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 “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.

2. Les següents dades corresponen als pesos (en lliures) que 12 voluntaris van ser capaços d’aixecar abans i després d’un programa de 8 setmanes d’entrenament muscular. Per un nivell de significació del 5%, determina si l’entrenament va millorar la capacitat per a aixecar pes, i quantifica aquesta millora en forma d’un interval de confiança al 95%.

Abans	14.4	15.9	14.4	13.9	16.6	17.4	18.6	20.4	20.4	15.4	15.4	14.1
Després	20.4	22.9	19.4	24.4	25.1	20.9	24.6	24.4	24.9	19.9	21.4	21.4

3. Volem comparar dues mostres independents de mides $n_1 = 2$ i $n_2 = 3$, respectivament, sense cap empat, mitjançant el test de Mann-Whitney. Volem fer una prova bilateral amb un nivell de significació 0,05, però ens trobem que per aquestes mides mostrals i aquest nivell de significació, el corresponent valor crític no apareix a la taula (comprova-ho). Determina el més petit nivell de significació pel qual podríem realitzar la prova, i el valor crític de l’estadístic U associat a aquest nivell de significació.

4. Per les mateixes mides mostrals anteriors, ara la nostra mostra presenta 2 valors empatats. Concretament, els valors que haurien de correspondre als rangs 3 i 4 estan empatats. Seguint el procediment habitual, aquests dos valors rebrien els rangs 3,5 i 3,5. El mínim nivell de significació aplicable i el valor crític trobats al problema anterior, són ara vàlids? (En altres paraules, una taula calculada pel cas sense empats, és vàlida en general quan hi ha empats?)

5. En un estudi sobre un nou mètode d’ensenyament de les matemàtiques elementals, es van seleccionar 10 nens d’un grup de 21 per a seguir el nou mètodes d’aprenentatge, mentre que els altres 11 serviren de control. Troba la probabilitat que, per pur atzar, el grup que seguí el nou mètode correspongués a les 10 millors qualificacions en matemàtiques.

6. En un estudi per a determinar l’efectivitat de la vitamina B_1 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 B_1 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 B_1 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
>
```

7. Es va comparar l'efectivitat de la vitamina C en forma de suc de taronja amb l'àcid ascòrbic sintètic, en 20 conillets d'índies 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):

```
> 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)
> 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

> 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

> 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

> 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

> 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
```

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:

Parella	*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
Parella	*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.

```
> 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
> 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
> 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
> 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
> 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
```

9. Se quiere saber si existen diferencias respecto del peso entre tres poblaciones del insecto *Tribolium castaneum* de diferentes localidades geográficas de las que se recogen muestras, obteniéndose los siguientes resultados:

Localidad 1 : 198, 186, 201, 190

Localidad 2: 200, 203, 205, 190, 187

Localidad 3: 215, 187, 212, 190

(peso expresado en mg).

Analiza estadísticamente las diferencias de los parámetros de localización entre localidades.

10. La següent taula mostra els pesos assolits per 25 rates que van créixer sota diferents dietes:

Dieta	Pes final							
A_1	257	205	206	164	190	214	228	203

A_2	201	231	197	185				
A_3	248	265	187	220	212	215	281	
A_4	202	276	205	204	230	227		

Rangs:

Dieta	rangs de Pes final							
A_1	1	4	8	10	11	14	18	22
A_2	2	5	6	20				
A_3	3	13	15	16	21	23	25	
A_4	7	9	12	17	19	24		

Atès que les sumes de rangs per dietes són 88, 33, 116 i 88 respectivament, determina si hi ha diferències significatives entre dietes.

11. Disposem d'una mostra de 8 unitats de paper procedents de cadascuna de 4 fàbriques de paper, A, B, C i D (per tant, un total de 32 unitats de paper). Amb un procediment físic determinem el grau de suavitat de cada una d'aquestes unitats, amb els següents resultats:

	unitat							
	1	2	3	4	5	6	7	8
A	38.7	41.5	43.8	44.5	45.5	46	47.7	58
B	39.2	39.3	39.7	41.4	41.8	42.9	43.3	45.8
C	34	35	39	40	43	43	44	45
D	34	34.8	34.8	35.4	37.2	37.8	41.2	42.8

a) Indica el disseny experimental seguit en aquest estudi. Es pot afirmar que la suavitat no és la mateixa segons la fàbrica de procedència?
b) Suposa ara que no tenim 32 unitats sinó solament 8, i que cada unitat passa, en ordre aleatori, per diversos laboratoris de certificació, A, B, C i D, cada un dels quals en mesura la suavitat. Indica el disseny experimental. Amb les mateixes dades anteriors, es pot afirmar que aquests laboratoris tendeixen a donar mesures de suavitat sistemàticament diferents?

12. Al problema 2 s'ha d'haver considerat un disseny en dades aparellades. Sota aquest disseny és d'esperar que les observacions de la variable d'interès (en aquest cas, capacitat per a aixecar pes) sota cada una de les condicions que es volen comparar (en aquest cas, abans i després de l'entrenament) siguin estocàsticament dependents. Sovint, aquesta dependència esperarem que sigui "positiva", en el sentit que si un individu ja tenia una capacitat gran (o petita) respecte de la resta per aixecar pes abans de l'entrenament, també sembla que l'hauria de tenir després de l'entrenament. Utilitzant els índexs de Kendall i de Spearman estudia si aquesta hipòtesi de dependència positiva queda confirmada per les dades del problema 2.

13. Al fitxer Law_School.txt hi ha dades sobre 82 facultats de dret dels Estats Units. Per cada facultat tenim la nota mitjana dels alumnes de nou ingrés per dos criteris utilitzats en el procés de selectivitat: LSAT i GPA. Analitza la possible dependència entre aquestes dues variables, utilitzant els índexs de Kendall i de Spearman.

14. Algú, utilitzant R però sense gaires coneixements d'estadística, ha aplicat (correctament) la funció 'wilcox.test' per comparar dues mostres independents, de mides $n_1 = 8$ i $n_2 = 7$ respectivament, sense empats. Com que li ha semblat que era el més important, només ha apuntat el valor de l'estadístic W, que ha estat 45. Per un nivell de significació 0,05 i utilitzant la taula de valors crítics per l'estadístic $T = min\{R_1 - n_1(n_1 + 1)/2, R_2 - n_2(n_2 + 1)/2\}$, determina si es pot afirmar que les medianes de les poblacions de les que procedeixen les mostres són diferents.

Wilcoxon Rank Sum and Signed Rank Tests

Description

Performs one- and two-sample Wilcoxon tests on vectors of data; the latter is also known as 'Mann-Whitney' test.

Usage

