

6. En un estudi per a determinar l'efectivitat de la vitamina B₁ per estimular el creixement dels fongs *Trichophyton album*, d'una mostra de 24 plaques de cultiu que contenien el fong, a 13 seleccionades a l'atzar els fou aplicada vitamina B₁ mentre que les restants serviren de control. El pes final dels fongs a cada placa va ser:

Control	18	14.5	13.5	12.5	23	24	21	17	18.5	9.5	14		
Vitamina B ₁	27	34	20.5	29.5	20	28	20	26.5	22	24.5	34	35.5	19

A un nivell 0.01, es pot afirmar que la vitamina B₁ incrementa significativament el creixement d'aquests fongs? Utilitza l'aproximació normal a partir dels següents càlculs R (els que siguin adequats):

```
> control = c(18, 14.5, 13.5, 12.5, 23, 24, 21, 17, 18.5, 9.5, 14)
> vitamina = c(27, 34, 20.5, 29.5, 20, 28, 20, 26.5, 22, 24.5, 34,
35.5, 19)
> mostra = c(control, vitamina)
> n1 = length(control)
> n2 = length(vitamina)
> N = n1 + n2
> n1 * (n1 + 1) / 2
[1] 66
> n2 * (n2 + 1) / 2
[1] 91
> N * (N + 1) / 2
[1] 300
> n1 * n2
[1] 143
> n1 * n2 * (n1 + n2 + 1)
[1] 3575
> n1 * n2 * (n1 + n2 - 1)
[1] 3289
> rangs = rank(mostra)
> rangs
[1] 7.0 5.0 3.0 2.0 15.0 16.0 13.0 6.0 8.0 1.0 4.0 19.0 22.5
12.0 21.0 10.5 20.0 10.5 18.0 14.0 17.0 22.5 24.0 9.0
> sum(rangs[1:n1])
[1] 80
>
```

És clarament una situació a la qual volem comparar les medianes a partir de dues mostres independents: les plaques control són diferents i totalment independents de les plaques amb vitamina B₁. Aplicarem doncs la prova de Mann-Whitney-Wilcoxon.

Primerament calcularem l'estadístic U.

$$U = \min \left\{ R_{1.} - \frac{n_1(n_1 + 1)}{2}, R_{2.} - \frac{n_2(n_2 + 1)}{2} \right\} = \min \{ 80 - 66, 220 - 91 \} = 14.$$

Fixem-nos que $1+2+\dots+N = \text{"suma de rangs"} = N(N+1)/2$ (no importa que hi hagi empats, per exemple 22,5 i 22,5 sumen 45, igual que 22+23). Per tant $R_{1.} = N(N+1)/2 - R_{2.} = 300 - 80 = 220$.

El contrast plantejat és unilateral, $H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 < \mu_2$ on μ_1 representa la mediana poblacional sota "control" i μ_2 la mediana poblacional sota "vitamina B₁".

L'enunciat demana que fem servir l'aproximació normal. Per tant, si no hi hagués empats, calcularíem:

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 - 1)}{12}}} = \frac{14 - \frac{143}{2}}{\sqrt{\frac{3575}{12}}} = -3,3314.$$

Com que volem fer una prova unilateral a un nivell $\alpha = 0,01$, hem de buscar el valor crític que deixaria 0,01 de probabilitat a una sola cua, és a dir, $2 \times 0,01 = 0,02$ en total a totes dues cues. A la taula de la distribució normal per dues cues aquest valor és 2,326347. Com que $|Z| = 3,3314 > 2,326347$ i la mitjana dels rangs del grup "control" és menor que el corresponent valor pel grup "vitamina B₁" podem rebutjar H_0 . En altres paraules afirmem que la vitamina B₁ té un efecte positiu en els creixement del fong.

En realitat hi ha empats. Observant el vector de rangs veiem que hi ha 2 sèries d'empats (10,5, 10,5 i 22,5, 22,5) totes dues de llargada 2. En altres paraules, $s = 2$, $t_1 = 2$, $t_2 = 2$. Per tant cal restar

$$\frac{n_1 n_2 \left(\sum_{i=1}^{s=2} (t_i^3 - t_i) \right)}{12 (n_1 + n_2) (n_1 + n_2 - 1)} = \frac{143 \left((2^3 - 2) + (2^3 - 2) \right)}{12 \times 24 \times 23} = 0,259057971$$

de l'expressió de la variància al denominador de Z:

$$Z = \frac{14 - \frac{143}{2}}{\sqrt{\frac{3575}{12} - 0,259}} = -3,3328$$

i lògicament la conclusió seria la mateixa.

7. Es va comparar l'efectivitat de la vitamina C en forma de suc de taronja amb l'àcid ascòrbic sintètic, en 20 porcs de Guinea repartits a l'atzar en 2 grups de 10, respecte de la llargada dels odontoblasts després de 6 setmanes de tractament. Els resultats foren:

Suc	8.2	9.4	9.6	9.7	10.0	14.5	15.2	16.1	17.6	21.5
A. ascòrbic	4.2	5.2	5.8	6.4	7.0	7.3	10.1	11.2	11.3	11.5

Determina amb un nivell de significació de 0,01, si el suc de taronja està associat a valors més grans de la llargada dels odontoblasts. Utilitza els càlculs R que siguin adequats (nota: l'ajuda de la funció 'wilcox.test' s'adjunta al final d'aquesta llista de problemes).

