

Зад. 1 Проведено е допитване с въпрос: приемате ли увеличаване на цената на цигарите с 25%, като начин за намаляване на тютюнопушенето. 351 от 605 непушачи са отговорили с да, докато при пушачите 71 от 195 отговарят с да. Може ли да се приеме, че мнението на пушачите и непушачите съвпада?

Реш:

	Непушачи	Пушачи
Да:	351	71
Не:	254	124
Общо	605	195

```
# H0: p1-p2=0
# HA: p1-p2<>0
# Двустранен тест
```

```
I н/н:
n1=605; n2=195
p1_hat=351/n1; p2_hat=71/n2
p_hat=(71+351) / (n1+n2)
z=((p1_hat-p2_hat)-0)/sqrt(p_hat*(1-p_hat)*(1/n1+1/n2))
z
# [1] 5.255544
```

```
#critical area
alpha <- 0.05
left=qnorm(alpha / 2, 0, 1); left
# [1] -1.959964
right=qnorm(1 - alpha/2, 0, 1); right
# [1] 1.959964
```

```
if(left <= z && z <= right) print("no evidence to reject H0") else print("reject H0 in favour of HA")
# [1] "reject H0 in favour of HA"
```

```
II начин:
res=prop.test(c(351, 71), c(605, 195))
if(res$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HA")
# [1] "reject H0 in favour of HA"
```

Зад. 2 Измерено е времето (в дни) за излекуване от дадена болест след прием на ново лекарство 15 10 13 7 9 8 21 9 14 8. Извършено е измерване и на контролна група приемащи плацебо: 15 14 12 8 14 10 7 16 10 15 12. Може ли да се приеме, че новото лекарство подобрява състоянието на пациентите?

Реш:

```
treatment=scan()
15 10 13 7 9 8 21 9 14 8
# double enter
placebo=scan()
15 14 12 8 14 10 7 16 10 15 12
# double enter
```

Първо проверяваме дали данните са нормално разпределени (за да знаем какво ще използваме за сравнение):

```
par(mfrow=c(2,2))
boxplot(treatment, horizontal=T, main="treatment")
boxplot(placebo, horizontal=T, main="placebo")
qqnorm(treatment); qqline(treatment)
qqnorm(placebo); qqline(placebo)
```

```
library(tseries); library(nortest)
```

```
sh1=shapiro.test(treatment); sh2=shapiro.test(placebo)
jb1=jarque.bera.test(treatment); jb2=jarque.bera.test(placebo)
ad1=ad.test(treatment); ad2=ad.test(placebo)
```

```
if(sh1$p.value && jb1$p.value && ad1$p.value) print("treatment is normal") else print("treatment is not normal")
# [1] "treatment is normal"
```

```
if(sh2$p.value > 0.05 && jb2$p.value > 0.05 && ad2$p.value > 0.05) print("placebo is normal") else
print("placebo is not normal")
# [1] "placebo is normal"
```

```
# Може допълнително да видим, че данните в опашките НЕ излизат от доверителните интервали за норм. разпределение по следния начин:
qqplot.das(treatment, "norm")
qqplot.das(placebo, "norm")
```

```
# Данните са с различна дължина, което автоматично ги прави независими
# Данните са нормално разпределени и не знаем тяхното стандартно отклонение (t-test)
```

```
# H0:  $\mu_1 - \mu_2 = 0$ 
# HA:  $\mu_1 - \mu_2 > 0$ 
# Едностраниен тест
res=t.test(treatment, placebo, alternative = "greater", conf.level=0.95)
if(res$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HA")
# [1] "no reason to reject H0"
```

Зад. 3 Сравняват се два радара за определяне на скоростта на автомобил. Направени са десет наблюдения, измерванията на първия са: 70 85 63 54 65 80 75 95 52 55, а на втория: 72 86 62 55 63 80 78 90 53 57. Да се провери дали двата радара са еднакви.

Реш:

```
radar1=scan()
70 85 63 54 65 80 75 95 52 55
# double enter
radar2=scan()
72 86 62 55 63 80 78 90 53 57
# double enter
```

Първо проверяваме дали данните са нормално разпределени (за да знаем какво ще използваме за сравнение):

```
par(mfrow=c(2,2))
boxplot(radar1, horizontal=T, main="radar1")
boxplot(radar2, horizontal=T, main="radar2")
qqnorm(radar1); qqline(radar1)
qqnorm(radar2); qqline(radar2)
```

```
library(tseries); library(nortest)
```

```
sh1=shapiro.test(radar1); sh2=shapiro.test(radar2)
jb1=jarque.bera.test(radar1); jb2=jarque.bera.test(radar2)
ad1=ad.test(radar1); ad2=ad.test(radar2)
```

```
if(sh1$p.value > 0.05 && jb1$p.value > 0.05 && ad1$p.value > 0.05) print("radar1 is normal") else
print("radar1 is not normal")
# [1] "radar1 is normal"
```

```
if(sh2$p.value > 0.05 && jb2$p.value > 0.05 && ad2$p.value > 0.05) print("radar2 is normal") else
print("radar2 is not normal")
# [1] "radar2 is normal"
```

