección por contrastes múltiples, como Bonferroni o Holm). Probemos la homoscedasticidad:

```

by (x, d, sd)
## d: 1
## [1] 2.602883
## -----
## d: 2
## [1] 4.571652
## -----
## d: 3
## [1] 7.27782
## -----
## d: 4
## [1] 3.904058

bartlett.test (x ~ d)
##
## Bartlett test of homogeneity of variances
##
## data: x by d
## Bartlett's K-squared = 4.8954, df = 3, p-value = 0.1796

```

Aunque los desvíos típicos varían notablemente, con algunos duplicando a otros, no se detectan diferencias significativas. Procedemos a hacer un anova:

```
summary (aov (x ~ d))
##           Df Sum Sq Mean Sq F value Pr(>F)
## d           1  630.2   630.2   26.73 3.5e-05 ***
## Residuals   22  518.8    23.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

El p-valor es muy pequeño por lo que se confirma que hay diferencias significativas. Identifiquemos cuáles:

```
pairwise.t.test (x, d)
##
## Pairwise comparisons using t tests with pooled SD
##
## data:  x and d
##
##   1     2     3
## 2 0.5276 -     -
## 3 0.0042 0.0354 -
## 4 0.0014 0.0133 0.5816
##
## P value adjustment method: holm
do.call (rbind, by (x, d, summary))
##   Min. 1st Qu. Median   Mean 3rd Qu. Max.
## 1  2.0   4.250   5.75  5.75000   7.625  9.0
## 2  1.0   7.375  10.50  9.00000  11.750 13.5
## 3  3.0  16.000  19.50 16.83333  20.750 23.0
## 4 13.5  16.250  17.50 18.41667  21.000 24.0
```

Con la corrección de Holm y $\alpha = 0,05$ los grupos 1 y 2 no son significativamente distintos entre sí, pero sí respecto a los grupos 3 y 4, y viceversa. En principio¹ no sería necesario, pero vamos a comprobar si el resultado sería el mismo aplicando una prueba no paramétrica, en este caso, el anova no paramétrico (Kruskal-Wallis):

```
kruskal.test (x, d)
##
## Kruskal-Wallis rank sum test
##
## data:  x and d
## Kruskal-Wallis chi-squared = 13.403, df = 3, p-value = 0.003841
pairwise.wilcox.test (x, d)
## Warning in wilcox.test.default(xi, xj, paired = paired, ...): cannot compute exact p-value with ties
## Warning in wilcox.test.default(xi, xj, paired = paired, ...): cannot compute exact p-value with ties
##
## Pairwise comparisons using Wilcoxon rank sum test with continuity correction
##
## data:  x and d
##
##   1     2     3
## 2 0.298 -     -
## 3 0.165 0.165 -
## 4 0.013 0.032 1.000
##
## P value adjustment method: holm
```

En principio, el contraste de Wilcoxon no detecta diferencias significativas entre los grupos 2 y 3, pero el *warning* nos informa de que su implementación no tiene en cuenta los empates. Como la de Kruskal-Wallis sí, y es equivalente para contrastes bilaterales, probamos con esos grupos:

¹Por otro lado, ya que hemos observado un atípico, podría pensarse que no se rechaza la gaussianidad porque las muestras son demasiado pequeñas.

```

kruskal.test (x ~ d, subset (data.frame(x,d), d %in% c(2,3))) # ¡p-valor crudo!

##
## Kruskal-Wallis rank sum test
##
## data:  x by d
## Kruskal-Wallis chi-squared = 4.3333, df = 1, p-value = 0.03737

## p-valor corregido por Holm:
pvalores <-
  unlist (combn (4, 2,
                function (ij, i=ij[1], j=ij[2])
                  setNames (kruskal.test (x ~ d,
                                          subset (data.frame(x,d),
                                                    d %in% c(i,j))) $ p.value,
                              paste (i, j)),
                simplify = FALSE)) # para conservar setNames
p.adjust (pvalores) # corrección de Holm