Novament tenim un disseny amb dues mostres independents i un contrast unilateral: $H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 > \mu_2$ on μ_1 representa la mediana poblacional sota "suc" i μ_2 la mediana poblacional sota "A. ascòrbic". Per tant una prova de rangs adient serà la de M-W-W.

```
> suc = c(8.2, 9.4, 9.6, 9.7, 10.0, 14.5, 15.2, 16.1, 17.6, 21.5)
> ascorbic = c(4.2, 5.2, 5.8, 6.4, 7.0, 7.3, 10.1, 11.2, 11.3, 11.5)
```

Examinem els diferents llistats que ens ofereix l'enunciat del problema:

```
> wilcox.test(suc, ascorbic)
```

Wilcoxon rank sum test

```
data: suc and ascorbic
W = 80, p-value = 0.02323
alternative hypothesis: true location shift is not equal to 0
```

El llistat anterior correspon a un test de M-W-W, però el descartem ja que correspon a una alternativa bilateral.

```
> wilcox.test(suc, ascorbic, paired = TRUE)
```

Wilcoxon signed rank test

```
data: suc and ascorbic
V = 55, p-value = 0.001953
alternative hypothesis: true location shift is not equal to 0
```

Aquest, en indicar 'paired = TRUE', correspondria a una prova de Wilcoxon dels rangs amb signe, per dades aparellades, i bilateral. Totalment inadequada per aquesta situació. El descartem.

```
> wilcox.test(suc, ascorbic, paired = TRUE, alternative = "less")
```

Wilcoxon signed rank test

data: suc and ascorbic

V = 55, p-value = 1

alternative hypothesis: true location shift is less than 0

Prova de Wilcoxon dels rangs amb signe, per dades aparellades (i amb una alternativa unilateral, la contrària a la nostra situació). El descartem.

```
> wilcox.test(suc, ascorbic, paired = TRUE, alternative = "greater")
```

Wilcoxon signed rank test

data: suc and ascorbic

V = 55, p-value = 0.0009766

alternative hypothesis: true location shift is greater than 0

Prova de Wilcoxon dels rangs amb signe, per dades aparellades (amb una alternativa unilateral en el mateix sentit que el nostre planteig). El descartem.

```
> wilcox.test(suc, ascorbic, alternative = "greater")
```

Wilcoxon rank sum test

data: suc and ascorbic

W = 80, p-value = 0.01162

alternative hypothesis: true location shift is greater than 0

Prova de M-W-W unilateral en el mateix sentit que hem plantejat.

Ens quedem amb aquest llistat. El p-valor és 0,01162 > 0,01, per tant no podem rebutjar H_0 si volem treballar sota el nivell de significació que demana l'enunciat del problema; no hem aconseguit demostrar la hipotètica superioritat del suc natural sobre l'àcid ascòrbic sintètic.

Recalquem que en cap cas hem demostrat que H_0 sigui certa (això no pot ser mai la conclusió d'un contrast d'hipòtesis) sinó que no hem obtingut prou evidències a favor de la hipòtesi alternativa. Si alguna cosa "es demostra" en certa manera en un contrast d'hipòtesis és la hipòtesi alternativa quan rebutgem la nul·la.

8. En un estudi sobre l'efecte que la familiaritat amb l'examinador pot tenir sobre el comportament d'un infant en un test d'intel·ligència, 48 alumnes de primària es van agrupar en 24 parelles molt semblants, i un infant de cada parella es va triar a l'atzar per formar part del grup experimental. Cada infant d'aquest grup va passar 20 minuts amb l'investigador, en un ambient distès, el qual va acabar passant-li el test. Amb el grup control es va realitzar exactament el mateix test, però sense cap preparatiu. Dins cada parella, les diferències "experimental – control" en la mesura del coeficient d'intel·ligència, IQ, van ser:

<i>Parella</i>	1	2	3	4	5	6	7	8	9	10	11	12
Diferència	-17	-15	-8	-7	-10	4	-10	4	12	1	21	2
<i>Parella</i>	13	14	15	16	17	18	19	20	21	22	23	24
Diferència	-20	2	0	5	-7	19	-8	9	-3	-7	-1	34

Escollint els llistats adequats, decideix si la familiaritat tendeix a millorar els resultats en el test, i determina una interval de confiança per a la mediana d'aquesta millora.