Може допълнително да видим, че данните в опашките НЕ излизат от доверителните интервали за норм. разпределение по следния начин:

```
qqplot.das(radar1, "norm")
qqplot.das(radar2, "norm")
```

Радарите са измервали едни и същи обекти, следователно данните са зависими
Данните са нормално разпределени и не знаем тяхното стандартно отклонение (t-test)

```
# H0:  $\mu_1 - \mu_2 = 0$ 
# HА:  $\mu_1 - \mu_2 < > 0$ 
# Двустранен тест
```

```
res=t.test(radar1, radar2, alternative = "two.sided", paired=T, conf.level=0.95)
if(res$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HА")
```

Зад. 4 Разгледайте данните "ewr" от пакета "UsingR". Сравнете времето за напускане на летището от такси обслужващо компаниите "American airlines" и "Northwest airlines".

Реш:

```
X=ewr$AA  
Y=ewr$NW
```

Първо проверяваме дали данните са нормално разпределени (за да знаем какво ще използваме за сравнение):

```
par(mfrow=c(2,2))  
boxplot(X, horizontal=T, main="AA")  
boxplot(radar2, horizontal=T, main="NW")  
qqnorm(X); qqline(X)  
qqnorm(Y); qqline(Y)
```

```
library(tseries); library(nortest)
```

```
sh1=shapiro.test(X); sh2=shapiro.test(Y)  
jb1=jarque.bera.test(X); jb2=jarque.bera.test(Y)  
ad1=ad.test(X); ad2=ad.test(Y)
```

```
if(sh1$p.value > 0.05 && jb1$p.value > 0.05 && ad1$p.value > 0.05) print("AA is normal") else  
print("AA is not normal")  
# [1] "AA is not normal"
```

```
if(sh2$p.value > 0.05 && jb2$p.value > 0.05 && ad2$p.value > 0.05) print("NW is normal") else  
print("NW is not normal")  
# [1] "NW is not normal"
```

Може допълнително да видим, че данните в опашките излизат от доверителните интервали за норм. разпределение по следния начин:

```
qqplot.das(radar1, "norm")  
qqplot.das(radar2, "norm")
```

Данните не са нормално разпределени и не знаем тяхното стандартно отклонение. Може да направим Wilcoxon test за медианата.

```
# H0: me1-me2=0  
# HA: me1-me2<>0  
# Двустранен тест  
res=wilcox.test(X, Y, alternative = "two.sided", conf.level=0.95)  
if(res$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HA")
```

Зад. 6 Измервани са напреженията на пробив на диоди от две партии и са получени следните наблюдения: 39, 50, 61, 67, 40, 40, 54 за първата партида и 60, 53, 42, 41, 40, 54, 63, 69 за втората. Може ли да се приеме че диодите имат еднакво напрежение за пробив?

Реш:

```
diod1=c(39, 50, 61, 67, 40, 40, 54)
diod2=c(60, 53, 42, 41, 40, 54, 63, 69)
```

Първо проверяваме дали данните са нормално разпределени (за да знаем какво ще използваме за сравнение):

```
par(mfrow=c(2,2))
boxplot(diod1, horizontal=T, main="diod1")
boxplot(diod2, horizontal=T, main="diod2")
qqnorm(diod1); qqline(diod1)
qqnorm(diod2); qqline(diod2)
```

```
library(tseries); library(nortest)
```

```
sh1=shapiro.test(diod1); sh2=shapiro.test(diod2)
jb1=jarque.bera.test(diod1); jb2=jarque.bera.test(diod2)
```

Не може да направим Anderson-Darling test, тъй като за него са необходими поне 7 наблюдения, а за "diod1" имаме само 6.

```
if(sh1$p.value > 0.05 && jb1$p.value > 0.05) print("diod1 is normal") else print("diod1 is not normal")
# [1] "diod1 is normal"
if(sh2$p.value > 0.05 && jb2$p.value > 0.05) print("diod2 is normal") else print("diod2 is not normal")
# [1] "diod2 is normal"
```

Може допълнително да видим, че данните в опашките НЕ излизат от доверителните интервали за норм. разпределение по следния начин:

```
qqplot.das(diod1, "norm")
qqplot.das(diod2, "norm")
```

