

---

Московский государственный университет имени М. В. Ломоносова  
Факультет Вычислительной математики и кибернетики

Никонов Максим Викторович  
316 группа

2020

---

**Задание 2:** Рассмотреть односторонние варианты t-теста, когда проверяемая нулевая гипотеза заключается в том, что одно из сравниваемых средних значений больше (или меньше) другого: использовать `t.test()` в сочетании с аргументом `alternative` (может принимать одно из трех значений – «two.sided» (по умолчанию), «greater» или «less»).

```
num <- random <- sample(1:2, nrow(head), replace=TRUE)
head <- cbind(head, num)
```

Так как у меня в `data` нет двухуровневой факторной переменной, сгенерируем ее с помощью функции `sample`

```
mean(data$Close)
tapply(data$Close, data$Volume, mean)
t.test(data$Close, mu = 44)
t.test(data$Close ~ num)
t.test(data$Close ~ num, var.equal = TRUE)
t.test(data$Close ~ num, alternative = "greater")
t.test(data$Close ~ num, alternative = "less")
t.test(data$Close ~ num, alternative = "two.sided")
```

Реализация.

Параметр 44 в функции `t.test` берётся из `mean`, данные из протокола:

```
> mean(data$Close)
[1] 43.44185
```

## Вывод.

```
> t.test(data$Close, mu = 44)

One Sample t-test

data: data$Close
t = -1.1227, df = 1509, p-value = 0.2618
alternative hypothesis: true mean is not equal to 44
95 percent confidence interval:
 42.46664 44.41706
sample estimates:
mean of x
 43.44185

> t.test(data$Close ~ num)

Welch Two Sample t-test

data: data$Close by num
t = 1.9231, df = 1505.6, p-value = 0.05466
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.03819567  3.85895413
sample estimates:
mean in group 1 mean in group 2
 44.39577      42.48539

> t.test(data$Close ~ num, var.equal = TRUE)

Two Sample t-test

data: data$Close by num
t = 1.923, df = 1508, p-value = 0.05467
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.0383039  3.8590624
sample estimates:
mean in group 1 mean in group 2
 44.39577      42.48539

> t.test(data$Close ~ num, alternative = "greater")

Welch Two Sample t-test

data: data$Close by num
t = 1.9231, df = 1505.6, p-value = 0.02733
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.2753922      Inf
sample estimates:
mean in group 1 mean in group 2
 44.39577      42.48539

> t.test(data$Close ~ num, alternative = "less")

Welch Two Sample t-test

data: data$Close by num
t = 1.9231, df = 1505.6, p-value = 0.9727
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf 3.545366
sample estimates:
mean in group 1 mean in group 2
 44.39577      42.48539

> t.test(data$Close ~ num, alternative = "two.sided")

Welch Two Sample t-test

data: data$Close by num
t = 1.9231, df = 1505.6, p-value = 0.05466
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.03819567  3.85895413
sample estimates:
mean in group 1 mean in group 2
 44.39577      42.48539
```

Используемые packages. dplyr, HSAUR2, nortest

Тестирование. Не требует

Неразрешенные вопросы. Нет

Новые функции. t.test

Статус компиляции. ОК. Данные из протокола: см.ниже

**Задание 3: Найти выражения для критериев Фишера, Левене, Бартлетта, Флигнера-Килина и изучить проверку гипотез об однородности дисперсий в R ( на примерах).**

Фишер

```
var.test(data$Close ~ num, alternative = c("two.sided"), conf.level = 0.95)
```

Левене

```
leveneTest(data$Close ~ num, data, location="median")
```

Бартлетт

```
bartlett.test(data$Close ~ num)
```

Фригнер-Килин

```
fligner.test(data$Close ~ num, data = InsectSprays)
```

Проверка гипотез об однородности дисперсий

```
aov(data$Close ~ num, data)  
lm(data$Close ~ num)
```

```
> var.test(data$Close ~ num, alternative = c("two.sided"), conf.level = 0.95)
```

F test to compare two variances

```
data: data$Close by num
F = 1.0896, num df = 755, denom df = 753, p-value = 0.2389
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.9445559 1.2569480
sample estimates:
ratio of variances
 1.089621

> leveneTest(data$Close ~ num, data, location="median")
Levene's Test for Homogeneity of Variance (center = median: "median")
  Df F value Pr(>F)
group 1  2.501 0.114
    1508

> bartlett.test(data$Close ~ num)
```

Bartlett test of homogeneity of variances

```
data: data$Close by num
Bartlett's K-squared = 1.3872, df = 1, p-value = 0.2389
```

```
> fligner.test(data$Close ~ num, data = InsectSprays)
```

Fligner-Killeen test of homogeneity of variances

```
data: data$Close by num
Fligner-Killeen:med chi-squared = 4.5208, df = 1, p-value = 0.03348
```

```
> aov(data$Close ~ num, data)
Call:
aov(formula = data$Close ~ num, data = data)
```

Terms:

```
      num Residuals
Sum of Squares 1377.7 561831.9
Deg. of Freedom    1    1508
```

```
Residual standard error: 19.30201
Estimated effects may be unbalanced
> lm(data$Close ~ num)
```

```
Call:
lm(formula = data$Close ~ num)
```

```
Coefficients:
(Intercept)      num
      46.31      -1.91
```

**Используемые packages.** dplyr, HSAUR2, car

**Тестирование.** Не требует

**Неразрешенные вопросы.** Нет

**Новые функции.** var.test, leveneTest, bartlett.test, fligner.test, aov

**Статус компиляции.** ОК. Данные из протокола:

```
>
[1m[7m%[27m[1m[0m
]7;file://MBP-Nikon.Dlink/Users/Nikon/Desktop/CMC%20MSU/MC/5%20sem/R/tz7
[0m[27m[24m[JNikon@MBP-Nikon tz7 % [K[?2004heexit[?2004l
```

Script done on Tue Oct 27 21:27:15 2020