```
wilcox.test(x, ...)

## Default S3 method:
wilcox.test(x, y = NULL,
            alternative = c("two.sided", "less", "greater"),
            mu = 0, paired = FALSE, exact = NULL, correct = TRUE,
            conf.int = FALSE, conf.level = 0.95, ...)

## S3 method for class 'formula'
wilcox.test(formula, data, subset, na.action, ...)
```

Arguments

`x`	numeric vector of data values. Non-finite (e.g. infinite or missing) values will be omitted.
`y`	an optional numeric vector of data values: as with `x` non-finite values will be omitted.
`alternative`	a character string specifying the alternative hypothesis, must be one of `"two.sided"` (default), `"greater"` or `"less"`. You can specify just the initial letter.
`mu`	a number specifying an optional parameter used to form the null hypothesis. See 'Details'.
`paired`	a logical indicating whether you want a paired test.
`exact`	a logical indicating whether an exact p-value should be computed.
`correct`	a logical indicating whether to apply continuity correction in the normal approximation for the p-value.
`conf.int`	a logical indicating whether a confidence interval should be computed.
`conf.level`	confidence level of the interval.
`formula`	a formula of the form `lhs ~ rhs` where `lhs` is a numeric variable giving the data values and `rhs` a factor with two levels giving the corresponding groups.

`data`	an optional matrix or data frame (or similar: see `model.frame`) containing the variables in the formula `formula`. By default the variables are taken from `environment(formula)`.
`subset`	an optional vector specifying a subset of observations to be used.
`na.action`	a function which indicates what should happen when the data contain `NA`s. Defaults to `getOption("na.action")`.
`...`	further arguments to be passed to or from methods.

Details

The formula interface is only applicable for the 2-sample tests.

If only `x` is given, or if both `x` and `y` are given and `paired` is `TRUE`, a Wilcoxon signed rank test of the null that the distribution of `x` (in the one sample case) or of `x - y` (in the paired two sample case) is symmetric about `mu` is performed.

Otherwise, if both `x` and `y` are given and `paired` is `FALSE`, a Wilcoxon rank sum test (equivalent to the Mann-Whitney test: see the Note) is carried out. In this case, the null hypothesis is that the distributions of `x` and `y` differ by a location shift of `mu` and the alternative is that they differ by some other location shift (and the one-sided alternative `"greater"` is that `x` is shifted to the right of `y`).

By default (if `exact` is not specified), an exact p-value is computed if the samples contain less than 50 finite values and there are no ties. Otherwise, a normal approximation is used.

Optionally (if argument `conf.int` is true), a nonparametric confidence interval and an estimator for the pseudomedian (one-sample case) or for the difference of the location parameters `x-y` is computed. (The pseudomedian of a distribution F is the median of the distribution of $(u+v)/2$, where u and v are independent, each with distribution F. If F is symmetric, then the pseudomedian and median coincide. See Hollander & Wolfe (1973), page 34.) Note that in the two-sample case the estimator for the difference in location parameters does **not** estimate the difference in medians (a common misconception) but rather the median of the difference between a sample from `x` and a sample from `y`.

If exact p-values are available, an exact confidence interval is obtained by the algorithm described in Bauer (1972), and the Hodges-Lehmann estimator is employed. Otherwise, the returned confidence interval and point estimate are based on normal approximations. These are continuity-corrected for the interval but *not* the estimate (as the correction depends on the `alternative`).