Наблюдаваните измервания са различен брой, което автоматично ги прави независими
Данните са нормално разпределени и не знаем тяхното стандартно отклонение (t-test)

```
# H0: mu1-mu2=0
# HA: mu1-mu2<>0
# Двустранен тест
res1=t.test(diod1, diod2, alternative = "two.sided", conf.level=0.95)
if(res1$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HA")
# [1] "no evidence to reject H0"
# следователно средните са равни
```

```
# H0: var1-var2=0
# HA: var1-var2<>0
# Двустранен тест
res2=var.test(diod1, diod2, alternative = "two.sided", conf.level=0.95)
if(res2$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HA")
# [1] "no evidence to reject H0"
# следователно вариациите също са равни
```

```
# Правим отново по-стриктен t-test, знаейки че вариациите са равни:
res3=t.test(diod1, diod2, alternative = "two.sided", var.equal=T, conf.level=0.95)
if(res3$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HA")
# [1] "no evidence to reject H0"
```

Зад. 11.1 Заредете в паметта данните "homework". Тя съдържа измервания за навиците на студенти от частни и публични гимназии. Направете boxplot един до друг на двете колекции от данни. Използвайте подходящ тест за проверка на равенство между центровете.

Реш:

```
X=homework$Private  
Y=homework$Public
```

Първо проверяваме дали данните са нормално разпределени (за да знаем какво ще използваме за сравнение):

```
par(mfrow=c(2,2))  
boxplot(X, horizontal=T, main="Private")  
boxplot(Y, horizontal=T, main="Public")  
qqnorm(X); qqline(X)  
qqnorm(Y); qqline(Y)
```

```
library(tseries); library(nortest)
```

```
sh1=shapiro.test(X); sh2=shapiro.test(Y)  
jb1=jarque.bera.test(X); jb2=jarque.bera.test(Y)  
ad1=ad.test(X); ad2=ad.test(Y)
```

```
if(sh1$p.value > 0.05 && jb1$p.value > 0.05 && ad1$p.value > 0.05) print("Private is normal") else  
print("Private is not normal")  
# "Private is normal"
```

```
if(sh2$p.value > 0.05 && jb2$p.value > 0.05 && ad2$p.value > 0.05) print("Public is normal") else  
print("Public is not normal")  
# "Public is normal"
```

Може допълнително да видим, че данните в опашките НЕ излизат от доверителните интервали за норм. разпределение по следния начин:

```
qqplot.das(X, "norm")  
qqplot.das(Y, "norm")
```

Данните са нормално разпределени и не знаем тяхното стандартно отклонение (t-test)

```
# H0: mu1-mu2=0  
# HA: mu1-mu2<>0  
# Двустранен тест  
res1=t.test(X, Y, alternative = "two.sided", conf.level=0.95)  
if(res1$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HA")  
# [1] "no evidence to reject H0"  
# следователно средните са равни
```

```
# H0: var1-var2=0  
# HA: var1-var2<>0  
# Двустранен тест  
res2=var.test(X, Y, alternative = "two.sided", conf.level=0.95)  
if(res2$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HA")  
# [1] "no evidence to reject H0"  
# следователно вариациите също са равни
```

Правим отново по-стриктен t-test, знаейки че вариациите са равни:

```
res3=t.test(X, Y, alternative = "two.sided", var.equal=T, conf.level=0.95)  
if(res3$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HA")  
# [1] "no evidence to reject H0"
```

Зад. 11.2 Заредете в паметта данните "corn". Дванадесет парцела земя са разделени на две и след това едната половина от всяка е засадена с ново семе на царевица, а другата със стандартно семе. Направете двустранен t-test на данните. Изглеждат ли изпълнени предположенията? Коментирайте защо съответстващия пробен тест е по-подходящ и след това извършете теста.

Реш:

X=homework\$New

Y=homework\$Standart

Първо проверяваме дали данните са нормално разпределени (за да знаем какво ще използваме за сравнение):

```
par(mfrow=c(2,2))
boxplot(X, horizontal=T, main="New")
boxplot(Y, horizontal=T, main="Standart")
qqnorm(X); qqline(X)
qqnorm(Y); qqline(Y)
```

```
library(tseries); library(nortest)
```

```
sh1=shapiro.test(X); sh2=shapiro.test(Y)
jb1=jarque.bera.test(X); jb2=jarque.bera.test(Y)
ad1=ad.test(X); ad2=ad.test(Y)
```

```
if(sh1$p.value > 0.05 && jb1$p.value > 0.05 && ad1$p.value > 0.05) print("New is normal") else
print("New is not normal")
# "Private is normal"
```

```
if(sh2$p.value > 0.05 && jb2$p.value > 0.05 && ad2$p.value > 0.05) print("Standart is normal")
else print("Standart is not normal")
# "Public is normal"
```