##          1 2          1 3          1 4          2 3          2 4          3 4
## 0.25507542 0.14949195 0.02368651 0.14949195 0.02499062 1.00000000

```

El p-valor entre las dietas 2 y 3 ha bajado un poco respecto al obtenido mediante `wilcox.test` pero no cambia la conclusión. Téngase en cuenta también que estos p-valores se calculan mediante distribución asintótica.

En conclusión, los efectos de las cuatro dietas presentan diferencias significativas, en dos bloques: las dietas 1 y 2 tienen efectos menores que los de las 3 y 4. Se aconseja en todo caso intentar ampliar la muestra para cerciorarse porque una prueba no paramétrica potente como el contraste de Wilcoxon no detecta diferencias entre las dietas 2 y 3, que parecen pertenecer a bloques distintos.

- Para estudiar la efectividad de 3 drogas anticoagulantes en la disolución de coágulos, 10 individuos recibieron las 3 drogas (en orden aleatorio y con espacio de tiempo suficiente). Los datos siguientes corresponden al tiempo en segundos necesario para que un determinado tamaño de coágulo sea disuelto:

Indivi.	Antic. 1	Antic. 2	Antic. 3
1	127,5	129	135,5
2	130,6	129,1	138
3	118,3	111,7	110,1
4	155,5	144,3	162,3
5	180,7	174,4	181,1
6	176,7	173,4	182,1
7	28,5	131	136,5
8	123,6	128,1	140
9	128,3	112,7	113,1
10	153,5	142,3	161,3

¿Qué conclusiones pueden obtenerse a partir de estos datos?

Ayuda para R:

```

127.5, 129, 135.5, 130.6, 129.1, 138, 118.3, 111.7, 110.1, 155.5, 144.3, 162.3,
180.7, 174.4, 181.1, 176.7, 173.4, 182.1, 28.5, 131, 136.5, 123.6, 128.1, 140,
128.3, 112.7, 113.1, 153.5, 142.3, 161.3

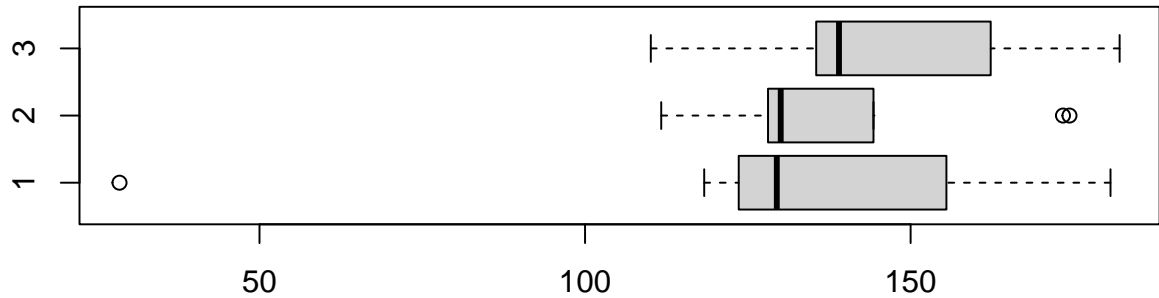
```

Representando los datos gráficamente:

```

x <- c(127.5, 129, 135.5, 130.6, 129.1, 138, 118.3, 111.7, 110.1, 155.5, 144.3, 162.3,
       180.7, 174.4, 181.1, 176.7, 173.4, 182.1, 28.5, 131, 136.5, 123.6, 128.1, 140,
       128.3, 112.7, 113.1, 153.5, 142.3, 161.3)
d <- t(matrix(x, 3))
boxplot(d, horizontal=TRUE)

```



No está claro si las diferencias observadas son² importantes; las dos primeras medianas son casi iguales, y la tercera es algo más alta. Para comprobarlo inferencialmente, al ser muestras dependientes podremos usar la prueba de Friedman o, en caso de haber gaussianidad, podríamos aplicar por parejas un contraste t para muestras pareadas.

```
friedman.test(d)
##
## Friedman rank sum test
##
## data: d
## Friedman chi-squared = 8.6, df = 2, p-value = 0.01357
```

El p-valor es menor que 0,05 por lo que, al nivel de significación habitual, rechazamos H_0 y sí hay diferencia significativa entre las drogas. Vamos a determinar entre cuáles. Probaremos con el contraste de Wilcoxon:

```
g <- rep(1:ncol(d), each=nrow(d)) # indica grupo
pairwise.wilcox.test(x, g, paired=TRUE)
##
## Pairwise comparisons using Wilcoxon signed rank exact test
##
## data: x and g
##
## 1 2
## 2 0.39 -
## 3 0.43 0.39
##
## P value adjustment method: holm
```

No se detectan diferencias. Podemos probar un contraste de signos, pero en general es mucho menos potente, así que a priori no esperamos resultados significativos:

```
prueba.signos <- function(x, y) binom.test(sum(x < y), length(x))$p.value
p.adjust(apply(combn(3,2), 2, function(ij) prueba.signos(d[,ij[1]], d[,ij[2]])))
## [1] 0.34375000 0.21875000 0.06445313
```

En efecto, ninguna pareja ofrece resultados significativos. ¿Podemos usar una prueba t pareada? Comprémoslo primero la gaussianidad de las diferencias:

²Hay un valor atípico en la primera muestra. Podríamos eliminar esa fila de la matriz de datos, o podríamos suponer que se olvidó el 1 de la centena que tiene el resto de datos. Aquí optaremos al inicio por dejarlo tal cual, ya que los métodos no paramétricos son robustos ante valores atípicos.

```
p.adjust(apply (combn(3,2), 2, function(ij) shapiro.test(d[,ij[1]]-d[,ij[2]])$p.value))
## [1] 3.076482e-05 1.618656e-04 6.210441e-01
```

Hay diferencias claramente no gaussianas. Comprobemos si se debe al valor atípico:

```
ati <- boxplot(d[,1], plot=FALSE) $ out
d[d[,1]==ati,1] <- NA
p.adjust(apply (combn(3,2), 2, function(ij) shapiro.test(d[,ij[1]]-d[,ij[2]])$p.value))
## [1] 1.0000000 0.6170623 1.0000000
```

Sin el atípico no hay problemas de gaussianidad, por lo que procedemos a usar pruebas t sin él:

```
pairwise.t.test(c(d), g, paired=TRUE)

##
## Pairwise comparisons using paired t tests
##
## data: c(d) and g
##
## 1 2
## 2 0.0688 -
## 3 0.3444 0.0093
##
## P value adjustment method: holm
```

Se detectan diferencias significativas al $\alpha = 0,05$ entre las drogas 2 y 3, lo que coincide con las diferencias entre medianas del diagrama de cajas inicial.

3. Se sospecha que el tinte T tiene un efecto negativo sobre la resistencia de ciertos tejidos. Para estudiarlo se proponen dos métodos estadísticos:

a) Seleccionar una muestra de tamaño n de piezas de tejidos, y partir cada pieza en dos mitades; aplicar a una de ellas el tinte dejando la otra como control; someter a los $2n$ trozos a una prueba de resistencia y anotar los resultados. En un estudio con $n = 11$ se ha obtenido:

tejidos	1	2	3	4	5	6	7	8	9	10	11
grupo tratado	13,8	16,1	14,9	15,2	15,9	16,0	15,8	16,1	16,3	16,5	15,5
grupo control	15,4	16,0	15,6	15,7	16,6	16,3	16,4	16,8	15,2	16,9	15,1

b) Seleccionar dos muestras independientes de tamaño n de piezas de tejidos, aplicar a una de las muestras el tinte dejando la otra como control y someter a todas las piezas de tejido a la prueba de resistencia anotando los resultados. Se lleva a cabo un estudio y los resultados obtenidos coinciden con las dos últimas filas de la tabla anterior.

Analiza el efecto del tinte en ambas circunstancias.

Ayuda para R:

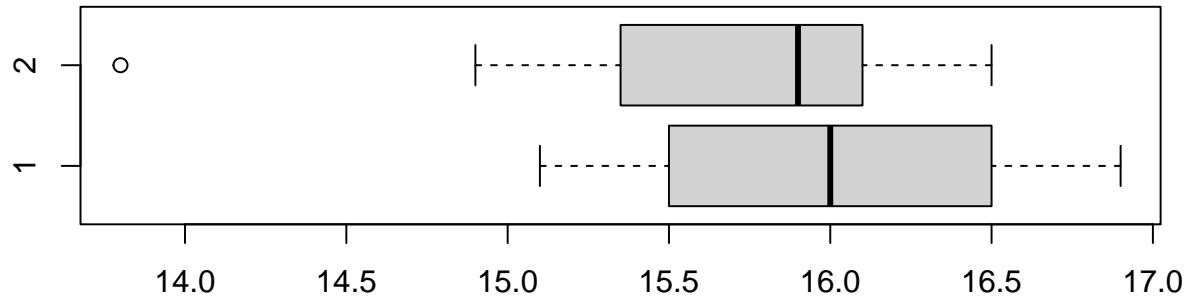
```
13.8,16.1,14.9,15.2,15.9,16.0,15.8,16.1,16.3,16.5,15.5
15.4,16.0,15.6,15.7,16.6,16.3,16.4,16.8,15.2,16.9,15.1
```

En primer lugar, describimos los datos:

```
tratado <- c(13.8,16.1,14.9,15.2,15.9,16.0,15.8,16.1,16.3,16.5,15.5)
control <- c(15.4,16.0,15.6,15.7,16.6,16.3,16.4,16.8,15.2,16.9,15.1)
summary(cbind(control, tratado))

## control tratado
## Min. :15.1 Min. :13.80
## 1st Qu.:15.5 1st Qu.:15.35
## Median :16.0 Median :15.90
## Mean :16.0 Mean :15.65
## 3rd Qu.:16.5 3rd Qu.:16.10
## Max. :16.9 Max. :16.50

boxplot(control, tratado, horizontal=TRUE)
```



Las medianas parecen indicar un pequeño descenso.

Consideraremos primero la primera circunstancia, en que las muestras son pareadas. Para ver si el descenso es significativo, comprobamos en primer lugar la gaussianidad de la diferencia:

```
shapiro.test(tratado - control)
##
## Shapiro-Wilk normality test
##
## data: tratado - control
## W = 0.92726, p-value = 0.3838
```

Como no se puede rechazar la gaussianidad, podemos contrastar

$$H_0 : \mu_{\text{tratado}} = \mu_{\text{control}} \quad H_1 : \mu_{\text{tratado}} < \mu_{\text{control}}$$

empleando una prueba t para muestras pareadas:

```
t.test(tratado, control, alternative="less", paired=TRUE)
##
## Paired t-test
##
## data: tratado and control
## t = -1.682, df = 10, p-value = 0.06174
## alternative hypothesis: true mean difference is less than 0
## 95 percent confidence interval:
##      -Inf 0.02749174
## sample estimates:
## mean difference
##      -0.3545455
```

La diferencia no es significativa al nivel habitual $\alpha = 0,05$.

En la segunda circunstancia, las hipótesis son las mismas,

$$H_0 : \mu_{\text{tratado}} = \mu_{\text{control}} \quad H_1 : \mu_{\text{tratado}} < \mu_{\text{control}}$$

pero las muestras son independientes, por lo que comprobaremos gaussianidad en cada muestra (y homoscedasticidad):

```
shapiro.test(tratado)
##
## Shapiro-Wilk normality test
```

```
##
## data:  tratado
## W = 0.87077, p-value = 0.07931
shapiro.test(control)
##
## Shapiro-Wilk normality test
##
## data:  control
## W = 0.93829, p-value = 0.5006
bartlett.test(c(control, tratado), rep(1:2, c(length(control), length(tratado))))
##
## Bartlett test of homogeneity of variances
##
## data:  c(control, tratado) and rep(1:2, c(length(control), length(tratado)))
## Bartlett's K-squared = 0.32487, df = 1, p-value = 0.5687
t.test(tratado, control, alternative="less", var.equal=TRUE)
##
## Two Sample t-test
##
## data:  tratado and control
## t = -1.1708, df = 20, p-value = 0.1277
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf  0.1677409
## sample estimates:
## mean of x mean of y
## 15.64545 16.00000
```

Tampoco hay diferencias significativas.

No es necesario, pero vamos a realizar los correspondientes contrastes no paramétricos:

```
wilcox.test(tratado, control, alternative="less", paired=TRUE)
## Warning in wilcox.test.default(tratado, control, alternative = "less", paired = TRUE): cannot compute exact p-value with ties
##
## Wilcoxon signed rank test with continuity correction
##
## data:  tratado and control
## V = 15, p-value = 0.05977
## alternative hypothesis: true location shift is less than 0
wilcox.test(tratado, control, alternative="less")
## Warning in wilcox.test.default(tratado, control, alternative = "less"): cannot compute exact p-value with ties
##
## Wilcoxon rank sum test with continuity correction
##
## data:  tratado and control
## W = 46.5, p-value = 0.1874
## alternative hypothesis: true location shift is less than 0
```

Los p-valores son mayores que $\alpha = 0,05$ y, por tanto, el resultado es el mismo: en ninguna circunstancia hay evidencias suficientes para afirmar que el tinte T tiene un efecto negativo.

4. En un estudio de comparación de funcionamiento de métodos automáticos de diagnóstico de enfermedades basados en síntomas, signos físicos e información de laboratorio con las opiniones médicas se estudiaron los resultados en 11 pacientes hipotiroideos. Estos resultados están recogidos en la tabla siguiente:

Paciente	D. médico	D. método 1	D. método 2	D. método 3
1	1	0	0	0
2	1	1	1	1
3	0	0	0	0
4	0	1	1	1
5	1	1	1	1
6	1	0	0	1
7	1	0	1	1
8	1	0	0	1
9	1	0	0	0
10	1	0	0	0
11	1	1	1	1

Analizar los datos anteriores.

Ayuda para R:

```
1,0,0,0, 1,1,1,1, 0,0,0,0, 0,1,1,1, 1,1,1,1, 1,0,0,1,
1,0,1,1, 1,0,0,1, 1,0,0,0, 1,0,0,0, 1,1,1,1
```

Empezamos describiendo las relaciones mediante correlaciones, dado que se trata de muestras relacionadas:

```
d <- t(matrix(c(1,0,0,0, 1,1,1,1, 0,0,0,0, 0,1,1,1, 1,1,1,1, 1,0,0,1,
                1,0,1,1, 1,0,0,1, 1,0,0,0, 1,0,0,0, 1,1,1,1),
              4))
cor(d)
##           [,1]      [,2]      [,3]      [,4]
## [1,] 1.00000000 -0.1336306 -0.04303315 0.1336306
## [2,] -0.13363062 1.0000000 0.82807867 0.5714286
## [3,] -0.04303315 0.8280787 1.00000000 0.6900656
## [4,] 0.13363062 0.5714286 0.69006556 1.0000000
summary(d)
##           V1           V2           V3           V4
## Min.      :0.0000   Min.      :0.0000   Min.      :0.0000   Min.      :0.0000
## 1st Qu.:1.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000
## Median :1.0000   Median :0.0000   Median :0.0000   Median :1.0000
## Mean    :0.8182   Mean    :0.3636   Mean    :0.4545   Mean    :0.6364
## 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000
## Max.    :1.0000   Max.    :1.0000   Max.    :1.0000   Max.    :1.0000
```

Los métodos tienen entre sí correlaciones positivas bastante altas, y casi cero respecto al diagnóstico médico. Respecto a sus distribuciones, el médico tiene la mayor proporción de unos, pero tan alejada del método que la tiene más alta como están las proporciones de los métodos entre sí. Para comprobar la igualdad de distribuciones en variables binarias como éstas, el contraste más adecuado sería la Q de Cochran pero, como no está en R básico, usaremos al efecto la prueba de Friedman con tratamiento de empates, que sí está:

```
friedman.test(d)
##
## Friedman rank sum test
##
## data: d
## Friedman chi-squared = 7.6957, df = 3, p-value = 0.05274
```

Al nivel de significación habitual $\alpha = 0,05$ no hay diferencias significativas ni entre los métodos ni entre éstos y el médico. Por tanto, no es necesario hacer un análisis a posteriori. Sin embargo, ya que el p-valor está próximo a α , lo haremos igualmente para cerciorarnos:

```
apply(combn(ncol(d),2), 2, function(ij) friedman.test(d[,ij]))
```

```

## [[1]]
##
## Friedman rank sum test
##
## data: d[, ij]
## Friedman chi-squared = 3.5714, df = 1, p-value = 0.05878
##
##
## [[2]]
##
## Friedman rank sum test
##
## data: d[, ij]
## Friedman chi-squared = 2.6667, df = 1, p-value = 0.1025
##
##
## [[3]]
##
## Friedman rank sum test
##
## data: d[, ij]
## Friedman chi-squared = 1, df = 1, p-value = 0.3173
##
##
## [[4]]
##
## Friedman rank sum test
##
## data: d[, ij]
## Friedman chi-squared = 1, df = 1, p-value = 0.3173
##
##
## [[5]]
##
## Friedman rank sum test
##
## data: d[, ij]
## Friedman chi-squared = 3, df = 1, p-value = 0.08326
##
##
## [[6]]
##
## Friedman rank sum test
##
## data: d[, ij]
## Friedman chi-squared = 2, df = 1, p-value = 0.1573

```

Ninguna diferencia es significativa, incluso sin corrección por contrastes múltiples.

5. En un estudio sobre la influencia de las condiciones de vida en la tensión diastólica, se selecciona al azar una muestra de adultos con edades comprendidas entre 30 y 50 años y de distintos países. Se obtienen los siguientes datos:

Tensión	50-60	60-70	70-80	80-90	90-100	100-110	110-160
Occidente	2	33	213	458	460	211	65
Tercer mundo	5	40	110	159	177	48	9

¿Qué conclusiones pueden obtenerse a partir de este estudio?

Ayuda para R:

"50-60", "60-70", "70-80", "80-90", "90-100", "100-110", "110-160"

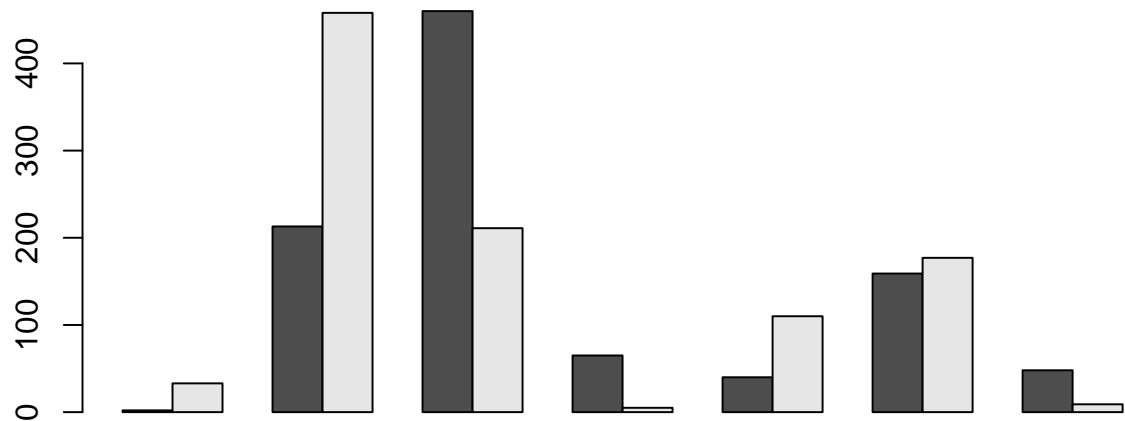
2,33,213,458,460,211,65,5,40,110,159,177,48,9

Hacemos una primera descripción mediante frecuencias relativas:

```
ten <- c("50-60", "60-70", "70-80", "80-90", "90-100", "100-110", "110-160")
fre <- matrix(c(2,33,213,458,460,211,65,5,40,110,159,177,48,9), 2)
round(100 * prop.table(fre, 1))

##      [,1] [,2] [,3] [,4] [,5] [,6] [,7]
## [1,]    0  22  47   7   4  16   5
## [2,]    3  46  21   0  11  18   1

barplot(fre, beside=TRUE)
```



Hay cierto parecido en las dos distribuciones, pero en las categorías segunda y tercera la diferencia es notable, aunque se podría sospechar que se debe a la agrupación en intervalos de la variable Tensión, que es continua. Haremos una primera inferencia mediante una prueba χ^2 de homogeneidad:

```
(ji2 <- chisq.test(fre))

##
## Pearson's Chi-squared test
##
## data:  fre
## X-squared = 320.95, df = 6, p-value < 2.2e-16

resid(ji2)

