Enunciat

1. D'una mostra de 9 rates, 5 es van seleccionar a l'atzar i van rebre entrenament consistent en imitar el comportament de rates l·líder", ja entrenades. La imitació permetia evitar una descàrrega elèctrica. Posteriorment es van barrejar amb les altres 4 rates no entrenades i, per cada rata, es va comptar el nombre d'intents necessaris fins a obtenir 10 respostes correctes seguides. Els resultats obtinguts van ser:

Entrenades 78 64 75 45 82

Control 110 70 53 51

Segons la prova de Mann-Whitney, i amb un nivell de significació de 0,05, determina si el nombre d'intents és significativament menor en el grup entrenat.

Solució

Aquest problema planteja una prova d'hipòtesis unilateral: $H_0 : \mu_{entrenades} = \mu_{control}$ vs $H_1 : \mu_{entrenades} < \mu_{control}$ on μ aquí correspondria a la mediana poblacional de la variable "nombre d'intents".

També es podria plantejar com $H_0 : \delta = 0$ vs $\delta < 0$ si $\delta = \mu_{entrenades} - \mu_{control}$.

Encara que es tracta d'una variable discreta, el concepte de rang està perfectament definit, la prova de Mann–Whitney–Wilcoxon seria apropiada tot i que hi hauria una probabilitat no nul·la d'empats (no presents a la mostra).

Intuïtivament, sembla que la hipòtesi alternativa que es planteja no s'hauria de poder demostrar ja que la simple observació de les medianes mostrals de cada grup:

```
> entrenades = c(78, 64, 75, 45, 82)
> control = c(110, 70, 53, 51)
> median(entrenades)

[1] 75

> median(control)

[1] 61.5
```

mostra que la mediana del grup "control" en realitat és inferior, i per tant la diferència no és negativa. Com ja sabem, el millor estimador de δ no és la diferència de medianes mostrals sinó la mediana mostral de totes les mútues diferencies:

```
> # Totes les diferències entre Entrenades i Control:
> d = outer(entrenades, control, "-")
> d
```

```
     [,1] [,2] [,3] [,4]
[1,]  -32    8   25   27
[2,]  -46   -6   11   13
[3,]  -35    5   22   24
[4,]  -65  -25   -8   -6
[5,]  -28   12   29   31
```

```
> # Estimació puntual de la diferència de medianes entre 'entrenades' i 'control':
> median(d)
```

```
[1] 6.5
```

que també és positiva, no suggereix res a favor de la hipòtesi alternativa.

La pregunta del problema es podria respondre d'una forma directa simplement fent:

```
> wilcox.test(entrenades, control, alternative = "less")
```

```
        Wilcoxon rank sum test

data:  entrenades and control
W = 11, p-value = 0.6349
alternative hypothesis: true location shift is less than 0
```

Atès el p-valor obtingut, no podem rebutjar H_0.

Aquest resultat seria compatible amb l'interval de confiança que ens proporcionaria la mateixa funció R, si li demanéssim:

```
> wilcox.test(entrenades, control, alternative = "less", conf.int = "TRUE")
```

```
        Wilcoxon rank sum test

data:  entrenades and control
W = 11, p-value = 0.6349
alternative hypothesis: true location shift is less than 0
95 percent confidence interval:
 -Inf   27
sample estimates:
difference in location
                   6.5
```

Veiem que l'interval de confiança 95% unilateral conté el valor zero i per tant no contradiu H_0. Intuïtivament, suggereix que δ és "com a màxim" 27.

L'extrem superior d'aquest interval s'hauria pogut obtenir "a mà" com el valor a la posició $\nu = nm - \lambda + 1$ del vector ordenat de totes les diferències abans obtingudes, $d_{(1)}, \dots, d_{(nm)}$ on $\lambda = u^*_{0.05}(n, m) + 1$, $u^*_{0.05}(n, m) = 2$ és el valor crític **unilateral** per un nivell de significació 0.05 de l'estadistic U de Mann–Witney (ara necessitem que tota la probabilitat 0.05 estigui acumulada a una cua) i $n = 5$, $m = 4$ són les mides mostrals:

```
> n = length(entrenades)
> m = length(control)
> d = sort(d)
> d

 [1] -65 -46 -35 -32 -28 -25  -8  -6  -6   5   8  11  12  13  22  24  25  27  29
[20]  31

> lambda = 2 + 1  # u*0.05(5,4) + 1
> nu = n * m - lambda + 1
> # Extrem superior de l'interval:
> d[nu]

[1] 27
```

L'interval de confiança anterior és el que correspondria al test unilateral realitzat. Si volguéssim obtenir l'interval de confiança bilateral caldria calcular $\lambda = u_{0.05}(n, m) + 1$ i $\nu = nm - \lambda + 1$ i l'interval correspondria a $[d_{(\lambda)}, d_{(\nu)}]$ on ara $u_{0.05}(n, m) = 1$ correspondria al valor crític bilateral (probabilitat 0.05 repartida entre totes dues cues simètriques):

```
> lambda = 1 + 1  # u0.05(5,4) + 1
> nu = n * m - lambda + 1
> # Interval de confiança:
> c(d[lambda], d[nu])

[1] -46  29
```

que lògicament ha de coincidir amb:

```
> wilcox.test(entrenades, control, alternative = "two.sided", conf.int = "TRUE")$conf.int

[1] -46  29
attr(,"conf.level")
[1] 0.95
```