Може допълнително да видим, че данните в опашките НЕ излизат от доверителните интервали за норм. разпределение по следния начин:

```
qqplot.das(X, "norm")
qqplot.das(Y, "norm")
```

Данните са нормално разпределени и не знаем тяхното стандартно отклонение (t-test)

```
# H0:  $\mu_1 - \mu_2 = 0$ 
```

```
# HА:  $\mu_1 - \mu_2 < > 0$ 
```

```
# Двустранен тест
```

```
res1=t.test(X, Y, alternative = "two.sided", paired=True, conf.level=0.95)
```

```
if(res1$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HА")
```

```
# [1] "no evidence to reject H0"
```

```
# следователно средните са равни
```

Зад. 11.3 Заредете в паметта данните "blood". направете тест за съответствие на центровете. Кой тест използвахте и защо? Каква стойност получихте за p-value?

Реш:

```
X=blood$Machine  
Y=blood$Expert
```

Първо проверяваме дали данните са нормално разпределени (за да знаем какво ще използваме за сравнение):

```
par(mfrow=c(2,2))  
boxplot(X, horizontal=T, main="Machine")  
boxplot(Y, horizontal=T, main="Expert")  
qqnorm(X); qqline(X)  
qqnorm(Y); qqline(Y)
```

```
library(tseries); library(nortest)
```

```
sh1=shapiro.test(X); sh2=shapiro.test(Y)  
jb1=jarque.bera.test(X); jb2=jarque.bera.test(Y)  
ad1=ad.test(X); ad2=ad.test(Y)
```

```
if(sh1$p.value > 0.05 && jb1$p.value > 0.05 && ad1$p.value > 0.05) print("Machine is normal")  
else print("Machine is not normal")  
# "Machine is normal"
```

```
if(sh2$p.value > 0.05 && jb2$p.value > 0.05 && ad2$p.value > 0.05) print("Expert is normal") else  
print("Public is not normal")  
# "Expert is normal"
```

Може допълнително да видим, че данните в опашките НЕ излизат от доверителните интервали за норм. разпределение по следния начин:

```
qqplot.das(X, "norm")  
qqplot.das(Y, "norm")
```

Данните са нормално разпределени и не знаем тяхното стандартно отклонение (t-test)
Измерванията са направени на едни и същи обекти, следователно данните са зависими

```
# H0: mu1-mu2=0
```

```
# HA: mu1-mu2<>0
```

```
# Двустранен тест
```

```
res1=t.test(X, Y, alternative = "two.sided", paired=T, conf.level=0.95)
```

```
if(res1$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HA")
```

```
# [1] "no evidence to reject H0"
```

```
# следователно средните са равни
```

```
res1$p.value
```


Зад. 11.4 Направете тест за съответствие на медианите на данните "cabinet". Защо този тест ще е по-подходящ от теста за съответствие на средните или пък няма да е?

Реш:

```
X=cabinet$est.dividend.cg  
Y=cabinet$est.tax.savings
```

Първо проверяваме дали данните са нормално разпределени (за да знаем какво ще използваме за сравнение):

```
par(mfrow=c(2,2))  
boxplot(X, horizontal=T, main="Machine")  
boxplot(Y, horizontal=T, main="Expert")  
qqnorm(X); qqline(X)  
qqnorm(Y); qqline(Y)
```

```
library(tseries); library(nortest)
```

```
sh1=shapiro.test(X); sh2=shapiro.test(Y)  
jb1=jarque.bera.test(X); jb2=jarque.bera.test(Y)  
ad1=ad.test(X); ad2=ad.test(Y)
```

```
if(sh1$p.value > 0.05 && jb1$p.value > 0.05 && ad1$p.value > 0.05) print("est.dividend.cg is  
normal") else print("est.dividend.cg is not normal")  
# "est.dividend.cg is not normal"
```

```
if(sh2$p.value > 0.05 && jb2$p.value > 0.05 && ad2$p.value > 0.05) print("est.tax.savings is  
normal") else print("est.tax.savings is not normal")  
# "est.tax.savings is not normal"
```

Може допълнително да видим, че данните в опашките излизат от доверителните интервали за норм. разпределение по следния начин:

```
qqplot.das(X, "norm")  
qqplot.das(Y, "norm")
```

Данните НЕ са нормално разпределени и не знаем тяхното стандартно отклонение (Wilcoxon test)

```
# H0: me1-me2=0
```

```
# HA: me1-me2<>0
```

```
# Двустранен тест
```

```
res1=wilcox.test(X, Y, alternative = "two.sided", conf.level=0.95)
```

```
if(res1$p.value > 0.05) print("no evidence to reject H0") else print("reject H0 in favour of HA")
```

```
# [1] "reject H0 in favour of HA"
```

```
# следователно медианите не са равни
```