Aquest problema planteja una situació sota un disseny en dades aparellades. Si $D = \text{experimental} - \text{control}$, s'està plantejant un contrast unilateral $H_0: \mu_D = 0$ vs $H_1: \mu_D > 0$, on μ_D seria la mediana poblacional de D , pel qual la prova de Wilcoxon dels rangs amb signe seria adequada.

```
> parella = 1:24
> millora = c(-17,-15,-8,-7,-10,4,-10,4,12,1,21,2,-20,2,0,5,
-7,19,-8,9,-3,-7,-1,34)
> signe = factor(ifelse(millora < 0, "negatiu", "positiu"))
> datos = data.frame(signo = signe, mejora = millora)
> wilcox.test(parella, millora, conf.int = TRUE)
```

Wilcoxon rank sum test with continuity correction

```
data:  parella and millora
W = 478, p-value = 9.247e-05
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 8.000089 19.000010
sample estimates:
difference in location
13.99993
```

Mensajes de aviso perdidos

```
1: In wilcox.test.default(parella, millora, conf.int = TRUE) :
  cannot compute exact p-value with ties
2: In wilcox.test.default(parella, millora, conf.int = TRUE) :
  cannot compute exact confidence intervals with ties
```

Descartat, correspondria a una prova de M-W-W per a dues mostres independents, bilateral.

```
> wilcox.test(mejora ~ signo, data = datos, conf.int = TRUE)
```

Wilcoxon rank sum test with continuity correction

```
data:  mejora by signo
```

```

W = 0, p-value = 3.546e-05
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -24.99997 -10.99998
sample estimates:
difference in location
      -16.45679

```

Mensajes de aviso perdidos

```

1: In wilcox.test.default(x = c(-17, -15, -8, -7, -10, -10, -20, -7, :
   cannot compute exact p-value with ties
2: In wilcox.test.default(x = c(-17, -15, -8, -7, -10, -10, -20, -7, :
   cannot compute exact confidence intervals with ties

```

Descartat, aquesta crida a la funció R fa una cosa bastant absurda: compara, com a dos grups independents (test de M-W-W), totes les diferències negatives amb les diferències positives.

```

> wilcox.test(mejora ~ signo, data = datos, conf.int = TRUE,
alternative = "greater")

```

Wilcoxon rank sum test with continuity correction

```

data: mejora by signo
W = 0, p-value = 1
alternative hypothesis: true location shift is greater than 0
95 percent confidence interval:
 -22.00007      Inf
sample estimates:
difference in location
      -16.45679

```

Mensajes de aviso perdidos

```

1: In wilcox.test.default(x = c(-17, -15, -8, -7, -10, -10, -20, -7, :
   cannot compute exact p-value with ties
2: In wilcox.test.default(x = c(-17, -15, -8, -7, -10, -10, -20, -7, :
   cannot compute exact confidence intervals with ties

```

Descartat, igual que l'anterior, encara que fa una prova unilateral.

```

> wilcox.test(millora, conf.int = TRUE, alternative = "greater")

```

Wilcoxon signed rank test with continuity correction

```

data: millora
V = 125.5, p-value = 0.6538
alternative hypothesis: true location is greater than 0
95 percent confidence interval:
 -5.499995      Inf
sample estimates:
(pseudo)median
      -1.499954

```

Mensajes de aviso perdidos

```

1: In wilcox.test.default(millora, conf.int = TRUE, alternative = "greater") :
   cannot compute exact p-value with ties
2: In wilcox.test.default(millora, conf.int = TRUE, alternative = "greater") :
   cannot compute exact confidence interval with ties
3: In wilcox.test.default(millora, conf.int = TRUE, alternative = "greater") :
   cannot compute exact p-value with zeroes

```

```
4: In wilcox.test.default(millora, conf.int = TRUE, alternative = "greater") :  
  cannot compute exact confidence interval with zeroes
```

Aquesta crida a 'wilcox.test' és correcta. La prova dels rangs amb signe de Wilcoxon equival a estudiar una sola mostra, la de les diferències.

L'alternativa especificada correspon al planteig del problema.

Per tant, atès que el p-valor és $0,6538 > 0,05$ (com que l'enunciat no en diu res, considerem aquest nivell de significació) no podem rebutjar H_0 , no hem pogut demostrar que la familiaritat millori els resultats del test.

L'interval de confiança és $[-5,499995, +\infty)$ que no descarta el valor 0 per la mediana de les diferències. Fixeu-vos que és un interval unilateral, en correspondència amb el caràcter de la hipòtesi alternativa.

```
> wilcox.test(millora, conf.int = TRUE)
```

```
Wilcoxon signed rank test with continuity correction
```

```
data: millora
```

```
V = 125.5, p-value = 0.715
```

```
alternative hypothesis: true location is not equal to 0
```

```
95 percent confidence interval:
```

```
-6.500001  5.000026
```

```
sample estimates:
```

```
(pseudo)median
```

```
-1.499954
```

```
Mensajes de aviso perdidos
```

```
1: In wilcox.test.default(millora, conf.int = TRUE) :  
  cannot compute exact p-value with ties
```

```
2: In wilcox.test.default(millora, conf.int = TRUE) :  
  cannot compute exact confidence interval with ties
```

```
3: In wilcox.test.default(millora, conf.int = TRUE) :  
  cannot compute exact p-value with zeroes
```

```
4: In wilcox.test.default(millora, conf.int = TRUE) :  
  cannot compute exact confidence interval with zeroes
```

Aquest darrer llistat correspon també a una prova de Wilcoxon pels rangs amb signe, però bilateral. No correspon exactament a la situació plantejada.