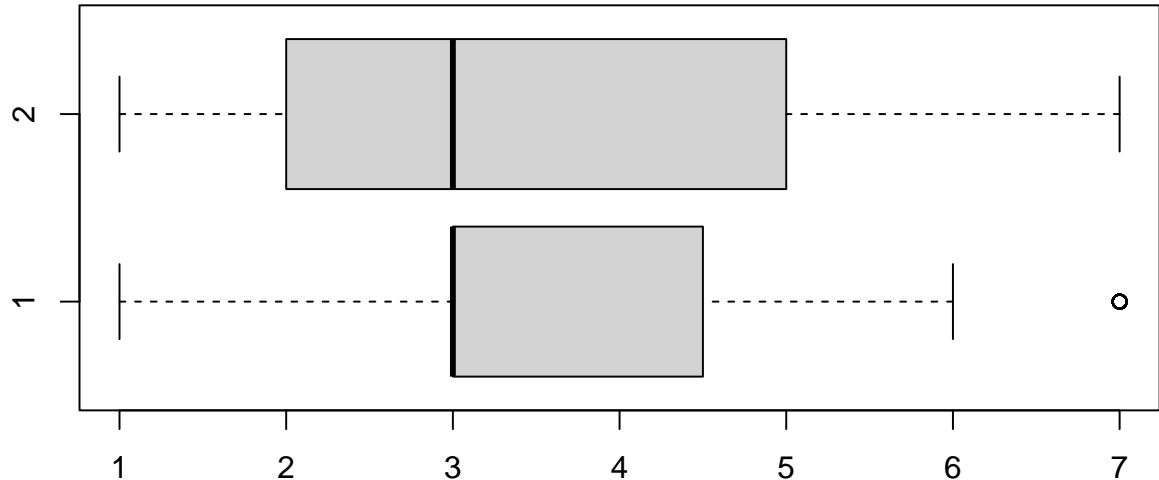
##      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]
## [1,] -3.686424 -6.567085  6.972447  5.139194 -3.987886 -0.5925390  3.710554
## [2,]  3.656902  6.514495 -6.916611 -5.098038  3.955950  0.5877939 -3.680840
```

Hay rechazo de la hipótesis de homogeneidad a cualquier nivel de significación y, como habíamos visto en la descripción, los mayores³ residuos corresponden con las categorías 2 y 3.

No es necesario, pero vamos a abordar el problema también considerando la variable Tensión como ordinal.

³En valor absoluto.

```
ten <- factor(ten, levels=ten) # levels para que no use orden alfabético
occidente <- as.numeric(rep(ten, fre[1,]))
tercermun <- as.numeric(rep(ten, fre[2,]))
boxplot(occidente, tercermun, horizontal=TRUE)
```



```
summary(occidente)
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      1.000  3.000   3.000   3.605  4.500   7.000

summary(tercermun)
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      1.000  2.000   3.000   3.267  5.000   7.000

wilcox.test(occidente, tercermun)
##
## Wilcoxon rank sum test with continuity correction
##
## data:  occidente and tercermun
## W = 598404, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

Los tamaños muestrales son tan grandes que, aunque la diferencia en tendencia central es pequeña, resulta significativa a cualquier α .

6. Se dispone de los siguientes datos sobre el número de horas de trabajo perdidas al año por 18 obreros correspondientes a 2 niveles profesionales:

Nivel A	39	21,5	38	24	30,5	44,5	38	29	35
Nivel B	22,5	21	22,5	21,5	23,5	21	23,5	18,5	14,5

- a) ¿Tiene influencia la categoría profesional en el número de horas de trabajo perdidas?

Ayuda para R:

39,21.5,38,24,30.5,44.5,38,29,35, 22.5,21,22.5,21.5,23.5,21,23.5,18.5,14.5

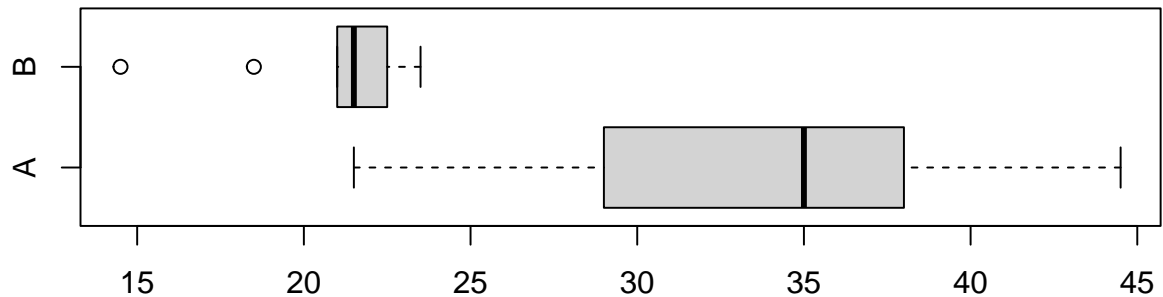
- b) Entre los obreros de la categoría A se efectuó un plan de prevención de accidentes laborales, y al año siguiente las horas perdidas por los mismos obreros son respectivamente:

24.5, 18, 31, 13, 20.5, 12.5, 31, 24.5, 21

Contrasta a nivel de significación 0,05 si el plan de prevención ha resultado efectivo.

