
ЛАБОРАТОРНАЯ РАБОТА №5

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Критерий Колмогорова-Смирнова

Теория

Напомним суть проблемы.

Критерий Колмогорова-Смирнова проверяет гипотезу о том, что две выборки извлечены из одного и того же теоретического распределения. Так, если выборка X_1, \dots, X_m имеет функцию распределения $F_x(t)$, а выборка Y_1, \dots, Y_n имеет функцию распределения $F_y(t)$, то $H_0 : F_x(t) = F_y(t)$ для всех $t \in \mathbb{R}$, а конкурирующая гипотеза может быть сформулирована одним из способов: $H_1 : F_x(t) < F_y(t)$, $H_1 : F_x(t) > F_y(t)$ или $H_1 : F_x(t) \neq F_y(t)$ для всех $t \in \mathbb{R}$.

Синтаксис команды (разложен на строки для удобства описания):

```
ks.test(  
  x,  
  y,  
  alternative=c("two.sided","less","greater"),  
  ...)
```

`x` вектор-выборка, распределённая по закону с непрерывной функцией распределения.

`y` вектор-выборка, распределённая по закону с непрерывной функцией распределения; также тут может быть строка символов, именуемая кумулятивную функцию некоторого теоретического распределения.

`alternative=c("two.sided","less","greater")` вид альтернативной гипотезы.

Задания

Задание. Изучите работу с тестом Колмогорова-Смирнова на материале из прилагаемого файла `ks.xlsx` (оба листа, по всем альтернативным гипотезам). Постройте эмпирические кумулятивные функции распределения по каждой выборке с помощью функции `ecdf()`. Сделайте так, чтобы графики кумулятивных функций пар исследуемых выборок располагались на том же рисунке. Вам может помочь функция `lines()`.