With small samples it may not be possible to achieve very high confidence interval coverages. If this happens a warning will be given and an interval with lower coverage will be substituted.

Value

A list with class "htest" containing the following components:

statistic	the value of the test statistic with a name describing it.
parameter	the parameter(s) for the exact distribution of the test statistic.
p.value	the p-value for the test.
null.value	the location parameter mu.
alternative	a character string describing the alternative hypothesis.
method	the type of test applied.
data.name	a character string giving the names of the data.
conf.int	a confidence interval for the location parameter. (Only present if argument conf.int = TRUE.)
estimate	an estimate of the location parameter. (Only present if argument conf.int = TRUE.)

Warning

This function can use large amounts of memory and stack (and even crash **R** if the stack limit is exceeded) if exact = TRUE and one sample is large (several thousands or more).

Note

The literature is not unanimous about the definitions of the Wilcoxon rank sum and Mann-Whitney tests. The two most common definitions correspond to the sum of the ranks of the first sample with the minimum value subtracted or not: **R** subtracts and S-PLUS does not, giving a value which is larger by $m(m+1)/2$ for a first sample of size m. (It seems Wilcoxon's original paper used the unadjusted sum of the ranks but subsequent tables subtracted the minimum.)

R's value can also be computed as the number of all pairs (x[i], y[j]) for which y[j] is not greater than x[i], the most common definition of the Mann-Whitney test.

References

David F. Bauer (1972), Constructing confidence sets using rank statistics. *Journal of the American Statistical Association* **67**, 687–690.

Myles Hollander and Douglas A. Wolfe (1973), *Nonparametric Statistical Methods.* New York: John Wiley & Sons. Pages 27–33 (one-sample), 68–75 (two-sample). Or second edition (1999).

See Also

psignrank, pwilcox.

wilcox_test in package **coin** for exact, asymptotic and Monte Carlo *conditional* p-values, including in the presence of ties.

kruskal.test for testing homogeneity in location parameters in the case of two or more samples; t.test for an alternative under normality assumptions [or large samples]

Examples

```
require(graphics)
## One-sample test.
## Hollander & Wolfe (1973), 29f.
## Hamilton depression scale factor measurements in 9 patients with
##  mixed anxiety and depression, taken at the first (x) and second
##  (y) visit after initiation of a therapy (administration of a
##  tranquilizer).
x <- c(1.83,  0.50,  1.62,  2.48, 1.68, 1.88, 1.55, 3.06, 1.30)
y <- c(0.878, 0.647, 0.598, 2.05, 1.06, 1.29, 1.06, 3.14, 1.29)
wilcox.test(x, y, paired = TRUE, alternative = "greater")
wilcox.test(y - x, alternative = "less")    # The same.
wilcox.test(y - x, alternative = "less",
            exact = FALSE, correct = FALSE) # H&W large sample
                                            # approximation

## Two-sample test.
## Hollander & Wolfe (1973), 69f.
## Permeability constants of the human chorioamnion (a placental
##  membrane) at term (x) and between 12 to 26 weeks gestational
##  age (y).  The alternative of interest is greater permeability
##  of the human chorioamnion for the term pregnancy.
x <- c(0.80, 0.83, 1.89, 1.04, 1.45, 1.38, 1.91, 1.64, 0.73, 1.46)
y <- c(1.15, 0.88, 0.90, 0.74, 1.21)
wilcox.test(x, y, alternative = "g")        # greater
wilcox.test(x, y, alternative = "greater",
            exact = FALSE, correct = FALSE) # H&W large sample
                                            # approximation

wilcox.test(rnorm(10), rnorm(10, 2), conf.int = TRUE)

## Formula interface.
boxplot(Ozone ~ Month, data = airquality)
wilcox.test(Ozone ~ Month, data = airquality,
            subset = Month %in% c(5, 8))
```