En el primer apartado se trata de dos muestras independientes. Descriptivos:

```
x <- cbind(A=c(39,21.5,38,24,30.5,44.5,38,29,35),
          B=c(22.5,21,22.5,21.5,23.5,21,23.5,18.5,14.5))
boxplot(x, horizontal=TRUE)
```



Parece que hay una asociación clara entre categoría profesional y número de horas perdidas. Inferencialmente resolveremos el contraste:

$$H_0 : \mu_A = \mu_B \quad H_1 : \mu_A \neq \mu_B$$

```
(gaus <- apply(x, 2, shapiro.test))
## $A
##
## Shapiro-Wilk normality test
##
## data:  newX[, i]
## W = 0.95574, p-value = 0.7528
##
##
## $B
##
## Shapiro-Wilk normality test
##
## data:  newX[, i]
## W = 0.82407, p-value = 0.0383
p.adjust(sapply(gaus, function(g) g$p.value))
##           A           B
## 0.75281017 0.07659804
```

Uno de los p-valores queda en principio por debajo de $\alpha = 0,05$, pero no al aplicarle la corrección por contrastes múltiples, por lo que podemos suponer gaussianidad:

```
t.test(x[, "A"], x[, "B"])
##
```

```
## Welch Two Sample t-test
##
## data: x[, "A"] and x[, "B"]
## t = 4.5771, df = 10.254, p-value = 0.0009515
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 6.349524 18.317142
## sample estimates:
## mean of x mean of y
## 33.27778 20.94444

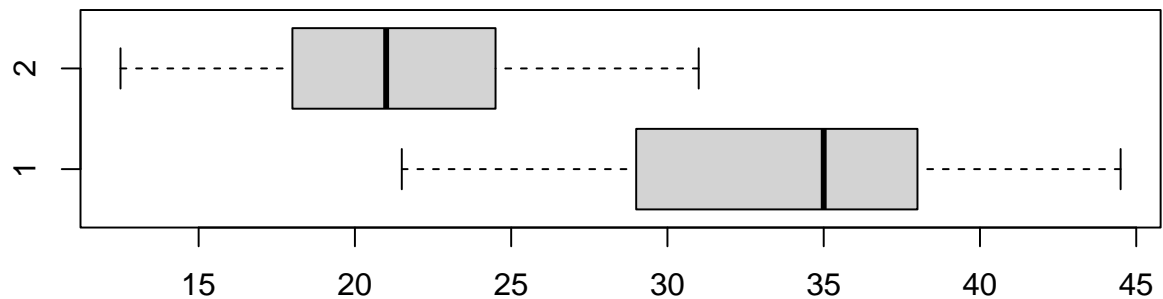
wilcox.test(x[, "A"], x[, "B"])

## Warning in wilcox.test.default(x[, "A"], x[, "B"]): cannot compute exact p-value with ties
##
## Wilcoxon rank sum test with continuity correction
##
## data: x[, "A"] and x[, "B"]
## W = 76.5, p-value = 0.001673
## alternative hypothesis: true location shift is not equal to 0
```

En ambos casos se rechaza H_0 con claridad, por lo que sí hay evidencias de asociación clara entre categoría profesional y número de horas perdidas. Los de nivel *A* pierden en promedio un 50% más de horas que los de nivel *B*.

En cuanto al último apartado, donde se recogen datos de los mismos individuos de nivel *A*, se trataría de una comparación entre muestras pareadas:

```
Apost <- c(24.5, 18, 31, 13, 20.5, 12.5, 31, 24.5, 21)
boxplot(x[, "A"], Apost, horizontal=TRUE)
```



```
shapiro.test(x[, "A"] - Apost)

##
## Shapiro-Wilk normality test
##
## data: x[, "A"] - Apost
## W = 0.80252, p-value = 0.02182

## se rechaza gaussianidad
wilcox.test(x[, "A"], Apost, paired=TRUE, alternative="greater")

## Warning in wilcox.test.default(x[, "A"], Apost, paired = TRUE, alternative = "greater"): cannot
compute exact p-value with ties
```

```
##  
## Wilcoxon signed rank test with continuity correction  
##  
## data: x[, "A"] and Apost  
## V = 45, p-value = 0.004545  
## alternative hypothesis: true location shift is greater than 0
```

En este caso se estaba contrastando si el plan de prevención ha sido efectivo, es decir,

$$H_0 : \mu_{Aante} = \mu_{Apost} \quad H_1 : \mu_{Aante} > \mu_{Apost}$$

por lo que se concluye que el plan de prevención ha sido efectivo.